IDENTIFICATION AND EVALUATION OF YELLOW-BILLED CUCKOO HABITAT USING ACOUSTIC MONITORING AND SPECIES DISTRIBUTION MODEL METHODS

By Nicholas Beauregard

A Dissertation

Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

in Biology

Northern Arizona University

December 2023

Approved:

Tad Theimer, Ph.D., Chair

Bret Pasch, Ph.D.

Jeffrey Foster, Ph.D.

Susan Sferra, M.Sc.

ABSTRACT

IDENTIFICATION AND EVALUATION OF YELLOW-BILLED CUCKOO HABITAT USING ACOUSTIC MONITORING AND SPECIES DISTRIBUTION MODEL METHODS

NICHOLAS BEAUREGARD

The western distinct population segment of the Yellow-billed Cuckoo (cuckoo) was federally listed as threatened in 2014. Long considered riparian obligate, populations have declined range-wide as riparian habitat has been lost or degraded due to conversion to agriculture, dams, or other development. Recent surveys, however, have documented cuckoos in southeastern Arizona using xeroriparian vegetation in ephemeral and intermittent drainages not previously known to be occupied. We (Northern Arizona University, in conjunction with U.S. Fish and Wildlife Service and U.S. Geological Survey) confirmed breeding in 23 of 24 sites that were classified as occupied according to standardized playback surveys, providing support for the use of established survey methods to estimate breeding status in xeroriparian habitat. Combining our data with previously collected data, we produced a novel map of cuckoo distribution in southeast Arizona which includes 100 occupied sites in xeroriparian habitat. These results indicate cuckoos in xeroriparian habitat represent a regionally significant and previously undescribed population. We also developed a species distribution model (SDM) for cuckoos in southern Arizona and southwestern New Mexico using topographic, climatic, and phenological variables. Our resulting model had an AUC score of 0.971, indicating high model performance. Evaluation of variable contribution suggested cuckoos are associated with warm drainages that experience high productivity during summer monsoon season. The SDM output map highlighted a broader spectrum of riparian and xeroriparian habitat than previous models, including key areas lacking survey data. Finally, we evaluated methods for the use of autonomous recording units (ARUs) to identify and monitor cuckoo habitat. Our semi-autonomous classifier for detecting cuckoo calls had a recall score of 0.33, precision score of 0.06, and an F-score ($\beta = 1$) of 0.102, consistent with studies of other cryptic birds. We demonstrated acoustic monitoring can determine occupancy rates comparable to traditional survey protocols with a minimum of 2 hours of daily acoustic recording. Our results provide important new information on cuckoo habitat associations, distribution, and monitoring methods which may inform conservation and management decisions, while our ARU methods provide an important new approach for increasing survey coverage.

ACKNOWLEDGEMENTS

- Dr. Tad Theimer. My advisor and mentor, who I am fortunate to have crossed paths with and walked alongside for the past several years. His perspectives and intellect never cease to inspire me, and his guidance and support has been invaluable.
- All of my committee members: Susan Sferra, Jeff Foster, and Bret Pasch. All of whom have provided me with unique perspectives, guidance, and challenges to help shape a rich and rewarding graduate school experience.
- The many other scientists, conservation biologists, and colleagues who have supported this work. In particular Charles Drost, Angela Dahlby, Don Swan, Edwin Juarez, Tracy McCarthy, Jennie MacFarland, Jonathan Horst, Steve Prager, Cathy Wise, Shannon McNeil, Diane Tracy, John Stanek, Mary Whitfield, and every one of the many volunteers who has contributed time in field collecting data.
- My parents Laura and David, and my sister Megan. I am grateful for their support, encouragement, and patience as I labored for several years on an obscure project thousands of miles away from home.
- Ashlee Wolf. I am grateful for her unwavering support and patience, and for her unique perspectives on the natural world that have helped shape my own over the past eight years.
- I am deeply grateful for each of the scholarships, grants, and awards I have received, which have enabled me to work towards my goals as a conservation biologist. In particular, the USGS-USFWS Science Support Partnership (SSP) grant, which initiated my research.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

DEDICATION

This dissertation is dedicated to the memory of Glen E. "Gooch" Goodwin.

You are always with me in the mountains.

Chapter 1 Breeding by western Yellow-billed Cuckoos in a novel and important habitat in southeastern Arizona

ABSTRACT

The identification of occupied habitat is an important component of recovery efforts for threatened and endangered species. The western population of the Yellow-billed Cuckoo (*Coccyzus americanus*) is federally listed as a threatened distinct population segment and has long been considered riparian-obligate, yet recent survey efforts in southeastern Arizona have documented cuckoos occurring in xeroriparian habitat during the nesting season. If these cuckoos are breeding, birds in xeroriparian habitat could represent a substantial addition to the species' numbers in western North America. We (Northern Arizona University, in conjunction with U.S. Fish and Wildlife Service and U.S. Geological Survey) investigated the distribution and breeding status of cuckoos in southeastern Arizona xeroriparian habitat by comparing the results of standardized call-playback surveys to the results of nest searching efforts in the same sites from 2018 to 2020. We then used this information to interpret more extensive survey data from 2013 to 2020 and develop an updated breeding distribution map for southern Arizona. We confirmed breeding in 94% of sites categorized as occupied according to survey results, confirming regular breeding in these xeroriparian habitats. Combining our data with previous survey data, we estimated a minimum of 100 occupied areas in southeastern Arizona xeroriparian habitat, representing a substantial increase in the known breeding population in Arizona. Occupied sites were concentrated in southern and western “Sky Island” mountain and foothill drainages, from 600–1800 m, with xeroriparian vegetation variously surrounded by Madrean-evergreen woodland, semi-desert grassland, or desert scrub. Cuckoos breeding in

southeast Arizona xeroriparian habitat may be important for conservation efforts, but face potential threats from grazing, climate change, and development.

INTRODUCTION

Accurately assessing population distribution, numbers, and habitat associations are important for the management and recovery of threatened and endangered species (Joseph et al. 2006, Hughes 2015, Camaclang et al. 2015). Obtaining accurate and complete survey data may be challenging for rare or cryptic species, however, due to low or unknown detection probability (Gu & Swihart 2004, MacKenzie 2005, Martin et al. 2022), poor understanding of habitat preferences (Rosenfeld & Hatfield 2006), or temporal or spatial variation in occupancy (Wiens et al. 1987, Durso et al. 2011, Hayes & Monfils 2015). These factors all affect current understanding for the western distinct population segment (DPS) of the Yellow-billed Cuckoo (*Coccyzus americanus*) (hereafter, “cuckoo”), which was federally listed as threatened in 2014 (USFWS 2014).

Yellow-billed Cuckoos are Neotropical migrants with a widespread distribution across sub-boreal North America (Hughes 2015). While cuckoo numbers have decreased throughout their range, the species has declined to a fraction of its former range and numbers across the western United States and extreme southwestern Canada (the "western DPS"; Gaines & Laymon 1984, Dettling et al. 2015, Hughes 2015), with most of the remaining population now in Arizona, New Mexico, southern California, and northern Sonora (Hughes 2015, USFWS 2021). Cuckoos are notably cryptic, exhibiting low call rates, large home ranges, a delayed breeding season, and rapid nesting cycle (Hamilton & Hamilton 1965, Halterman 2009, McNeil et al. 2013, Sechrist et al. 2013). They have also been hypothesized to have a nomadic period before nesting during which they may wander widely to assess prey availability (Sechrist et al. 2012, McNeil et al. 2015), making it difficult to determine whether detections in atypical (e.g., non-riparian) vegetation or early in the season are breeding or transient birds. These behaviors present challenges in assessing site occupancy, breeding status, population numbers, distributional patterns, and habitat preferences.

Cuckoos in the western DPS commonly breed in low-mid elevation riparian areas in broad, low-gradient floodplains, typically dominated by obligate-phreatophyte trees including cottonwood (*Populus spp.)* and willow (*Salix spp.*), often with multi-story or early successional structure and adjacent facultative-phreatophytes such as mesquite (Anderson & Laymon 1989, Ahlers et al. 2016, Johnson et al. 2017, McNeil et al. 2013, Wohner et al. 2021). This is referred to as “Rangewide Habitat” (hereafter “rangewide riparian habitat”) by the recent critical habitat rule issued by the U.S. Fish and Wildlife Service (USFWS 2021). These habitat conditions are supported by perennial surface or ground water (Snyder 2000), and the recruitment of riparian trees and resulting vegetation structure depend on seasonal flood regimes (Stromberg 1993, 2001, Lytle et al. 2017). In the southwestern United States, cuckoo survey efforts, ecological studies, and habitat modeling, have focused primarily on these cottonwood- and willow-dominated rangewide riparian habitats (Halterman 2009, Johnson et al. 2010, McNeil et al. 2013, Dettling et al. 2015, Wohner et al. 2021). Recent cuckoo surveys in mountain and foothill drainages in southeastern Arizona (Corman & Magill 2000, USFWS 2014, MacFarland & Horst 2015, 2016, 2017, 2019, Corson 2018), however, have documented cuckoo occurrence in mid- to upper-watershed drainages that are generally dry, with only intermittent or ephemeral surface water. The dominant trees along these drainage courses are typically a mix of Velvet Mesquite (*Prosopis velutina*), Arizona Ash (*Fraxinus velutina*), Net-leaf Hackberry (*Celtis reticulata*), and various oak species (*Quercus* spp.), depending on the elevation and aridity of the area. Tree density ranges from scattered individuals to small clumps to continuous bands along the edges of the drainage course, but the bands are typically narrow (one to a few trees in width). Cottonwood, willow, or other obligate phreatophytes are sometimes present but only in low numbers (this study; see also MacFarland & Horst 2015). Adjacent uplands may be comprised of Madrean-evergreen woodland, semi-desert grassland, or desert scrub (Brown 1994, MacFarland & Horst 2015), again depending on aridity of the region. These areas are referred to as “Southwestern Habitat” in the critical habitat designation rule (USFWS 2021) and throughout this paper we refer to them as “SE AZ xeroriparian habitat” or simply "xeroriparian." Although breeding behavior has been observed anecdotally at some of these sites (MacFarland & Horst 2015, 2017), the extent to which detections in these SE AZ xeroriparian habitats reflected breeding birds rather than transients or migrants was unknown.

Breeding status of cuckoos in rangewide riparian habitat has typically been evaluated based on results of a standardized survey protocol using repeat-visit, call-playback methods along a survey transect (Halterman et al. 2015). To avoid unnecessary stress to cuckoos, nest searching is not a component of these protocol surveys, and few nests are incidentally detected due to the cuckoo’s cryptic behavior and nest concealment. Instead, breeding status and site occupancy is estimated at the end of the survey season by reviewing the location and timing of cuckoo detections within the survey site. Although this approach has been shown to accurately reflect breeding status in rangewide riparian habitat (Halterman 2009, McNeil et al. 2013), its applicability has not been assessed in SE AZ xeroriparian habitat. Interpretation of cuckoo occurrence and breeding status in SE AZ xeroriparian habitats is further complicated by geolocator studies indicating cuckoos regularly use southeastern Arizona for migration (Sechrist et al. 2012, McNeil et al. 2015) and historical accounts of cuckoos in upland vegetation in California during migration (Shelton, 1911). Therefore, the spatial extent of breeding in this novel and potentially significant habitat in SE AZ cannot be reliably assessed because of limited surveys and lack of careful assessment of breeding status in SE AZ xeroriparian habitat.

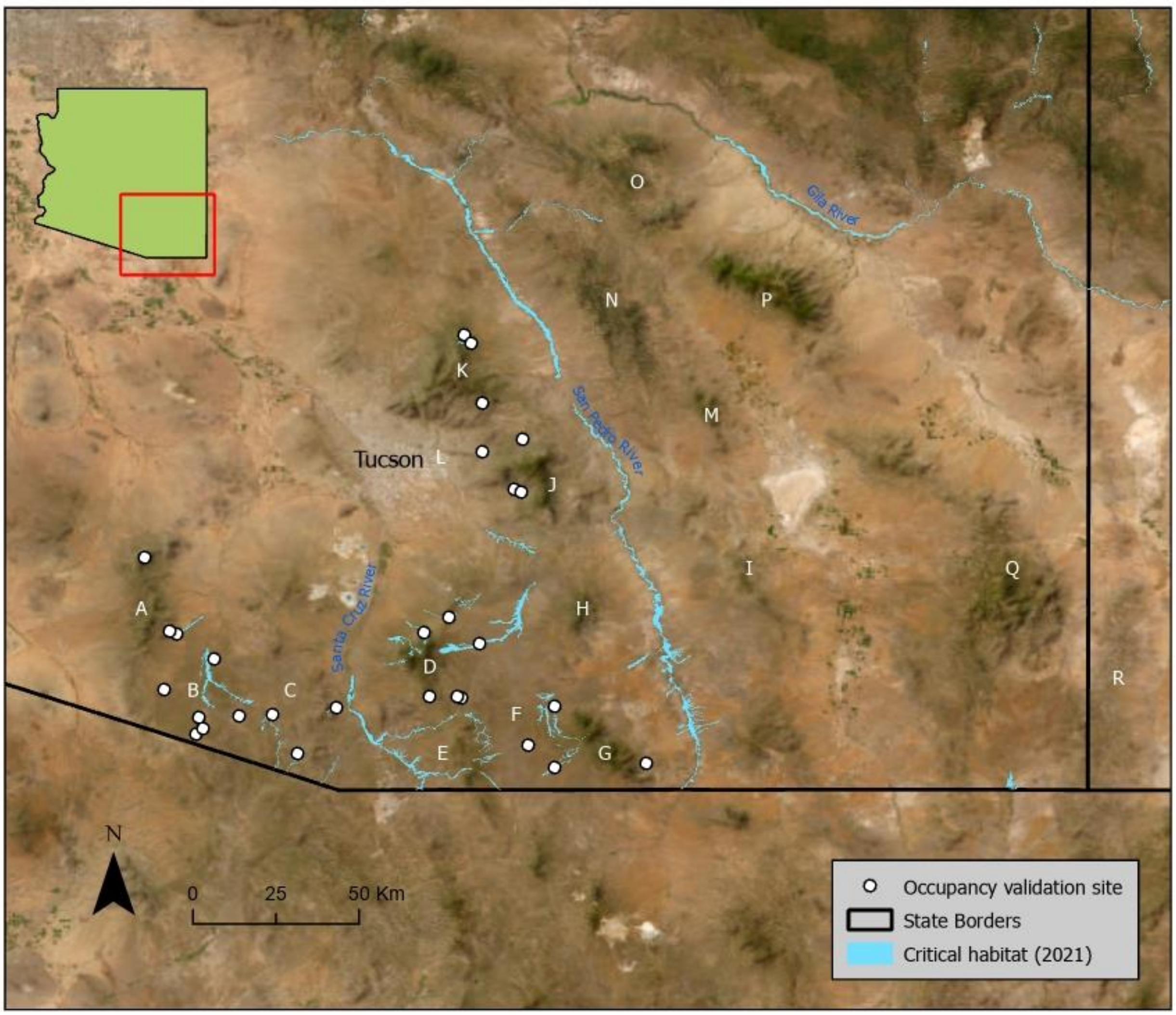
To evaluate the breeding status and distribution of cuckoos in SE AZ xeroriparian habitat, we (Northern Arizona University, in conjunction with U.S. Fish and Wildlife Service and U.S. Geological Survey) first tested whether site occupancy and breeding status as estimated using the standard USFWS-accepted survey protocol (Halterman et al. 2015) accurately reflected breeding status by conducting both protocol surveys and intensive nest searching at a subset of sites in the region. We also conducted opportunistic nest searching in additional sites being informally surveyed. We then reevaluated previously collected survey data and developed a map of known breeding distribution and occupancy in SE AZ xeroriparian habitat.

METHODS

I. Study area

Our study area included major mountain ranges and foothills between the San Pedro River and the Baboquivari Mountains, but additional historic data collected east of the San Pedro River were compiled and included for analyses of distribution (Figure 1.1). The study area lies within the broader, international “Madrean Sky Island Archipelago” region (hereafter, Sky Islands), which consists of prominent mountain ranges separated by desert valleys that extend from the southern terminus of the Colorado Plateau in Arizona and New Mexico to the northern Sierra Madre in Sonora and Chihuahua, Mexico (Brown 1994).

Figure 1.1 Study area in southeastern Arizona, with locations of 37 sites used for assessment of Yellow-billed Cuckoo occupancy and breeding, 2018–2020. Blue indicates designated critical habitat for the species (USFWS 2021). White letters denote localities mentioned in the text: A. Baboquivari Mountains, B. Altar Valley, C. Atascosa Mountains, D. Santa Rita Mountains, E. Patagonia Mountains, F. Canelo Hills, G. Huachuca Mountains, H. Whetstone Mountains, I. Dragoon Mountains, J. Rincon Mountains, K. Santa Catalina Mountains, L. Tucson Lowlands, M. Winchester Mountains, N. Galiuro Mountains, O Santa Teresa Mountains, P. Pinaleno Mountains, Q. Chiricahua Mountains, R. Peloncillo Mountains.



II. Site selection

We sampled 83 sites within our study area between 2018 and 2020. To maximize our ability to test whether current survey protocols accurately reflect breeding status, 61 of these sites were selected non-randomly by choosing sites where cuckoos had been previously detected. The remaining 22 sites were selected using a stratified random sampling approach, which was initiated in 2019. Our 2018 survey results, together with eBird (2021) data for southeastern Arizona, indicated that the major vegetation associations where cuckoos occurred in July and August (peak nesting period in Arizona; Hamilton & Hamilton 1965, Hughes 2015) were in Apacherian-Chihuahuan Mesquite Upland Scrub, Madrean Encinal, and North American Warm Desert Riparian Forest & Woodland (associations from LANDFIRE 2016). We then randomly selected grid cells having over 50% coverage of one or more of these LANDFIRE vegetation associations from a GIS coverage of the overall study area, with a 2 km x 2 km grid overlay. In each random grid cell, we selected a vegetated drainage bottom to survey. A total of 54 random sites were initially selected, and the first five or six drawn for each mountain range were used for survey sites, with the remaining selections serving as backups in the event we were unable to access the primary sites.

We selected 37 of the 83 sites to be used in our occupancy validation analysis, where we conducted protocol surveys in tandem with intensive nest searching between 2018 and 2020. Seven of these occupancy validation sites were from the randomly selected subset of sites while the rest were sites where cuckoos had been reported previously. In 46 additional sites (15 of which were from the randomly selected subset of sites), we either opportunistically documented breeding or conducted surveys without follow-up nest searching (see Mapping Distribution below), but these data were not used in the occupancy validation analysis.

III. Protocol surveys and nest searching

We surveyed for cuckoos using the currently established USFWS-accepted protocol (Halterman et al. 2015). Surveys consisted of call-playback at 100 m intervals along a pre-established transect that followed the main drainage at each site. All detection locations were recorded using handheld GPS, and the total number of individual cuckoos detected (accounting for possible instances of double counting) was estimated upon completion of each survey visit.

Following positive detections within a site, we conducted additional nest searching surveys to determine breeding status. Surveyors returned to the area of previous cuckoo detections with the goal of locating an active nest or observing other evidence of breeding including copulation, nest building, fledglings, or distraction displays. To avoid disturbance to potentially breeding birds, playback was used minimally, and we maintained a minimum distance of 10 m from birds while tracking individuals. Locations of nests or other breeding evidence were recorded using handheld GPS units. Breeding surveys were undertaken only after cuckoos were detected in a site during protocol surveys. If no cuckoos were detected in a site, no follow-up visits for breeding surveys were made.

The standard protocol (Halterman et al. 2015) calls for a minimum of 4 surveys per site, spaced 10–15 days apart, with the season lasting from June 15 – August 15. We followed the protocol when conducting a survey on a single morning, referred to as a “survey visit”. However, at some sites we deviated from the standard protocol by shifting the surveys two weeks later, and at some sites we ceased survey efforts after 2 or 3 survey visits if both positive occupancy status and positive breeding status had been established (see Occupancy Validation, below). We refer to any of these survey methods as “protocol surveys”.

IV. Occupancy validation

Upon completion of each survey season (2018-2020), data collected from sites were evaluated to estimate occupancy and breeding status according to an established protocol rubric (Halterman et al. 2015). "Occupancy" as used here refers to an evaluation of whether cuckoos were present at a site consistently during the breeding season; it is not formal statistical occupancy analysis (cf. MacKenzie et al. 2002). We identified “occupied” territories as areas where cuckoos were detected during 2 or more survey periods, with survey visits separated by at least 10 days, with detection locations between surveys no greater than 500 m apart and/or where evidence of breeding was observed incidentally during a survey visit or a follow-up to a survey visit. An occupied site may consist of one or more of these cuckoo territories. Sites were classified as “unoccupied”, and therefore not assigned a breeding estimate, under three scenarios: 1) if no cuckoos were detected during any of at least four survey visits, 2) if cuckoos were detected during only one of at least four survey visits, or 3) if cuckoos were detected in two or more of at least four survey visits but no two detections from separate survey visits were <500 m apart. We categorized scenarios 2 and 3 as “unoccupied with detection”. These criteria allowed for sites to be included if they were visited less than four times only if occupancy status had been established and breeding had been confirmed with fewer than four survey visits, as those results constituted a positive validation of the occupancy estimate without full effort. We did not include in our occupancy validation analysis either unoccupied sites with fewer than four survey visits or sites where cuckoos were detected but breeding was not confirmed because no additional follow-up nest searching was conducted.

The primary purpose of this study was to determine and document the occurrence and extent of nesting in these xeroriparian habitats, so we therefore assumed that occupancy of a site by a pair of cuckoos over the course of the nesting season was likely to be associated with attempted breeding. We further assumed that intensive nest-searching at a site where cuckoos were breeding would reliably confirm the breeding attempt (through finding an active nest or fledged young, or nest-building, repeated carrying of food, or other confirmatory evidence). On this basis, each site was classified according to the standard survey protocol designation (Halterman et al. 2015) and its breeding status based on nest searching efforts. This resulted in five possible classifications:

1) Occupied based on protocol surveys; breeding confirmed based on nest searching.
2) Occupied based on protocol surveys; breeding not confirmed based on nest searching.
3) Unoccupied based on protocol surveys; no breeding surveys due to unoccupied status.
4) Unoccupied with detection based on protocol surveys; breeding confirmed based on nest searching.
5) Unoccupied with detection based on protocol surveys; breeding not confirmed based on nest searching.

We then calculated the total number of each of these five classifications to determine overall proportion of sites in which occupancy status according to standard protocol survey designations accurately reflected attempted breeding, and proportions of sites where standard protocol survey designations of "occupied" or "unoccupied" did not agree with nest searching results (i.e., false positive or false negative). This evaluation, in turn, would allow us to evaluate whether site occupancy, as determined through protocol surveys, could reliably be used to assess extent and distribution of breeding by cuckoos in xeroriparian habitats.

V. Mapping and evaluating distribution

Results from the occupancy validation were used to reevaluate occupancy estimation results from additional sites we surveyed between 2018 and 2020, as well as previous survey data collected between 2013 and 2020 and contributed by other entities, including Tucson Audubon Society, Audubon Southwest, Saguaro National Park, Buenos Aires National Wildlife Refuge, Coronado National Forest, and the consulting firms Moore Biological Services, Archeological Consulting Services, and WestLand Resources Inc. Sites were classified according to their occupancy status based on standard protocol surveys. For sites with multiple years of data, the highest level of occupancy status was used to represent that site's "occupancy potential" in our map. In some sites, only three survey visits were conducted rather than the four required by the protocol. In these cases, we excluded sites that contained detections on only 1 of 3 survey visits and were classified as unoccupied, as results from a 4^{th} survey visit could possibly yield detections and elevate the site's status to occupied. However, we did retain sites classified as unoccupied with three negative survey visits, under the assumption that 4^{th} survey visit would not yield results sufficient for occupied status. Additionally, some surveyors shifted their survey window two weeks later to ensure their surveys extended into late August. We retained data for all of these sites given that this shifted schedule is intended to capture more of the breeding season in southeastern Arizona (Hamilton & Hamilton 1965). Except for locations of confirmed breeding not associated with protocol surveys, incidental detections were not included.

All sites were then mapped using ArcGIS (ESRI) according to their final occupancy classifications with symbols differentiating locations of confirmed breeding sites, occupied sites, and unoccupied sites. We also included polygons of designated critical habitat (USFWS 2021), representing known cuckoo populations which primarily consists of rangewide riparian habitat in perennial drainages. If multiple territories were documented in a single site, only one location was used to represent that site.

RESULTS

I. Occupancy validation

A total of 37 sites were included for our occupancy validation analysis, with three sites having two years of survey data, resulting in 40 survey records (Table 1.1). Twenty-four (60%) of those 40 survey records resulted in an occupied status, with breeding evidence documented in 23 of them. Cuckoos were detected in 2 of 16 surveys classified as unoccupied, with no additional detections or breeding evidence documented in follow-up surveys. No cuckoos were detected in the remaining 14 unoccupied surveys. Of the three sites with two years of survey data, one site was occupied in both years, one was unoccupied in both years, and one site was occupied in one year and unoccupied in the subsequent year. Taken together, these results reflected only a single potential false-positive site (a site designated as occupied based on protocol surveys, but no evidence of breeding based on follow-up nest searching) and no false-negative occupancy estimates. Protocol surveys at the lone potential false-positive site (Brown Wash, Baboquivari Mountains), estimated three territories present based on detections in all six protocol surveys conducted between June 26 and September 6, suggesting the lack of breeding confirmation at this site may have been failure of nest searchers to find the nests rather than absence of breeding. Regardless, these results demonstrated high confidence in the use of existing protocol survey criteria for estimating occupancy and breeding status.

Table 1.1 Survey results (n = 40) of occupancy validation analysis, comparing the results of Yellow-billed Cuckoo standard protocol surveys to the actual breeding status determined through nest searching at field sites (n = 37 with three of these surveyed in two years) in SE AZ xeroriparian habitat, 2018–2020.

Survey Occupancy Category	Totals	Breeding confirmed	No evidence of breeding
Occupied	24	23	1
Unoccupied with detection	2	0	2
Unoccupied, no detections	14	0	0
Totals	40	23	3

II. Breeding surveys

In addition to confirmed breeding at the 23 sites used in the Occupancy validation analysis, we opportunistically documented breeding at an additional 22 sites. Multiple breeding territories were documented at several of these sites, resulting in a total of 55 known breeding territories (Supplemental Material Table 1.1). We documented 24, 26, and 5 breeding territories in 2018, 2019, and 2020, respectively (survey efforts were reduced in 2020). Evidence of breeding included 39 active nests, 11 observations of fledglings/juveniles, 3 observations of copulation, and 4 observations of distraction displays. Nests were placed in several tree species, including oak (*Quercus* sp.; 14), hackberry (*Celtis reticulata*; 12), mesquite (*Prosopis* sp.; 5), juniper (*Juniperus* sp.; 3), acacia (*Senegalia greggii*; 2), ash (*Fraxinus velutina*; 2), yew-leaf willow (*Salix taxifolia*; 1), and cottonwood (*Populus fremontii*; 1). Nest monitoring was not a specific goal of our study, but we documented confirmed nesting success in 15 locations (based on observations of fledglings), nest failure in 2 locations, and were unable to determine nest fate in 38 locations where nests were not revisited. The earliest and latest dates of nesting activity were July 3 (active nest) and September 11 (nest with nestling), respectively. The earliest fledgling observation was July 29, and the latest observation of copulation was August 15.

III. Random sites

We surveyed 22 random sites across 7 mountain ranges in our study area. Of these, 68% (15) were occupied, including 6 of 7 random sites used for our Occupancy validation. We were unable to survey sufficient random sites per mountain range to evaluate spatial trends in occupancy. However, our observed overall random site occupancy rate of 68% was consistent with our 61 non-random sites, where 70% of sites were occupied (Table 1.2). When we removed non-random sites from ranges with no random sites, the occupancy rate of non-random sites was 73%. These results indicate targeted survey efforts tracked closely with results from randomly selected sites.

Table 1.2 Comparison of Yellow-billed Cuckoo occupancy status results from random sites (n=22) to non-random sites (n=61), separated by survey area/mountain range in SE AZ xeroriparian habitat, 2018–2020.

	Random sites		Non-random sites		Total sites	
Area/Range	Sites	Occupied (%)	Sites	Occupied (%)	Sites	Occupied (%)
Altar Valley	n/a	n/a	7	7 (100)	7	7 (100)
Atascosa Highlands	4	4 (100)	6	5 (80)	10	9 (90)

Baboquivari Mountains	n/a	n/a	3	2 (67)	3	2 (67)
Canelo Hills	5	3 (60)	9	7 (78)	14	10 (71)
Chiricahua Mountains	n/a	n/a	2	0 (0)	2	0 (0)
Dragoon Mountains	n/a	n/a	2	1 (50)	2	1 (50)
Huachuca Mountains	1	1 (100)	2	1 (50)	3	2 (67)
Patagonia Mountains	2	1 (50)	5	5 (100)	7	6 (86)
Rincon Mountains	5	3 (60)	8	2 (25)	13	5 (38)
Santa Catalina Mountains	1	0 (0)	3	1 (33)	4	1 (25)
Santa Rita Mountains	4	3 (75)	12	12 (100)	16	15 (94)
Whetstone Mountains	n/a	n/a	2	0 (0)	2	0 (0)
	22	15 (68)	61	43 (70)	83	58(70)

IV. Mapping distribution

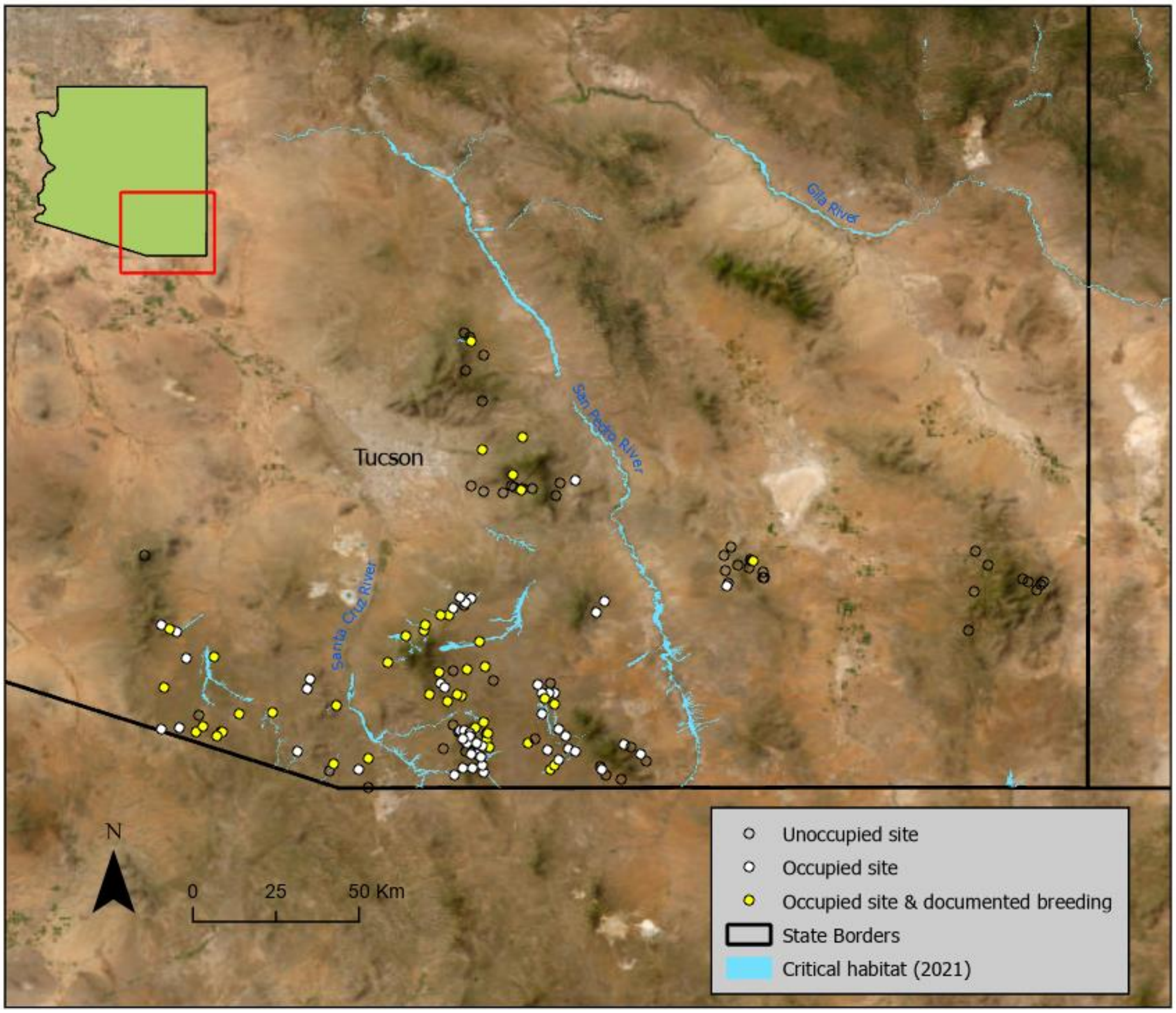
We combined data we collected between 2018 and 2020 with available survey and nest searching data collected by other entities between 2013 and 2020 in SE AZ xeroriparian habitat, resulting in data for 163 sites in our southeastern Arizona study area (Figure 1.2; Table 1.3; Supplemental Material Table 1.2). We documented breeding at 46 sites, either based on nest-searching or anecdotal evidence (serendipitous sighting of fledglings, copulation, etc.), 54 sites were considered occupied based on surveys but were not revisited for nest searching and had no anecdotal evidence of breeding, 18 sites were considered unoccupied sites with detections, and 45 were considered unoccupied sites with no detections (Figure 1.2).

Table 1.3 Summary of compiled Yellow-billed Cuckoo survey results for sites in SE AZ xeroriparian habitat in ephemeral and intermittent drainages, 2013-2020. In sites with multiple years of data, the highest level of occupancy documented is reported.

Survey Occupancy Category	Totals	Breeding confirmed	Breeding not confirmed	No cuckoos present
Occupied	100	46	54	0
Unoccupied with detection	18	0	18	0
Unoccupied, no detections	45	0	0	45
Totals	163	46	72	45

Figure 1.2 Results of Yellow-billed Cuckoo surveys and nest searching at 163 sites in southeastern Arizona between 2013 and 2020. Open circles are unoccupied sites, white circles are occupied sites (based on surveys), yellow circles are sites where breeding has been

documented. Due to the scale of the map, sites within 1–2 km of each other have overlapping symbols.



