EARTHJUSTICE

December 21, 2022

Via E-mail

Jade Florence
U.S. Fish and Wildlife Service
Austin Ecological Services Field Office
10711 Burnet Road, Suite 200
Austin, TX 78758-4460
Jade_Florence@fws.gov

Re: Golden-cheeked warbler five-year status review

Dear Dr. Florence:

On behalf of Bexar Audubon Society, Travis Audubon Society, Defenders of Wildlife, Greater Edwards Aquifer Alliance, San Marcos River Foundation, Save Barton Creek Association, Save Our Springs Alliance, Sierra Club Lone Star Chapter, and Wimberley Valley Watershed Association, we submit the following comments regarding the U.S. Fish and Wildlife Service's ongoing five-year status review for the golden-cheeked warbler (*Setophaga chrysoparia*) ("warbler"). *See* Endangered and Threatened Wildlife and Plants; Initiation of 5-Year Status Reviews of 23 Species in the Southwest, 86 Fed. Reg. 23,976 (May 5, 2021). The Service must base its determination regarding the warbler's Endangered Species Act ("ESA") listing status on the "best scientific . . . data available." 16 U.S.C. § 1533(b)(1)(A), (c)(2). As set forth below, recent and forthcoming scientific research shows that the golden-cheeked warbler faces unabating and even accelerating threats from urbanization, habitat fragmentation, wildfires, and extreme weather due to climate change. Further, recent projections of high warbler abundance on the species' breeding range, under scientific scrutiny, appear to overestimate the warbler's population size. The cumulative weight of this evidence demonstrates that the golden-cheeked warbler continues to face significant threats to its continued existence and therefore should remain protected under the Endangered Species Act.

I. Recent scientific information indicates that the warbler faces significant threats from habitat loss and fragmentation due to urbanization and climate change and should remain listed as endangered.

Recent and forthcoming research highlights the rapid rate of actual and threatened warbler habitat loss, and consequent impacts on warbler population viability. Dreiss et al. (2022) recently found that there was a 42% decrease in suitable habitat within the warbler's breeding range between 1985 and 2018, and a decrease in mean patch size caused by habitat fragmentation during this same period. *See* Lindsay Dreiss et al., *Spatiotemporal Patterns in Golden-cheeked Warbler Breeding Habitat Quantity and Suitability*, 17 Avian Conservation & Ecology 1 (2022). Smaller patch sizes translate to more "forest edge," which can leave nests "less concealed and more vulnerable to nest predation" and has been linked to decreased nest survival generally. *Id.* at 5 (collecting studies). Dreiss et al. (2022) further pointed out that at least four Texas counties that include suitable habitat on the southeastern portion of warbler breeding range are projected to see human population increases of over 100% by 2050. *Id.* The authors forecasted that "[i]ncreased development pressures in the Texas Hill Country will likely continue to drive the trends of [golden-cheeked warbler] habitat disturbance and degradation." *Id.*

The warbler also faces major threats from habitat loss due to droughts and wildfires that are exacerbated by climate change. Polley et al. (2018) discussed the historic drought of 2011–13 in central Texas which killed an estimated 300 million trees. H. Wayne Polley et al., *Projected Drought Effects on the Demography of Ashe Juniper Populations Inferred from Remote Measurements of Tree Canopies*, 219 Plant Ecol. 1259, 1260 (2018). They wrote that "mortality was particularly high in Ashe juniper . . . trees, which experienced a 27% canopy dieback in the Edwards Plateau region"—a region that substantially overlaps with warbler habitat. *Id.; see also* Sarah E. Crouchet et al., *Tree Mortality After a Hot Drought: Distinguishing Density-Dependent and -Independent Drivers and Why it Matters*, 2 Front. For. Glob. Change 1, 3 (2019). Droughts of this nature—and consequent tree mortality—are expected to increase in frequency "as climate change promotes more frequent or severe water limitation." *Id.* at 1259; *see also* John Nielsen-Gammon et al., *Assessment of Historic and Future Trends of Extreme Weather in Texas, 1900-2036*, Texas A&M University, Office of the Texas State Climatologist (March 5, 2020) at 16. Polley et al. (2018) ultimately found that a severe drought could kill "18–85% . . . of intermediate- to large-sized Ashe juniper trees in central Texas savannas." *Id.* at 1264.