As summarized in Table 1.2, occupied sites were numerous in the western portion of the Sky Islands, particularly in the region extending from the Baboquivari Mountains in the west, east through the Altar Valley, Atascosa Highlands, Santa Rita Mountains, Patagonia Mountains, Canelo Hills, San Rafael Valley, and finally the western side of the Huachuca Mountains. Some occupied sites were documented in the Whetstone Mountains, Rincon Mountains, and Santa Catalina Mountains, but we observed a higher proportion of unoccupied sites and surveyed fewer overall sites in these ranges. East of the San Pedro River, data contributed from other surveyors resulted in two occupied sites in the Dragoon Mountains and no occupied sites in the Chiricahua Mountains.

DISCUSSION

Previous studies of cuckoo habitat in the western United States stress the importance of native riparian vegetation along major rivers and streams (Gaines 1974, Girvetz & Greco 2009, Johnson et al. 2017), while information on cuckoo use of SE AZ xeroriparian habitat has remained largely anecdotal. Our results provide robust documentation of widespread breeding of cuckoos in SE AZ xeroriparian habitat, and offer insights into previously undescribed habitat associations, distribution, and conservation implications. By verifying that breeding occurred in 97% of sites designated as "occupied" using standard survey protocol estimates at our intensively studied Occupancy validation sites, we demonstrated strong support for the use of established survey protocol (Halterman et al. 2015) to determine breeding status in SE AZ xeroriparian habitat at those sites. Moreover, the percentage of randomly selected sites that were occupied was similar to that of non-random sites, suggesting our estimates of sites occupancy were not biased by a-priori site selection. This allowed us to extrapolate the results from the occupancy validation set of sites to other sites and to surveys carried out in other years, providing a larger temporal and spatial scale for assessing cuckoo presence across the Sky Islands region.

The 100 occupied sites we documented between 2013 and 2020, including many sites with multiple nests or territories, represent a significant addition to the estimated population of cuckoos in Arizona given previous population estimates. Corman and Magill (2000) conducted statewide surveys in 145 sites with riparian habitat between 1998 and 1999, with detections in 84 sites. At 62 sites on the Lower Colorado River (LCR), McNeil et al. (2013) reported a 70% increase in breeding territory estimates, from 47 in 2008 to 80 in 2012, attributing increases to newly available habitat in restoration sites. In 2021, only 41 of the 62 sites on the LCR were surveyed, but cuckoo populations remained consistent with 64 estimated breeding territories (Tracy & Squibb 2022). Elsewhere in the western DPS of the cuckoo's range, populations on the Sacramento River in California have declined precipitously, with only 8 and 10 individual detections in 2012 and 2013, respectively (Dettling et al. 2015). The Kern River in California has a stable but fluctuating population, with annual individual detection totals from standard protocol surveys ranging from 13 to 89 (Stanek 2014, 2017). Surveys along ~82 river miles on the Middle Rio Grande in New Mexico in 2019 resulted in 75 estimated breeding territories (White et al. 2020), which is consistent with 2015 results reporting 110 estimated breeding territories along 129 river miles (Ahlers et al. 2016). Although these previous studies provide some context for our results, direct comparisons are not possible because most previous studies used different methods and occurred prior to 2015 when current standardized survey protocols (Halterman et al. 2015) were adopted. Timing, total number of surveys per site, methods for estimating pairs and occupancy status, and reporting of area surveyed have all varied among studies. Despite these caveats, our results suggest cuckoo total abundance in SE AZ xeroriparian habitat is likely similar to that of the total abundance of areas previously surveyed on the San Pedro River or Lower Colorado River, possibly higher than areas surveyed on the Gila River, Salt River, or Santa Cruz River, and higher than most populations in other western DPS states, with the exception of large populations along the Rio Grande in New Mexico. Furthermore, while we did not specifically monitor nest fate, at 20% of the 55 breeding locations we surveyed we recorded fledglings or juveniles, indicating that breeding not only occurred but was successful. Taken in this context, our results confirm that SE AZ supports a significant cuckoo breeding population in the DPS.

Our results indicate cuckoos in SE AZ use a range of vegetative communities different from the "extensive riverbottom vegetation" (Gaines 1974) or "riparian vegetation with substantial canopy cover provided by native riparian trees, particularly willows (*Salix* spp.) and cottonwoods (*Populus* spp.)" described for the species in the western part of their range (Johnson et al. 2017). Some of our survey sites had typical riparian trees such as cottonwood or willow, but these were sparse or absent altogether at most sites. Instead, occupied sites represented a continuum of riparian conditions, with local vegetation ranging from isolated or narrow patches of obligate phreatophytes (cottonwood and willow), to xeric ephemeral drainages with only facultative phreatophytes such as oak, mesquite, and hackberry. The contrast in the xeroriparian vegetative communities occupied by cuckoos as documented here, compared to the cottonwood-willow riparian areas used by cuckoos, is also reflected in our list of nest tree species, with 78% (31 of 40) nests found in oak, hackberry, or mesquite. In most cases, adjacent uplands contained either Madrean encinal woodland, semi-desert grassland (typically with interspersed shrubs and trees), or desert scrub (Brown 1994), all of which were used for foraging. Our occupied sites also represented diverse physiographies, ranging from alluvial flats to narrow canyons, thus expanding the range of vegetation communities traditionally considered habitat for cuckoos.

Geographically, our combined results indicate a lower cuckoo occupancy in the northern and eastern portions of the Sky Islands region, with most occupied sites and breeding locations in the southern and western portions of the region, which coincides with the wettest region of the North American monsoon in the United States (Forzieri et al. 2014, Higgins et al. 1997, Wallace et al 2013). In particular, cuckoos were detected in most drainages surveyed in the Patagonia Mountains, Santa Rita Mountains, Canelo Hills, Atascosa Highlands, and Altar Valley. Many of these sites were located in watersheds of rivers that support riparian habitat known to contain high numbers of cuckoos (e.g., San Pedro River, Santa Cruz River, Sonoita Creek). Whether and how often cuckoos move between these riparian and xeroriparian habitats, both within and between years, remains an important question. Importantly, our mapped distribution (Figure 1.2) depicts known breeding locations and occupied survey sites and is not an estimate of actual distributional limits. Large geographic gaps in available data include the Baboquivari Mountains (west side), Galiuro Mountains, Winchester Mountains, Santa Teresa Mountains, Pinaleño Mountains, and Peloncillo Mountains, as well as lower-elevation drainages in sub-ranges, foothills, and valleys between these major mountain ranges. Many of the gaps in survey coverage are either on private property or are remote and difficultto access. While survey data from Mexico are limited (Macías-Duarte et al. 2015, 2023) and not included in our analyses, they are consistent with our findings in southeastern Arizona, suggesting use of xeroriparian habitat in ephemeral drainages may extend into the Sierra Madre of Mexico. Therefore, we recommend additional survey efforts be made in under-surveyed areas of the Sky Islands region in both the United States and Mexico to further refine distribution and population estimates.

Although our results support using existing protocols for estimating occupancy and breeding status based on survey results in xeroriparian habitat, our nesting data indicate cuckoos may often breed through August and as late as September in some sites (Supplemental Material Table 1.2). This is consistent with Hamilton & Hamilton’s (1965) findings of cuckoos in southern Arizona nesting later than cuckoos in southern California. Under current protocols, the fourth and final survey could hypothetically be completed on August 1st, potentially resulting in a nesting cycle occurring after surveys have ended. We suggest that for xeroriparian sites in southeastern Arizona, a more accurate breeding window may be captured with an additional late-season survey between August 15 – 31 or shifting the 4 required surveys approximately two weeks later (July 1 – August 31). We caution, however, that multiple years of survey data may be necessary to estimate occupancy at any given site. For example, in Upper Box Canyon in the Santa Rita Mountains we discovered 3 nests in 2019, while no cuckoos were detected during any protocol surveys in 2020. The latter year had a notably dry monsoon season and underscores the need to identify factors influencing interannual variation in occupancy.

Cuckoo populations are believed to have declined in rangewide riparian habitat primarily due to loss and degradation of bottomland riparian vegetation (USFWS 2014). Southwestern xeroriparian habitat, often occurring above these stressors in the watershed, may serve as important refugia for cuckoos. However, these drainages have undergone varied and often significant anthropogenic changes, many of which may present unique threats to habitat quality and resilience. For example, although livestock grazed many of our sites, and cuckoos have previously been documented utilizing actively grazed riparian habitat (Hamilton & Hamilton 1965), livestock grazing may result in degradation to riparian and xeroriparian habitat (Stromberg 1993, Fleischner 1994, Brock & Green 2003, Goodrich et al. 2018). Alternative grazing practices and exclusion of cattle from riparian and xeroriparian drainages in southern Arizona have resulted in improvements to hydrological and ecological function (Beard 2004, Krueper 2003), and may benefit cuckoo habitat. Likewise, although cuckoos in our area occupied sites that had experienced historical mining activity, modern industrial mining often occurs on a much larger scale with greater potential impacts to watershed hydrology, geochemistry, and habitat quality through dewatering of aquifers or otherwise redirecting or altering flows in drainages (Lewis & Burraychak 1979, Brock & Green 2003). Finally, climate change may act to increase aridification of southwestern uplands and exacerbate the risk of fire, posing a threat to the resilience of SE AZ xeroriparian habitat (Bock & Bock 2014, Friggens et al. 2014). In spite of these potential threats, the number of breeding cuckoos we documented in these varied xeroriparian habitats, and the fact that many of those birds bred successfully, indicates these areas represent an important addition to known habitat for western Yellow-billed Cuckoo conservation.

ACKNOWLEDGEMENTS

Work presented in this chapter was supported by contributions from Dr. Tad Theimer, Susan Sferra, and Charles Drost. This research was conducted under U.S. Fish & Wildlife Service Recovery Permit #TE81148C-0 and Arizona Game & Fish Department Special Collecting License permit #SP407168. The findings and conclusions in this article are those of the authors and do not necessarily represent the views of the U.S. Fish and Wildlife Service. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the United States Government. Data for this study are available as a U.S. Geological Survey data release.

LITERATURE CITED

Ahlers, D. D., Moore, D., & White, M. (2016). Yellow-billed Cuckoo Study Results – 2015 Middle Rio Grande from Los Lunas to Elephant Butte Reservoir, New Mexico (p. 65). U.S. Department of the Interior, Bureau of Reclamation, Fisheries and Wildlife Resources Group.

Anderson, B. W., & Laymon, S. A. (1989). Creating habitat for the yellow-billed cuckoo (Gen. Tech. Rep. PSW-110; Proceedings of the California Riparian Systems Conference: Protection, Management, and Restoration for the 1990's; September 22-24, 1988 Davis, California, pp. 468–472). USDA Forest Service.

Beard, R. (2004). Stream Channel Change in Response to Cattle Exclosures in Semi-Arid Central Arizona. Journal of the Arizona-Nevada Academy of Science. 36(2) 81–87.

Bock, C. E., & Bock, J. H. (2014). Effects of Wildfire on Riparian Trees in Southeastern Arizona. The Southwestern Naturalist, 59(4), 570–576. https://doi.org/10.1894/JEM-08.1

Brock, J. H., & Green, D. M. (2003). Impacts of Livestock Grazing, Mining, Recreation, Roads, and Other Land Uses on Watershed Resources. Journal of the Arizona-Nevada Academy of Science. 35(1) 11-22..

Brown, D. E. (Ed.). (1994). Biotic communities: Southwestern United States and Northwestern Mexico. Univ. of Utah Press.

Camaclang, A. E., Maron, M., Martin, T. G., & Possingham, H. P. (2015). Current practices in the identification of critical habitat for threatened species: Identifying Critical Habitat. Conservation Biology, 29(2), 482–492. https://doi.org/10.1111/cobi.12428

Corman, T. E., & Magill, R. T. (2000). Western Yellow-billed Cuckoo in Arizona: 1998 and 1999 Survey Report (Nongame and Endangered Wildlife Program Technical Report 150; p. 55). Arizona Game and Fish Department.

Dettling, M. D., Seavy, N. E., Howell, C. A., & Gardali, T. (2015). Current Status of Western Yellow-Billed Cuckoo along the Sacramento and Feather Rivers, California. PLOS ONE, 10(4), e0125198. https://doi.org/10.1371/journal.pone.0125198

Durso, A. M., Willson, J. D., & Winne, C. T. (2011). Needles in haystacks: Estimating detection probability and occupancy of rare and cryptic snakes. Biological Conservation, 144(5), 1508–1515. https://doi.org/10.1016/j.biocon.2011.01.020

Fleischner, T. L. (1994). Ecological Costs of Livestock Grazing in Western North America. Conservation Biology, 8(3), 629–644. https://doi.org/10.1046/j.1523-1739.1994.08030629.x

Forzieri, G., Feyen, L., Cescatti, A., & Vivoni, E. R. (2014). Spatial and temporal variations in ecosystem response to monsoon precipitation variability in southwestern North America: Ecosystem Response to Monsoon Rainfall. Journal of Geophysical Research: Biogeosciences, 119(10), 1999–2017. https://doi.org/10.1002/2014JG002710

Friggens, M. M., Loehman, R., Holsinger, L., & Finch, D. (2014). Vulnerability of riparian obligate species to the interactive effect of fire, climate and hydrological change (p. 213) [Final Report for Interagency Agreement #13-IA-11221632-006.]. U.S. Department of Agriculture, Forest Service Rocky Mountain Research Station.

Gaines, D. (1974). Review of the Status of the Yellow-Billed Cuckoo in California: Sacramento Valley Populations. The Condor, 76(2), 204. https://doi.org/10.2307/1366731

Gaines, D., & Laymon, S. A. (1984). Decline, Status and Preservation of the Yellow-Billed Cuckoo in California. Western Birds, 15(2), 49–80.

Girvetz, E. H., & Greco, S. E. (2009). Multi-Scale Predictive Habitat Suitability Modeling Based on Hierarchically Delineated Patches: An Example for Yellow-Billed Cuckoos Nesting in Riparian Forests, California, Usa. Landscape Ecology, 24(10), 1315–1329. https://doi.org/10.1007/s10980-009-9384-2

Goodrich, D. C., Kepner, W. G., Levick, L. R., & Wigington, P. J. (2018). Southwestern Intermittent and Ephemeral Stream Connectivity. JAWRA Journal of the American Water Resources Association, 54(2), 400–422. https://doi.org/10.1111/1752-1688.12636

Gu, W., & Swihart, R. K. (2004). Absent or undetected? Effects of non-detection of species occurrence on wildlife–habitat models. Biological Conservation, 116(2), 195–203. https://doi.org/10.1016/S0006-3207(03)00190-3

Halterman, M. D., Johnson, M. J., Holmes, J. A., & Laymon, S. A. (2015). A Natural History Summary and Survey Protocol for the Western Distinct Population Segment of the Yellow-billed Cuckoo. U.S. Fish and Wildlife Techniques and Methods.

Halterman, M. M. (2009). Sexual Dimorphism, Detection Probability, Home Range, and Parental Care in the Yellow-billed Cuckoo [PhD Dissertation]. University of Nevada.

Hamilton, W. J., III, & Hamilton, M. E. (1965). Breeding characteristics of yellow-billed cuckoos in Arizona. Proceedings of the California Academy of Sciences, 32(14), 405–432.

Hayes, D. B., & Monfils, M. J. (2015). Occupancy modeling of bird point counts: Implications of mobile animals: Occupancy Modeling of Mobile Animals. The Journal of Wildlife Management, 79(8), 1361–1368. https://doi.org/10.1002/jwmg.943

Higgins, R. W., Yao, Y., & Wang, X. L. (1997). Influence of the North American Monsoon System on the U.S. Summer Precipitation Regime. Journal of Climate, 10(10), 2600–2622. https://doi.org/10.1175/1520-0442(1997)010<2600:IOTNAM>2.0.CO;2

Hughes, J. M. (2015). Yellow-billed Cuckoo (Coccyzus americanus), version 2.0. In P. G. Rodewald (Ed.), The Birds of North America Online. Cornell Lab of Ornithology. https://birdsoftheworld.org/bow/historic/bna/yebcuc/2.0/introduction

Johnson, M. J., Hatten, J. R., Holmes, J. A., & Shafroth, P. B. (2017). Identifying western yellow-billed cuckoo breeding habitat with a dual modelling approach. Ecological Modelling, 347, 50–62. https://doi.org/10.1016/j.ecolmodel.2016.12.010

Johnson, M., Magill, R. T., & van riper, C. (2010). Yellow-billed cuckoo distribution and habitat associations in Arizona, 1998-1999. The Colorado Plateau IV: Integrating Research and Resources Management for Effective Conservation, 197–212.

Joseph, L. N., Field, S. A., Wilcox, C., & Possingham, H. P. (2006). Presence-Absence versus Abundance Data for Monitoring Threatened Species. Conservation Biology, 20(6), 1679–1687. https://doi.org/10.1111/j.1523-1739.2006.00529.x

Krueper, D. J. (2003). Effect of land use practices on western riparian ecosystems. In D. M. Finch & P. W. Stangel (Eds.), Status and Management of Neotropical Migratory Birds: Vol. Gen Tech Rep RM-229 (p. 422). U.S. Department of Agriculture, Forest Service, Rocky Mountain Forest and Range Experimental Station.

Lewis, M., & Burraychak, R. (1979). Impact of Copper Mining on a Desert Intermittent Stream in Central Arizona: A Summary. Journal of the Arizona-Nevada Academy of Science. 14(1) 22-29.

Lytle, D. A., Merritt, D. M., Tonkin, J. D., Olden, J. D., & Reynolds, L. V. (2017). Linking river flow regimes to riparian plant guilds: A community-wide modeling approach. Ecological Applications, 27(4), 1338–1350. https://doi.org/10.1002/eap.1528

MacFarland, J., & Horst, J. (2017). Yellow-billed Cuckoo Surveys on the Coronado National Forest within Five Sky Island Mountain Ranges in Southeast Arizona in 2017 (p. 145) [Final Report for the Coronado National Forest]. Tucson Audubon Society.

Macías-Duarte, A., Alvarado-Castro, J. A., Ortega Rosas, C. I., Gutiérrez-Ruacho, O. G., Villarruel-Sahagún, L., & Sánchez-Murrieta, E. (2015). Assess status and habitat use of yellow-billed cuckoo populations in Sonora, Mexico (Technical Report UES-PIE-15-UAH-LE-03; p. 25). Sonora State University.

Macías-Duarte, A., Juárez, E., Murrieta, E. S., Perales-Hoeffer, E. L., & Rosas, C. I. O. (2023). Abundance and occupancy of the western yellow-billed cuckoo (Coccyzus americanus) in Sonora, Mexico. Canadian Journal of Zoology, cjz-2022-0096. https://doi.org/10.1139/cjz-2022-0096

MacKenzie, D. I. (2005). What are the issues with presence-absence data for wildlife managers? Journal of Wildlife Management, 69(3), 849–860. https://doi.org/10.2193/0022-541X(2005)069[0849:WATIWP]2.0.CO;2

Martin, M. E., Delheimer, M. S., Moriarty, K. M., Early, D. A., Hamm, K. A., Pauli, J. N., Mcdonald, T. L., & Manley, P. N. (2022). Conservation of rare and cryptic species: Challenges of uncertainty and opportunities for progress. Conservation Science and Practice, 4(11), e12809. https://doi.org/10.1111/csp2.12809

McNeil, S. E., Tracy, D., & Capello, C. D. (2015). Loop Migration by a Western Yellow-Billed Cuckoo Wintering in the Gran Chaco. Western Birds, 46, 24–215.

McNeil, S. E., Tracy, D., Stanek, J. R., Stanek, J. E., & Station, S. S. R. (2013). Yellow-billed Cuckoo Distribution, Abundance, and Habitat Use on the Lower Colorado River and Tributaries 2008-2012 Summary Report (p. 331). Lower Colorado River Multi-Species Conservation Program, U.S. Bureau of Reclamation.

Rosenfeld, J. S., & Hatfield, T. (2006). Information needs for assessing critical habitat of freshwater fish. Canadian Journal of Fisheries and Aquatic Sciences, 63(3), 683–698. https://doi.org/10.1139/f05-242

Sechrist, J., Ahlers, D. D., Zehfuss, K. P., Doster, R. H., Paxton, E. H., & Ryan, V. M. (2013). Home Range and Use of Habitat of Western Yellow-billed Cuckoos On the Middle Rio Grande, New Mexico. The Southwestern Naturalist, 58(4), 411–419. https://doi.org/10.1894/0038-4909-58.4.411

Sechrist, J. D., Paxton, E. H., Ahlers, D. D., Doster, R. H., & Ryan, V. M. (2012). One Year Of Migration Data For A Western Yellow-Billed Cuckoo. Western Birds, 43(1), 12.

Shelton, A. C. (1911). Nesting of the California Cuckoo. The Condor, 13(1), 19–22. https://doi.org/10.2307/1362213

Snyder, K. (2000). Water sources used by riparian trees varies among stream types on the San Pedro River, Arizona. Agricultural and Forest Meteorology, 105(1–3), 227–240. https://doi.org/10.1016/S0168-1923(00)00193-3

Stanek, J. (2014). Yellow-billed Cuckoo South Fork Kern River Valley 2014 Annual Report (p. 28) [Annual Report]. Southern Sierra Research Station.

Stanek, J. (2017). Yellow-billed Cuckoo Annual Report 2011. Annual Report. Southern Sierra Research Station.

Stromberg, J. C. (1993). Frémont Cottonwood-Goodding Willow Riparian Forests: A Review of Their Ecology, Threats, and Recovery Potential. Journal of the Arizona-Nevada Academy of Science, 27(1), 97–110.

Stromberg, J. C. (2001). Restoration of Riparian Vegetation in the South-Western United States: Importance of Flow Regimes and Fluvial Dynamism. Journal of Arid Environments, 49(1), 17–34. https://doi.org/10.1006/jare.2001.0833

Tracy, D., & Squibb, C. L. (2022). Yellow-billed Cuckoo Surveys in Conservation Areas on the Lower Colorado River and Bill Williams River, 2021 Annual Report (contract No. 140R3019C0004; p. 142). Southern Sierra Research Station.

U.S. Fish and Wildlife Service (USFWS). (2021). Endangered and Threatened Wildlife and Plants: Designation of Critical Habitat for the Western Distinct Population Segment of the Yellow-billed Cuckoo; Final rule. Federal Register, 86(75), 20798–21005.

USFWS. (2014). Endangered and Threatened Wildlife and Plants; Determination of Threatened Status for the Western Distinct Population Segment of the Yellow-billed Cuckoo (Coccyzus americanus); Final Rule. Federal Register, 79(129), 59992–60038.

USFWS. (2021). Endangered and Threatened Wildlife and Plants; Designation of Critical Habitat for the Western Distinct Population Segment of the Yellow- Billed Cuckoo. Federal Register.

Wallace, C. S. A., Villarreal, M. L., & van Riper III, C. (2013). Influence of monsoon-related riparian phenology on yellow-billed cuckoo habitat selection in Arizona. Journal of Biogeography, 40(11), 2094–2107. https://doi.org/10.1111/jbi.12167

White, M. C., Moore, D., & Siegle, R. (2020). Yellow-billed Cuckoo Study Results – 2019, Middle Rio Grande from San Acacia to Elephant Butte Reservoir, New Mexico. (p. 67). U.S. Bureau of Reclamaiton, Technical Service Center Fisheries and Wildlife Resources Group.

Wiens, J. A., Rotenberry, J. T., & Horne, B. V. (1987). Habitat Occupancy Patterns of North American Shrubsteppe Birds: The Effects of Spatial Scale. Oikos, 48(2), 132. https://doi.org/10.2307/3565849

Wohner, P. J., Laymon, S. A., Stanek, J. E., King, S. L., & Cooper, R. J. (2021). Early Successional Riparian Vegetation Is Important for Western Yellow-Billed Cuckoo Nesting Habitat. Restoration Ecology, 29(5). https://doi.org/10.1111/rec.13376

Supplemental Materials Table 1.1 List of 55 Yellow-billed Cuckoo breeding locations documented between in 45 sites in southeastern Arizona xeroriparian habitat, 2018–2020.

Site	Drainage	Range/Locality	Year	Date Found	Evidence	Nest Outcome [a]	Nest Tree
Canoa Wash 1	Canoa Wash 1	Altar Valley	2018	8/4/2018	Nest	UNK	Hackberry
Canoa Wash 2	Canoa Wash 2	Altar Valley	2018	8/15/2018	Distraction display	UNK	n/a
Cuadro Wash	Cuadro Wash	Altar Valley	2018	8/8/2018	Juvenile	S	n/a
Las Guijas Wash 1	Las Guijas Wash 1	Altar Valley	2019	7/31/2019	Nest	UNK	Acacia
Las Guijas Wash 2	Las Guijas Wash 2	Altar Valley	2019	8/15/2019	Juvenile	S	n/a
Alamito Wash	Alamito Wash	Atascosa Highlands	2020	8/15/2020	Nest	UNK	Hackberry
Arivaca Lake (Random 108)	Arivaca Lake	Atascosa Highlands	2019	8/29/2019	Nest	S	Cottonwood
Arrieta Wash (Random 111)	Arrieta Wash	Atascosa Highlands	2019	8/14/2019	Nest	UNK	Mesquite
Dry Well 2 Dry Well (Random 104)	Fraguita Wash	Atascosa Highlands	2020	8/10/2020	Nest	UNK	Oak
Fresnal Wash North	Fresnal Wash	Atascosa Highlands	2018	8/17/2018	Nest	UNK	Hackberry
Fresnal Wash	Fresnal Wash	Atascosa Highlands	2020	8/22/2020	Nest	F	Hackberry
Rock Corral Canyon	Rock Corral Canyon	Atascosa Highlands	2018	8/11/2018	Juvenile	S	n/a
Sycamore Canyon	Sycamore Canyon	Atascosa Highlands	2018	9/1/2018	Nest	UNK	Yew-leaf Willow
Brown Canyon - Lower	Lower Brown Canyon	Baboquivari Mountains	2018	8/19/2018	Nest	S	Hackberry

Cherry Creek	Cherry Creek	Canelo Hills	2018	8/15/2019	Nest	UNK	Oak
Jones Canyon	Jones Canyon - Parker Canyon tributary	Canelo Hills	2019	8/12/2019	Nest	UNK	Oak
Lyle Canyon	Lyle Canyon	Canelo Hills	2018	8/2/2018	Distraction display	UNK	n/a
Lyle Canyon	Lyle Canyon	Canelo Hills	2019	9/3/2019	Juvenile	S	n/a
O'donnell Canyon	O'donnell Canyon	Canelo Hills	2019	8/6/2019	Fledgling	S	n/a
Parker Canyon / Parker Canyon 2 (Random 210)	Parker Canyon	Canelo Hills	2019	8/12/2019	Copulation	UNK	n/a
Casa Arroyo	Sonoita Creek - Unnamed Tributary	Canelo Hills	2020	7/30/2020	Nest	S	Oak
Halfmoon Ranch	Stronghold Canyon East	Dragoon Mountains	2019	7/31/2019	Fledgling	S	n/a
Goldbaum Canyon	Goldbaum Canyon	Patagonia Mountains	2018	9/2/2018	Fledgling	S	n/a
Harshaw Creek Nest 1	Harshaw Creek	Patagonia Mountains	2018	7/23/2018	Nest	UNK	Mesquite
Harshaw Creek Nest 1 2	Harshaw Creek	Patagonia Mountains	2019	8/5/2019	Nest	UNK	Oak
SE of Red Mtn/Lead Queen Nest 2 (Random 166)	Lead Queen (Unnamed Canyon) - Harshaw Creek Complex	Patagonia Mountains	2020	7/27/2020	Nest	UNK	Oak
SE of Red Mtn/Lead Queen Nest 1 (Random 166)	Lead Queen (Unnamed Canyon) - Harshaw Creek Complex	Patagonia Mountains	2019	8/5/2019	Nest	UNK	Juniper
Willow Springs Canyon	Willow Springs Canyon	Patagonia Mountains	2018	8/3/2018	Nest	UNK	Oak
Chiminea Canyon - Upper (Random 44)	Chiminea Canyon	Rincon Mountains	2019	7/30/2019	Fledgling	S	n/a
Rincon Creek - North	Rincon Creek	Rincon Mountains	2019	7/30/2019	Nest	UNK	Acacia
Italian Trap (Random 36)	Tanque Verde Creek	Rincon Mountains	2019	8/6/2019	Nest	UNK	Hackberry
Tanque Verde Wash - La Cebadilla	Tanque Verde Creek	Rincon Mountains	2018	8/16/2018	Nest	UNK	Mesquite
Peppersauce Canyon	Peppersauce Canyon	Santa Catalina Mountains	2018	8/23/2018	Nest	UNK	Oak

Adobe Canyon 1	Adobe Canyon	Santa Rita Mountains	2019	8/14/2019	Copulation	UNK	n/a
Adobe Canyon 2	Adobe Canyon	Santa Rita Mountains	2019	8/18/2019	Fledgling	S	n/a
Box Canyon 3 - Upper	Box Canyon	Santa Rita Mountains	2019	7/3/2019	Nest	UNK	Oak
Box Canyon 4 - Upper	Box Canyon	Santa Rita Mountains	2019	7/25/2019	Nest	UNK	Juniper
Box Canyon 5 - Upper	Box Canyon	Santa Rita Mountains	2019	7/24/2019	Nest	UNK	Oak
Box Canyon 1 - Lower	Box Canyon	Santa Rita Mountains	2018	7/29/2018	Nest	UNK	Hackberry
Box Canyon 2 - Lower	Box Canyon	Santa Rita Mountains	2018	7/26/2018	Nest	UNK	Juniper
Chino Canyon	Chino Canyon	Santa Rita Mountains	2018	7/29/2018	Nest	UNK	Mesquite
Florida Canyon 1	Florida Canyon	Santa Rita Mountains	2018	7/27/2018	Nest	UNK	Hackberry
Florida Canyon 2	Florida Canyon	Santa Rita Mountains	2018	8/23/2018	Nest	UNK	Hackberry
Gardner Canyon 1	Gardner Canyon	Santa Rita Mountains	2018	8/2/2018	Copulation	UNK	n/a
Gardner Canyon 2	Gardner Canyon	Santa Rita Mountains	2019	8/15/2019	Nest	S	Ash
Madera Canyon - Proctor Road	Madera Canyon	Santa Rita Mountains	2019	7/24/2019	Nest	UNK	Hackberry
Montosa Canyon 2	Montosa Canyon	Santa Rita Mountains	2020	8/16/2020	Nest	UNK	Hackberry
Smith Canyon	Smith Canyon	Santa Rita Mountains	2018	8/17/2018	Nest	UNK	Oak
Squaw Gulch (Random 124)	Squaw Gulch	Santa Rita Mountains	2019	8/7/2019	Nest	UNK	Oak
Stevens Canyon	Stevens Canyon	Santa Rita Mountains	2019	7/29/2019	Fledgling	S	n/a
Temporal Gulch 1 - Upper	Temporal Gulch	Santa Rita Mountains	2018	8/9/2018	Nest	UNK	Oak
Temporal Gulch 2 - Middle	Temporal Gulch	Santa Rita Mountains	2018	7/27/2018	Nest	UNK	Ash
Temporal Gulch 3 - Lower (Random 125)	Temporal Gulch	Santa Rita Mountains	2019	7/14/2019	Nest	UNK	Oak

W. Sawmill Canyon N 1 (Random 73)	W. Sawmill Canyon	Santa Rita Mountains	2019	7/20/2019	Juvenile	S	Mesquite
W. Sawmill Canyon S 2 (Random 73)	W. Sawmill Canyon	Santa Rita Mountains	2019	7/28/2019	Nest	F	Hackberry

[a]S=successful, F=failed, UNK=unknown

Supplemental Materials Table 1.2 List of Yellow-billed Cuckoo survey results collected from 163 sites in southeastern Arizona xeroriparian habitat, 2013–2020. In sites with multiple years of data, the highest recorded detection, occupancy, or breeding status is reported. In the last five columns, 0=no, 1=yes.

Site	Drainage	Range	Surveyor	Years Surveyed	Detection	Occupied	Breeding detected	Random Site	Occupancy Validation
Harshaw-Flux West	Alum Gulch	Patagonia Mountains	ACS	2018, 2019	1	0	0	0	0
Flux Canyon	Flux Canyon	Patagonia Mountains	ACS	2018	1	0	0	0	0
Flux Canyon Road	Flux Canyon and unnamed wash	Patagonia Mountains	ACS	2019	1	0	0	0	0
Harshaw Creek-FR4701	Harshaw Creek Complex	Patagonia Mountains	ACS	2018	0	0	0	0	0
Humboldt Canyon	Humboldt Canyon - Alum Gulch Complex	Patagonia Mountains	ACS	2018, 2019	0	0	0	0	0
Harshaw-FR 4701 East	unnamed wash near Humbolct C	Patagonia Mountains	ACS	2019	0	0	0	0	0
FR215 and Flux Canyon	Alum Gulch Complex	Patagonia Mountains	ACS	2019	1	1	0	0	0
Harshaw-Flux East	Alum Gulch Complex	Patagonia Mountains	ACS	2018, 2019	1	1	0	0	0
Harshaw Creek-FR4701 West	Harshaw Creek Complex	Patagonia Mountains	ACS	2019	1	1	0	0	0
La Osa Wash	La Osa Wash	Atascosa Highlands	AtoZ	2018	1	1	0	0	0

Telles Tank	Telles Tank	Canelo Hills	AWRR	2015, 2016, 2017, 2018	1	0	0	0	0
Lyle Canyon	Lyle Canyon	Canelo Hills	AWRR	2015, 2016, 2017, 2018, 2019, 2020	1	1	1	0	1
Lower Lyle Canyon (AWRR)	Lyle Canyon	Canelo Hills	AWRR	2016	1	1	0	0	0
Research Ranch HQ	O'Donnell Canyon	Canelo Hills	AWRR	2015, 2016, 2017,2018	1	1	0	0	0
Post Canyon	Post Canyon	Canelo Hills	AWRR	2016, 2017, 2018	1	1	0	0	0
Post Canyon (AWRR HQ)	Post Canyon	Canelo Hills	AWRR	2016	1	1	0	0	0
Vaughn Canyon	Vaughn Canyon	Canelo Hills	AWRR	2015, 2016, 2017, 2018	1	1	0	0	0
Fourr Canyon	Fourr Canyon	Dragoon Mountains	Moore	2017	0	0	0	0	0
Grapevine Canyon	Grapevine Canyon	Dragoon Mountains	Moore	2017	0	0	0	0	0
Jordan Canyon	Jordan Canyon	Dragoon Mountains	Moore	2017	0	0	0	0	0
Kerwin Canyon	Kerwin Canyon	Dragoon Mountains	Moore	2017	0	0	0	0	0
Noonan Canyon	Noonan Canyon	Dragoon Mountains	Moore	2017	0	0	0	0	0
Northfork Noonan Canyon	Noonan Canyon	Dragoon Mountains	Moore	2017	0	0	0	0	0
Upper Slavin Gulch	Slavin Gulch	Dragoon Mountains	Moore	2017	0	0	0	0	0
Stronghold East	Stronghold Canyon East	Dragoon Mountains	Moore	2017	0	0	0	0	0
Stronghold East Campground	Stronghold Canyon East	Dragoon Mountains	Moore	2017	0	0	0	0	0
Lower Stronghold West	Stronghold Canyon West	Dragoon Mountains	Moore	2017	0	0	0	0	0
Upper Stronghold West	Stronghold Canyon West	Dragoon Mountains	Moore	2017, 2020	0	0	0	0	0
Lower Slavin Gulch	Slavin Gulch	Dragoon Mountains	Moore	2017	1	1	0	0	0

Madrona Canyon	Madrona Canyon	Rincon Mountains	SAGU	2019, 2020	0	0	0	0	1
Box Canyon (Rincon)	Box Canyon	Rincon Mountains	SAGU	2019	0	0	0	0	0
Chiminea Canyon - Lower	Chiminea Canyon	Rincon Mountains	SAGU	2019	0	0	0	0	0
Ruiz Trail	Coyote Wash	Rincon Mountains	SAGU	2019, 2020	1	0	0	0	0
Rincon Creek - Sentinel Butte	Rincon Creek	Rincon Mountains	SAGU	2019, 2020	1	0	0	0	0
Rincon Creek Grotto	Rincon Creek	Rincon Mountains	SAGU	2020	1	0	0	0	0
Rincon Creek - North	Rincon Creek	Rincon Mountains	SAGU	2019, 2020	1	1	1	0	1
Ash Creek / Happy Valley (Random 52)	Ash Creek	Rincon Mountains	SAGU	2019	0	0	0	1	0
Rincon Creek (Random 47)	Rincon Creek	Rincon Mountains	SAGU	2019	0	0	0	1	0
Italian Trap (Random 36)	Tanque Verde Creek	Rincon Mountains	SAGU	2017, 2019	1	1	1	1	1
Chiminea Canyon - Upper (Random 44)	Chiminea Canyon	Rincon Mountains	SAGU	2019	1	1	1	1	0
Paige Creek (Random 48)	Paige Creek	Rincon Mountains	SAGU	2016, 2019	1	1	0	1	0
Las Guijas Wash	Las Guijas Wash	Altar Valley	SSP	2018, 2019	1	1	1	0	1
Brown Wash	Brown Wash	Altar Valley	SSP	2018	1	1	0	0	1
Vineyard Bosque	Arroyo Sasabe - Unnamed Tributary	Altar Valley	SSP	2018	1	1	0	0	0
Rock Corral Canyon	Rock Corral Canyon	Atascosa Highlands	SSP	2015, 2018	1	1	1	0	1
Sycamore Canyon	Sycamore Canyon	Atascosa Highlands	SSP	2015, 2018	1	1	1	0	1
Alamito Wash	Alamito Wash	Atascosa Highlands	SSP	2020	1	1	1	0	0
Fresnal Wash North	Fresnal Wash	Atascosa Highlands	SSP	2018, 2020	1	1	1	0	0

Arivaca Lake (Random 108)	Arivaca Lake	Atascosa Highlands	SSP	2015, 2019	1	1	1	1	1
Fresnal Wash South (Random 133)	Fresnal Wash	Atascosa Highlands	SSP	2018, 2019	1	1	1	1	0
Dry Well (Random 104)	Fraguita Wash	Atascosa Highlands	SSP	2020	1	1	1	1	1
Brown Canyon - Upper	Brown Canyon	Baboquivari Mountains	SSP	2018	1	1	0	0	0
Cherry Creek	Cherry Creek	Canelo Hills	SSP	2016, 2018	1	1	1	0	1
Jones Canyon	Jones Canyon - Parker Canyon tributary	Canelo Hills	SSP	2019	1	1	1	0	0
O'Donnell Canyon	O'Donnell Canyon	Canelo Hills	SSP	2019	1	1	1	0	0
Casa Arroyo	Sonoita Creek - Unnamed Tributary	Canelo Hills	SSP	2020	1	1	1	0	0
Lone Mountain Canyon (Random 214)	Lone Mountain Canyon	Canelo Hills	SSP	2019	0	0	0	1	0
Parker Canyon/ Park Canyon 2 (Random 210)	Parker Canyon	Canelo Hills	SSP	2019	1	1	1	1	1
O'Donnell Canyon Trib/N Pauline Cyn (Random 169)	O'Donnell Canyon - Unnamed Tributary	Canelo Hills	SSP	2019	1	1	0	1	0
Parker Canyon/ Park Canyon 1 (Random 208)	Parker Canyon	Canelo Hills	SSP	2020	1	1	0	1	0
Halfmoon Ranch, Stronghold Canyon	Stronghold Canyon	Dragoon Mountains	SSP	2019	1	1	1	0	0
Hunter Canyon	Hunter Canyon	Huachuca Mountains	SSP	2015, 2019	1	0	0	0	1
Ramsey Canyon (Random 204)	Ramsey Canyon	Huachuca Mountains	SSP	2015, 2020	1	1	0	1	0
Corral Canyon (FS 58)- San Rafael	Corral Canyon - Harshaw Creek Complex	Patagonia Mountains	SSP	2013, 2018	1	1	1	0	0

Valley Rd / Harshaw Creek									
Goldbaum Canyon	Goldbaum Canyon - Harshaw Creek Complex	Patagonia Mountains	SSP	2013, 2017, 2018, 2019	1	1	1	0	0
Harshaw Creek Nest 1	Harshaw Creek	Patagonia Mountains	SSP	2018	1	1	1	0	0
Harshaw Creek Nest 2	Harshaw Creek	Patagonia Mountains	SSP	2019	1	1	1	0	0
Cumero Canyon (Random 186)	Cumero Canyon	Patagonia Mountains	SSP	2020	0	0	0	1	0
Lead Queen Canyon (Random 166)	Lead Queen Canyon - Harshaw Creek Complex	Patagonia Mountains	SSP	2018, 2019, 2020	1	1	1	1	0
Tanque Verde Wash - La Cebadilla	Tanque Verde Creek	Rincon Mountains	SSP	2018	1	1	1	0	1
Bear Canyon (Random 21)	Bear Canyon	Santa Catalina Mountains	SSP	2019	0	0	0	1	1
Box Canyon - Upper	Box Canyon	Santa Rita Mountains	SSP	2015, 2017, 20192020	1	1	1	0	1
Florida Canyon	Florida Canyon	Santa Rita Mountains	SSP	2015, 2017, 2018	1	1	1	0	1
Gardner Canyon	Gardner Canyon	Santa Rita Mountains	SSP	2018, 2019	1	1	1	0	1
Smith Canyon	Smith Canyon	Santa Rita Mountains	SSP	2017, 2018, 2020	1	1	1	0	1
Stevens Canyon	Stevens Canyon	Santa Rita Mountains	SSP	2019	1	1	1	0	1
Adobe Canyon	Adobe Canyon	Santa Rita Mountains	SSP	2019, 2020	1	1	1	0	0
Box Canyon - Lower	Box Canyon	Santa Rita Mountains	SSP	2018	1	1	1	0	0
Madera Canyon (Proctor Road)	Madera Canyon	Santa Rita Mountains	SSP	2015, 2019, 2020	1	1	1	0	0
Montosa Canyon	Montosa Canyon	Santa Rita Mountains	SSP	2015, 2020	1	1	1	0	0

Temporal Gulch 1 - Upper	Temporal Gulch	Santa Rita Mountains	SSP	2018	1	1	1	0	0
Temporal Gulch 2 - Middle	Temporal Gulch	Santa Rita Mountains	SSP	2018	1	1	1	0	0
Big Casa Blanca Canyon (Random 99)	Casa Blanca Canyon - Unnamed Tributary	Santa Rita Mountains	SSP	2019	0	0	0	1	0
Squaw Gulch (Random 124)	Squaw Gulch	Santa Rita Mountains	SSP	2019	1	1	1	1	1
W Sawmill Canyon (Random 73)	Sawmill Canyon	Santa Rita Mountains	SSP	2019	1	1	1	1	0
Temporal Gulch 3 - Lower (Random 125)	Temporal Gulch	Santa Rita Mountains	SSP	2019	1	1	1	1	0
Canoa Wash 1 & 2	Canoa Wash	Altar Valley	SSP	2018	1	1	1	0	1
Cuadro Wash	Cuadro Wash	Altar Valley	SSP	2018	1	1	1	0	1
Santa Margarita Wash - Lower	Santa Margarita wash	Altar Valley	SSP	2018	1	1	0	0	0
Carpenter Tank	Lopez Wash	Atascosa Highlands	SSP	2018	0	0	0	0	1
Arrieta Wash (Random 111)	Arrieta Wash	Atascosa Highlands	SSP	2019	1	1	1	1	1
Kitt Peak	N/A	Baboquivari Mountains	SSP	2018	0	0	0	0	1
Brown Canyon - Lower	Brown Canyon	Baboquivari Mountains	SSP	2018	1	1	1	0	1
N Tributary /N of Cherry Creek (Random 191)	Cherry Creek - Unnamed tributary	Canelo Hills	SSP	2019	0	0	0	1	0
Willow Springs Canyon	Willow Springs Canyon - Harshaw Creek Complex	Patagonia Mountains	SSP	2013, 2018, 2019	1	1	1	0	0
Mariposa Canyon	Mariposa Canyon	Atascosa Highlands	TAS	2017	1	0	0	0	0
Pena Blanca Canyon	Pena Blanca Canyon	Atascosa Highlands	TAS	2015	1	0	0	0	0

Bellatosa Canyon	Bellatosa Canyon	Atascosa Highlands	TAS	2017	1	1	1	0	0
Pena Blanca Lake	Pena Blanca Canyon	Atascosa Highlands	TAS	2015	1	1	1	0	0
Pesquiera Canyon	Pesquiera Canyon	Atascosa Highlands	TAS	2017	1	1	0	0	0
Sardinia Canyon - 1	Sardinia Canyon	Atascosa Highlands	TAS	2017	1	1	0	0	0
Sardinia Canyon - 2	Sardinia Canyon	Atascosa Highlands	TAS	2017	1	1	0	0	0
Sycamore Canyon (Patagonia)	Sycamore Canyon	Atascosa Highlands	TAS	2015	1	1	0	0	0
Alamo Canyon	Alamo Canyon	Canelo Hills	TAS	2016	0	0	0	0	0
Collins Canyon	Collins Canyon	Canelo Hills	TAS	2015	1	1	0	0	0
Dove Canyon	Dove Canyon	Canelo Hills	TAS	2016	1	1	0	0	0
Korn Canyon	Korn Canyon	Canelo Hills	TAS	2015	1	1	0	0	0
Upper Lyle Canyon	Lyle Canyon	Canelo Hills	TAS	2015	1	1	0	0	0
Merritt Canyon	Merritt Canyon	Canelo Hills	TAS	2015	1	1	0	0	0
Cave Creek Canyon - SWRS	Cave Creek Canyon	Chiricahua Mountains	TAS	2015	0	0	0	0	0
Cave Creek Canyon (1) - Stewart Campground	Cave Creek Canyon	Chiricahua Mountains	TAS	2015	0	0	0	0	0
Cave Creek Canyon (1) - Stewart Campground	Cave Creek Canyon	Chiricahua Mountains	TAS	2017	0	0	0	0	0
Cave Creek Canyon (3) - Upper Canyon	Cave Creek Canyon	Chiricahua Mountains	TAS	2017	0	0	0	0	0
South Fork Cave Creek Canyon (2)	Cave Creek Canyon	Chiricahua Mountains	TAS	2015, 2017	0	0	0	0	0
Pine Canyon	Pine Canyon	Chiricahua Mountains	TAS	2017	0	0	0	0	0
Pinery Canyon	Pinery Canyon	Chiricahua Mountains	TAS	2017	1	0	0	0	0

East Turkey Creek	Turkey Creek	Chiricahua Mountains	TAS	2017	0	0	0	0	0
West Turkey Creek	West Turkey Creek	Chiricahua Mountains	TAS	2017	0	0	0	0	0
Carr Canyon	Carr Canyon	Huachuca Mountains	TAS	2015	0	0	0	0	0
Copper Canyon	Copper Canyon	Huachuca Mountains	TAS	2016	1	0	0	0	0
Ida Canyon	Ida Canyon	Huachuca Mountains	TAS	2016	0	0	0	0	0
Bear Canyon	Bear Canyon	Huachuca Mountains	TAS	2016	1	1	0	0	0
Miller Canyon	Miller Canyon	Huachuca Mountains	TAS	2015, 2018	1	1	0	0	0
Washington Gulch	Washington Gulch	Patagonia Mountains	TAS	2015	1	1	0	0	0
Turkey Creek	Turkey Creek	Rincon Mountains	TAS	2016	0	0	0	0	0
Campo Bonito	Bonito Canyon	Santa Catalina Mountains	TAS	2017, 2018, 2019	1	0	0	0	1
Alder Canyon	Alder Canyon	Santa Catalina Mountains	TAS	2017	0	0	0	0	0
Southern Belle Mine	Bonito Canyon	Santa Catalina Mountains	TAS	2018, 2019	1	0	0	0	0
Sabino Canyon	Sabino Canyon	Santa Catalina Mountains	TAS	2015	0	0	0	0	0
Stratton Canyon	Stratton Canyon	Santa Catalina Mountains	TAS	2017	0	0	0	0	0
Peppersauce Canyon	Peppersauce Canyon	Santa Catalina Mountains	TAS	2015, 2016, 2017, 2018, 2019	1	1	1	0	1
Mansfield Canyon	Mansfield Canyon	Santa Rita Mountains	TAS	2017	1	1	0	0	0
French Joe Canyon	French Joe Canyon	Whetstone Mountains	TAS	2015, 2016, 2018, 2019	1	1	0	0	0

Guindani Canyon	Guindani Canyon	Whetstone Mountains	TAS	2015, 2016, 2018, 2019	1	1	0	0	0
East South Fork Cave Creek Canyon	Cave Creek Canyon	Chiricahua Mountains	USFS	2020	1	0	0	0	0
West South Fork Cave Creek Canyon	Cave Creek Canyon	Chiricahua Mountains	USFS	2020	1	0	0	0	0
Temporal Gulch 2 - Middle / Mansfield East	Temporal Gulch	Santa Rita Mountains	USFS	2020	1	1	0	0	0
Corral Canyon (5) - N Tributary	Corral Canyon	Patagonia Mountains	WestLand	2013	1	0	0	0	0
Corral Canyon (6) - NW Tributary	Corral Canyon	Patagonia Mountains	WestLand	2013	0	0	0	0	0
Endless Chain Canyon	Endless Chain Canyon	Patagonia Mountains	WestLand	2015	0	0	0	0	0
Upper Harshaw Creek	Harshaw Creek	Patagonia Mountains	WestLand	2017	0	0	0	0	0
Hermosa Hill	Harshaw Creek - Unnamed Tributary	Patagonia Mountains	WestLand	2016, 2017, 2018, 2019, 2020	1	0	0	0	0
Lower Alum Gulch	Alum Gulch	Patagonia Mountains	WestLand	2017, 2018, 2019	1	1	0	0	0
Corral Canyon	Corral Canyon	Patagonia Mountains	WestLand	2013, 2018, 2019	1	1	0	0	0
Lower Finley & Adams Canyon	Finley & Adams Canyon	Patagonia Mountains	WestLand	2015, 2017, 2018, 2019	1	1	0	0	0
Upper Finley & Adams Canyon	Finley & Adams Canyon	Patagonia Mountains	WestLand	2017, 2018, 2019	1	1	0	0	0
Harshaw C-FS 4701 to Guajolote Flat	Harshaw Creek	Patagonia Mountains	WestLand	2017, 2018, 2019, 2020	1	1	0	0	0
Harshaw Creek (1, 2)	Harshaw Creek	Patagonia Mountains	WestLand	2013, 2016, 2017, 2018, 2019, 2020	1	1	0	0	0
Basin Mine	Harshaw Creek - Unnamed Tributary	Patagonia Mountains	WestLand	2018, 2019, 2020	1	1	0	0	0
Great Silver Mine	Harshaw Creek - Unnamed Tributary	Patagonia Mountains	WestLand	2018, 2019, 2020	1	1	0	0	0

Unnamed Canyon E of Hermosa	Harshaw Creek - Unnamed Tributary	Patagonia Mountains	WestLand	2017, 2018, 2019	1	1	0	0	0
Hermosa Canyon	Hermosa Canyon	Patagonia Mountains	WestLand	2013, 2017, 2018, 2019, 2020	1	1	0	0	0
Humboldt and Upper Alum	Humboldt Canyon	Patagonia Mountains	WestLand	2016, 2017, 2018, 2019, 2020	1	1	0	0	0
East Mowry Wash	Mowry Wash	Patagonia Mountains	WestLand	2019	1	1	0	0	0
West Mowry Wash	Mowry Wash	Patagonia Mountains	WestLand	2018, 2019	1	1	0	0	0
Paymaster Canyon - Flying R Ranch	Paymaster Canyon	Patagonia Mountains	WestLand	2015, 2016, 2018, 2019	1	1	0	0	0
Lower Sycamore Canyon	Sycamore Canyon	Patagonia Mountains	WestLand	2017, 2018, 2019	1	1	0	0	0
Upper Sycamore Canyon	Sycamore Canyon	Patagonia Mountains	WestLand	2017, 2018, 2019	1	1	0	0	0
Rosemont Springs	Barrel Canyon	Santa Rita Mountains	WestLand	2013, 2014, 2018	0	0	0	0	0
Lower Barrel Canyon	Barrel Canyon	Santa Rita Mountains	WestLand	2013, 2014, 2015, 2017, 2018	1	1	0	0	0
Upper Barrel Canyon	Barrel Canyon	Santa Rita Mountains	WestLand	2013, 2014, 2015, 2017, 2018	1	1	0	0	0
McCleary Canyon	McCleary Canyon	Santa Rita Mountains	WestLand	2013, 2014, 2016, 2017, 2018	1	1	0	0	0
Wasp Canyon	Wasp Canyon	Santa Rita Mountains	WestLand	2013, 2014, 2015, 2016, 2017, 2018	1	1	0	0	0

Chapter 2 Increasing classifier efficiency for acoustic monitoring of rare and cryptic species: using Kaleidoscope to detect Yellow-billed Cuckoos

ABSTRACT

Use of autonomous recording units for monitoring birds has increased dramatically in recent years, along with improvements in signal classification software designed to semi-autonomously process data and isolated target signals. Classifier performance varies among studies, however, indicating a need for improved methods in classifier development. We (Northern Arizona University, in conjunction with U.S. Fish and Wildlife Service and U.S. Geological Survey) reviewed studies that use a popular software platform, Kaleidoscope Pro, to develop classifiers for birds. We identified inconsistencies in methods and data reporting and used this information to guide development of a classifier for western Yellow-billed Cuckoos, demonstrating a reiterative approach to increasing classifier performance. Our literature review yielded 13 studies and revealed highly variable methodology and data reporting, with generally poor classifier performance. Our classifier demonstration relied on multiple iterations of manual review and subsequent reprocessing of data to increase true positive rates. This resulted in precision increases of 0.2 to 0.42 and recall increases of 0.36 to 0.60, with a top F-score of 0.52 (β=1.5). Our study illustrates a novel approach to improving classifier performance which can be readily adapted to other scenarios.

INTRODUCTION

Acoustic monitoring of birds using autonomous recording units (ARUs) is a technique increasingly used by wildlife researchers and managers (Shonfield & Bayne 2017, Gibb et al.

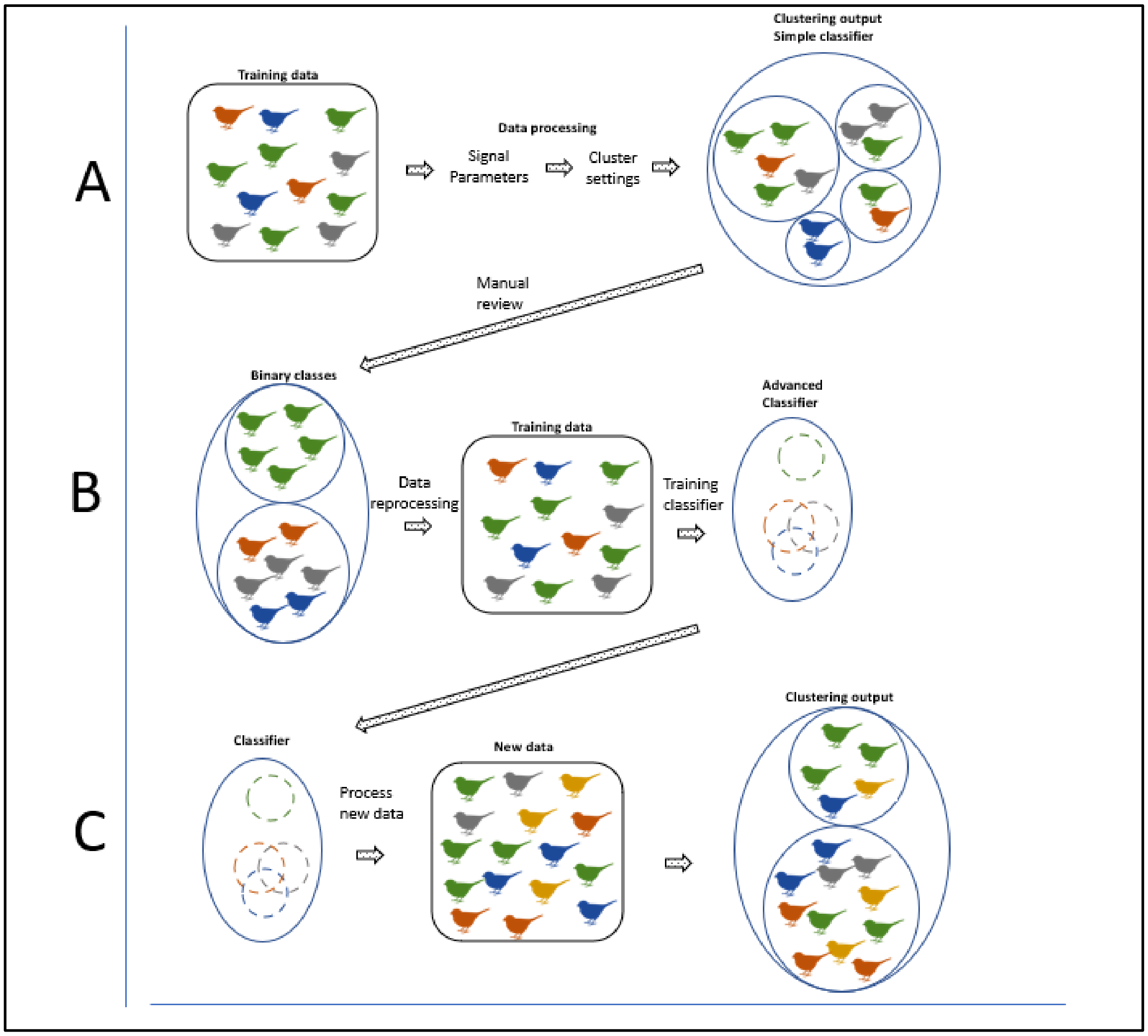
2019) and may be especially useful for monitoring rare or threatened species (Bobay et al. 2018, Schroeder & McRae 2020, Marchal et al. 2022). ARUs offer advantages over traditional human-observer survey methods by allowing large amounts of data to be recorded over long durations, and by storing those data as a permanent record that can be subsequently reviewed and verified. In avian research, ARUs have been demonstrated to be as effective as traditional human-observer point counts (Hutto & Stutzman 2009), and useful for estimating bird occupancy (Abrahams & Geary 2020), density (Pérez-Granados & Traba 2021), and richness (Bateman & Uzal 2022). While application of ARU technology has rapidly increased, challenges remain for managers in determining the most effective methods for efficient processing of acoustic data once it is acquired. This is particularly true of studies seeking to identify target signals (e.g., specific songs or calls) in large acoustic datasets (Gibb et al. 2019). Such studies typically require the development and use of a classifier (alternatively called a "recognizer"), which consists of a statistical model of target signal characteristics and other competing non-target signals that can be used to scan acoustic data and automatically isolate statistically similar signals (Bardeli et al. 2010). Classifiers can be developed using many software platforms currently available, offering a range of features and applications. Most software options use a combination of user-defined signal parameters (e.g., signal frequency range and signal duration) to pre-filter undesired acoustic signals and machine learning methods to scan data and sort detected signals based on statistical similarities. In most cases, the process is semi-automated, with users manually reviewing results and adjusting software settings to maximize true positives ("hits") and minimize false negatives ("misses") and false positives. Each available platform uses a different statistical framework and workflow to develop classifiers, but all require users to supervise the process by providing acoustic data inputs, parameter selection, and reviewing classifier output (Knight et al. 2017, Priyadarshani et al. 2018).

Standardized methods for the assessment of classifier performance have been established and adapted for use in acoustic analysis (Sokolova et al. 2006, Sokolova & Lapalme 2009, Potamitis et al. 2014, Knight et al. 2017). This provides an important framework for evaluating classifier performance based on project goals as well as comparing results across studies. Important among these metrics are recall (the proportion of target signals that were accurately classified as true positives), precision (the proportion of detections classified as the target signal that are true positives), and F-score (see formula in Methods) which incorporates both recall and precision, and allows users to assess classifiers by alternately prioritizing either recall or precision (Knight et al. 2017). Recall (and therefore F-score) requires a known rate of false negatives. For small datasets, an exact false negative of a target signal rate may be determined by auditorily listening back to data and/or visually scanning spectrograms, and then comparing the total number of target signals in the dataset to those detected by the classifier. In larger datasets, full manual review of data is often impractical due to the time investment required, and false negatives instead can be estimated by using a manually verified "benchmark" dataset (Knight et al. 2017), which is a subset of manually-processed acoustic data (often randomly selected) with all target-signals identified and annotated. Although classifier performance assessment is important for interpretation of results, many studies fail to report these metrics or the formulae used to calculate them, making comparison across studies difficult (Knight et al. 2017).

Although increased clarity in classifier performance reporting is necessary across software platforms, an additional avenue for improving classifier performance is to more thoroughly investigate techniques for classifier development in individual software programs. Of several available software platforms, Kaleidoscope Pro (Wildlife Acoustics, Inc) is an increasingly popular choice for avian researchers and managers (Knight et al. 2017, Priyadarshani et al. 2018), as it provides a user-friendly interface and user guides, and requires little prior knowledge of machine learning techniques. This software includes several workflow options (Figure 2.1) for classifier development, all based on cluster analysis methods and Hidden Markov Models (HMMs). A “simple” classifier can be developed by processing data through user defined signal parameters that filter the acoustic data and cluster analysis settings that influence the statistical models. The software then sorts detected signals into multiple “clusters” based on statistical similarities (Figure 2.1A). The simple classifier is the resulting statistical model based on the cluster analysis of the input files. An “advanced” classifier goes a step further and is developed by manually reviewing detected signals from the initial cluster analysis output, and manually reassigning all target signals into the same cluster, and non-target signals into one or more other manually defined clusters. The data is then reprocessed, thereby restructuring the statistical model of the classifier based on these manual edits (Figure 2.1B). Both simple and advanced classifiers can be used to sort detections in new data (Figure 2.1C).

The Kaleidoscope workflow also typically includes the use of a training dataset to develop classifiers, which is a subset of data containing examples of target signals (often including examples acquired from other “library” sources) as well as “background” data containing other signals present in the acoustic environment. These training data are used to help form the statistical models by which the classifier sorts target signals from non-target signals in new datasets (as in Figure 2.1C). Training data are used to improve classifier development efficiency and are most appropriate for users applying classifiers to larger datasets or seeking to identify uncommon or rare target signals.

Figure 2.1 Conceptual diagram of three modes of operation in Kaleidoscope Pro. Different color birds represent different types of signals in an acoustic environment; green birds represent target signals. Mode A processes data and classifies detections into clusters based on statistical similarity. In Mode B, results of Mode A are manually reviewed and changed to a binary classification of target and non-target signals. Data is then reprocessed using these manual edits to produce a classifier containing the statistical algorithm. In Mode C, the classifier is used to process a new dataset, with detected signals being reported in one of the two clusters (target/non-target). Misclassification may occur, as demonstrated by three species occurring in the target cluster output.



As with most classifier development platforms, studies have shown mixed results in the use of Kaleidoscope Pro's classifiers to accurately detect target signals (Knight et al. 2017, Marchal et al. 2022), especially for species that are rare or vocalize infrequently (Bobay et al. 2018, Schroeder & McRae 2020). Authors have reported high rates of false positives and uncertainty in data interpretation, particularly when false negative rates (and therefore recall scores) are unknown (Bobay et al. 2018). In such scenarios with rare species, recall may be a more important performance metric than precision, because with fewer calls there are fewer opportunities to detect an individual over a given sampling period. Increasing recall, however, may decrease precision scores due to higher rates of false positives, requiring additional time to manually review detections. For rare or threatened species, this additional cost in time may be worth the increased ability to detect the presence of the species of interest.

Poor classifier performance may in part be attributed to characteristics of both the target signals and background signals in acoustic data being processed, as these characteristics can influence the detectability of target signals (Buxton & Jones 2012, Sidie-Slettedahl et al. 2015, Priyadarshani et al. 2018). However, a lack of literature documenting the use of established performance metrics to evaluate and guide classifier development for specific software platforms may also partially account for inconsistent results across studies. To investigate how refinement of the classifier development process could improve classifier performance, especially for monitoring rare and cryptic species, we (Northern Arizona University, in conjunction with U.S. Fish and Wildlife Service and U.S. Geological Survey) reviewed papers that used Kaleidoscope Pro and its classifiers to detect bird vocalizations from passively recorded acoustic data. We then applied the results of our review to the development of classifiers for the western Yellow-billed Cuckoo, a federally threatened distinct population segment (DPS) (USFWS 2014). This species exhibits low call rates and often goes undetected by human surveyors (Hamilton & Hamilton 1965, Halterman 2009), making it an ideal species for investigating the use of classifiers to detect a rare and cryptic species. Finally, we synthesize our findings into general workflow recommendations for Kaleidoscope classifier development, assessment, and application in avian research and management settings.

METHODS: LITERATURE REVIEW

We reviewed recent use of Kaleidoscope software in the scientific literature using Web of Science and searching for published peer-reviewed articles between 2014–2022 (as 2014 is when

Kaleidoscope first became available) using the following terms: “birds kaleidoscope”, “kaleidoscope classifier”, “kaleidoscope acoustic”, “acoustic birds”, and “kaleidoscope ARU”, “classifier ARU”. Our focus was on single-species classifiers typically used for studying species that are rare, threatened, or of other management concern, but we also included studies that either studied a single species but used data obtained through a multi-species classifier, or developed a classifier for three or fewer co-occurring species with similar vocalization characteristics. For each paper, we recorded 1) classifier type, 2) data inputs, 3) signal parameters and cluster analysis settings used, 4) evaluation of classifier performance and metrics used to assess performance, and 5) whether time invested to develop classifiers was reported.

RESULTS: LITERATURE REVIEW

I. Classifier type

In total, we identified 13 papers (Supplementary Materials Table 2.1) using Kaleidoscope Pro classifiers for the semi-automated detection of target bird vocalizations. Of these, 6 papers investigated a single species, while 5 papers investigated 2 or 3 species, with a total of 20 individual species being studied among all 13 papers. Eight papers developed a single-species classifier for each species within their study, with 9 classifiers total. Of these, 4 were binary-cluster advanced classifiers, 1 was a multi-cluster simple classifier with a single target cluster, and the remaining 4 (1 advanced and 3 simple) did not specify how many clusters were included in the classifier. Five papers used data derived from multi-species classifiers to investigate 11 individual species. Of these 5 papers, 1 developed a simple classifier with 55 clusters, 1 used an advanced classifier with 2 target species clusters and a single non-target cluster, and 3 did not specify whether classifiers were simple or advanced or the number of clusters.

II. Data inputs

Two papers did not report any information on data used for developing classifiers. Of the remaining 10 papers, data inputs varied between studies and were largely related to the specific research goals. Six papers used training to develop classifiers which were in turn used to process larger datasets, with training data ranging from several dozen call examples combined with 10 minutes of background data to several hundred hours of data. The remaining 4 papers typically processed entire acoustic datasets to sort signals without training an initial classifier.

III. Parameters and settings

Signal parameters are user-defined settings that act as a pre-filter on the acoustic data being processed, allowing only signals that meet specified criteria to be included in cluster analysis output. All signal parameters were reported in all papers except two papers which did not report any signal parameters for one classifier and another paper that did not report the maximum inter-syllable gap (ms) for two classifiers used (Supplemental Materials S2).

Cluster analysis settings were less frequently reported across studies we reviewed (Supplemental Materials S2). These settings affect the sensitivity of the statistical algorithms that cluster and report detected signals. Explanation and guidance for each of these settings is provided in the Kaleidoscope Pro user manual. The same two papers that did not report any signal parameters also did not report any cluster analysis settings. The most-often reported setting was the maximum distance to cluster center to include in .csv file, which was only omitted from 3 papers (5 classifiers). This setting determines how many signals within the acoustic data will be reported in final output. The fast-Fourier transform (FFT) window, max states, max distance from cluster center for building clusters, and max number of clusters were all reported in 5 papers (7 classifiers). The FFT window influences the resolution of signal frequency versus time, and the latter three settings influence the number of clusters formed and distribution of signals included in clusters.