The implications of this research for golden-cheeked warblers are significant. Of course, the warbler breeds exclusively in Ashe juniper-oak woodlands in central Texas. *See, e.g.*, Dreiss et al. (2022) at 1. But within this limited breeding habitat, the warbler relies in particular on an abundance of *mature* Ashe juniper trees: warblers construct nests from Ashe juniper bark, and Ashe junipers do not begin shedding bark until at least 20 years of age. *See* U.S. Fish & Wildlife Serv., Golden-cheeked Warbler Recovery Plan (1992) ["Recovery Plan"] at 8. Within a given species, larger trees tend to be older trees—so the "intermediate- to large-sized" Ashe junipers that are threatened with widespread mortality in a severe drought, as forecasted by Polley et al. (2018), will, in many instances, be the older trees that warblers need to construct nests. Polley et al. (2018) at 1260. In other words, a severe drought—an event rendered more likely by global climate change—threatens to kill many of the specific type of tree that golden-cheeked warblers need to fulfill a basic life function. *Id.*

Wildfires have also increased in frequency on the Edwards Plateau, representing another growing threat to warbler breeding habitat. Donovan et al. (2017) reported, "For the entire [Great Plains] biome, an average of only 33.4 ± 5.6 large wildfires occurred each year from 1985 to 1994, whereas mean annual large wildfire number increased to 116.8 ± 28.8 wildfires per year from 2005 to 2014." Victoria M. Donovan et al., *Surging Wildfire Activity in a Grassland Biome*, 44 Geophys. Res. Lett. 5986, 5988 (2017). Furthermore, "[t]his significantly increasing trend over time . . . was primarily driven by increases in the number of wildfires in southern and west central portions of the . . . biome," which includes the Edwards Plateau. *Id.* at 5988, 5990 fig. 2 & tbl. 2.

Crown fire, in particular, presents a serious threat to the golden-cheeked warbler population. Reemts et al. (2008) noted that, while all five species of oak that occur in central Texas "resprout after being top-killed, Ashe juniper does not." Charlotte M. Reemts et al., *Slow Recolonization of Burned Oak-Juniper Woodlands by Ashe Juniper (*Juniperus ashei*): Ten Years of Succession After Crown Fire*, 255 Forest Ecology & Management 1057, 1058 (2008). Following three grass fires in February 1996 at the Fort Hood Military Reservation—which "consumed more than 4000 [hectares] of woodland, including 2108 [hectares] of golden-cheeked warbler habitat"—the authors monitored the burned areas over a nine-year period. *Id.* They found that, in the tenth growing season after the fires, "Ashe juniper abundance in the burned areas was still very low." *Id.* at 1062. They further found that, in the second warbler breeding season after the fires, "the relative abundance of [warblers] in one of the burned areas [had] declined by more than 80%," and eight years later, warbler density in that area was largely unchanged. *Id.* at 1064. Reemts et al. (2008) concludes, "Because of the slow recovery of Ashe juniper, the burned areas will likely not host large numbers of golden-cheeked warblers for several more decades." *Id.* As climate change drives an increase in wildfire frequency, the threat of crown fire, which can destroy essential warbler breeding habitat for decades at a time, is even more dire.

Climate change also threatens to increase the frequency of strong El Niño phases, which in turn have been linked to lower survival rates, lower bodily fitness pre-migration, and lower reproductive success for migratory birds breeding in the eastern United States and wintering in Central and South America or the Caribbean, such as the golden-cheeked warbler. In a forthcoming paper, Jennifer L. Reidy, Emily A. Sinnott, Frank R. Thompson, and Lisa O'Donnell describe an integrated population model that shows that an increase in frequency of strong El Niño events increases the warbler's probability of quasi-extinction, defined as fewer than 10 breeding males across all study plots, within 25 years. Strong El Niño phases are characterized by severe drought in the warbler's wintering range. The Reidy et al. study also shows that an increase in urban land cover around breeding habitat can increase the warbler's probability of quasi-extinction to an even greater degree than strong climate events.

Finally, a forthcoming study by Dr. Giri Athrey appears likely to provide valuable information regarding the golden-cheeked warbler's current genetic status. Dr. Athrey and his team are assembling the reference genome of the species, making the warbler the first endangered songbird in the contiguous United States which will have a fully assembled genome. This research builds on Dr. Athrey's prior work, which showed steep declines in the warbler's genetic diversity, and concurrent increases in genetic differentiation, caused by habitat fragmentation over a 100-year period. *See* Giridhar Athrey et al., *Crumbling Diversity: Comparison of Historical Archived and Contemporary Natural Populations Indicate Reduced Genetic Diversity and Increasing Genetic Differentiation in the Golden-cheeked Warbler*, 12 Conserv. Genet. 1345 (2011). Dr. Athrey has explained that this represents one of the greatest documented genetic declines in a bird species, and there is currently no evidence to suggest that the warbler has recovered genetically. At this writing, Dr. Athrey and his team are analyzing DNA sequencing data received from a core facility.