IV. Performance metrics

Six papers did not report whether a benchmark dataset was used to assess classifier performance and were therefore assumed to have not used one. Of the remaining 7 papers, benchmark datasets ranged from a 2-hour subset of data to entire acoustic datasets.

Recall and precision were reported for 6 classifiers in 4 papers, and F-score was reported for only 2 classifiers in 2 papers (Supplemental Materials S3). The rate of false positives was reported for 8 classifiers in 6 papers. However, for one paper, false positives were calculated as the percentage of all detected signals among all clusters that were reported incorrectly as the target species, rather than the percentage of detections in the target species cluster that were not actually the target species. This resulted in extremely low false positive rates of 0.1% for each of two classifiers, as opposed to 24.75% and 8% when we calculated these rates with reported raw data using standard calculations. One paper reported species that most commonly accounted for false positives. False negative rates were reported for 10 classifiers in 7 papers, with methods for calculation varying between papers. Three papers used an arbitrarily chosen number of classifier detections from non-target signal clusters to manually review for false negatives. Notably, no performance metrics were reported for 7 classifiers in 4 papers.

V. Time investment

Only two papers reported specific information on time investment for classifier development. Knight et al. (2017) reported 9.6 minutes required to manually review each hour's worth of classifier detections for Kaleidoscope. They also reported a software "learning" time of eight hours and classifier build time of four hours and determined that use of Kaleidoscope became more efficient than human listening after 19 hours of data. Marchal et al. (2021) reported two hours required for each for detection (initial data processing) and classification (classifier training) step. Several other papers discussed the relative efficiency of acoustic data processing but did not provide specifics from their studies.

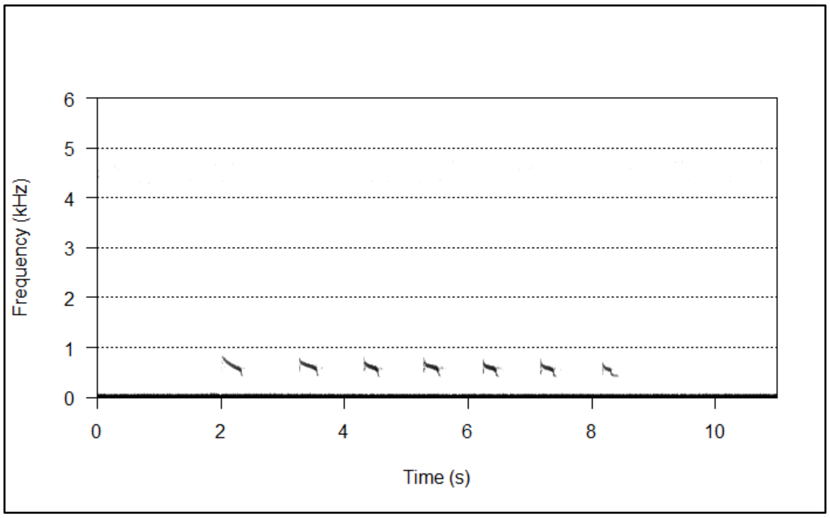
METHODS: YELLOW-BILLED CUCKOO CLASSIFIER DEMONSTRATION

I. Species & vocalizations

We chose the western DPS of the Yellow-billed Cuckoo (hereafter, "YBCU") as a model species to develop Kaleidoscope classifiers for several reasons. First, the YBCU is federally listed as threatened (USFWS 2014), making it a species of high conservation priority for much of western North America. Second, YBCU is a cryptic species and analysis of human-observer survey results has shown that approximately 90% of positive detections are aural-only (McNeil et al. 2013) meaning visual observation is rare during surveys. Finally, despite this large proportion of aural detections, YBCU vocalize infrequently relative to most other avian species (approximately once per hour; Halterman 2009), which would suggest the need for classifiers with high recall scores.

YBCU give several call types throughout the breeding season. We chose to use the "coo" call for our classifier demonstration (Figure 2.2). This call is given at a frequency of approximately 350–1000 Hz, a range that overlaps with species such as doves and owls but is generally lower than most other avian species. The call consists of a series of several individual "coo" notes separated by approximately 0.5–1 s. YBCU often produce this call in extended bouts ranging from several minutes to over an hour, with each series of coo notes usually separated by several seconds or more (Hughes 2015). Therefore, for the purposes of our study we define a “call” as a distinct series of coo notes, rather than the individual notes themselves.

Figure 2.2 Spectrogram of the Yellow-billed Cuckoo coo call, consisting of 7 individual coo notes.



II. Classifier workflow demonstration

We used the results of our review to aid in the development of a simple and readily adaptable framework for classifier development. Our approach was to first demonstrate an efficient technique for the use of a training dataset to develop a binary advanced classifier for the coo call, and then to apply this classifier on a larger test dataset, which represents a hypothetical dataset of interest. We then used a technique for refining the classifier and improving performance by iteratively reviewing and reprocessing the test data, thereby retraining the classifier for the dataset of interest multiple times. We used established classifier performance metrics to assess each classifier iteration and identify best performing classifiers (Knight et al. 2017). Kaleidoscope Pro employs a unique user interface with several workflow options for semi-automated data processing. Below we describe our process and refer to steps A–C illustrated in Figure 2.1.

III. Acoustic data collection

For training and test datasets, we used ARU data collected in 2018 and 2019 at sites in xeroriparian habitat in southeastern Arizona, as part of a broader study on YBCU in this region (Beauregard et al., in prep). Data was collected using Wildlife Acoustics SM4 and SM2+ units (Wildlife Acoustics, Inc., Maynard MA). ARUs were affixed to small-diameter trees at a height of approximately 2m from the ground. They were deployed a minimum of 500m apart from each other (to avoid detecting the same YBCU calls on multiple ARUs), and set to record continuously each day from 0400-1100, using a minimum sampling rate of 16 kHz (range: 16 kHz–24 kHz) at 16-bit resolution.

IV. Training data

We used a training dataset to demonstrate the development of an initial classifier using pre-selected "library" examples of YBCU coo calls. In this approach, library examples are combined with additional "background data" of other signals in the acoustic environment, which the classifier is subsequently trained to discriminate against. To obtain high quality examples of YBCU calls for our "library" we used RavenPro (version 1.6.4, Cornell University, Ithaca NY) to visually scan spectrograms of acoustic data from ARUs deployed near active nests or in known-occupied territories during 2018. We manually extracted complete examples of calls but excluded examples that were obstructed by vocalizations of other species. However, we included examples of faint calls and calls that were partially obstructed by some forms of anthrophony (drones, vehicle noise, etc.) and geophony (rain, wind, etc.). All YBCU call examples were exported into a single library folder as individual .wav files.

For background data, we used data collected in 2019 from 16 ARUs deployed across four sites (4 ARUs per site) known to be occupied by YBCU. These sites were different than 2018 sites in which we obtained YBCU library data but were in the same region and contained similar xeroriparian habitat. Importantly, background data came from the same sites as our test dataset (see below). We selected 4 random hours of data for each ARU between June 15 and September 15, resulting in 64 hours of acoustic data (16 hours per site). We manually reviewed spectrograms for all background data to ensure that no YBCU vocalizations were present. If YBCU vocalizations were present, we replaced that individual hour of data with another. All background data was exported into a single background data folder. This data preparation process ensured that our combined library and background training dataset consisted of a known quantity of call examples partitioned into a unique folder, and a known amount of background data containing no additional YBCU calls in its own unique folder (see Discussion for alternative approaches).

V. Test dataset

For test data, we randomly selected 4 full days (7 hours per day) of data from each ARU at each site (excluding days used for background data), resulting in 448 hours of data. Test data were manually reviewed by scanning spectrograms of all data in RavenPro. All signals that visually resembled the YBCU coo call were reviewed auditorily, and all verified calls were annotated with the date, time, and location. This resulted in a fully verified dataset with a known quantity of calls, allowing us to apply performance metrics and assess classifiers (see Performance Assessment below).

VI. Processing training data

In Kaleidoscope Pro 5, we first processed the training data using Mode A, which performs an initial cluster analysis and sorts detections into multiple distinct clusters based on statistical similarity. For initial data processing, we used the following parameters and settings: 350-900 Hz frequency range, 1–12 s signal duration, 1s maximum inter-syllable gap, 2.0 maximum distance from cluster center to include in cluster.csv, 21.33 ms FFT window, 12 max states, 0.5 maximum distance to cluster center for building clusters, and 500 max clusters. We chose these settings based on guidance provided in user manuals.

We then reviewed the results, which consisted of a list of detections, sorted by the software-assigned cluster, with other associated information such as time, date, and file location. Importantly, all YBCU library examples were contained within a single unique folder separate from background data, so we only needed to review the spectrograms of signals located in this folder because all other detections located in the background data folder were known to not be cuckoo vocalizations. To define a cluster for the coo call, we manually renamed the cluster name of high quality YBCU examples to "Coo" (regardless of their initial software-assigned cluster). For detected signals containing incomplete coo calls, or signals containing coo calls as well as non-target signals, we manually cleared the cluster name. This step removes these detection data from the classifier algorithm during a subsequent reprocessing of the data. To define a non-target signal cluster, we changed the names of all other detections to "Not Coo".

We then reprocessed the training data using Mode B. This step uses the manually defined cluster names to reanalyze the training data and restructure the statistical algorithms, resulting in a binary advanced "Coo/Not Coo" classifier. For simplicity, we hereafter refer to these clusters as "target signal" and "non-target signal" clusters for "coo" and "not coo", respectively.

VII. Test data processing

We processed the test dataset using Mode C, which uses the binary classifier (the .kcs output file from the previous step) to sort detected signals according to the predefined binary clusters. All detected signals classified in the target-signal cluster were manually reviewed, and all detections containing complete, unobstructed examples of target signals were left unedited, whereas incomplete and/or obstructed target signal examples were renamed “target signal - clear”. This indicated that the detection was a true positive (and therefore counted as such for performance assessment) but that its name should later be cleared for subsequent data reprocessing (see below). All false positives (non-target calls assigned to the target signal cluster) were manually reassigned to the non-target signal cluster.

We then manually reviewed a subset of the detected signals assigned to the non-target signal cluster equal to the number of detections originally reported in the target signal cluster (prior to review). We chose this method because the non-target cluster is meant to represent all other similar detected signals in the acoustic data and therefore contains multitudes more detections than the target signal cluster. While a full review of the non-target cluster would hypothetically reveal most (if not all) false negative target signals, it is an inefficient process and undermines the purpose of using a classifier to detect signals in large datasets. Therefore, we employed an “equal effort” approach for each cluster to balance time investment in data review. We reviewed detections in the non-target cluster in descending order according to the “Distance to cluster center” score, as it is the default output order of reported signals. For all non-target signals we reviewed (i.e., true negatives), we left the “non-target signal” cluster name unchanged. If target signals were observed in the non-target cluster (i.e., false negatives), we reassigned it by naming it either “target signal” or “target signal - clear” according to our previous description. We then cleared the cluster name of all remaining signals in the non-target cluster. This ensured that no target signals were mistakenly incorporated into the non-target cluster (by means of not being manually reviewed).

The final step in reviewing the results was to count the number of true-positive and false-positive detections in the target signal cluster, and then clearing the names of all calls that we labeled “target signal - clear”. The true-positive total was then subtracted from the total known calls in the test dataset to determine the number of false-negatives.

These edited results were then used in Mode B to reprocess the test dataset. This step forms a new classifier using only the manually named detections from the edited results of the test data, overwriting the classifier information from the training data. This restructures the classifier’s statistical algorithms to be based on the test data itself. This new classifier was then used in Mode C to reprocess the test data. In this case, Kaleidoscope is treating the test data as “new” data, even though the classifier’s algorithms are based on the test data itself (so it is therefore not actually new data). This has the effect of automatically detecting calls used to develop this classifier (as they are exact matches), while also using this information to find new coo calls that the first classifier missed. Therefore, the total number of true positives should not be lower than the number of examples manually identified as such in the previous review of results. We again reviewed the results, inspecting all detected signals in the target signal cluster and naming detections as either “target signal”, “target signal - clear”, or “non-target signal”, and reviewing a subset of non-target signal detections proportional to the detections in the target signal cluster. We repeated these steps of Modes B and C until the test dataset had been processed with 5 separate classifiers (1 based on the training data, 4 based on test data reprocessing).

VIII. Performance assessment

To assess the performance of all classifiers, we used methods outlined in Knight et al. (2017). We first calculated the number of true positives, false positives, and false negatives (as described above). We then used these totals to calculate precision, recall, and F-score of the two classifiers using the following formulae:

(1) Precision = tp/tp+fp

(2) Recall = tp/tp +fn

(3) F = (β^2+1) * Precision * Recall / β^2 * Precision + Recall

Where tp is the number of true positives (the number of coo calls accurately classified in the coo cluster), fp is the number of false positives (signals that are not coo calls classified as coo), and fn is the number of false negatives (the number of coo calls not accurately classified in the coo cluster). β is a user-defined metric that allows for the prioritization of either precision or recall based on specific analysis goals. Precision is favored when $\beta > 1$ and recall is favored when $\beta < 1$, and precision and recall are balanced when $\beta = 1$ (Sokolova et al. 2006). We calculated F-scores for $\beta = 0.5$, $\beta = 1$, and $\beta = 1.5$.

IX. Time investment

All classifier-based approaches require some human supervision of the process, which may vary based on software platform, amount of data, project goals, and methods used. To compare the time investment required to develop our classifier with that of other studies, we recorded the amount of time spent actively working on classifier development at each step of the process. This includes manually scanning spectrograms of acoustic data (for background data and the test/benchmark dataset) and reviewing results of software outputs. We did not record the amount of time it took for the software to process data, as this is an automated process requiring no human supervision, and previous studies have shown it to not require significantly more time than other software platforms (Knight et al. 2017, Marchal et al. 2022).

RESULTS: YELLOW-BILLED CUCKOO CLASSIFIER DEMONSTRATION

I. Classifier demonstration

The performance results of our classifier demonstration are summarized in Table 1. Our test dataset contained 638 manually identified coo call examples. When using the first classifier developed from the training data to process the test dataset, the classifier identified 229 coo vocalizations, with 911 false positives in the coo cluster. This resulted in a precision score of 0.20, a recall score of 0.36, and an F-score (β=1) of 0.26.

Our second classifier, developed by restructuring the classifier using results of the first classifier, detected 324 coo calls, a 15% increase over the first classifier. The second classifier also resulted in far fewer false positives (397). This classifier resulted in a precision score of 0.45, a recall score of 0.51, and an F-score (β=1) of 0.48.

Subsequent re-scans of the data continued to increase the number of true positive detections, reaching a peak of 328 true positives (Recall = 0.60) with the 4th classifier. Using the 5th classifier, one less true positive was detected than the 4th. This is possible because multiple examples were “cleared” from manual renaming as described in the methods, meaning fewer examples were used to form the subsequent classifier than were actually recorded as true positives (363 examples in the case of classifier 4).

F-score and precision did not demonstrate consistent improvement with each iteration due to higher proportions of false positives in classifiers 3 and 4, with a maximum of 1252 false positives in classifier 4. However, classifier 5 resulted in a decreasing false positive rate, thereby achieving our highest rates of recall and precision and our highest overall F-score (0.52 where β=1.5).

Table 2.1 Results of 5 iterations of Yellow-billed Cuckoo classifiers developed using a manually verified test dataset.

	Test Dataset	Initial Classifier	Classifier 2	Classifier 3	Classifier 4	Classifier 5
# True Positive	638	229	324	366	382	381
# False Positive	0	911	397	531	1252	533
# False Negative	0	409	314	272	256	257
Precision	1.00	0.20	0.45	0.41	0.23	0.42
Recall	1.00	0.36	0.51	0.57	0.60	0.60
F (β=1)	n/a	0.26	0.48	0.48	0.34	0.49
F (β=1.5)	n/a	0.29	0.49	0.51	0.40	0.53
F (β=0.5)	n/a	0.22	0.46	0.43	0.27	0.44

Figure 2.3 Precision and Recall scores for Yellow-billed Cuckoo coo calls over five classifier iterations.

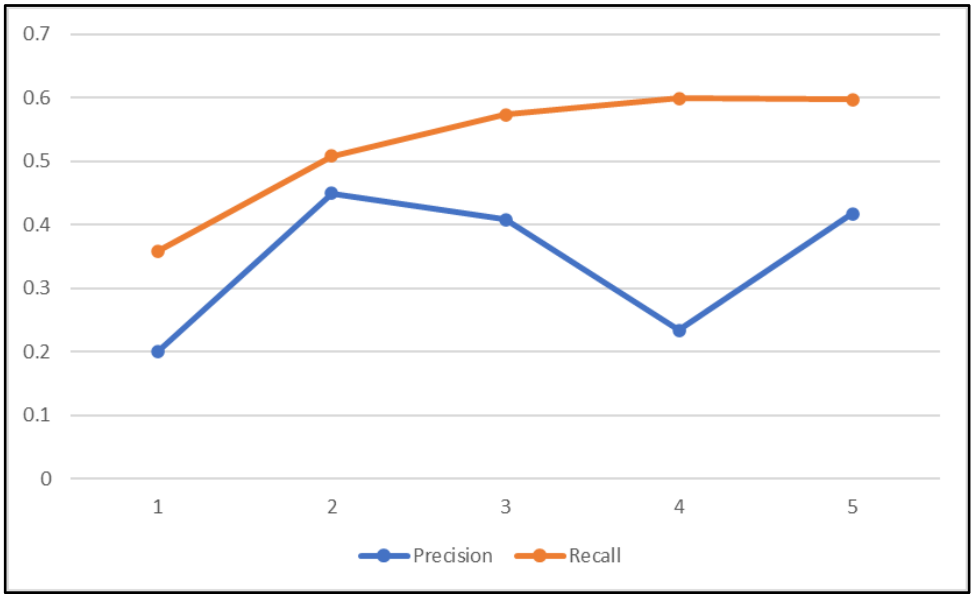
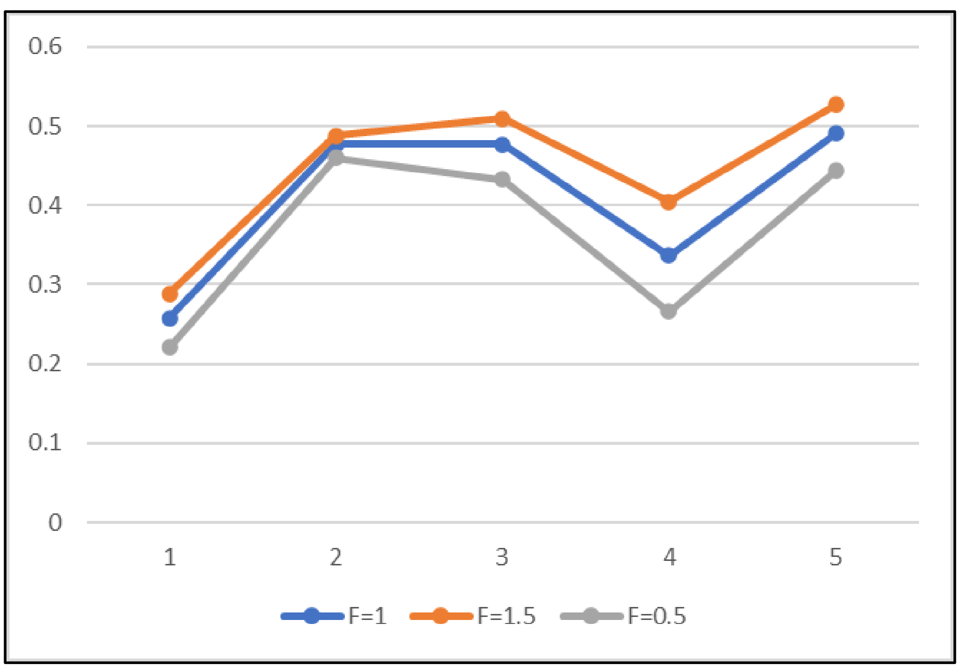


Figure 2.4 F-scores over five Yellow-billed Cuckoo classifier iterations, calculated with 3 different β values.



II. Time investment

We estimated that visual scanning of spectrograms for coo calls required approximately 5 minutes per 1 hour of acoustic data. For our 64 hours of background training data, this amounted to 5 hours and 20 minutes of time investment. For our 448 hours of test data, this amounted to 37 hours and 20 minutes of time investment. Combined, this was 42 hours and 40 minutes to develop our two manually scanned datasets.

For every 1000 detections reported by Kaleidoscope, we estimated approximately 12 minutes were required for manual review. For 5306 detections among all target signal cluster results and 7479 detections among all non-target signal cluster results (12,785 total), this required 2 hours and 31 minutes of manual review. Therefore, the total time invested in developing our classifiers was estimated to be 45 hours and 11 minutes.

DISCUSSION

We have shown that a reiterative refinement of classifiers resulted in improvements in several important classifier performance metrics. In our demonstration, we developed our initial classifier with training data and used it to perform an initial scan of the test dataset. This resulted in an F-score of 0.26, which is less than Knight et al. (2017) reported for Common Nighthawks (F=0.34), but much better than Schroeder and McRae (2020) reported for King Rails (0.04). These initial results provided us with sufficient examples of target signals as well as novel non-target signals reported as false positives from which we were able to re-scan the data and form a new classifier based on the test dataset itself. By reprocessing our test dataset and repeating the steps of restructuring the classifier after each iteration, we were able to track our increases in recall scores at each step to a maximum of 0.60. While this metric is rarely reported (Knight et al. 2017), our scores are on par with some of the highest reported from Kaleidoscope (Abrahams 2019, Marchal et al. 2022). With our final iteration we observed our highest F-score of 0.53 (β =1.5), which is higher than any we identified in our review. Moreover, our F-score of 0.46 (β =0.5) is higher than some other state-of-the art approaches prioritizing precision over recall (Kahl et al. 2021). Importantly, the extra time invested to perform additional classifier iterations was minimal, and we observed our largest increases in recall and precision with only one extra iteration. These results demonstrate an efficient approach to both classifier development and assessment, for which we provide guidance for adaptation below.

We recommend users develop classifiers using a methodical workflow for data selection and preparation to increase overall efficiency of the process. To build our initial classifier with training data, we used library examples because we assumed our calls to be scarce in the broader dataset (which we confirmed through manual review for background data), and therefore insufficient examples would be present to effectively form a robust classifier if we chose training data randomly without library examples. Importantly, we partitioned target signal library examples and background data into different folders, allowing us to efficiently review our initial cluster analysis results and sort target signals from non-target signals. This was achieved by ensuring that our background data was free of target signals. However, we note that if target signals are known to be more common in randomly selected training data, researchers could forego the use of library examples and instead manually review and annotate all call examples in the training data, and then locate these signals in the results through date and time stamps (though this may be less efficient than using separate folders). We also note that such data preparation methods may be applied to other classifier development platforms.

The background data of our training dataset was approximately 20% the size of our test dataset and consisted of random data from all four sites and across the entire season. This ratio of training to test data size proved successful in allowing us to efficiently develop a robust and effective classifier. Although we did not investigate other ratios, we recommend users carefully consider the proportion of data that may be necessary to train classifiers, particularly with regards to differing acoustic environments between training data and broader datasets of interest, which may influence false positive rates.

Kaleidoscope Pro provides many user-defined settings and parameters that will affect the outcomes of data processing. Many of these settings act as filters designed to exclude undesired signals from data output. However, inappropriate setting selections may also have the effect of excluding target signals by being too narrow or increasing false positives by being too wide. In our demonstration, we maintained consistent signal parameters and cluster analysis settings to more clearly examine the effect of multiple classifier iterations. However, as described in user guides, different settings and parameters may perform better depending on the characteristics of the target signal or the dataset of interest. We recommend that users consider an *a priori* evaluation of their acoustic data and their target signals in particular. Users may identify settings that maximize the software's ability to discern target signals from undesired signals by first ensuring target signals will successfully pass through signal parameter filters, and then identifying competing non-target signals with similar characteristics (e.g., syllables, frequency range).

The setting of the "Distance to cluster center to include in cluster.csv output" is used to filter signals highly dissimilar from any of the clusters. Lowering this setting has the effect of producing clustering results of fewer, more similar signals, but may exclude uncommon, faint, or obstructed signals, which may be desired in studies such as ours seeking to identify rare vocalizations. Therefore, users should consider whether a maximum distance setting of 2.0 is warranted to maximize the potential of detecting target signals (higher recall), despite the possibility of increasing false positives (lower precision), or if target signals are common enough in the dataset to risk missing some (lower recall) in exchange for fewer false positives (higher precision) and thus less time investment in manual review of results. In our review, most studies used a maximum setting, while one study used 1.0 (default), and four studies did not report this setting. We emphasize that the settings and parameters step may strongly influence classifier performance, and therefore users should consider which metrics are of highest priority. We also encourage authors to report settings used to aid in interpretation of results, and suggest future research investigate the effects of changing settings on classifier performance.

One inconsistency we observed among studies was the number of clusters used in classifiers. Some studies used simple classifiers composed of the number of clusters automatically generated by the software (reported range of 12–55), whereas others used advanced classifiers with manually defined binary clusters. We chose the latter approach for our study, forcing detections into either target signal or non-target signal clusters, which proved to be effective and efficient. However, Kaleidoscope’s clustering algorithms can handle any number of manually defined combinations (up to 500), which users may be able to experiment with based on specific needs or data characteristics. For instance, a single classifier may be used with individual clusters for each of several vocalization types, as in Schroeder & McRae (2020). Moreover, a multi-cluster classifier may contain several non-target signal clusters, any of which may contain false negative examples of target signals (Abrahams & Denny 2018). By adapting our iterative approach to a multi-cluster classifier, one might more readily locate these false negatives in similar clusters and then re-assign them to the target signal cluster (we note that in each review of the non-target signal cluster, we identified no false negatives). Under this scenario, efficient adoption of our iterative approach may mean selecting a pre-determined number of detections from each non-target signal cluster to review prior to data reprocessing (as opposed to a number of detections proportional to the target signal cluster as we demonstrate with our binary classifier). Finally, we note that classifiers with multiple target-signal clusters may be most effective in scenarios where the multiple target signals share similar characteristics. A single classifier used for multiple dissimilar signals will require broader signal parameters and therefore introduce many additional non-target signals that may result in poor performance. In such cases, we recommend developing individual classifiers for each target signal.

In reviewing the literature, we observed inconsistencies in performance metrics reporting and the formulas used to calculate them similar to those reported previously by Knight et al (2017), including in the years subsequent to this 2017 publication. While most studies did not seek to specifically investigate classifier performance, many did contextualize their use of ARUs and classifiers as a novel approach. Therefore, we echo Knight et al (2017) in suggesting that proper reporting of classifier assessment metrics can improve the interpretation of results and aid others in the application of methods.

The use of a manually processed subset of raw "benchmark" data is the most reliable way to estimate performance metrics (Knight et al. 2017). We strongly caution against using only Kaleidoscope's detection output to calculate false negative rates, as the software automatically filters out some amount of acoustic data during processing and it cannot be assumed that the software did not mistakenly filter out target signals from the clustering output. With no verified or established methods to calculate false negatives, we recommend users employ the use of a randomly selected benchmark dataset of raw acoustic data. While we used our entire test dataset as the benchmark to calculate exact metrics, we suggest that an estimate based on a subset of raw data will be more comparable to (if not also more accurate than) other in-software methods mentioned. The size of the benchmark dataset relative to the broader dataset may depend on the needs of the study and practicality of time investments. For species with low call rates, a larger benchmark dataset will be appropriate. However, we highlight the need for further assessment of benchmark datasets to determine the appropriate size to ensure accuracy, as well as the influence of call characteristics and rates, and features of the acoustic environments.

A major emphasis of acoustic analysis methods is the potential to provide high quality data with minimal human time investment (Digby et al. 2013). However, all software platforms require substantial time to develop a robust and effective classifier (Digby et al. 2013, Knight et al. 2017). We invested over 42 hours to manually review a large quantity of data (512 hours) to more effectively investigate classifier performance across large datasets, and our review rate was consistent with that of other studies (Digby et al. 2013). For many applications, benchmark datasets of this size may not be necessary to determine reliable estimates of performance metrics. For example, Schroeder (2020) used a 41-hour benchmark, and Ethier and Wilson (2020) used a 2-hour benchmark. However, we encourage researchers to use a benchmark dataset and consider the time required to reliably review it based on research objectives.

Manually reviewing detection results required comparatively much less time than reviewing the benchmark dataset. This suggests that the iterative process of improving the classifier is not a substantial time commitment aside from the unsupervised time required for the automated processing of data between manual reviews. While our “equal effort” approach to the target and non-target clusters was successful in helping to increase recall, it is possible that increasing the amount of time spent manually reviewing the non-target clusters could have aided in accounting for fluctuating false positive rates and therefore improve precision and F-scores. We also note that our classifier was used in a relatively small frequency range, and other users may experience higher rates of false positives in broader frequency ranges or more diverse acoustic datasets. Therefore, we recommend that users track their time investment at each classifier iteration and consider the potential benefits of increasing the time spent on manual review of non-target signal clusters to reduce false positive rates.

In conclusion, we identified numerous inconsistencies in classifier development methods and assessment used in the literature. To address these inconsistencies, we recommend future users of Kaleidoscope adopt a workflow similar to that of our classifier development process. This includes data preparation, the use of a benchmark dataset to properly derive performance metrics, and additional classifier iterations to improve classifier performance. As we have demonstrated, this workflow can result in substantial performance improvements with little additional time investment. These methods provide a much-needed framework for others to adapt and modify based on specific project goals.

ACKNOWLEDGEMENTS

Work presented in this chapter was supported by contributions from Dr. Tad Theimer, Susan Sferra, and Dr. Bret Pasch. This research was conducted under U.S. Fish & Wildlife Service Recovery Permit #TE81148C-0 and Arizona Game & Fish Department Special Collecting License permit #SP407168. The findings and conclusions in this article are those of the authors and do not necessarily represent the views of the U.S. Fish and Wildlife Service. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by

the United States Government. Data for this study are available as a U.S. Geological Survey data release.

LITERATURE CITED

Abrahams, C. (2019). Comparison between lek counts and bioacoustic recording for monitoring Western Capercaillie (Tetrao urogallus L.). Journal of Ornithology, 160(3), 685–697. https://doi.org/10.1007/s10336-019-01649-8

Abrahams, C., & Denny, M. J. H. (2018). A first test of unattended, acoustic recorders for monitoring Capercaillie Tetrao urogallus lekking activity. Bird Study, 65(2), 197–207. https://doi.org/10.1080/00063657.2018.1446904

Abrahams, C., & Geary, M. (2020). Combining bioacoustics and occupancy modelling for improved monitoring of rare breeding bird populations. Ecological Indicators, 112, 106–131. https://doi.org/10.1016/j.ecolind.2020.106131

Bardeli, R., Wolff, D., Kurth, F., Koch, M., Tauchert, K.-H., & Frommolt, K.-H. (2010). Detecting bird sounds in a complex acoustic environment and application to bioacoustic monitoring. Pattern Recognition Letters, 31(12), 1524–1534. https://doi.org/10.1016/j.patrec.2009.09.014

Bateman, J., & Uzal, A. (2022). The relationship between the Acoustic Complexity Index and avian species richness and diversity: A review. Bioacoustics, 31(5), 614–627. https://doi.org/10.1080/09524622.2021.2010598

Bobay, L. R., Taillie, P. J., & Moorman, C. E. (2018). Use of autonomous recording units increased detection of a secretive marsh bird. Journal of Field Ornithology, 89(4), 384–392. https://doi.org/10.1111/jofo.12274

Buxton, R. T., & Jones, I. L. (2012). Measuring nocturnal seabird activity and status using acoustic recording devices: Applications for island restoration. Journal of Field Ornithology, 83(1), 47–60. https://doi.org/10.1111/j.1557-9263.2011.00355.x

Digby, A., Towsey, M., Bell, B. D., & Teal, P. D. (2013). A practical comparison of manual and autonomous methods for acoustic monitoring. Methods in Ecology and Evolution, 4(7), 675–683. https://doi.org/10.1111/2041-210X.12060

Ethier, J. P., & Wilson, D. R. (2020). Using microphone arrays to investigate microhabitat selection by declining breeding birds: Oroignal. Ibis, 162(3), 873–884. https://doi.org/10.1111/ibi.12785

Gibb, R., Browning, E., Glover-Kapfer, P., & Jones, K. E. (2019). Emerging opportunities and challenges for passive acoustics in ecological assessment and monitoring. Methods in Ecology and Evolution, 10(2), 169–185. https://doi.org/10.1111/2041-210X.13101

Halterman, M. (2009). Sexual dimorphism, detection probability, home range, and parental care in the Yellow-billed Cuckoo. University of Nevada, Reno.

Hamilton, W. J., III, & Hamilton, M. E. (1965). Breeding characteristics of yellow-billed cuckoos in Arizona. Proceedings of the California Academy of Sciences, 32(14), 405–432.

Hughes, J. M. (2015). Yellow-billed Cuckoo (Coccyzus americanus), version 2.0. In P. G. Rodewald (Ed.), The Birds of North America Online. Cornell Lab of Ornithology. https://birdsoftheworld.org/bow/historic/bna/yebcuc/2.0/introduction

Hutto, R. L., & Stutzman, R. J. (2009). Humans versus autonomous recording units: A comparison of point-count results. Journal of Field Ornithology, 80(4), 387–398. https://doi.org/10.1111/j.1557-9263.2009.00245.x

Kahl, S., Wood, C. M., Eibl, M., & Klinck, H. (2021). BirdNET: A deep learning solution for avian diversity monitoring. Ecological Informatics, 61, 101236. https://doi.org/10.1016/j.ecoinf.2021.101236

Knight, E. C., Hannah, K. C., Foley, G. J., Scott, C. D., Brigham, R. M., & Bayne, E. (2017). Recommendations for acoustic recognizer performance assessment with application to five common automated signal recognition programs. Avian Conservation and Ecology, 12(2), art14. https://doi.org/10.5751/ACE-01114-120214

Marchal, J., Fabianek, F., & Aubry, Y. (2022). Software performance for the automated identification of bird vocalisations: The case of two closely related species. Bioacoustics, 31(4), 397–413. https://doi.org/10.1080/09524622.2021.1945952

McNeil, S. E., Tracy, D., Stanek, J. R., Stanek, J. E., & Station, S. S. R. (2013). Yellow-billed Cuckoo Distribution, Abundance, and Habitat Use on the Lower Colorado River and Tributaries 2008-2012 Summary Report. Lower Colorado River Multi-Species Conservation Program, U.S. Bureau of Reclamation.

Pérez-Granados, C., Schuchmann, K., & Marques, M. I. (2020). Vocal behavior of the Undulated Tinamou (Crypturellus undulatus) over an annual cycle in the Brazilian Pantanal: New ecological information. Biotropica, 52(1), 165–171. https://doi.org/10.1111/btp.12742

Pérez-Granados, C., & Schuchmann, K.-L. (2020a). Illuminating the nightlife of two Neotropical nightjars: Vocal behavior over a year and monitoring recommendations. Ethology Ecology & Evolution, 32(5), 466–480. https://doi.org/10.1080/03949370.2020.1753117

Pérez-Granados, C., & Schuchmann, K.-L. (2020b). Monitoring the annual vocal activity of two enigmatic nocturnal Neotropical birds: The Common Potoo (Nyctibius griseus) and the

Great Potoo (Nyctibius grandis). Journal of Ornithology, 161(4), 1129–1141. https://doi.org/10.1007/s10336-020-01795-4

Pérez-Granados, C., Schuchmann, K.-L., & Marques, M. I. (2021). Passive acoustic monitoring of the Ferruginous Pygmy-Owl (Glaucidium brasilianum) over a complete annual cycle: Seasonality and monitoring recommendations. Studies on Neotropical Fauna and Environment, 1–8. https://doi.org/10.1080/01650521.2021.1933699

Pérez-Granados, C., & Traba, J. (2021). Estimating bird density using passive acoustic monitoring: A review of methods and suggestions for further research. Ibis, 163(3), 765–783. https://doi.org/10.1111/ibi.12944

Potamitis, I., Ntalampiras, S., Jahn, O., & Riede, K. (2014). Automatic bird sound detection in long real-field recordings: Applications and tools. Applied Acoustics, 80, 1–9. https://doi.org/10.1016/j.apacoust.2014.01.001

Priyadarshani, N., Marsland, S., & Castro, I. (2018). Automated birdsong recognition in complex acoustic environments: A review. Journal of Avian Biology, 49(5), jav-01447. https://doi.org/10.1111/jav.01447

Schroeder, K. M., & McRae, S. B. (2020). Automated auditory detection of a rare, secretive marsh bird with infrequent and acoustically indistinct vocalizations. Ibis, 162(3), 1033–1046. https://doi.org/10.1111/ibi.12805

Shonfield, J., & Bayne, E. M. (2017). Autonomous recording units in avian ecological research: Current use and future applications. Avian Conservation and Ecology, 12(1), art14. https://doi.org/10.5751/ACE-00974-120114

Sidie-Slettedahl, A. M., Jensen, K. C., Johnson, R. R., Arnold, T. W., Austin, J. E., & Stafford, J. D. (2015). Evaluation of autonomous recording units for detecting 3 species of secretive marsh birds: Marsh Bird Surveys. Wildlife Society Bulletin, 39(3), 626–634. https://doi.org/10.1002/wsb.569

Sokolova, M., Japkowicz, N., & Szpakowicz, S. (2006). Beyond Accuracy, F-Score and ROC: A Family of Discriminant Measures for Performance Evaluation. In A. Sattar & B. Kang (Eds.), AI 2006: Advances in Artificial Intelligence (Vol. 4304, pp. 1015–1021). Springer Berlin Heidelberg. https://doi.org/10.1007/11941439_114

Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. Information Processing & Management, 45(4), 427–437. https://doi.org/10.1016/j.ipm.2009.03.002

USFWS. (2014). Endangered and Threatened Wildlife and Plants; Determination of Threatened Status for the Western Distinct Population Segment of the Yellow-billed Cuckoo (Coccyzus americanus); Final Rule. Federal Register, 79(129), 59992–60038.

Wilhite, N. G., Howell, P. E., & Martin, J. A. (2020). Evaluation of Acoustic Recording Devices to Survey Northern Bobwhite Populations. Wildlife Society Bulletin, 44(1), 200–207. https://doi.org/10.1002/wsb.1061

Supplemental Materials Table 2.1 List of studies reviewed that use Kaleidoscope Pro classifiers for birds (2014–2022), with classifier development summarized.

Author	Species	Time investment	Classifier Type	Training data (call examples)	Validation audio
Wilhite et al. (2020)	Northern Bobwhite	Unspecified	Single-species, Unspecified classes	22 recordings (16 1 h recordings containing thousands of "koi-lee" calls)	Yes
Ethier & Wilson (2020)	Boreal Chickadee	Unspecified	Multi-species, Multi-class, Unspecified classes	1632 h	Yes (2 h subset)

Ethier & Wilson (2020)	Cape-May Wabler	Unspecified	Multi-species, Multi-class, Unspecified classes	1632 h	Yes (2 h subset)
Schroeder & McRae (2020)	King Rail	Unspecified	Single-species Unspecified classes	Two 5 m ARU recordings (not scanned) and 30 kek/10 grunts.	Yes (41 h subset)
Marchal et al. (2022)	Bicknell's Thrush	Processing: 2hr detection and class step	Multi-species, Multi-class, Unspecified classes	142 h (5284 calls)	Yes (all data)
Marchal et al. (2022)	Gray-cheeked Thrush	Processing: 2hr detection and class step	Multi-species, Multi-class, Unspecified classes	3.3 h (1471 calls)	Yes (all data)
Knight et al. (2017)	Common Nighthawk	9.6m/hr validation,	Single-species, Binary	400 m (200/200 with/without nighthawk)	Yes (all data)
Pérez-Granados & Schuchmann (2020b)	Common Potoo	Unspecified	Multi-species, Unspecified classes	Not given	Yes (168 15 m recordings)
Pérez-Granados & Schuchmann (2020b)	Great Potoo	Unspecified	Multi-species, Unspecified classes	Not given	Yes (168 15 m recordings)
Pérez-Granados et al. (2021)	Ferruginous Pygmy-owl	Unspecified	Single species, Unspecified classes	Not given	Not given
Pérez-Granados & Schuchmann (2020a)	Little Nightjar	Unspecified	Single species, Unspecified classes	1825 15 m recordings	Not given
Pérez-Granados et al. (2020)	Pauraque	Unspecified	Single species, Unspecified classes	1825 15 m recordings	Not given
Pérez-Granados et al. (2020)	Undulated Tinamou	Unspecified	Single species, Unspecified classes	Not given	Not given
Abrahams & Geary (2020)	European Nightjar	Unspecified	Multi-species, Multi-class, 55 classes	All data processed (288 m/site - 18 sites)	Not given
Abrahams & Geary (2020)	Woodlark	Unspecified	Multi-species, Multi-class, 55 classes	All data processed (288 m/site - 18 sites)	Not given
Abrahams & Geary (2020)	Dartfod warbler	Unspecified	Multi-species, Multi-class, 55 classes	All data processed (288 m/site - 18 sites)	Not given

Abrahams & Denny (2018)	Capercaillie	Unspecified	Single-species, Multi-class, Unspecified classes	5401 phrases, 258 target specie phrases	Not given
Abrahams, (2019)	Capercaillie	Unspecified	Single-species, Binary	28.1 GB of data (20493 phrases, 2114 target species phrases)	Random 3755 phrases
Bobay et al., (2018)	Least Bittern	Unspecified	Multi-species, Multi-class, 3 classes	Not given	Not given
Bobay et al., (2018)	Black Rail	Unspecified	Multi-species, Multi-class, 3 classes	Not given	Not given

Supplemental Materials Table 2.2 List of studies reviewed that use Kaleidoscope Pro classifiers for birds (2014–2022), with classifier settings summarized.

Author	Freq-Range	Length (s)	Inter-syllable Gap (s)	Max distance to include	FFT Window	Max states	Max distance for building	Max clusters
Wilhite et al. (2020)	Not given	Not given	Not given	Not given	Not given	Not given	Not given	Not given
Ethier & Wilson (2020)	2000-10000 Hz	0.1-4.0	0.35	2	256	12	0.5	500
Ethier & Wilson (2020)	2000-10000 Hz	0.1-4.0	0.35	2	256	12	0.5	500
Schroeder & McRae (2020)	1500-7000 Hz	0.1-6	0.3	Not given	Not given	Not given	Not given	Not given
Marchal et al. (2022)	2000-6000 Hz	0.1-0.5	Not given	2	256	12	0.5	500
Marchal et al. (2022)	2000-6000 Hz	0.1-0.5	Not given	2	256	12	0.5	500
Knight et al. (2017)	1000-7000 Hz	0.1-0.7	0	2	256	12	1	2
Pérez-Granados & Schuchmann (2020b)	400-1400 Hz	0.8-6	0.3	2	Not given	Not given	Not given	Not given
Pérez-Granados & Schuchmann (2020b)	400-1400 Hz	0.8-6	0.3	2	Not given	Not given	Not given	Not given
Pérez-Granados et al. (2021)	1150-1450 Hz	2-24s	0.7	2	Not given	Not given	Not given	Not given
Pérez-Granados & Schuchmann (2020a)	600-2700 Hz	0.8-3	0.2	2	Not given	Not given	Not given	Not given
Pérez-Granados et al. (2020)	600-2700 Hz	0.3-0.8	0.2	2	Not given	Not given	Not given	Not given

Pérez-Granados et al. (2020)	1150-1350 Hz	1.3-3	0.6	2	Not given	Not given	Not given	Not given
Abrahams & Geary (2020)	1500-1700 Hz	20-Feb	1	Not given	Not given	Not given	Not given	Not given
Abrahams & Geary (2020)	1500-1700 Hz	20-Feb	1	Not given	Not given	Not given	Not given	Not given
Abrahams & Geary (2020)	1500-1700 Hz	20-Feb	1	Not given	Not given	Not given	Not given	Not given
Abrahams & Denny (2018)	1500-4000 Hz	6-Feb	1	1	5.33 ms	12	0.5	500
Abrahams, (2019)	1500-4000 Hz	6-Jan	1	1	5.33 ms	12	0.5	500
Bobay et al., (2018)	Not given	Not given	Not given	Not given	Not given	Not given	Not given	Not given
Bobay et al., (2018)	Not given	Not given	Not given	Not given	Not given	Not given	Not given	Not given

Supplemental Materials Table 2.3 List of studies reviewed that use Kaleidoscope Pro classifiers for birds (2014–2022), with performance metrics summarized.

Author	False-positive (total or percent given)	False-negative	Recall	Precision	F-score
Wilhite et al. (2020)	Not given	68.50%	Not given	Not given	Not given
Ethier & Wilson (2020)	0.1*	75.30%	Not given	Not given	Not given
Ethier & Wilson (2020)	0.1*	17.90%	Not given	Not given	Not given
Schroeder & McRae (2020)	96%	88.6% grunt/77.9% kek	0.11(grunt)/0.22(kek)	0.07(grunt) 0.04(kek)	0.04(grunt)/0.03(kek)
Marchal et al. (2022)	73%	49%	0.51	0.25	not given
Marchal et al. (2022)	70%	57%	0.43	0.33	not given
Knight et al. (2017)	Not given	Not given	0.34	0.7	0.34

Pérez-Granados & Schuchmann (2020b)	Not given	0.015%*	85.2	9.4	Not given
Pérez-Granados & Schuchmann (2020b)	Not given	0.09%*	73.6	28.6	Not given
Pérez-Granados et al. (2021)	Not given	Not given	Not given	Not given	Not given
Pérez-Granados & Schuchmann (2020a)	Not given	Not given	Not given	Not given	Not given
Pérez-Granados et al. (2020)	27.60%	Not given	Not given	72.4	Not given
Pérez-Granados et al. (2020)	22.70%	Not given	Not given	77.3	Not given
Abrahams & Geary (2020)	Not given	Not given	Not given	Not given	Not given
Abrahams & Geary (2020)	Not given	Not given	Not given	Not given	Not given
Abrahams & Geary (2020)	Not given	Not given	Not given	Not given	Not given
Abrahams & Denny (2018)	16.40%	11%*	Not given	Not given	Not given
Abrahams, (2019)	21%	6%*	0.81	0.61	0.7
Bobay et al., (2018)	3915	Not given	Not given	Not given	Not given
Bobay et al., (2018)	11781	Not given	Not given	Not given	Not given

Chapter 3 Using autonomous recording units to identify and monitor western Yellow-billed Cuckoos in southwestern xeroriparian habitat

ABSTRACT

Autonomous recording units (ARUs) paired with signal classification software can be used to detect species-specific songs and calls, making them useful for evaluating patterns in avian activity and occupancy. Application of such methods may be challenging for rare and cryptic species that vocalize infrequently or that occur in low numbers, and by inaccurate classification of target signals by classifiers. We (Northern Arizona University, in conjunction with U.S. Fish and Wildlife Service and U.S. Geological Survey) tested the use of ARUs to identify occupied habitat for a cryptic and federally threatened distinct population segment, the western Yellow-billed Cuckoo (*Coccyzus americanus*). Using Kaleidoscope Pro, we developed a call-classifier and processed acoustic data collected in sites also surveyed using traditional human-observer methods, applying the same spatial and temporal detection criteria to estimate site occupancy for each method. The classifier detected a total of 4061 calls in four sites, and had an overall precision score of 0.06, recall score of 0.33, and F-score (beta = 1) of 0.102, indicating many calls were not detected by the classifier. Our results were consistent with other ARU studies of rare and cryptic species, however, and ARUs estimated occupancy as effectively as human surveys with as little as two hours of daily recording. Total detections varied between sites, likely due to differences in cuckoo population densities and the interaction between topography and ARU detection space. Our results suggest that despite performance shortcomings of call-classifiers, ARUs can be effective for monitoring cuckoos, with potential for providing higher resolution temporal and spatial information on activity and habitat use.

INTRODUCTION

Use of autonomous recording units (ARUs) to passively survey and monitor birds has rapidly increased over the past decade (Shonfield & Bayne 2017). With long battery life, customizable recording schedules, and high data storage capacity, modern ARU hardware offers the ability to obtain large amounts of acoustic data over long durations with minimal maintenance. Recent advances in data processing software allow for semi-automated detection and classification of target signals in large datasets (Potamitis et al. 2014). Together, ARUs and data processing software provide innovative methods for avian monitoring, which can complement or even outperform traditional human-observer methods such as point counts or playback surveys (Hutto & Stutzman 2009, Alquezar & Machado 2015, Darras et al., 2018). These benefits have led to the successful application of ARUs for occupancy studies of birds (Campos-Cerqueira & Aide 2016, Abrahams & Geary 2020) but have also revealed limitations and challenges (Shonfield & Bayne 2017). For example, signal classification performance varies between software platforms, and may require time-consuming manual verification of datasets (Knight et al. 2017). In addition, the amount of data generated by ARU's can become very large depending upon how long ARU's are programmed to record, but few studies have assessed how reduced sampling schemes may affect detection capability.

The use of ARUs to evaluate occupancy of rare or cryptic species may be especially important given those species often have the greatest need for improved monitoring programs to aid in conservation and recovery efforts (Sidie-Slettedahl et al. 2015, Bobay et al. 2018, Schroeder & McRae 2020, Manzano-Rubio et al. 2022). This is the case for the federally threatened distinct population segment of the western Yellow-billed Cuckoo (*Coccyzus* *americanus*). Historically found from northern Mexico to southern British Columbia (Hughes 1999), populations have declined due to loss and degradation of habitat (USFWS 2014), with most remaining populations located in Arizona, New Mexico, California, and Sonora (USFWS 2021). Traditional call-playback protocols have been developed for surveying Yellow-billed Cuckoos and estimating site occupancy based on the temporal and spatial patterns of detected cuckoos (Halterman et al. 2015), and this information has been used to estimate population size and designate critical habitat (USFWS 2021). These surveys have focused primarily on riparian areas, but recent studies in southern Arizona have demonstrated that cuckoos are also breeding xeroriparian habitats (Beauregard (Chapter 1). Populations in these xeroriparian habitats likely represent a significant addition to the estimated western population segment, but the geographical extent of use of these habitats remains to be investigated. If ARU methods can be developed to survey for cuckoos and estimate occupancy in these often remote and difficult to access xeroriparian areas, this would provide biologists and land managers an important alternative to the current human-based surveys that can be resource intensive and may not be feasible in many remote locations. Although Yellow-billed Cuckoos vocalize relatively infrequently, approximately once per hour (Halterman 2009), their distinctive vocalizations and large home ranges (Halterman 2009, Sechrist et al. 2013), particularly while advertising for mates prior to nest initiation and during pair interactions (Hamilton & Hamilton 1965), make them a likely candidate for ARU-based survey methods.

We (Northern Arizona University, in conjunction with U.S. Fish and Wildlife Service and U.S. Geological Survey) investigated the utility of using ARUs to identify occupied western Yellow-billed Cuckoo habitat by comparing the efficacy of ARUs to detect cuckoos to that of human-observers at five field sites in southern Arizona. In the process, we developed and tested the performance of a semi-autonomous classifier for cuckoo calls using a popular and easily accessible data processing software, Kaleidoscope Pro (Wildlife Acoustics, Inc.). Finally, we tested subsets of our full acoustic dataset to simulate commonly used ARU deployment schedules to identify the optimal timing and duration of recording that would most efficiently estimate cuckoo occupancy.

METHODS

I. Study area

Our study was conducted in southeastern Arizona, USA, within the Madrean Sky Islands Archipelago region (Warshall 1995). We chose five field sites located on the Coronado National Forest and Buenos Aires National Wildlife Refuge. Sites were chosen non-randomly to ensure accessibility for surveyors and at locations where cuckoos had been documented previously, while also representing the variation in geomorphology and vegetative structure and composition typical of areas used by cuckoos in xeroriparian habitat. All sites were located in mid-elevation (3300–5700 feet) ephemeral drainages and ranged from a broad, shallow valley to a narrow, sinuous, steep-walled canyon. Habitat consisted primarily of Madrean-oak woodland, semi-desert grassland (Brown 1994), with widely spaced xeroriparian and riparian species along the drainage bottom, consistent with southwestern breeding habitat as previously described (USFWS 2021).

II. Protocol call-back surveys and nest searching

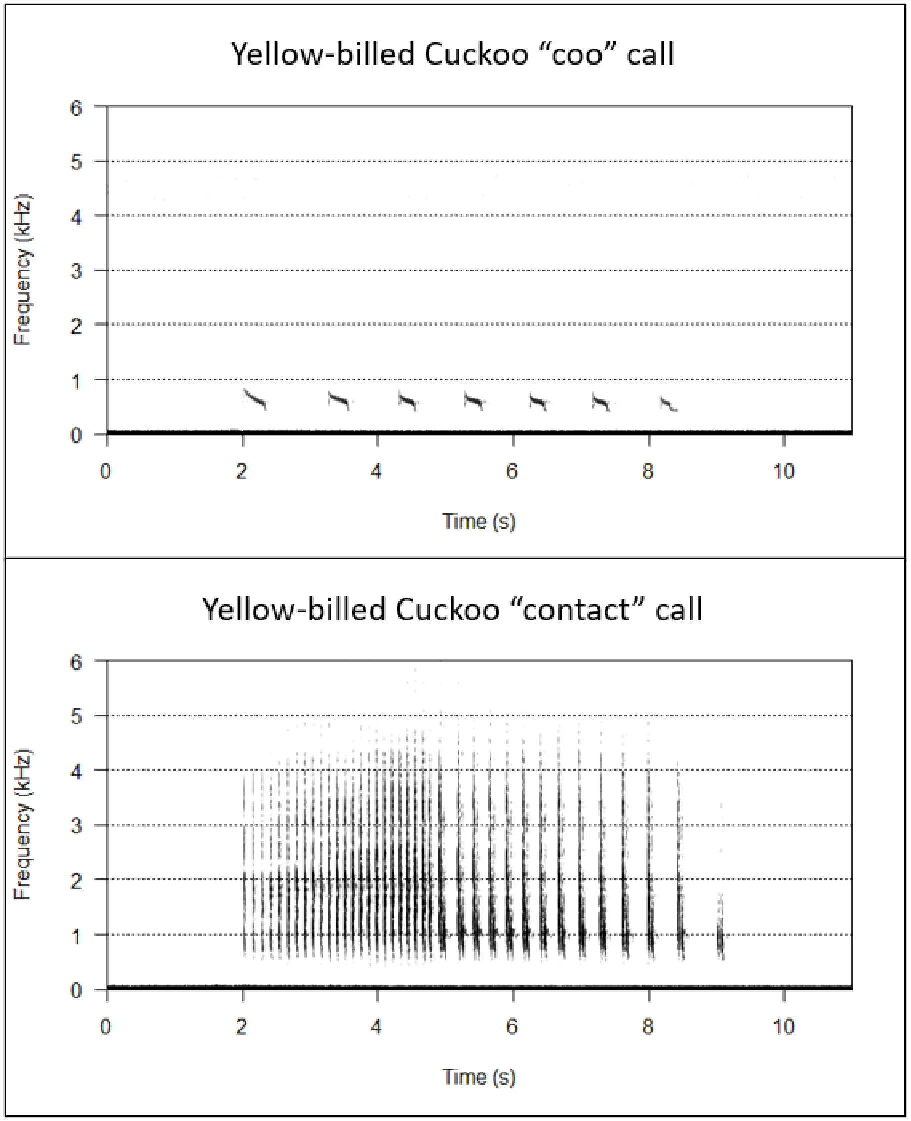
We surveyed each site using a standard survey protocol developed to monitor western Yellow-billed Cuckoos that requires four survey visits per site, spaced 10-15 days apart, within a survey season lasting from June 15 – August 15 (Halterman et al. 2015). However, we also included additional surveys between August 15 – September 15 to account for potential late-season nesting activity. Each survey requires the use of call-playback to elicit responses from cuckoos at 100-meter intervals along a survey transect that follows a drainage bottom. Transects in our study were 2 km in length and for our analyses divided into 4 quadrants of 0.5 km each. If cuckoos were detected during these survey visits, we made additional visits to confirm breeding status by searching for nests or other evidence of breeding including copulation, nest building, fledglings, or distraction displays. Nests were not monitored intensively and nest searching effort varied between sites and over the course of the field season. Therefore, we were not able to determine a definitive number of nests in a site, or the nest fate unless fledglings were observed. At the end of the field season, locations of cuckoo detections for each protocol survey were plotted in ArcGIS (ESRI version 3.2). Following protocol guidelines, we identified occupied territories as areas where cuckoos were detected during two or more survey periods, with survey visits separated by at least 10 days, with detection locations between surveys no greater than 500 m apart (Halterman et al. 2015, see Occupancy Estimation below).

III. Yellow-billed cuckoo vocalizations

Adult Yellow-billed Cuckoos produce several call types (Hamilton & Hamilton 1965, Hughes 2015) but we focused on the two most commonly observed: the *Coo* and *Contact* (Figure 3.1). Coo calls are hypothesized to be produced primarily by females advertising for mates (McNeil et al. 2013). Coos are often given as a series of 5–11 individual coo notes. A cooing bout may last for only one series of coo notes, but more typically lasts for several minutes or more with repeated series of coos separated by a brief pause (Hughes 2015). For this analysis, each individual series of coo notes was counted as a single full “coo call”, rather than each individual coo note. Thus, the number of “coo calls” recorded in a cooing bout that included three series of 5 coo notes each, was considered three “coo calls” in our analyses.

Contact calls are the call type used for playback during protocol surveys and are commonly observed but rarely given more than once per 10 minute period (Hughes 2015). More variable in structure than coo calls, contact calls typically consist of several introductory “kuk” notes followed by a series of “kowlp” notes (Figure 3.1), but variations are also given consisting of only “kuk” (often considered an alarm call) or only “kowlp” notes (Hamilton & Hamilton 1965, Hughes 2015).

Figure 3.1 Spectrogram of the western Yellow-billed Cuckoo's "coo call” (top) and "contact call” (bottom). Because coos are typically given as a series of repeated individual notes, the “coo call” shown here, consisting of a series of seven individual coo notes, would have been considered a single “coo call” in our analysis. Likewise, the multiple notes given in rapid succession in the contact call shown here would be considered a single “contact call” in our analysis. Both vocalizations recorded at Cherry Creek in the Canelo Hills in the Coronado National Forest, Arizona, USA in 2018.



IV. Autonomous recording unit deployment

Within each site, four ARUs (Model SM4, Wildlife Acoustics, Inc.) were deployed along the 2 km survey transect at 500 m intervals, beginning 250 m past the initial start point of the transect and ending 250 m before the stop point. Thus, each transect was divided into 4 quadrants (hereafter “quads”) of 500 m in length with the ARU in the center of each. ARUs were deployed prior to the start of the survey season (June 15) and were retrieved in October. Each unit was affixed to a small-diameter tree (<10” DBH) using nylon webbing at a height of approximately 2 m from the ground, and no further than 25 m from the drainage bottom. To avoid tampering or theft, ARUs were deployed away from trails or roads. All ARUs were set to record continuously every day from 04:00-11:00 at a sampling rate of 22050 Hz. ARUs were inspected by surveyors during each site visit, and batteries and memory cards were changed as needed.

V. Determining ARU range

Detection probability varies among ARUs deployed at different locations due to the attenuating effects of environmental factors on sound detection space (Darras et al. 2016). Therefore, acoustic monitoring programs must account for this potential variation in acoustic range when comparing results from different sites (Darras et al. 2016, Manzano-Rubio et al. 2022). To determine the approximate range of each individual ARU at our sites, we used a portable speaker (JBL Clip 3) to broadcast the recording of the cuckoo’s contact call used for standardized playback surveys at 70 dB SPL (verified with a handheld decibel meter) as recommended by the survey protocol (Halterman et al. 2015) from distances of 1 m, 100 m, 200 m, 300 m, and 400 m from the actively-recording ARUs. We then viewed the spectrograms of the recordings, isolated the clips of the recorded playback, and used the Kaleidoscope Pro sound pressure level (SPL) analysis feature to measure the maximum sound pressure level (in dB) and identified the distance at which playback calls were no longer detected. These levels were plotted against the distance interval at which they were recorded, resulting in a distance attenuation curve for each individual ARU.

VI. Classifier development

Kaleidoscope Pro software (Wildlife Acoustics, Inc., Maynard MA) uses cluster analysis algorithms to process acoustic data and sort detected signals into individual clusters of statistically similar signals. Users can develop classifiers for target signals by reviewing the cluster analysis output and manually sorting target signals and non-target signals into separate clusters. Subsequent reprocessing of the data trains the algorithms based on these manual verifications, which results in a classifier that can then be used to process new datasets. User-defined signal parameters such as frequency window, minimum and maximum duration of signals, and maximum inter-syllable gap also act as data processing filters.

To evaluate the use of Kaleidoscope Pro to detect cuckoo vocalizations in acoustic data, we developed a classifier to include clusters for both Coo and Contact calls. Building a classifier requires training data containing numerous examples of the target signals from which clusters can be formed, as well as “background data” containing examples of other signals present in the acoustic environment which are discriminated against. Cuckoos vocalize relatively infrequently, so we assumed that a given time period of continuous acoustic data may contain very few cuckoo vocalizations relative to other signals, which may inhibit the software’s ability to form distinct clusters around vocalizations. To address this, we used RavenPro (Cornell University, Ithaca NY) to view spectrograms from a subset of acoustic data from ARUs deployed near cuckoo nests in 2018 during a pilot study and manually isolated 879 clear, unobstructed examples of cuckoo vocalizations, which we then combined with 64 hours of background data audio files from our field sites to be used as a training dataset.

To inform our choices of signal parameters for data processing, we manually inspected examples of spectrograms of Coo and Contact calls as well as the background data. We noted that contact calls occupy a large frequency range (approximately 600–4000 Hz), whereas coo calls occupy a narrower frequency range (approximately 500–1000 Hz). However, both calls exhibit their peak amplitudes in the range of approximately 500–1000 Hz. Furthermore, while some other species' calls and non-biological signals were present at these lower ranges, many more species' vocalizations were present at frequencies above 1250 Hz, which led to poor clustering of partially obstructed cuckoo calls in test runs of the data. Therefore, we used a 350–1250 Hz frequency range to process data. We also used 0.5–10 seconds length of detection, and 0.5 seconds maximum inter-syllable gap, as both call types also vary in their duration, and the gap between individual coo notes is often similar to that of the gap observed between kowlp notes of the contact call. For cluster analysis settings, we used a 1.5 maximum distance from cluster center for outputs, FFT window size of 512, 12 maximum stats, 0.5 maximum distance from cluster center, and 500 maximum clusters, following recommendations provided in the user manual. The cluster analysis resulted in 33 clusters, which were manually reviewed and edited before subsequent reprocessing of the data to train the classifier.

We then processed all 2019 ARU data for each site using the classifier. All signals that the classifier reported in the Coo or Contact clusters were manually reviewed to separate true-positive calls from false-positive signals. Manual review consisted of a visual inspection of the spectrogram of each reported signal in the Coo and Contact clusters. Additionally, all detected signals with spectrograms that superficially resembled cuckoo vocalizations (and which could not be confidently identified as another known signal by visual inspection) were reviewed auditorily to ensure that faint, obscure, or atypical cuckoo vocalizations were not mislabeled as false negatives. Reported signals confirmed to contain Yellow-billed Cuckoo vocalizations were counted as true positives and labeled either "Coo" or "Contact." Signals not containing Yellow-billed Cuckoo vocalizations were labeled "false positive".

VII. Classifier performance assessment

To assess the performance of our classifiers, we used methods outlined in Knight et al. (2017). We first calculated the number of true positives, false positives, and false negatives. To calculate false negatives, we manually reviewed a subset of data (often referred to as a "benchmark" dataset), which consisted of 4 randomly selected full days of data for each site. This allowed us to calculate the number of calls actually present within the benchmark dataset and compare against the results of the classifier for the same data, resulting in a rate of false negatives which can be used as an estimate for the overall dataset. We then calculated the number of true positives and false positives reported by the classifier for the Coo and Contact clusters. We then used these totals to calculate precision, recall, and F-score of the two call types using the following formulae:

(1) Precision = tp/tp+fp

(2) Recall = tp/tp +fn

(3) $F = (\beta^2+1) * \text{Precision} * \text{Recall} / \beta^2 * \text{Precision} + \text{Recall}$

Where *tp* is the number of true positives (the number of calls accurately classified in the Coo or Contact cluster by Kaleidoscope), *fp* is the number of false positives (non-cuckoo signals classified in the Coo or Contact cluster by Kaleidoscope), and *fn* is the number of false negatives (the number of calls not accurately classified in the Coo or Contact cluster by Kaleidoscope). β is a user-defined metric that allows for the prioritization of either precision or recall based on specific analysis goals. Precision is favored when $\beta > 1$ and recall is favored when $\beta < 1$, and precision and recall are balanced when $\beta = 1$ (Sokolova et al. 2006).

VIII. Exploring temporal patterns in detections

We used the results of our classifier to evaluate the temporal patterns of calls detected at occupied sites, emphasizing data collected from ARUs deployed near documented nests or breeding activity. Total detected calls were calculated for each call type at each ARU and plotted by day and time, with the goal of identifying consistent patterns in activity before, during, or after estimated nesting cycles. We also calculated the total number of each call type detected during each hour of the morning to identify trends in peak vocal activity.

IX. Estimating occupancy

To evaluate the ability of ARUs and Kaleidoscope Pro classifiers to estimate Yellow-billed Cuckoo occupancy, we applied guidelines for defining occupancy described in the human-based survey protocol. Our field season (June 15 – August 15) was broken into four survey periods, with survey periods beginning on the 1st and 16th of each month. As previously described, a quad was categorized as occupied if it contained portions of an occupied territory as determined through protocol surveys. We then applied this framework to estimate occupancy at each quad using our detection data obtained through ARUs and classifiers. Detected cuckoo calls were grouped according to the ARU at which they were detected, and the survey period in which they occurred. If cuckoos were detected by the classifier at the same ARU during two or more survey periods (with the detections a minimum of 10 days apart), that ARU's quad was estimated to be occupied according to the acoustic data. We then compared the estimated occupancy status results for each site and each quad according to both human-based protocol surveys and ARUs to determine whether ARU detections accurately reflected the results from protocol surveys.

X. Optimizing deployment parameters

To determine whether reduced amounts of ARU data could yield reliable occupancy estimates, we compared occupancy estimates from protocol surveys against multiple subsets of our acoustic dataset (Table 3.1). We first used the results from our call patterns analysis to identify peak hours of call activity and reduce our full, 7-hours daily recording dataset to two new datasets of three hours daily and two hours daily. To each of these three datasets, we then applied two customized, intermittent recording schedules commonly used in research and management to further subset the data. These included a 5-minute on-off recording interval, and a 15-minutes-on, 45-minutes-off interval beginning at the top of each hour.

Table 3.1 ARU recording schedule simulations used to test whether reduced recording times could yield accurate occupancy estimates of western Yellow-billed Cuckoos compared to occupancy determined by human observers.

Begin/end times	Sampling	Hours of daily recording	Total hours of data per ARU
04:00-11:00	Continuous	7	420
05:00-08:00	Continuous	3	180
05:00-07:00	Continuous	2	120
04:00-11:00	5-min on-off	3.5	210
05:00-08:00	5-min on-off	1.5	90
05:00-07:00	5-min on-off	1	60
04:00-11:00	First 15 min/hr	1.75	105
05:00-08:00	First 15 min/hr	0.75	45
05:00-07:00	First 15 min/hr	0.5	30

RESULTS

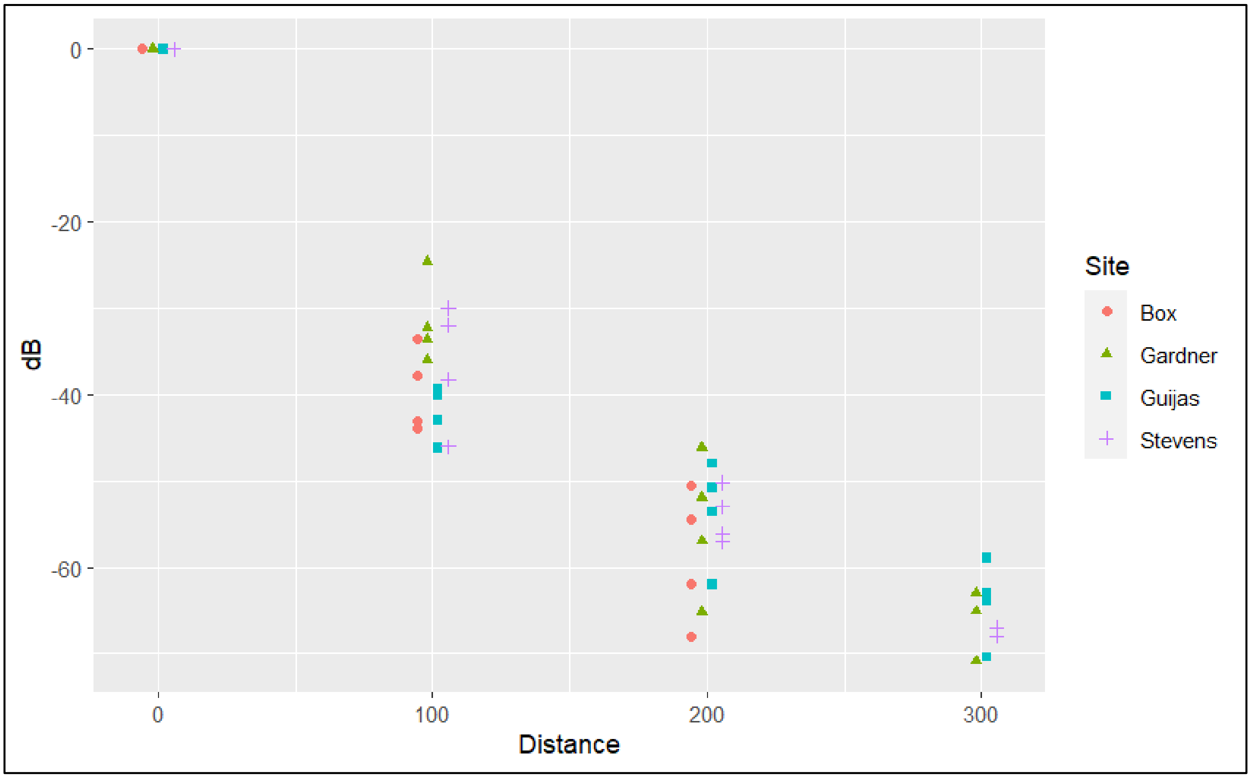
I. Protocol surveys & nest searching

Yellow-billed Cuckoos were detected in all 5 sites during protocol surveys, with a total of 46 detections during the 4 required protocol surveys between June 15 – August 15. Based on protocol guidelines, 4 of 5 sites were considered occupied, with a total of 13 occupied quads, while the fifth site had detections during one survey and no evidence of breeding. Follow-up nest searching confirmed breeding at all four occupied sites. Detections at the fifth site were presumed to be from non-breeding, migrant or transient birds.

II. ARU range

Contact calls played back at 70 dB were recorded at all ARUs in all 4 occupied sites when at 200 m distance, but the number of ARUs detecting signal calls varied from 0-4 across sites at 300 m distance, while no ARUs in any sites recorded playback calls at 400 m or greater distances (Figure 3.2).

Figure 3.2 Changes in sound pressure level (relativized to 70 dB at 1 m) of recorded Yellow-billed Cuckoo contact calls played back at 100 m interval distances at each ARU within each occupied site.



III. Acoustic analysis

All ARUs were functional for the entirety of the field season and recorded until at least September 15 in all sites. ARUs collected a total of 12,880 hours, which amounted to approximately 8 terabytes of data. Using the Kaleidoscope Pro classifier, a total of 89,759 detected signals were classified into the contact call cluster, and 8,449 signals were classified into the coo call cluster. Manual review of all detections in the contact and coo clusters took approximately 40 hours and resulted in 5,380 true-positive cuckoo detections (1772 contact and 3608 coo). Between June 15 – August 15, 4061 true-positive calls were detected (1609 contact and 2452 coo). The true positive rate was 38% for coo calls and 3% for contact calls, with an overall true positive rate of 6%. The majority of false positives in the contact call cluster were from Yellow-breasted Chats, Cactus Wrens, and Mexican Jays, while the majority of false positives in the coo call cluster were from several species of dove and nightjar, as well as cattle and anthropogenic sources.

Manual review of the benchmark dataset resulted in 544 contact calls and 638 coo calls. Of these, the classifier detected 69 contact calls and 325 coo calls, resulting in a false negative rate of 86.58% for contact/alarm calls and 49.06% for coo calls. The combined Coo and Contact clusters resulted in an overall precision score of 0.06 and recall score of 0.33, with an F-score ($\beta = 1$) of 0.102.

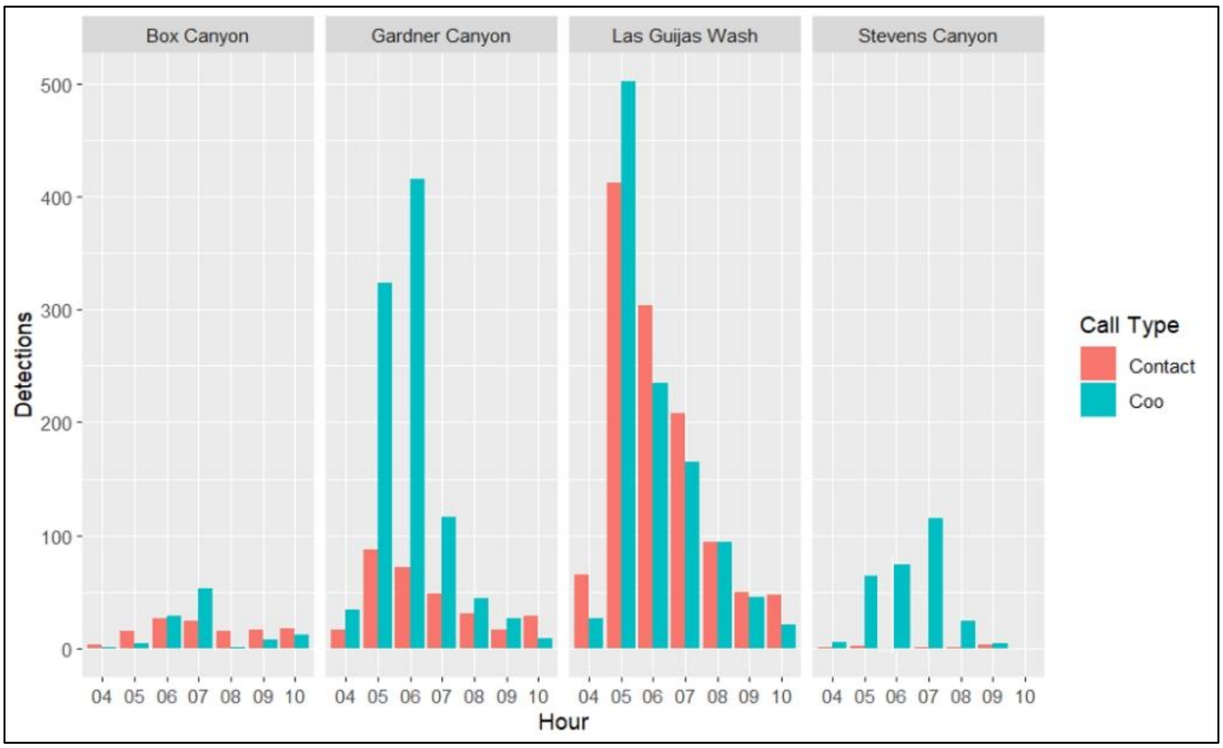
IV. Temporal call patterns

Across our 4 occupied sites, the total number of calls detected ranged from 225–2270 per site and the number of days on which calls were detected ranged from 17–61 days per site between June 15 and August 15. Our fifth site (Hunter Canyon) detected only two calls on a single day, which was consistent with protocol survey results of an unoccupied status, and we therefore removed this site from further analysis. Coo calls accounted for 60% of all detected calls. However, contact calls were detected on more days overall (62) than coo calls (55).

For both call types, calls were most commonly detected between 0500 and 0700 (Figure 3.3), consistent with previous observations by Hamilton & Hamilton (1965), with a downward trend in subsequent morning hours. Overall, rates of contact calls peaked in the first half of July and slowly declined through August. Coo calls peaked once in late July and again in late August but were largely driven by relatively few prolonged bouts of cooing.

Evidence of breeding was documented at 3 locations within 100 m of individual ARUs, each at different sites. We also documented an additional 3 breeding locations within 250 m of ARU’s, and 1 breeding location approximately 600 m from the nearest ARU. While the total number of calls detected varied between ARUs located near breeding territories, we observed spikes in coo calls detected during the estimated days of nest initiation and completion in all of these ARUs, particularly those < 100 m from nests. At two sites, contact calls were consistently detected during the entire nesting period, with a slight decrease in daily detections during the middle of the estimated breeding period. At another site, no contact calls were detected during the estimated incubation period of a nest < 100 m away from an ARU but contact calls were detected on other adjacent units during this time.

Figure 3.3 Total Yellow-billed Cuckoo contact and coo calls detected each hour across 4 occupied sites in 2018.



V. Occupancy

Protocol surveys resulted in 13 of 16 quads considered occupied while ARU data estimated 10 to 16 occupied quads depending upon recording schedule (Figure 3.4). The only two recording schedules which failed to identify the same quads occupied as the protocol surveys were those that used data recorded during 15 minutes of each hour during either 2 or 3 hours in the morning, a total of 30 and 45 minutes/day respectively. The schedule that drew data from every other 5 minutes for the 2 hours between 0500 and 0700, a total of 60 minutes/day, gave the same results as all ARU schedules requiring more time of recording (Figure 3.4).

Figure 3.4 Occupancy table comparing results of Yellow-billed Cuckoo protocol surveys against results of nine unique ARU recording schedules in each of four quads at four study sites with confirmed breeding activity in southeastern Arizona from June 15 – August 15, 2019.

		4 survey periods of acoustic data								
		Full Continuous ARU Dataset			5min On-Off Interval			First 15min/Hour		
Quad	Protocol Surveys	7 Hours	3 Hours	2 Hours	7 Hours (3.5 hrs daily)	3 Hours (1.5 hrs daily)	2 Hours (1 hr daily)	7 Hours (1.75 hrs daily)	3 Hours (0.75 hrs daily)	2 Hours (0.5 hrs daily)
Gardner 1	O	O	O	O	O	O	O	O	O	O
Gardner 2	O	O	O	O	O	O	O	O	O	O
Gardner 3	O	O	O	O	O	O	O	O	O	O
Gardner 4	O	O	O	O	O	O	O	O	O	O
Las Guijas 1	O	O	O	O	O	O	O	O	O	O
Las Guijas 2	O	O	O	O	O	O	O	O	O	O
Las Guijas 3	O	O	O	O	O	O	O	O	O	O
Las Guijas 4	O	O	O	O	O	O	O	O	O	O
Box 1	O	O	O	O	O	O	O	O	O	O
Box 2	O	O	O	O	O	O	O	O	O	O
Box 3	O	O	O	O	O	O	O	O	O	-
Box 4	O	O	O	O	O	O	O	O	-	-
Stevens 1	-	O	O	O	O	O	-	O	-	-
Stevens 2	-	O	O	O	O	-	-	O	-	-
Stevens 3	-	O	O	O	O	O	-	O	O	-
Stevens 4	O	O	O	O	O	O	O	O	O	-

O	Occupied
-	Unoccupied

DISCUSSION

We found that ARUs paired with our associated classifier identified occupied xeroriparian habitat of the western Yellow-billed Cuckoo as reliably as traditional human-observer, playback survey protocols. With known locations of breeding cuckoos near ARUs and repeat-visit protocol surveys, we were able to interpret acoustic data and identify key components of ARU monitoring programs for cuckoos.

I. Classifier performance

We developed a Kaleidoscope Pro classifier with two manually defined clusters for coo and contact calls, and 31 additional clusters defined by other background signals in the training data. Our classifier detected coo calls more efficiently than contact calls, with higher rates of both recall and precision. This may be due to the lower variability in the structure of the coo call as compared to the contact call, as well as the lower and narrower frequency range of the coo call. The natural variation in the structure of the contact call - including alarm and kowlp variations - resulted in a more loosely defined statistical algorithm used to form the contact call cluster, which therefore had increased overlap with other, similarly structured, non-cuckoo sounds and their associated clusters in the classifier.

As the analysis of our benchmark dataset demonstrated, low recall for the contact calls meant an estimated 86% of contact calls were missed by the classifier. This is consistent with other studies using Kaleidoscope Pro to study cryptic species (88.6% and 77.9% false negative rates, Schroeder & McRae 2020; 99.03% false negative rate, Bobay et al. 2018) but is much higher than some studies using other emerging software platforms such as BirdNet (6% and 24% false negative, Manzano-Rubio et al., 2022) or convolutional neural networks (25% false negative, Knight et al. 2017). Given that classifier development is a supervised process with many user-defined parameters at each step, further experimentation based on our results would likely lead to improvements in performance and efficiency. However, despite the high false negative rate, our classifier resulted in sufficient detections to estimate occupancy using temporal and spatial criteria adopted from the Halterman et al. (2015) survey protocol. This suggests that even a poor-performing classifier may be sufficient for obtaining biologically relevant occupancy information.

II. Interpretation of detection results

Overall, cuckoos were detected in as many or more survey quads by ARUs than protocol surveys, demonstrating the utility of ARUs to provide spatial and temporal detection patterns similar to traditional call-back surveys. Additionally, ARU data may offer the potential to confirm breeding behavior beyond occupancy. ARUs deployed near known nests consistently resulted in high rates of coo calls before and after the estimated nesting cycle, with generally reduced rates of all call types during the estimated incubation period. This suggests that a dual-peak of coo call activity separated by 1-3 weeks may be an important indicator of cuckoo nesting, and is consistent with the hypotheses that 1) females use coo calls to solicit copulation during the egg laying stage (Hughes 1999), and 2) that females may initiate new nests soon after young have hatched (Halterman 2009).

Variability in call detection rates among sites could be driven by a variety of factors that influenced the signal detection space of ARUs, such as dense vegetation, running water, anthropogenic noise or geomorphology (Darras et al. 2016). Geomorphology was likely important at our site where the fewest number of calls were detected despite visual confirmation of three nests at that site. The false negative rate was 83% (3% lower than the rate among all sites), and therefore low detection rates at this site were not due to site-specific low recall scores.

Instead, ARUs at this site had the shortest average range, likely due to the steep terrain and winding drainage of the site. This highlights the importance of evaluating local factors affecting ARU detection space when designing acoustic monitoring programs and analyzing results.

Despite poor classifier performance for the contact call cluster, contact calls were detected over more overall days than coo calls in all sites except one. Therefore, monitoring for contact calls may provide superior temporal resolution of occupancy despite the low rate of detection. Likewise, without the inclusion of coo calls at our Stevens Canyon site, where only 10 total contact calls were detected, occupancy would have been underestimated even though we had confirmed breeding at that site. These results highlight the importance of recording sufficient calls from either call types for effective monitoring programs and reliable occupancy estimation. Finally, at our fifth site that both human-based and ARU methods classified as unoccupied, and where subsequent nest searching failed to find any evidence of breeding, both survey methods detected cuckoos early in the season with no subsequent detections. This demonstrates the efficacy of both methods in detecting migrant and/or transient cuckoos and provides support for the use of occupancy criteria for both methods to require detections over multiple survey periods in order to avoid over-estimating occupancy.

III. Recommendations for ARU monitoring programs

Recording schedules for ARU monitoring programs ultimately depend on specific project goals and available resources (Blumstein et al. 2011, Shonfield & Bayne 2017, Gibb et al. 2019, Manzano-Rubio et al. 2022). For monitoring programs seeking to determine western Yellow-billed Cuckoo occupancy using classifiers with efficacy similar to ours, we recommend 1) recording schedules with a minimum of 30 minutes per hour between 0500-0700, 2) spacing of ARUs at 300m intervals unless complex terrain likely obstructs sound to distances less than

300m, 3) the use of classifiers for both coo and contact calls, and 4) further experimentation in classifier development, including splitting the contact call into distinct “kowlp” and “alarm” clusters to reduce signal variability within each cluster. Together, such efforts will provide reliable occupancy estimates while also providing advances in acoustic monitoring, which could be shared and adapted by others.

ACKNOWLEDGEMENTS

Work presented in this chapter was supported by contributions from Dr. Tad Theimer, Susan Sferra, and Dr. Bret Pasch. This research was conducted under U.S. Fish & Wildlife Service Recovery Permit #TE81148C-0 and Arizona Game & Fish Department Special Collecting License permit #SP407168. The findings and conclusions in this article are those of the authors and do not necessarily represent the views of the U.S. Fish and Wildlife Service. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the United States Government. Data for this study are available as a U.S. Geological Survey data release.

REFERENCES

Abrahams, C., & Geary, M. (2020). Combining bioacoustics and occupancy modelling for improved monitoring of rare breeding bird populations. Ecological Indicators, 112, 106–131. https://doi.org/10.1016/j.ecolind.2020.106131

Alquezar, R. D., & Machado, R. B. (2015). Comparisons Between Autonomous Acoustic Recordings and Avian Point Counts in Open Woodland Savanna. The Wilson Journal of Ornithology, 127(4), 712–723. https://doi.org/10.1676/14-104.1

Blumstein, D. T., Mennill, D. J., Clemins, P., Girod, L., Yao, K., Patricelli, G., Deppe, J. L., Krakauer, A. H., Clark, C., Cortopassi, K. A., Hanser, S. F., McCowan, B., Ali, A. M., & Kirschel, A. N. G. (2011). Acoustic monitoring in terrestrial environments using microphone arrays: Applications, technological considerations and prospectus: Acoustic monitoring. Journal of Applied Ecology, 48(3), 758–767. https://doi.org/10.1111/j.1365-2664.2011.01993.x

Bobay, L. R., Taillie, P. J., & Moorman, C. E. (2018). Use of autonomous recording units increased detection of a secretive marsh bird. Journal of Field Ornithology, 89(4), 384–392. https://doi.org/10.1111/jofo.12274

Brown, D. E. (Ed.). (1994). Biotic communities: Southwestern United States and Northwestern Mexico. Univ. of Utah Press.

Campos-Cerqueira, M., & Aide, T. M. (2016). Improving distribution data of threatened species by combining acoustic monitoring and occupancy modelling. Methods in Ecology and Evolution, 7(11), 1340–1348. https://doi.org/10.1111/2041-210X.12599

Darras, K., Batáry, P., Furnas, B., Celis-Murillo, A., Van Wilgenburg, S. L., Mulyani, Y. A., & Tscharntke, T. (2018). Comparing the sampling performance of sound recorders versus point counts in bird surveys: A meta-analysis. Journal of Applied Ecology, 55(6), 2575–2586. https://doi.org/10.1111/1365-2664.13229

Darras, K., Pütz, P., Fahrurrozi, Rembold, K., & Tscharntke, T. (2016). Measuring sound detection spaces for acoustic animal sampling and monitoring. Biological Conservation, 201, 29–37. https://doi.org/10.1016/j.biocon.2016.06.021

Gibb, R., Browning, E., Glover-Kapfer, P., & Jones, K. E. (2019). Emerging opportunities and challenges for passive acoustics in ecological assessment and monitoring. Methods in Ecology and Evolution, 10(2), 169–185. https://doi.org/10.1111/2041-210X.13101

Halterman, M. (2009). Sexual dimorphism, detection probability, home range, and parental care in the Yellow-billed Cuckoo. University of Nevada, Reno.

Halterman, M. D., Johnson, M. J., Holmes, J. A., & Laymon, S. A. (2015). A Natural History Summary and Survey Protocol for the Western Distinct Population Segment of the Yellow-billed Cuckoo. U.S. Fish and Wildlife Techniques and Methods.

Hamilton, W. J., III, & Hamilton, M. E. (1965). Breeding characteristics of yellow-billed cuckoos in Arizona. Proceedings of the California Academy of Sciences, 32(14), 405–432.

Hughes, J. M. (1999). Yellow-billed Cuckoo (Coccyzus americanus). In A. Poole & F. Gill (Eds.), The Birds of North America (p. 28). The Birds of North America, Inc.

Hughes, J. M. (2015). Yellow-billed Cuckoo (Coccyzus americanus), version 2.0. In P. G. Rodewald (Ed.), The Birds of North America Online. Cornell Lab of Ornithology. https://birdsoftheworld.org/bow/historic/bna/yebcuc/2.0/introduction

Hutto, R. L., & Stutzman, R. J. (2009). Humans versus autonomous recording units: A comparison of point-count results. Journal of Field Ornithology, 80(4), 387–398. https://doi.org/10.1111/j.1557-9263.2009.00245.x

Kahl, S., Wood, C. M., Eibl, M., & Klinck, H. (2021). BirdNET: A deep learning solution for avian diversity monitoring. Ecological Informatics, 61, 101236. https://doi.org/10.1016/j.ecoinf.2021.101236

Knight, E. C., Hannah, K. C., Foley, G. J., Scott, C. D., Brigham, R. M., & Bayne, E. (2017). Recommendations for acoustic recognizer performance assessment with application to five common automated signal recognition programs. Avian Conservation and Ecology, 12(2), art14. https://doi.org/10.5751/ACE-01114-120214

Manzano-Rubio, R., Bota, G., Brotons, L., Soto-Largo, E., & Pérez-Granados, C. (2022). Low-cost open-source recorders and ready-to-use machine learning approaches provide effective monitoring of threatened species. Ecological Informatics, 72, 101910. https://doi.org/10.1016/j.ecoinf.2022.101910

McNeil, S. E., Tracy, D., Stanek, J. R., Stanek, J. E., & Station, S. S. R. (2013). Yellow-billed Cuckoo Distribution, Abundance, and Habitat Use on the Lower Colorado River and Tributaries 2008-2012 Summary Report (p. 331). Lower Colorado River Multi-Species Conservation Program, U.S. Bureau of Reclamation.

Potamitis, I., Ntalampiras, S., Jahn, O., & Riede, K. (2014). Automatic bird sound detection in long real-field recordings: Applications and tools. Applied Acoustics, 80, 1–9. https://doi.org/10.1016/j.apacoust.2014.01.001

Schroeder, K. M., & McRae, S. B. (2020). Automated auditory detection of a rare, secretive marsh bird with infrequent and acoustically indistinct vocalizations. Ibis, 162(3), 1033–1046. https://doi.org/10.1111/ibi.12805

Sechrist, J., Ahlers, D. D., Zehfuss, K. P., Doster, R. H., Paxton, E. H., & Ryan, V. M. (2013). Home Range and Use of Habitat of Western Yellow-Billed Cuckoos On the Middle Rio Grande, New Mexico. The Southwestern Naturalist, 58(4), 411–419. https://doi.org/10.1894/0038-4909-58.4.411

Shonfield, J., & Bayne, E. M. (2017). Autonomous recording units in avian ecological research: Current use and future applications. Avian Conservation and Ecology, 12(1), art14. https://doi.org/10.5751/ACE-00974-120114

Sokolova, M., Japkowicz, N., & Szpakowicz, S. (2006). Beyond Accuracy, F-Score and ROC: A Family of Discriminant Measures for Performance Evaluation. In A. Sattar & B. Kang (Eds.), AI 2006: Advances in Artificial Intelligence (Vol. 4304, pp. 1015–1021). Springer Berlin Heidelberg. https://doi.org/10.1007/11941439_114

Stowell, D., Wood, M. D., Pamuła, H., Stylianou, Y., & Glotin, H. (2019). Automatic acoustic detection of birds through deep learning: The first Bird Audio Detection challenge. Methods in Ecology and Evolution, 10(3), 368–380. https://doi.org/10.1111/2041-210X.13103

USFWS. (2014). Endangered and Threatened Wildlife and Plants; Determination of Threatened Status for the Western Distinct Population Segment of the Yellow-billed Cuckoo (Coccyzus americanus); Final Rule. Federal Register, 79(129), 59992–60038.

USFWS. (2021). Endangered and Threatened Wildlife and Plants; Designation of Critical Habitat for the Western Distinct Population Segment of the Yellow- Billed Cuckoo. Federal Register.

Warshall, P. (1995). The Madrean sky island archipelago: A planetary overview (pp. 6-18). US Department of Agriculture, Forest Service, Rocky Mountain Forest and Range Experiment Station; Vol. 1995.

Chapter 4 Incorporating recently documented breeding habitat to improve species distribution models for the federally threatened western Yellow-billed Cuckoo in Arizona

ABSTRACT

Recent survey efforts for the federally-threatened western Yellow-billed Cuckoo have documented cuckoos breeding in xeroriparian habitat in many upper-watershed ephemeral and intermittent drainages in southeastern Arizona, conflicting with previous assumptions that western populations were riparian obligates. We (Northern Arizona University, in conjunction with U.S. Fish and Wildlife Service, U.S. Geological Survey, and Arizona Game and Fish Department) developed a MaxEnt species distribution model (SDM) for cuckoos using topographic, climate, and phenology variables derived from Normalized Difference Vegetation Index (NDVI) satellite imagery to account for this broader spectrum of riparian and xeroriparian habitat. Our final SDM contained 1 topographic, 4 climate, and 3 phenology variables, and had an AUC score of 0.976. Top variables included NDVI at the start of the season, mean annual temperature, timing of the end of the productive season, and topographic wetness index, with permutation importance of 51, 24.7, 9.3, and 7.4% respectively. Our model output map predicts high likelihood of occupancy of many major riparian systems consistent with previous studies, but also identified two new core areas: mid-elevation xeroriparian drainages in the "Sky Islands" region of southeastern Arizona, and broad low-elevation ephemeral washes primarily on the Tohono O'odham Nation in south-central Arizona. Our results highlight that many areas outside of cottonwood-willow-dominated riparian areas should be considered cuckoo habitat in the southern portion of the western Yellow-billed cuckoo's range. Riparian habitat in our study area is likely to decline due to continued and increasing human demand for water, and xeroriparian

habitat could potentially serve as a refugia for cuckoos as riparian conditions decline, unless climate change alters current monsoonal patterns that drive productivity in these xeroriparian uplands.

INTRODUCTION

Identifying occupied habitat for endangered species is important both for discovering previously unknown populations and for identifying areas of high conservation and management value (Foin et al. 1998, Camaclang et al. 2015). This process can be accelerated by identifying environmental variables that influence habitat conditions and applying them to biologically informed species distribution models (SDMs) (Scott et al. 2005, Franklin 2013, Tulloch et al. 2016). The environmental variables, quality of the presence data, and a priori assumptions about what areas should be excluded in SDMs all influence model output (Elith & Leathwick 2009, Merow et al. 2013, Guillera-Arroita et al. 2015). In developing SDMs, researchers often define the study area extent to exclude areas assumed to be unsuitable based on available knowledge of the study species (Acevedo et al. 2012). If subsequent study reveals those assumptions about habitat were invalid, however, models need to be updated to incorporate that new knowledge (Williams et al. 2009, Barve et al. 2011, Wisz et al. 2013).

The federally-threatened, western distinct population segment (DPS) of the Yellow-billed Cuckoo (*Coccyzus americanus*) has long been considered a riparian obligate, requiring large tracts of mesic cottonwood-willow woodland (Hamilton & Hamilton 1965, Girvetz & Greco 2009, Johnson et al. 2017, Wohner et al., 2021). This habitat is typically restricted to broad low-gradient floodplains with permanent surface or groundwater. Therefore, most field surveys, such as a 1998–1999 statewide effort in Arizona (Corman & Magill 2000), focused on these riparian habitats. Subsequent habitat and distribution models based on these surveys excluded consideration of habitat in drainages away from major streams and rivers (Villarreal et al. 2014, Johnson et al. 2017). However, recent surveys have documented cuckoos breeding in many ephemeral and intermittent mid- and upper-watershed drainages in the "Sky Islands" region of southeastern Arizona (MacFarland & Horst 2015,Westland Resources, Inc. 2015, MacFarland & Horst 2017, USFWS 2021). These drainages are typically narrower, steeper, and more xeric (i.e., ephemeral or intermittent), and lack large stands of cottonwood-willow woodland. They are instead best described as "xeroriparian" (Johnson & Haight 1985, Hardy et al. 2004, Beauchamp & Shafroth, 2011, Bateman & Riddle 2020) and characterized primarily by more arid-adapted riparian trees such as mesquite, hackberry, walnut, ash, sycamore, oak, and juniper, and adjacent to uplands comprised of Madrean-oak woodland, semi-desert grassland, or upper-Sonoran desert scrub (Brown 1994). Although documented use of xeroriparian habitat is currently confined to a relatively small portion of Arizona, numerous confirmed breeding locations (Beauregard (Chapter 1)) conflict with previous assumptions of strict riparian habitat requirements for cuckoos and indicates a potentially regionally significant population that has previously been unaccounted for. This additional breeding habitat has not been included in any previous models of potential habitat for Yellow-billed Cuckoos in the southwestern United States.

The influence of the North American monsoon on vegetation phenology has previously been linked to cuckoo habitat selection (Wallace et al. 2013) and timing of nesting in southeastern Arizona (Hamilton & Hamilton 1965), and this link may be particularly important in determining suitability of xeroriparian areas. Following a period of typically hot and dry conditions, the monsoon season extends from June to September, which coincides with the cuckoo's delayed breeding season (relative to other neotropical migrants), and accounts for a large proportion of annual precipitation in northwestern Mexico and the southwestern United States where its influence is strongest (Higgins et al. 1997, Forzieri et al. 2011, Crimmins et al. 2011). Increased precipitation and ambient humidity, coupled with cooler temperatures, supports a brief but pronounced green-up of vegetation in both drainages and uplands (Crimmins et al. 2011). In riparian habitat, monsoons effects may serve to increase foraging opportunities within and around core cuckoo breeding areas (Wallace et al. 2013) while cuckoos may be using xeroriparian woodlands in drainages where productivity is heavily augmented by monsoon precipitation, resulting in temporarily suitable nesting conditions in areas lacking cottonwood-willow woodland. If so, then cuckoo SDMs in the southwestern United States could be improved by including xeroriparian drainages in the study area extent and selecting variables that reliably characterize spatial and temporal patterns of vegetation phenology and productivity along a continuum of climatic and geophysical conditions.

Remotely sensed Normalized Difference Vegetation Index (NDVI) data are commonly used to approximate vegetation phenology metrics through measurements of vegetation "greenness" which can be interpreted as a proxy for productivity (Pettorelli et al. 2005, Wiegand et al. 2008). NDVI-derived metrics are widely used in habitat models for a variety of taxa (Wiegand et al. 2008, York et al. 2011, Wen et al. 2015). Villareal et al. (2012) used NDVI to improve MaxEnt (Phillips 2005, Elith et al. 2011) predictive SDMs for cuckoos in Arizona riparian habitat, and Wallace et al. (2013) described the influence of monsoon-related vegetation phenology on cuckoo habitat selection in southeastern Arizona. While these studies provide support for including NDVI and monsoonal influence into models of cuckoo habitat, these models were developed prior to the documentation of widespread use of xeroriparian habitat and relied heavily on assumptions of cuckoo dependence on cottonwood-willow habitat. Therefore, these models are not applicable for the identification and management of xeroriparian habitat.

The objective of our (Northern Arizona University, in conjunction with U.S. Fish and Wildlife Service, U.S. Geological Survey, and Arizona Game and Fish Department) study was to produce an improved Yellow-billed Cuckoo species distribution model that accounts for newly available location data in xeroriparian breeding habitat in the southwestern portion of its range. Our approach was to first identify spatial datasets characterizing geographic, bioclimatic, and phenological factors that may influence cuckoo distribution in both riparian and xeroriparian habitat. We then used these variables and presence data from an exhaustive survey effort in 2022 to produce a new MaxEnt species distribution model. Finally, we evaluated variable contributions and model output to identify implications for management and species recovery.

METHODS

I. Modeling approach

Species distribution models combine species location data with spatial datasets of ecologically relevant variables to produce predictive maps of potential species distribution and help identify variables that tend to be most important in predicting such habitat (Elith et al. 2011, Radosavljevic & Anderson 2014, Zeng et al. 2016). MaxEnt uses presence locations (also called training data) and randomly chosen pseudoabsence points (Phillips et al. 2009, Barbet-Massin et al. 2012) to predict species occurrences by identifying the distribution that is closest to uniform, while accounting for limits of the environmental variables measured at known locations (Elith et al. 2011). MaxEnt SDMs are robust to highly complex models with correlated variables but provide more useful results when fewer variables are selected based on hypothesized biological relevance (Warren & Seifert 2011, Bradie & Leung 2017, Feng et al. 2019). MaxEnt also relies on assumptions of random sampling (Elith et al. 2011, Yackulic et al. 2013) but allows methodology for accounting for spatial sampling bias within modeling area extents (Phillips et al. 2009, Kramer-Schadt et al. 2013). For these reasons, MaxEnt was an appropriate approach for our goal of producing a regional SDM based on standardized survey data and evaluating variable contribution to identify potential influences on distribution and habitat quality.

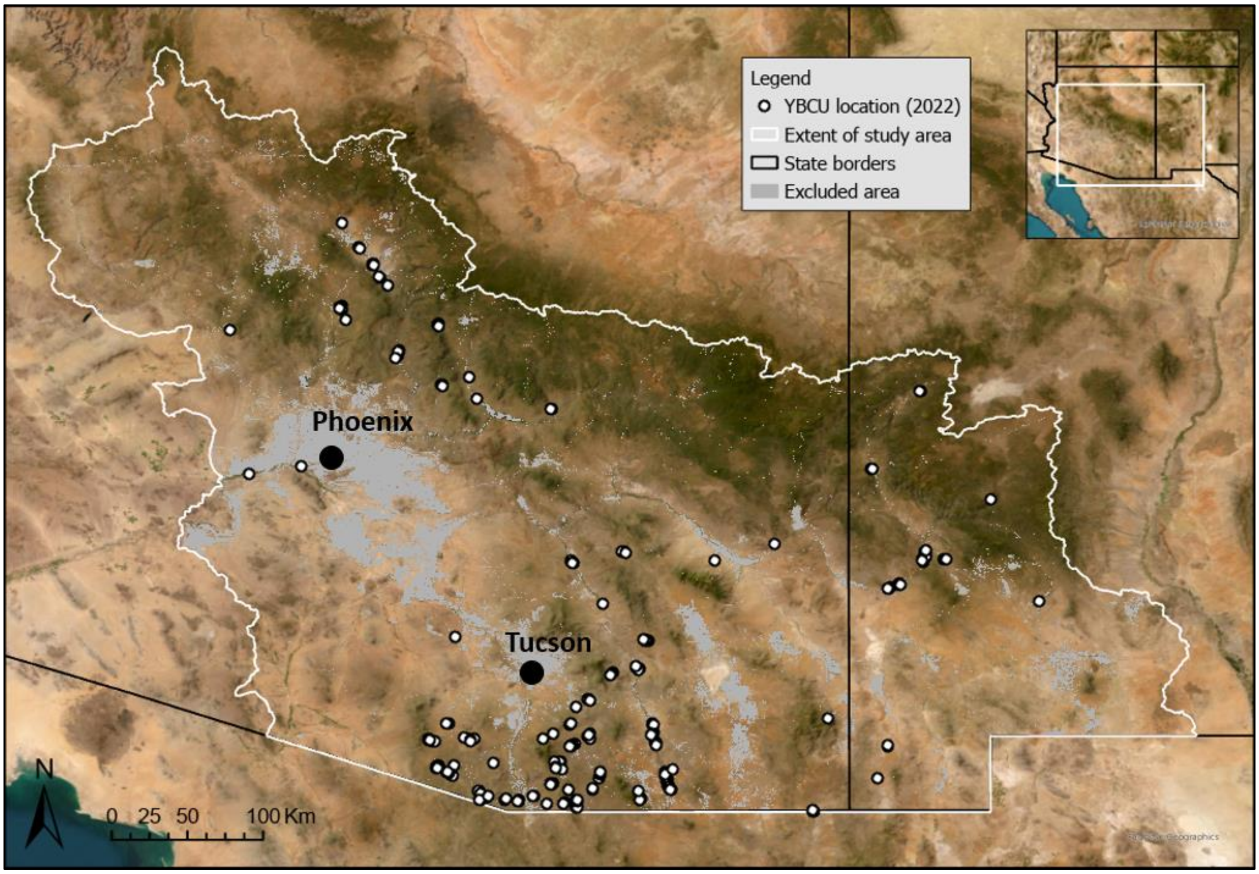
II. Study area

We chose a study area extent that encompasses the monsoon-influenced region of the southern Colorado Plateau and Madrean Sky Island Archipelago of Arizona and New Mexico (Figure 4.1). Our study area extent was delineated using the combined boundaries of multiple HUC-8 watersheds, with the southern boundary being the US-MX border. This region primarily consists of the middle- and upper-Gila River watershed, as well as small portions of the Mimbres River watershed in NM, and the Bill Williams River, San Simon Wash, and Rio De La Concepcion watersheds in AZ.

The study area contains high topographic, climatic, and ecological heterogeneity, with elevations ranging from ~500–3000 m. The northern portion of the study area is characterized by mountainous, mid-high elevation pinyon-juniper or mixed conifer woodland, interior chapparal, semi-desert grassland, and desert upland scrub (Brown 1994), with narrow riparian and xeroriparian drainages that generally drain south and west into the Gila River. The Sky Island region consists of multiple ranges extending from the Gila River near its confluence with the San Pedro River, south to the Mexican border (and further south to the northern Sierra Madre in Mexico), west to the Baboquivari Mountains, and east to the Peloncillo Mountains straddling the Arizona-New Mexico border. This topographically complex region is similar to the basin-and-range region but contains larger elevation gradients of ~750–3000 m, with diverse habitat types ranging from lower Sonoran and Chihuahuan deserts, semi-desert grassland, Madrean-oak woodland, pinyon-juniper woodland, and mixed conifer woodland, as well as lowland riparian and xeroriparian vegetation types (Brown 1994). The southwestern portion of the study area includes low-elevation (~250–1250 m) basin-and-range topography with primarily lower Sonoran and Mojave Desert vegetation, whereas the southeastern portion of the study area includes low-mid elevation (~750–2000 m) basin-and-range topography and Chihuahuan desert vegetation.

Many anthropogenic landscape features were present within the outer boundaries of our study area, including cities, irrigated agriculture, mines, highways, and other infrastructure. Although in some cases cuckoos may occupy habitat in or adjacent to some of these modified landscapes (Kingsley 1985, eBird 2023), we chose to exclude these areas given that we were predominantly interested in cuckoo distribution in natural areas with minimal disturbance. We used LANDFIRE (LANDFIRE 2022) physical land cover data in ArcPro (ESRI) to identify areas categorized as Quarries/Mines, Developed, Agriculture, Roads, or Open Water, which were combined into a single group and excluded from the study area. We first used the Resample tool to adjust the scale of the LANDFIRE data to the 250 m^2 resolution of our model rasters (see below), using Nearest Neighbor resampling technique which is appropriate for categorical variables. We then used the Reclassify tool to reprocess the LANDFIRE raster such that it only retained the remaining “natural” physical cover types as a single category. Finally, we used the Raster to Polygon tool to convert the raster to a polygon feature, which represented the final boundaries of the study area extent.

Figure 4.1 Study area and Yellow-billed Cuckoo presence data used for species distribution modeling of habitat in AZ and NM, USA.



III. Yellow-billed Cuckoo presence data

We developed our MaxEnt SDM using presence data collected in 2022 as part of a range-wide survey effort funded by a USFWS Competitive State Wildlife Grant (CSWG) (Stanek et al., in prep). This survey effort was coordinated by state wildlife agencies and several non-governmental organizations and included surveys in 12 western states within the cuckoo's western DPS. We used the subset of these results that were within our study area in Arizona and New Mexico.

All sites for the CSWG project were selected from a 1 km^2 grid, which was based on the boundaries of USFWS designated critical habitat (USFWS 2021), the locations of historic cuckoo detections (which were obtained through state agency databases), and from the model output of a preliminary SDM developed by Utah Division of Wildlife Resources (unpublished data). From this 1km^2 gridded sampling frame, 75% of sites were randomly selected from critical habitat or historic detection data, and 25% of sites were randomly selected from the preliminary SDM. Sites were manually assessed by experts and any sites that were deemed inaccessible or that contained no potential habitat (i.e., no riparian or xeroriparian woodland cover) were discarded.

Field sites were surveyed using a standardized survey protocol (Halterman et al. 2015) which uses call-playback methods along a survey transect that follows a drainage or patch of habitat. However, our survey differed from the standardized survey protocol in that each site was surveyed three times between June 15 and August 1 instead of four times between June 15-August 15. Survey transects were required to bisect 1 km^2 sampling unit grid cells such that at least one playback point could be established, but they could also continue away from grid cells to cover additional habitat. Locations of detected cuckoos were recorded using FieldMaps (ESRI) on smartphones. After spatially thinning the dataset (see below), 331 presence locations remained to be included in our dataset.

We also used data collected in Arizona riparian sites in 1998–1999 (Corman & Magill 2000) and data collected between 2018 and 2020 in xeroriparian habitat (Beauregard et al., in prep). These data were not used as training data for the model but were instead used as an independent test of model performance, from which we derived thresholds for interpreting model output (see below). In both cases, data were not collected according to the current standard survey protocol

(Halterman et al. 2015) but were collected as part of efforts focused specifically on documenting cuckoos in appropriate habitat. The 1998–1999 data consisted of 484 presence locations (after reducing to 1 point per 250 m^2 raster cell) across the state, and the 2018-2020 dataset consisted of 54 documented breeding locations in southeastern Arizona xeroriparian habitat.

IV. Sampling bias

MaxEnt species distribution modeling relies on the assumption that all locations within the study area extent have equal chance of being sampled (Yackulic et al. 2013). This assumption is usually unreasonable to meet for rare species or study areas with inaccessible habitat (Kramer-Schadt et al. 2013, Fourcade et al. 2014). In our case, our presence data were collected using a stratified random sampling approach. However, large areas of potential habitat were intentionally excluded due to inaccessibility, and many sites were based on historic location data. Therefore, to address sampling bias in our presence data, we first thinned our data points such that only one sampling point was present in any given 250m^2 raster cell (Kiedrzyński et al. 2017). We then employed the use of bias files in models, which informs the model on areas of sampling bias and increases emphasis of background data in these areas (Phillips et al. 2009, Kramer-Schadt et al. 2013). We developed a bias file by using the Kernal Density Estimator tool in ArcPro. Our resulting bias file was converted into a ASCI raster file at 250 m^2 resolution.

V. Environmental variables

Given our goals of evaluating the influence of seasonal vegetation phenology in riparian and xeroriparian habitat, we selected 1 topographic, 8 bioclimatic, and 10 phenological continuous variables for use in our models (Table 4.1). In both riparian and xeroriparian habitats, cuckoos are associated with drainage bottoms. Therefore, we used Topographic Wetness Index (Amatulli et al. 2020) to aid the model in differentiating drainage basins over upland areas. This variable is defined as the logarithm of the cumulative upstream catchment area divided by the tangent of the local slope angle, which results in a raster that distinguishes between lowland basins and upland, steep-sloped areas.

To investigate the influence of vegetation phenology on cuckoo models, we used 10 pre-processed NDVI datasets made available by USGS. These pre-processed variables are based on analysis of NDVI values through time-series imagery, resulting in annual temporal phenological curves from which ecologically relevant measurements can be estimated (Pettorelli et al. 2005, Hmimina et al. 2013). These include datasets for the estimated day of the year for the start, end, and peak of the phenological season, as well as the corresponding NDVI values measured at those times, and the amplitude value for the difference in peak NDVI over the baseline. The duration of the growing season (in total days) is also calculated, along with the total accumulated area under the phenological curve (time-integrated NDVI), representing the total photosynthetic activity across the growing season. All of these NDVI datasets are based on averages from 2001–2015.

We also included 8 bioclimatic variables obtained from WorldClim (Fick & Hijmans 2017), which are derived from monthly temperature and precipitation averages between 1970 and 2000. These variables are commonly used in MaxEnt modeling to evaluate climatic factors influencing species occurrence (Bradie & Leung 2017). We specifically selected variables related to conditions either during the cuckoo’s breeding season (precipwettestq, tempseasonality, precipseasonality, tempwarmestq) or annual conditions that could influence habitat conditions during the breeding season (meantemp, annualprecip, tempannualrange, meandiurnalrange) (Ortega-Huerta & Vega-Rivera 2017, Brambilla et al. 2022).

All variable raster data were imported into ArcPro (ESRI). We scaled all variables to match the 250m^2 resolution of the NDVI data using the Resample function in ArcPro, using bilinear interpolation for new raster values. We then used the Clip Raster function to clip all rasters to the dimensions and extent of our study area. Finally, we exported rasters as .ASCI files for use in MaxEnt.

Table 4.1 List of variables used in Yellow-billed Cuckoo species distribution models.

Variable	Type	Description
topowetness	Geography	Topographic Wetness Index, a proxy of the long-term soil moisture availability.
ndvisosn	NDVI	The NDVI at the start of the green-up season.
Ndvitin	NDVI	Time-integrated NDVI
ndvieost	NDVI	The day of the year of the end of the season.
ndviduration	NDVI	The duration (in total number of days) of the growing season.
ndvimaxn	NDVI	Maximum NDVI value.
ndvieosn	NDVI	The NDVI value at the end of the season.
ndvisost	NDVI	The day of the year of the start of the season.
ndvimaxt	NDVI	Day of the year with the maximum NDVI value.

ndviamp	NDVI	Amplitude of change in NDVI values between the start peak NDVI and the baseline.
meantemp	Climate	Mean annual temperature
tempseasonality	Climate	Temperature seasonality (temperature change throughout the season)
precipseasonality	Climate	Precipitation seasonality (precipitation change throughout the season)
tempwarmestq	Climate	Mean temperature of the warmest quarter.
tempannualrange	Climate	Temperature annual range.
precipwettestq	Climate	Mean precipitation of the wettest quarter.
meandiurnalrange	Climate	Mean diurnal range of temperature.
annualprecip	Climate	Mean annual precipitation.
meantemp	Climate	Mean annual temperature.

VI. Model development and selection

We used the MaxEntVariableSelection package in program R (Jueterbock 2015; R Core Team 2023) to select the best combination of variables and β multipliers, which is a penalty coefficient to reduce overfitting (Gastón & García-Viñas 2011). MaxEntVariableSelection runs multiple stepwise iterations of Maxent models through a series of user-defined settings and β multipliers to identify and remove correlated or redundant variables and minimize model complexity. Model iterations are then compared against each other using sample-size-corrected

Akaike Information Criteria (AICc; Akaike 1974, Burnham & Anderson 2004) to indicate the variables that fit the best model.

To process our data through MaxEntVariableSelection, we selected a correlation threshold of 0.95, a variable contribution threshold of 2.5%, and β multipliers of 1-10 by steps of 1. Upon completion of data processing, we identified the model that contained fewer than 18 variables with the lowest AIC_C scores to use in our final MaxEnt model. The final model was then processed using the variables and β values suggested by MaxEntVariableSelection, using autofeatures (excluding threshold). The model was run using a 33% random test percentage, with a random seed for each of 10 subsampled replicates. For all other parameters we used default MaxEnt settings.

VII. Model evaluation

MaxEnt results in several diagnostic output plots, tables, and maps useful for model evaluation (Phillips 2005, Merow et al. 2013). We first evaluated overall model performance by using the threshold-independent method Area Under the Curve of the Receiver Operating Characteristic plot (AUC) (Phillips 2005). We also examined the average omission and predicted area plot. The average omission represents the factional rate of test data omission over an increasing cumulative threshold of the model output (Phillips & Dudík 2008). Thus, a higher proportion of test data is omitted as the cumulative threshold increases and the model's predicted area becomes more discriminating. Accordingly, the predicted area represents the proportion of the study area that is predicted, which decreases with increases in the cumulative threshold. Finally, we extracted the habitat suitability score at each point for the riparian test dataset and the xeroriparian test datasets, and calculated the average, standard deviation, and mean for each of these datasets.

VIII. Mapping distribution

We reviewed the output map depicting average predicted probability from the 10 model replicates to evaluate trends in predicted distribution and compare against previous modeling efforts. To interpret the MaxEnt output map of predicted suitability, we selected three thresholds based on established guidance (Cao et al. 2013, Liu et al. 2013): equal training sensitivity and specificity (ETSS), maximizing training sensitivity and specificity (MTSS), and minimum training presence (MTP). We also chose two additional thresholds based on the average probability scores for each of the riparian and xeroriparian test datasets (APSR and APSX, respectively). In ArcPro, we created a binary raster file for each threshold, clipping out all cells falling below the threshold. We then stacked each of these rasters in order from lowest to highest threshold, creating a "heat map" of increasing predicted suitability.

IX. Assessing variable contribution

MaxEnt provides two measures for identifying important variables for fitting models: percent contribution and permutation importance (Phillips & Dudík 2008). Percent contribution for variables depends on the specific path MaxEnt uses to obtain the optimal solution, but many possible paths may exist which would result in different contribution values. Permutation importance depends on the final MaxEnt model output and normalized percentage values are obtained by randomly permuting the values of variables among training and background points and measuring the decrease training AUC, with larger decreases in AUC indicating higher variable importance (Elith et al. 2011). Given the potential for correlation among the variables selected for our models, we primarily relied on permutation importance to evaluate variables because percent contribution values may be less meaningful with correlated variables (Phillips 2005).

We also used MaxEnt's jackknife test plots to evaluate variable importance. This process runs multiple models including a standard model with all variables included, and a series of models in which each variable is excluded, creating a model with the remaining variables. Models are also created using each variable in isolation. The results allow users to evaluate which variables perform best when used in isolation, as well as which variables reduce model performance the most and thus contain the most unique information not accounted for by other variables (Shcheglovitova & Anderson 2013).

RESULTS

I. Model selection

The model with the lowest AICc score was model #1, with all 18 variables and a betamultiplier of 1 (Table 4.2). However, this model included many variables with low contribution rates. Therefore, to reduce model complexity and eliminate correlated and redundant variables we also selected the top-performing model that contained less than 18 variables. This included model #3 with 8 variables (topowetness, ndvisosn, ndvitin, meantemp, ndvieost, tempannualrange, precipseasonality, precipwettestq) and a betamultiplier of 1.

Table 4.2 List of 10 top-performing models based on AICc scores. Single asterisks indicate top-performing models, and double-asterisk indicates top-performing model with reduced variables, which were used for final model outputs.

Model	Beta multiplier	variables	samples	parameters	Log likelihood	AIC	AICc	BIC	AUC. Test	AUC. Train	AUC. Diff
1*	1	18	331	117	-3672.63	7579.25	7708.88	8024.10	0.9768	0.9805	0.003
30	5	18	331	34	-3820.8	7709.60	7717.64	7838.87	0.9736	0.976	0.002

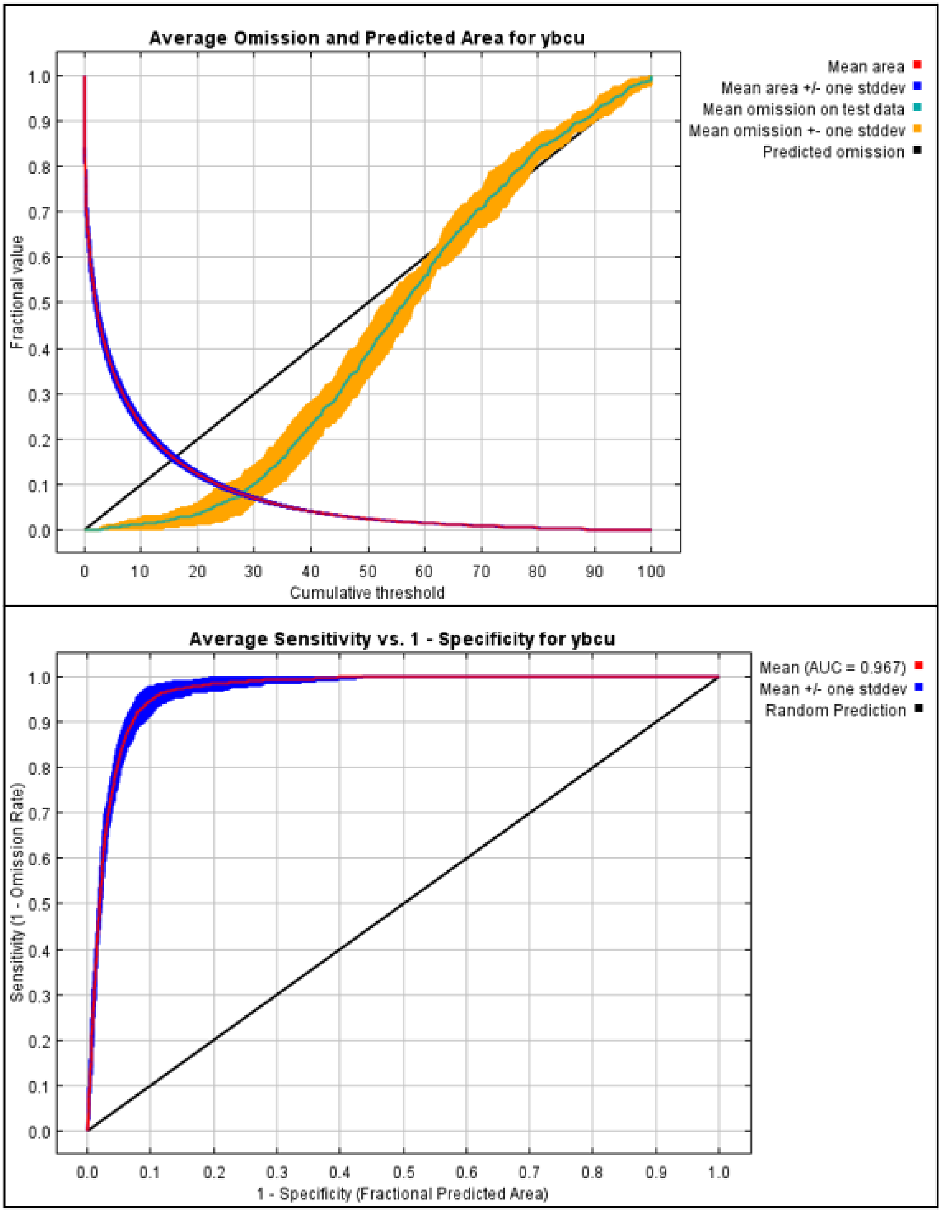
16	3	18	331	73	-3765.51	7677.01	7719.05	7954.56	0.9762	0.9774	0.001
23	4	18	331	54	-3796.67	7701.34	7722.86	7906.66	0.9738	0.977	0.003
37	6	18	331	27	-3836.24	7726.47	7731.46	7829.12	0.975	0.9746	0.000
3**	1	8	331	98	-3728.33	7652.65	7736.29	8025.26	0.9762	0.9788	0.002
4	1	8	331	98	-3728.33	7652.65	7736.29	8025.26	0.975	0.9798	0.004
5	1	8	331	98	-3728.33	7652.65	7736.29	8025.26	0.9758	0.9791	0.003
6	1	8	331	98	-3728.33	7652.65	7736.29	8025.26	0.9757	0.9793	0.003
7	1	8	331	98	-3728.33	7652.65	7736.29	8025.26	0.9756	0.9791	0.003

II. Model performance

The model with 18 variables resulted in an average AUC of 0.967, and model with 8 variables resulted in an average AUC of 0.966. Given the minor differences in performance between 18-variable model and reduced-variable model, we discarded the 18-variable model for further analysis and only evaluated the reduced-variable model (Figure 4.2).

For the 54 xeroriparian breeding locations, the average probability of suitability was 0.66 (SD=0.22) with a median of 0.72. Using 485 independent detections from riparian areas, the average probability of suitability was 0.82 (SD =0.24) and a median of 0.94.

Figure 4.2 Results of MaxEnt analysis of omission/commission and AUC for the final model based on an average of 10 replicate model runs.



III. Variable contribution & importance

Table 4.3 List of variables and their MaxEnt model contributions for the final Yellow-billed Cuckoo model.

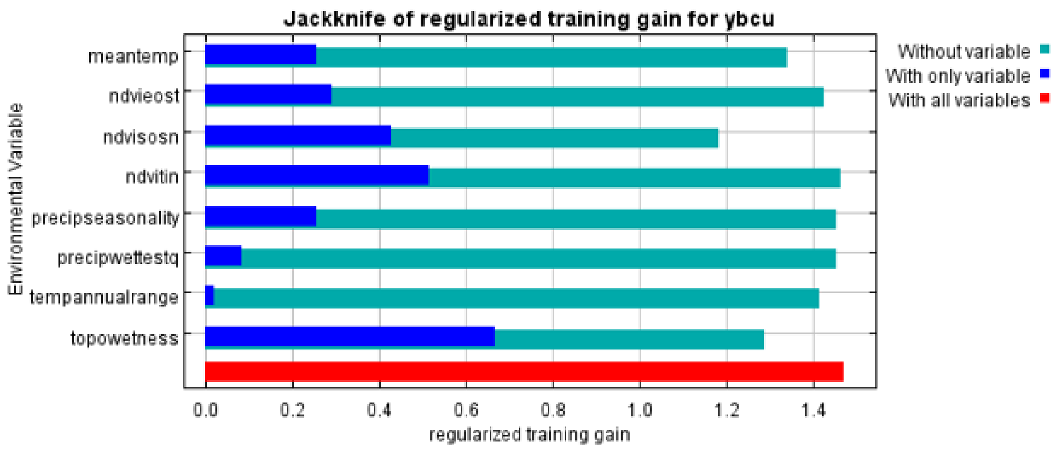
Variable	Percent contribution	Permutation importance

Topowetness	40.5	7.4
Ndvisosn	23.2	51
Ndvitin	12.7	0.3
Meantemp	9.2	24.7
Ndvieost	7.7	9.3
tempannualrange	3.5	3.9
precipseasonality	2.4	2
Precipwettestq	0.9	1.4

Topowetness, ndvisosn, ndvitin, and meantemp were the top contributing variables, with ndvieost, tempannualrange, precipseasonality, and precipwettestq each contributing relatively less to the model (Table 4.3). The permutation importance of ndvisosn was over 50%, more than double its percent contribution, whereas the permutation importance of topowetness ranked only 4^{th} and 5^{th}.

In our review of jackknife plots of variables (Figure 4.3), topowetness had the highest gain when used in isolation, indicating it had the most useful information by itself. The variable ndvisosn decreased the gain the most when it was omitted, indicating it has the most information that isn't present in other variables. Time-integrated NDVI (ndvitin) was the second most important variable on its own, but also decreased the gain the least, indicating that when excluded from the model, other variables may readily account for the information it contains.

Figure 4.3 Jacknife test plots for variables used in our final Yellow-billed Cuckoo MaxEnt model.



IV. Predicted area

Of the five thresholds, MTP was the lowest at 0.0594, which by definition omits no training data, and predicted over 29% of the study area. This was followed by MTSS, ETSS, APSX, and APSR, which respectively predicted approximately 8%, 6%, 2%, and 1% of the study area and omitted approximately 5, 8, 28, and 43% of training data (Table 4.4, Figure 4.4). Visual inspection of the output indicated that inclusion of the MTP and MTSS thresholds resulted in excessive over-prediction into unsuitable areas, so they were excluded from further review and interpretation of model output was based on the ETSS, APSX, and APSR thresholds (Figures 4.4 & 4.5).

Table 4.4 Results from Yellow-billed Cuckoo MaxEnt predicted suitability output at different thresholds.

Threshold type	Threshold value	Percent study area predicted	Percent omission of training data

MTP	0.0594	29.56%	0.00%
MTSS	0.269	7.61%	4.83%
ETSS	0.3331	5.71%	8.16%
APSX	0.6662	1.60%	28.10%
APSR	0.82	0.85%	43.50%

Figure 4.4 Predicted MaxEnt habitat suitability for Yellow-billed Cuckoos based on 3 threshold levels, using the average of 10 replicate model runs.

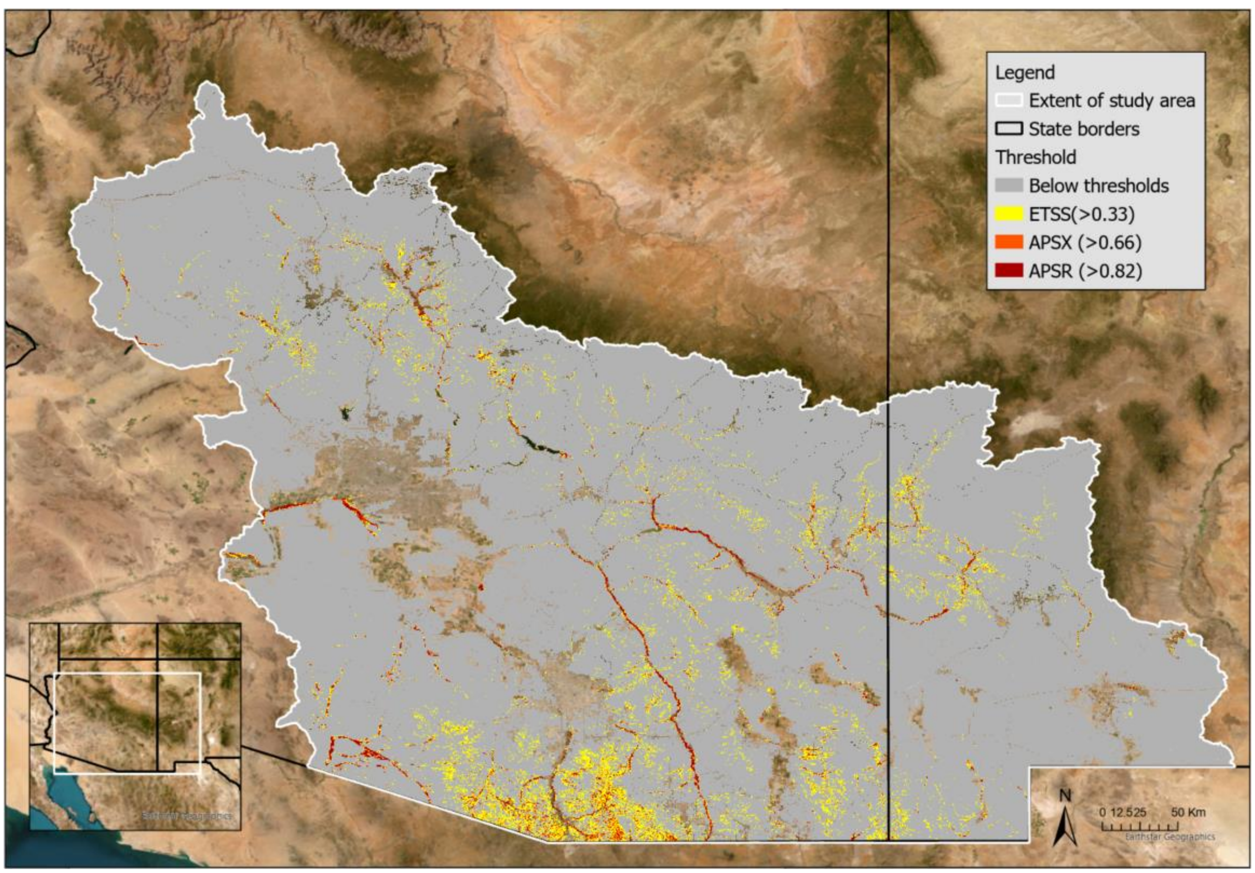
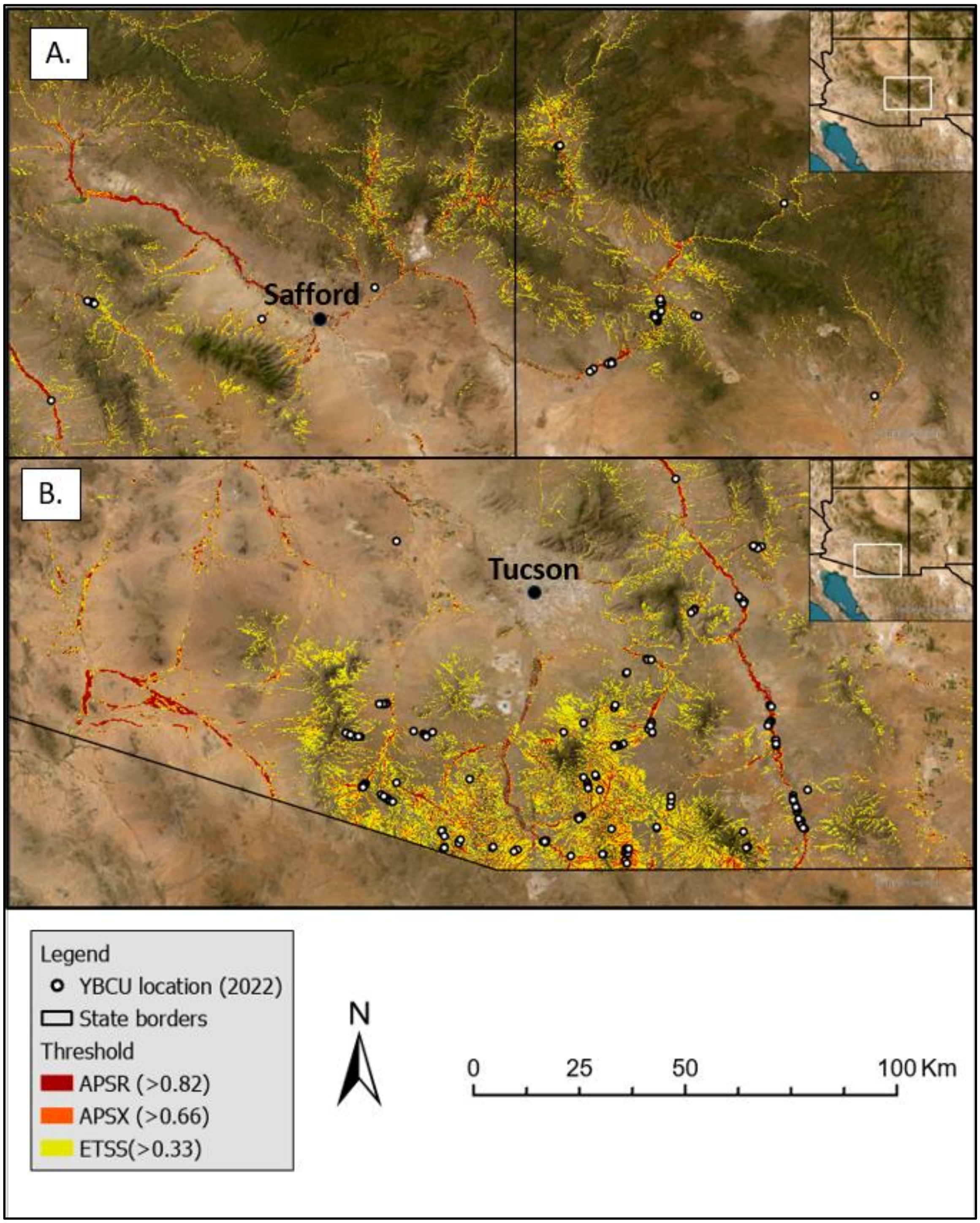


Figure 4.5 Zoomed-in view of MaxEnt output of predicted Yellow-billed Cuckoo habitat in (A.) the upper Gila River Watershed and (B.) the San Simon Wash Watershed, Madrean Sky Islands, and San Pedro River.



DISCUSSION

Our study coupled newly available survey data with ecologically relevant variables to model a broad spectrum of riparian and xeroriparian habitat for Yellow-billed Cuckoos in the southwestern USA. Our model AUC of 0.967 is above accepted standards for model performance (Elith et al. 2011, Syfert et al. 2013, Friggens et al. 2014, Ortega-Huerta & Vega-Rivera 2017), indicating high model specificity and sensitivity and demonstrating a strong fit of data to the selected environmental variables. Our results provide an important tool for identifying potential cuckoo habitat in the core of its extant range while offering insight into potential climatic, phenological, and topographic influences on distribution.

Previous studies of western Yellow-billed Cuckoos have heavily emphasized major rivers and streams containing cottonwood-willow riparian habitat (Hamilton & Hamilton 1965, Laymon & Halterman 1987, Girvetz & Greco 2009; Johnson et al. 2017, Wohner et al. 2021). Most surveys and conservation efforts prior to and following the listing of the western DPS of Yellow-billed Cuckoo as threatened in 2014 (USFWS 2014) focused on riparian habitat, and thus habitat modeling efforts have also primarily focused on riparian habitat. For example, in their “conflation” modeling approach, in which they combined independent land cover data sets to map habitat more accurately, Villareal et al. (2014) focused primarily on landcover identified as riparian. Accordingly, Wallace et al. (2013) couched their model findings about the importance of landscape “greenness” in terms of differences among “areas within their cottonwood–willow habitat”. Although our model generally agreed with this previous work, it also identified additional regions of predicted high suitability omitted from these previous models, much of which challenges these previous assumptions of strict riparian habitat requirements.

First, many mid- and upper-watershed drainages in the Sky Islands region were identified as suitable due to the inclusion of presence data from some of these drainages which had not been available for previous modeling efforts (Villarreal et al. 2014). This included riparian and xeroriparian habitat in narrow, low-mid elevation drainages in the Baboquivari Mountains, Atascosa Mountains, Santa Rita Mountains, Patagonia Mountains, Canelo Hills, and Huachuca Mountains. Additional habitat was predicted in other Sky Islands ranges east of the San Pedro River south of the Gila River but tended to be more sparsely distributed. Many narrow (e.g., one pixel in width) portions of drainages were included in the APSR and APSX thresholds, but additional habitat was predicted within the ETSS threshold, including some adjacent upland areas. This predicted distribution is consistent with recent research in southeastern Arizona (Beauregard (Chapter 1)) suggesting cuckoos in the southwestern Sky Islands (ranges named above) commonly use xeroriparian habitat for nesting and use adjacent uplands for foraging. Importantly, these xeroriparian areas of predicted high suitability occurred along the southern edge of our study area, indicating xeroriparian habitat likely extends into northern Mexico, consistent with reports of cuckoos using xeroriparian habitat in Sonora (Macías-Duarte et al., 2023). One variable that likely contributed to identifying these areas was topographic wetness index (topowetness), the most important variable when used in isolation in our model. Unlike previous approaches that defined the study area extent using buffered perennial stream layers under the assumption that western Yellow-billed Cuckoos were associated with cottonwood-willow riparian habitat (Villarreal et al. 2014), topographic wetness identified areas with high potential for water pooling based on topography regardless whether the drainage had perennial water (Amatulli et al. 2020).

Second, portions of the San Simon watershed (which is predominantly on the Tohono O'odham Nation), and to a lesser extent Aguirre Valley and Santa Rosa Wash, were also highly predicted despite no available presence data from this region (Figure 45.B). These broad, low-elevation ephemeral washes contain extensive mesquite bosque and xeroriparian habitat among Sonoran Desert uplands. These areas were not included in the Villareal et al. (2014) models and have historically received little survey effort, although Sutton & Phillips (1942) observed cuckoos in the San Simon watershed in June and presumed them to be migratory. In contrast with mountainous xeroriparian Sky Islands drainages to the east, most of the habitat in the San Simon watershed was predicted above the APSX and APSR thresholds, with relatively little adjacent habitat added at lower thresholds, likely due to a clear boundary between mesic riparian areas and adjacent xeric desert uplands.

Third, our model predicted abundant habitat in portions of the Lower Gila River (southwest of Phoenix), Middle Gila River (primarily on the San Carlos Apache Nation), and throughout the Upper Gila River watershed (primarily in New Mexico) to elevations of approximately 2200 m (Figure 4.5A). Wide, low elevation floodplains with extensive habitat in the Lower and Middle Gila River were consistently predicted at APSR and APSX thresholds. These areas generally are heavily modified and fragmented due to agricultural use and non-native species such as tamarisk (*Tamarix* spp.). While this may have the effect of reducing habitat quality (Johnson et al. 2017), cuckoos may utilize areas with high proportions of tamarisk (Sechrist et al. 2013, Stanek et al. 2021). Upper-watershed drainages in mountainous canyons tended to be predicted with the inclusion of the ETSS threshold, indicating potential habitat along many of the Gila River's upper tributaries. Cuckoos have previously been documented in much of the Gila River watershed, but gaps in data exist on many private and tribal lands, as well as remote portions of the upper watershed.

One caveat of our approach is that we excluded developed and agricultural areas, both because we considered them to be unsuitable and because their high or low NDVI profiles were due to human altered landscapes (e.g., irrigated agriculture resulted extremely high NDVI values, while urban areas were very low). In some cases, the resolution of our data resulted in pixels being removed that included natural habitat in addition to developed areas. Likewise, pixels containing agriculture, roads, and open water that we excluded may have been adjacent to areas where habitat is currently or was historically present. For example, both the Santa Cruz River Valley south of Tucson, AZ and the Middle Gila River Valley near Safford AZ contain a mosaic of developed, agricultural, and natural areas along the floodplain, which inevitably results in some areas excluded in our study that may contain isolated patches of habitat. Given this, we caution that some cuckoo habitat may exist in areas we excluded, although the sum of this habitat in excluded areas would be small relative to the overall area of predicted habitat.

Threshold selection is an important step in evaluating SDM output, and should be based on biologically meaningful parameters (Liu et al. 2016). As described above, our higher thresholds (APSR and APSX), which were based on independent test datasets, reliably predicted wide lowland riparian habitat throughout the study area but carried an omission rate of at least 28%. Including ETSS, however, accounted for more xeroriparian habitat and narrower riparian drainages, reducing omission rates to slightly over 8% without excessive overprediction around lowland riparian habitat. Lower thresholds (MTP and MTSS) resulted in overprediction in many clearly unsuitable areas, and we thus excluded these thresholds for analysis. We therefore suggest that the combined APSR, APSX, and ETSS thresholds are most appropriate for evaluating potential distribution of cuckoos. In some areas, such as southeastern Arizona where upland habitat adjacent to drainages is predicted, manual review at the drainage scale may be necessary for identification of priority areas for survey efforts.

Previous models of western Yellow-billed cuckoo habitat included measures of vegetation phenology, given the importance of plant productivity in driving insect prey availability important for migratory birds that require a reliable prey base during migration and on their breeding grounds (McGrath et al. 2009, La Sorte & Graham 2021). As an effective proxy for vegetation productivity and phenology, NDVI has been used extensively to model species habitats and differentiate vegetation classes (Clerici et al. 2012, He et al. 2015, Hoagland et al. 2018, Ponti & Sannolo 2023), and has been shown to correlate with insect abundance (Sweet et al. 2015, Lafage et al. 2014, Gebert et al. 2023). Villareal et al. (2014) suggested NDVI was important in their models because it differentiated between "productive floodplain vegetation and surrounding non-habitat upland desert". This approach was effective for identifying wide, extensive riparian floodplains, but did not account for ephemeral drainages and uplands, which are often narrow and primarily green up with the onset of the summer monsoon season, thereby exhibiting different NDVI phenological profiles (Birtwistle et al. 2016). Wallace et al. (2013) found that cuckoo detections were associated with cottonwood-willow areas that were "greener" than average and that reached peak greenness later in the season, presumably as a result of monsoon precipitation. Our model likewise found strong support for variables reflecting high NDVI values (ndvitin) and the timing of green-up, both its onset (ndvisosn) and a later end of the productive season (ndieost), likely driven by high summer precipitation (precipseasonality, precipwettestq). These results are consistent with the hypothesis that higher productivity is important for cuckoos because of their dependence on abundant, large insect prey such as caterpillars, grasshoppers, and cicadas (Hamilton & Hamilton 1965, Laymon 1980, Hughes 2015) that require extended periods of plant growth to complete their life cycles. We caution however, that more study is needed on interaction between prey availability and climate and vegetation phenology in cuckoo habitat.

The western Yellow-billed Cuckoo's threatened status is based largely on rangewide loss and degradation of riparian habitat (USFWS 2014), and riparian habitat continues to decline throughout this range due to dams and water diversions, invasive species, climate change, and land use change (Stromberg 1993, Serrat-Capdevila et al. 2007, Poff et al. 2011, Eastoe 2020). For example, the San Pedro River basin is an area of high suitability identified in our model and previous studies (Halterman 2009, Mcfarland et al. 2012, Wallace et al. 2013, Villarreal et al. 2014), but base flow into the river has been declining as a result of increased human demand due to urbanization, compounded by ongoing drought, and reduction in agriculture use of water has not stopped this decline (Eastoe 2020). In light of loss of riparian areas like these, xeroriparian habitat may become increasingly important refugia for cuckoos and our model identifies many xeroriparian areas that could potentially harbor cuckoos that have yet to be documented. In that sense, our model gives managers an important tool for guiding future efforts to determine the extent of xeroriparian habitat used by these birds. That said, the projected weakening of the monsoon as a result of climate change (Pascale et al. 2017, 2019) may disproportionately affect southeastern Arizona xeroriparian habitat, which relies on monsoon precipitation for the majority of annual vegetation productivity (Crimmins et al. 2011, Forzieri et al. 2011). Future models incorporating predictions of climate change on monsoon-driven productivity and phenology and effects of increasing human demand for water on cottonwood-willow riparian areas have the potential for projecting how the suitability model presented here may best be modified to account for these future challenges.

ACKNOWLEDGEMENTS

Work presented in this chapter was supported by contributions from Dr. Tad Theimer and Susan Sferra. This research was conducted under U.S. Fish & Wildlife Service Recovery Permit #TE81148C-0 and Arizona Game & Fish Department Special Collecting License permit #SP407168. The findings and conclusions in this article are those of the authors and do not necessarily represent the views of the U.S. Fish and Wildlife Service. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the United States Government. Data for this study are available as a U.S. Geological Survey data release.

LITERATURE CITED

Acevedo, P., Jiménez-Valverde, A., Lobo, J. M., & Real, R. (2012). Delimiting the geographical background in species distribution modelling. Journal of Biogeography, 39(8), 1383–1390. https://doi.org/10.1111/j.1365-2699.2012.02713.x

Akaike, H. (1974). A new look at the statistical model identification. IEEE Transactions on Automatic Control, 19(6), 716–723. https://doi.org/10.1109/TAC.1974.1100705

Amatulli, G., McInerney, D., Sethi, T., Strobl, P., & Domisch, S. (2020). Geomorpho90m, empirical evaluation and accuracy assessment of global high-resolution geomorphometric layers. Scientific Data, 7(1), 162. https://doi.org/10.1038/s41597-020-0479-6

Barbet-Massin, M., Jiguet, F., Albert, C. H., & Thuiller, W. (2012). Selecting pseudo-absences for species distribution models: How, where and how many? Methods in Ecology and Evolution, 3(2), 327–338. https://doi.org/10.1111/j.2041-210X.2011.00172.x

Barve, N., Barve, V., Jiménez-Valverde, A., Lira-Noriega, A., Maher, S. P., Peterson, A. T., Soberón, J., & Villalobos, F. (2011). The crucial role of the accessible area in ecological niche modeling and species distribution modeling. Ecological Modelling, 222(11), 1810–1819. https://doi.org/10.1016/j.ecolmodel.2011.02.011

Bateman, H. L., & Riddle, S. B. (2020). Reptile community responses to native and non-native riparian forests and disturbance along two rivers in Arizona. River Research and Applications, 36(3), 492–502. https://doi.org/10.1002/rra.3587

Beauchamp, V. B., & Shafroth, P. B. (2011). Floristic composition, beta diversity, and nestedness of reference sites for restoration of xeroriparian areas. Ecological Applications, 21(2), 465–476. https://doi.org/10.1890/09-1638.1

Birtwistle, A. N., Laituri, M., Bledsoe, B., & Friedman, J. M. (2016). Using NDVI to measure precipitation in semi-arid landscapes. Journal of Arid Environments, 131, 15–24. https://doi.org/10.1016/j.jaridenv.2016.04.004

Bradie, J., & Leung, B. (2017). A quantitative synthesis of the importance of variables used in MaxEnt species distribution models. Journal of Biogeography, 44(6), 1344–1361. https://doi.org/10.1111/jbi.12894

Brambilla, M., Rubolini, D., Appukuttan, O., Calvi, G., Karger, D. N., Kmecl, P., Mihelič, T., Sattler, T., Seaman, B., Teufelbauer, N., Wahl, J., & Celada, C. (2022). Identifying climate refugia for high-elevation Alpine birds under current climate warming predictions. Global Change Biology, 28(14), 4276–4291. https://doi.org/10.1111/gcb.16187

Brown, D. E. (Ed.). (1994). Biotic communities: Southwestern United States and Northwestern Mexico. Univ. of Utah Press.

Burnham, K. P., & Anderson, D. R. (2004). Multimodel Inference: Understanding AIC and BIC in Model Selection. Sociological Methods & Research, 33(2), 261–304. https://doi.org/10.1177/0049124104268644

Camaclang, A. E., Maron, M., Martin, T. G., & Possingham, H. P. (2015). Current practices in the identification of critical habitat for threatened species: Identifying Critical Habitat. Conservation Biology, 29(2), 482–492. https://doi.org/10.1111/cobi.12428

Cao, Y., DeWalt, R. E., Robinson, J. L., Tweddale, T., Hinz, L., & Pessino, M. (2013). Using Maxent to model the historic distributions of stonefly species in Illinois streams: The effects of regularization and threshold selections. Ecological Modelling, 259, 30–39. https://doi.org/10.1016/j.ecolmodel.2013.03.012

Clerici, N., Weissteiner, C. J., & Gerard, F. (2012). Exploring the Use of MODIS NDVI-Based Phenology Indicators for Classifying Forest General Habitat Categories. Remote Sensing, 4(6), 1781–1803. https://doi.org/10.3390/rs4061781

Corman, T. E., & Magill, R. T. (2000). Western Yellow-billed Cuckoo in Arizona: 1998 and 1999 Survey Report (Nongame and Endangered Wildlife Program Technical Report 150; p. 55). Arizona Game and Fish Department.

Crimmins, T. M., Crimmins, M. A., & Bertelsen, C. D. (2011). Onset of summer flowering in a ‘Sky Island’ is driven by monsoon moisture. New Phytologist, 191(2), 468–479. https://doi.org/10.1111/j.1469-8137.2011.03705.x

Eastoe, C. J. (2020). Sources of Perennial Water Supporting Critical Ecosystems, San Pedro Valley, Arizona. Environmental and Engineering Geoscience, 26(4), 463–479. https://doi.org/10.2113/EEG-D-20-00040

Elith, J., & Leathwick, J. R. (2009). Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. Annual Review of Ecology, Evolution, and Systematics, 40(1), 677–697. https://doi.org/10.1146/annurev.ecolsys.110308.120159

Elith, J., Phillips, S. J., Hastie, T., Dudík, M., Chee, Y. E., & Yates, C. J. (2011). A statistical explanation of MaxEnt for ecologists: Statistical explanation of MaxEnt. Diversity and Distributions, 17(1), 43–57. https://doi.org/10.1111/j.1472-4642.2010.00725.x

Feng, X., Park, D. S., Liang, Y., Pandey, R., & Papeş, M. (2019). Collinearity in ecological niche modeling: Confusions and challenges. Ecology and Evolution, 9(18), 10365–10376. https://doi.org/10.1002/ece3.5555

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. International Journal of Climatology, 37(12), 4302–4315. https://doi.org/10.1002/joc.5086

Foin, T. C., Riley, S. P. D., Pawley, A. L., Ayres, D. R., Carlsen, T. M., Hodum, P. J., & Switzer, P. V. (1998). Improving Recovery Planning for Threatened and Endangered Species. BioScience, 48(3), 177–184. https://doi.org/10.2307/1313263

Forzieri, G., Castelli, F., & Vivoni, E. R. (2011). Vegetation Dynamics within the North American Monsoon Region. Journal of Climate, 24(6), 1763–1783. https://doi.org/10.1175/2010JCLI3847.1

Fourcade, Y., Engler, J. O., Rödder, D., & Secondi, J. (2014). Mapping Species Distributions with MAXENT Using a Geographically Biased Sample of Presence Data: A Performance

Assessment of Methods for Correcting Sampling Bias. PLoS ONE, 9(5), e97122. https://doi.org/10.1371/journal.pone.0097122

Franklin, J. (2013). Species distribution models in conservation biogeography: Developments and challenges. Diversity and Distributions, 19(10), 1217–1223. https://doi.org/10.1111/ddi.12125

Friggens, M. M., Loehman, R., Holsinger, L., & Finch, D. (2014). Vulnerability of riparian obligate species to the interactive effect of fire, climate and hydrological change (p. 213) [Final Report for Interagency Agreement #13-IA-11221632-006.]. U.S. Department of Agriculture, Forest Service Rocky Mountain Research Station.

Gastón, A., & García-Viñas, J. I. (2011). Modelling species distributions with penalised logistic regressions: A comparison with maximum entropy models. Ecological Modelling, 222(13), 2037–2041. https://doi.org/10.1016/j.ecolmodel.2011.04.015

Gebert, F., Bollmann, K., Schuwirth, N., Duelli, P., Weber, D., & Obrist, M. K. (2023). Similar temporal patterns in insect richness, abundance and biomass across major habitat types. Insect Conservation and Diversity, icad.12700. https://doi.org/10.1111/icad.12700

Girvetz, E. H., & Greco, S. E. (2009). Multi-scale predictive habitat suitability modeling based on hierarchically delineated patches: An example for yellow-billed cuckoos nesting in riparian forests, California, USA. Landscape Ecology, 24(10), 1315–1329. https://doi.org/10.1007/s10980-009-9384-2

Guillera-Arroita, G., Lahoz-Monfort, J. J., Elith, J., Gordon, A., Kujala, H., Lentini, P. E., McCarthy, M. A., Tingley, R., & Wintle, B. A. (2015). Is my species distribution model fit for purpose? Matching data and models to applications. Global Ecology and Biogeography, 24(3), 276–292. https://doi.org/10.1111/geb.12268

Halterman, M. (2009). Sexual dimorphism, detection probability, home range, and parental care in the Yellow-billed Cuckoo. University of Nevada, Reno.

Halterman, M. D., Johnson, M. J., Holmes, J. A., & Laymon, S. A. (2015). A Natural History Summary and Survey Protocol for the Western Distinct Population Segment of the Yellow-billed Cuckoo. U.S. Fish and Wildlife Techniques and Methods.

Hamilton, W. J., III, & Hamilton, M. E. (1965). Breeding characteristics of yellow-billed cuckoos in Arizona. Proceedings of the California Academy of Sciences, 32(14), 405–432.

Hardy, P. C., Griffin, D. J., Kuenzi, A. J., & Morrison, M. L. (2004). Occurrence and habitat use of passage neotropical migrants in the Sonoran Desert. Western North American Naturalist, 64, 13.

He, K. S., Bradley, B. A., Cord, A. F., Rocchini, D., Tuanmu, M.-N., Schmidtlein, S., Turner, W., Wegmann, M., & Pettorelli, N. (2015). Will remote sensing shape the next generation of species distribution models? Remote Sensing in Ecology and Conservation, 1(1), 4–18. https://doi.org/10.1002/rse2.7

Higgins, R. W., Yao, Y., & Wang, X. L. (1997). Influence of the North American Monsoon System on the U.S. Summer Precipitation Regime. Journal of Climate, 10(10), 2600–2622. https://doi.org/10.1175/1520-0442(1997)010<2600:IOTNAM>2.0.CO;2

Hmimina, G., Dufrêne, E., Pontailler, J.-Y., Delpierre, N., Aubinet, M., Caquet, B., De Grandcourt, A., Burban, B., Flechard, C., Granier, A., Gross, P., Heinesch, B., Longdoz, B., Moureaux, C., Ourcival, J.-M., Rambal, S., Saint André, L., & Soudani, K. (2013). Evaluation of the potential of MODIS satellite data to predict vegetation phenology in different biomes: An investigation using ground-based NDVI measurements. Remote Sensing of Environment, 132, 145–158. https://doi.org/10.1016/j.rse.2013.01.010

Hoagland, S. J., Beier, P., & Lee, D. (2018). Using MODIS NDVI phenoclasses and phenoclusters to characterize wildlife habitat: Mexican spotted owl as a case study. Forest Ecology and Management, 412, 80–93. https://doi.org/10.1016/j.foreco.2017.12.017

Hughes, J. M. (2015). Yellow-billed Cuckoo (Coccyzus americanus), version 2.0. In P. G. Rodewald (Ed.), The Birds of North America Online. Cornell Lab of Ornithology. https://birdsoftheworld.org/bow/historic/bna/yebcuc/2.0/introduction

Johnson, M. J., Hatten, J. R., Holmes, J. A., & Shafroth, P. B. (2017). Identifying western yellow-billed cuckoo breeding habitat with a dual modelling approach. Ecological Modelling, 347, 50–62. https://doi.org/10.1016/j.ecolmodel.2016.12.010

Johnson, R. R., & Haight, L. T. (1985). Avian Use of Xeroriparian Ecosystems in the North American Warm Deserts (RM-GTR-120; p. RM-GTR-120). U.S. Department of Agriculture, Forest Service, Rocky Mountain Forest and Range Experiment Station. https://doi.org/10.2737/RM-GTR-120

Jueterbock, A. (2015). MaxentVariableSelection.

Kiedrzyński, M., Zielińska, K. M., Rewicz, A., & Kiedrzyńska, E. (2017). Habitat and spatial thinning improve the Maxent models performed with incomplete data. Journal of Geophysical Research: Biogeosciences, 122(6), 1359–1370. https://doi.org/10.1002/2016JG003629

Kingsley, K. J. (1985). Riparian Ecosystems and Their Management: Reconciling Conflicting Uses: the First North American Riparian Conference, Tucson, Arizona, U.S.A., April 16-18, 1985. University of Arizona.

Kramer-Schadt, S., Niedballa, J., Pilgrim, J. D., Schröder, B., Lindenborn, J., Reinfelder, V., Stillfried, M., Heckmann, I., Scharf, A. K., Augeri, D. M., Cheyne, S. M., Hearn, A. J., Ross, J., Macdonald, D. W., Mathai, J., Eaton, J., Marshall, A. J., Semiadi, G., Rustam, R., … Wilting, A. (2013). The importance of correcting for sampling bias in MaxEnt species distribution models. Diversity and Distributions, 19(11), 1366–1379. https://doi.org/10.1111/ddi.12096

La Sorte, F. A., & Graham, C. H. (2021). Phenological synchronization of seasonal bird migration with vegetation greenness across dietary guilds. Journal of Animal Ecology, 90(2), 343–355. https://doi.org/10.1111/1365-2656.13345

Lafage, D., Secondi, J., Georges, A., Bouzillé, J.-B., & Pétillon, J. (2014). Satellite-derived vegetation indices as surrogate of species richness and abundance of ground beetles in temperate floodplains. Insect Conservation and Diversity, 7(4), 327–333. https://doi.org/10.1111/icad.12056

Laymon, S. A. (1980). Feeding and Nesting Behavior of the Yellow-billed Cuckoo in the Sacramento Valley (p. 29) [Wildlife Management Branch Administrative Report 80-2]. State of California Department of Fish and Game.

Laymon, S. A., & Halterman, M. D. (1987). Distribution and Status of the Yellow-billed Cuckoo in California: 1986-87 (Final Report to the California Department of Fish and Game Contract #C-1845; p. 35).

Liu, C., Newell, G., & White, M. (2016). On the selection of thresholds for predicting species occurrence with presence-only data. Ecology and Evolution, 6(1), 337–348. https://doi.org/10.1002/ece3.1878

Liu, C., White, M., & Newell, G. (2013). Selecting thresholds for the prediction of species occurrence with presence-only data. Journal of Biogeography, 40(4), 778–789. https://doi.org/10.1111/jbi.12058

MacFarland, J., & Horst, J. (2015). Yellow-billed Cuckoo Surveys on the Coronado National Forest within Eight Sky Island Mountain Ranges in Southeast Arizona (p. 239) [Final Report for the Coronado National Forest Service]. Tucson Audubon Society.

MacFarland, J., & Horst, J. (2017). Yellow-billed Cuckoo Surveys on the Coronado National Forest within Five Sky Island Mountain Ranges in Southeast Arizona in 2017 (p. 145) [Final Report for the Coronado National Forest]. Tucson Audubon Society.

Macías-Duarte, A., Juárez, E., Murrieta, E. S., Perales-Hoeffer, E. L., & Rosas, C. I. O. (2023). Abundance and occupancy of the western yellow-billed cuckoo (Coccyzus americanus) in Sonora, Mexico. Canadian Journal of Zoology, cjz-2022-0096. https://doi.org/10.1139/cjz-2022-0096

Mcfarland, T. M., Van Riper, C., & Johnson, G. E. (2012). Evaluation of NDVI to assess avian abundance and richness along the upper San Pedro River. Journal of Arid Environments, 77, 45–53. https://doi.org/10.1016/j.jaridenv.2011.09.010

McGrath, L. J., Van Riper Iii, C., & Fontaine, J. J. (2009). Flower power: Tree flowering phenology as a settlement cue for migrating birds. Journal of Animal Ecology, 78(1), 22–30. https://doi.org/10.1111/j.1365-2656.2008.01464.x

Merow, C., Smith, M. J., & Silander Jr, J. A. (2013). A practical guide to MaxEnt for modeling species' distributions: What it does, and why inputs and settings matter. Ecography, 36(10), 1058–1069. https://doi.org/10.1111/j.1600-0587.2013.07872.x

Ortega-Huerta, M. A., & Vega-Rivera, J. H. (2017). Validating distribution models for twelve endemic bird species of tropical dry forest in western Mexico. Ecology and Evolution, 7(19), 7672–7686. https://doi.org/10.1002/ece3.3160

Pascale, S., Boos, W. R., Bordoni, S., Delworth, T. L., Kapnick, S. B., Murakami, H., Vecchi, G. A., & Zhang, W. (2017). Weakening of the North American monsoon with global warming. Nature Climate Change, 7(11), 806–812. https://doi.org/10.1038/nclimate3412

Pascale, S., Carvalho, L. M. V., Adams, D. K., Castro, C. L., & Cavalcanti, I. F. A. (2019). Current and Future Variations of the Monsoons of the Americas in a Warming Climate. Current Climate Change Reports, 5(3), 125–144. https://doi.org/10.1007/s40641-019-00135-w

Pettorelli, N., Vik, J. O., Mysterud, A., Gaillard, J.-M., Tucker, C. J., & Stenseth, N. Chr. (2005). Using the satellite-derived NDVI to assess ecological responses to environmental change. Trends in Ecology & Evolution, 20(9), 503–510. https://doi.org/10.1016/j.tree.2005.05.011

Phillips, S. J. (2005). A Brief Tutorial on Maxent.

Phillips, S. J., & Dudík, M. (2008). Modeling of species distributions with Maxent: New extensions and a comprehensive evaluation. Ecography, 31(2), 161–175. https://doi.org/10.1111/j.0906-7590.2008.5203.x

Phillips, S. J., Dudík, M., Elith, J., Graham, C. H., Lehmann, A., Leathwick, J., & Ferrier, S. (2009). Sample selection bias and presence-only distribution models: Implications for background and pseudo-absence data. Ecological Applications, 19(1), 181–197. https://doi.org/10.1890/07-2153.1

Poff, B., Koestner, K. A., Neary, D. G., & Henderson, V. (2011). Threats to Riparian Ecosystems in Western North America: An Analysis of Existing Literature1. JAWRA Journal of the American Water Resources Association, 47(6), 1241–1254. https://doi.org/10.1111/j.1752-1688.2011.00571.x

Ponti, R., & Sannolo, M. (2023). The importance of including phenology when modelling species ecological niche. Ecography, 2023(4), e06143. https://doi.org/10.1111/ecog.06143

Radosavljevic, A., & Anderson, R. P. (2014). Making better M AXENT models of species distributions: Complexity, overfitting and evaluation. Journal of Biogeography, 41(4), 629–643. https://doi.org/10.1111/jbi.12227

Scott, J. M., Goble, D. D., Wiens, J. A., Wilcove, D. S., Bean, M., & Male, T. (2005). Recovery of Imperiled Species under the Endangered Species Act: The Need for a New Approach. Frontiers in Ecology and the Environment, 3(7), 383–389. https://doi.org/10.2307/3868588

Sechrist, J., Ahlers, D. D., Zehfuss, K. P., Doster, R. H., Paxton, E. H., & Ryan, V. M. (2013). Home Range and Use of Habitat of Western Yellow-Billed Cuckoos On the Middle Rio Grande, New Mexico. The Southwestern Naturalist, 58(4), 411–419. https://doi.org/10.1894/0038-4909-58.4.411

Serrat-Capdevila, A., Valdés, J. B., Pérez, J. G., Baird, K., Mata, L. J., & Maddock, T. (2007). Modeling climate change impacts – and uncertainty – on the hydrology of a riparian system: The San Pedro Basin (Arizona/Sonora). Journal of Hydrology, 347(1–2), 48–66. https://doi.org/10.1016/j.jhydrol.2007.08.028

Shcheglovitova, M., & Anderson, R. P. (2013). Estimating optimal complexity for ecological niche models: A jackknife approach for species with small sample sizes. Ecological Modelling, 269, 9–17. https://doi.org/10.1016/j.ecolmodel.2013.08.011

Stanek, J. E., Mcneil, S. E., Tracy, D., Stanek, J. R., Manning, J. A., & Halterman, M. D. (2021). Western Yellow-Billed Cuckoo Nest-Site Selection and Success in Restored and Natural Riparian Forests. The Journal of Wildlife Management, 85(4), 782–793. https://doi.org/10.1002/jwmg.22020

Stromberg, J. C. (1993). Frémont Cottonwood-Goodding Willow Riparian Forests: A Review of Their Ecology, Threats, and Recovery Potential. Journal of the Arizona-Nevada Academy of Science, 27(1), 97–110.

Sutton, G. M., & Phillips, A. R. (1942). June Bird Life of the Papago Indian Reservation, Arizona. The Condor, 44(2), 57–65. https://doi.org/10.2307/1364225

Sweet, S. K., Asmus, A., Rich, M. E., Wingfield, J., Gough, L., & Boelman, N. T. (2015). NDVI as a predictor of canopy arthropod biomass in the Alaskan arctic tundra. Ecological Applications, 25(3), 779–790. https://doi.org/10.1890/14-0632.1

Syfert, M. M., Smith, M. J., & Coomes, D. A. (2013). The Effects of Sampling Bias and Model Complexity on the Predictive Performance of MaxEnt Species Distribution Models. PLoS ONE, 8(2), e55158. https://doi.org/10.1371/journal.pone.0055158

Tulloch, A. I. T., Sutcliffe, P., Naujokaitis-Lewis, I., Tingley, R., Brotons, L., Ferraz, K. M. P. M. B., Possingham, H., Guisan, A., & Rhodes, J. R. (2016). Conservation planners tend to ignore improved accuracy of modelled species distributions to focus on multiple threats and ecological processes. Biological Conservation, 199, 157–171. https://doi.org/10.1016/j.biocon.2016.04.023

USFWS. (2021). Endangered and Threatened Wildlife and Plants; Designation of Critical Habitat for the Western Distinct Population Segment of the Yellow- Billed Cuckoo. Federal Register.

Villarreal, M. L., van Riper, C., & Petrakis, R. E. (2014). Conflation and aggregation of spatial data improve predictive models for species with limited habitats: A case of the threatened yellow-billed cuckoo in Arizona, USA. Applied Geography, 47, 57–69. https://doi.org/10.1016/j.apgeog.2013.12.003

Wallace, C. S. A., Villarreal, M. L., & van Riper III, C. (2013). Influence of monsoon-related riparian phenology on yellow-billed cuckoo habitat selection in Arizona. Journal of Biogeography, 40(11), 2094–2107. https://doi.org/10.1111/jbi.12167

Warren, D. L., & Seifert, S. N. (2011). Ecological niche modeling in Maxent: The importance of model complexity and the performance of model selection criteria. Ecological Applications, 21(2), 335–342. https://doi.org/10.1890/10-1171.1

Wen, L., Saintilan, N., Yang, X., Hunter, S., & Mawer, D. (2015). MODIS NDVI based metrics improve habitat suitability modelling in fragmented patchy floodplains. Remote Sensing Applications: Society and Environment, 1, 85–97. https://doi.org/10.1016/j.rsase.2015.08.001

Westland Resources, Inc. (2015). Rosemont Project: Potential Effects to Yellow-Billed Cuckoo and Its Proposed Critical Habitat [Prepared for Hudbay Minerals, Inc.].

Wiegand, T., Naves, J., Garbulsky, M. F., & Fernández, N. (2008). Animal Habitat Quality And Ecosystem Functioning: Exploring Seasonal Patterns Using NDVI. Ecological Monographs, 78(1), 87–103. https://doi.org/10.1890/06-1870.1

Williams, J. N., Seo, C., Thorne, J., Nelson, J. K., Erwin, S., O'Brien, J. M., & Schwartz, M. W. (2009). Using species distribution models to predict new occurrences for rare plants. Diversity and Distributions, 15(4), 565–576. https://doi.org/10.1111/j.1472-4642.2009.00567.x

Wisz, M. S., Pottier, J., Kissling, W. D., Pellissier, L., Lenoir, J., Damgaard, C. F., Dormann, C. F., Forchhammer, M. C., Grytnes, J.-A., Guisan, A., Heikkinen, R. K., Høye, T. T., Kühn, I., Luoto, M., Maiorano, L., Nilsson, M.-C., Normand, S., Öckinger, E., Schmidt, N. M., … Svenning, J.-C. (2013). The role of biotic interactions in shaping distributions and realised assemblages of species: Implications for species distribution modelling. Biological Reviews, 88(1), 15–30. https://doi.org/10.1111/j.1469-185X.2012.00235.x

Wohner, P. J., Laymon, S. A., Stanek, J. E., King, S. L., & Cooper, R. J. (2021). Challenging our understanding of western Yellow-billed Cuckoo habitat needs and accepted management practices. Restoration Ecology, 29(3). https://doi.org/10.1111/rec.13331

Yackulic, C. B., Chandler, R., Zipkin, E. F., Royle, J. A., Nichols, J. D., Campbell Grant, E. H., & Veran, S. (2013). Presence-only modelling using MAXENT: When can we trust the inferences? Methods in Ecology and Evolution, 4(3), 236–243. https://doi.org/10.1111/2041-210x.12004

York, P., Evangelista, P., Kumar, S., Graham, J., Flather, C., & Stohlgren, T. (2011). A habitat overlap analysis derived from maxent for tamarisk and the south-western willow flycatcher. Frontiers of Earth Science, 5(2), 120–129. https://doi.org/10.1007/s11707-011-0154-5

Zeng, Y., Low, B. W., & Yeo, D. C. J. (2016). Novel methods to select environmental variables in MaxEnt: A case study using invasive crayfish. Ecological Modelling, 341, 5–13. https://doi.org/10.1016/j.ecolmodel.2016.09.019