The Service must incorporate the findings of the forthcoming Reidy et al. study, and the forthcoming results of Dr. Athrey's research, in its warbler five-year status review. Indeed, the Service itself has acknowledged that "additional information . . . includ[ing] predictions about future climate change and how future climate change may affect [warbler] populations and habitat," and "population genetics and genomics studies to identify existing population structure on breeding grounds"—namely, the very work Dr. Athrey was commissioned to complete—"is needed to quantify the progress of recovery for the species to date and inform the current condition of [warbler] populations in the context of a possible future [Species Status Assessment]." *See* U.S. Fish and Wildlife Serv. & U.S. Dep't of Def., Recovery and Sustainment Partnership Initiative: Golden-cheeked warbler (*Setophaga chrysoparia*; formerly *Dendroica chrysoparia*) Action Plan at 4 (October 30, 2018).

In sum, the new and forthcoming scientific research discussed above documents ongoing and future threats that support a finding that the golden-cheeked warbler should remain listed as endangered under the Endangered Species Act.

II. The Mueller et al. (2022) study on golden-cheeked warbler abundance does not support a change in the warbler's listing status.

Available information about the progress of the Service's Species Status Assessment process for the golden-cheeked warbler to date indicates that the Service is placing a high degree of reliance on a recent modeling study by Dr. James Mueller and other contributors that posits a high level of warbler abundance. However, available scientific information, including critiques of the Mueller study methodology as well as actual warbler monitoring data that conflicts with the Mueller study's estimates, calls the Mueller study's population estimate into question and counsels caution in basing any decisions about the warbler's listing status on the population estimate derived from that study.

On May 17, 2022, several dozen golden-cheeked warbler expert biologists received an email invitation from the Service to a June 2, 2022, golden-cheeked warbler Species Status Assessment Expert Meeting.[1] Pursuant to internal agency guidance, a Species Status Assessment forms the scientific underpinnings for the Service's status review, which itself culminates in an FWS decision to delist, "downlist," or preserve the ESA listing status quo for the relevant species—in this case, the golden-cheeked warbler. *See* 16 U.S.C. § 1533(c)(2); *see also* U.S. Fish & Wildlife Serv., Species Status Assessment Framework (Aug. 2016), https://www.fws.gov/sites/default/files/documents/species-status-assesment-framework-2016-08-10.pdf. The May 17 email stated that the Service's draft Species Status Assessment's analysis of "current conditions draws from a new publication by USFWS biologists Dr. Jim Mueller, Dr.

[1] A copy of this email was produced pursuant to an August 26, 2022, Texas Public Information Act request to the City of Austin and is available upon request.

Steve Sesnie, and Dr. Sarah Lehnen," among other contributors. The email included a "strong[] request" that invitees read the Mueller et al. (2022) study before the meeting.

This Mueller et al. (2022) study attempts to estimate golden-cheeked warbler population size using a multi-scale density model. *See* James M. Mueller et al., *Multi-Scale Species Density Model for Conserving an Endangered Songbird*, 86 J. Wildlife Management 1 (2022). Mueller et al. (2022) modifies the approach undertaken by prior modeling analyses by "describing the gradients of habitat features that influence warbler density . . . rather than creating a binary classification of habitat versus non-habitat." *Id.* at 14. Applying their new approach, the authors conclude that "[p]opulation estimates exceeded the targets established for a viable population in every recovery unit . . . and summed to a robust global population size of 217,444 males." *Id.* at 17.

The conclusions of the Mueller et al. (2022) study do not provide grounds for a decision to change the status of the golden-cheeked warbler. To begin, the study utilizes a model that overestimates warbler abundance and fails to validate its model using known populations of banded warblers. Expert warbler biologists Lisa O'Donnell and Jennifer Reidy discussed these and other critiques of Mueller et al. (2022) in letters to the authors, which are attached to this letter and summarized, in pertinent part, as follows.

Reidy offered three overarching critiques of Mueller et al. (2022): the sampling design was biased, and the authors do not account for this bias in a satisfactory manner; the model was not validated against known populations of banded warblers; and the authors did not provide biological reasons for the model selected. *See* Ex. A at 1. On the first point, Reidy commented that the study's sampling design resulted in oversampling in areas with higher predicted warbler density, and no sampling in areas of lower predicted density. *Id.* at 3. In that regard, Mueller et al. (2022) adopted a model from Steven E. Sesnie et al., *Airborne Laser Altimetry and Multispectral Imagery for Modeling Golden-cheeked Warbler* (Setophaga chrysoparia) *Density*, 7 Ecosphere 1, 3 (2016), but that model was validated only for the Balcones Canyonlands National Wildlife Refuge. Mueller et al. (2022) uses that model, from a small number of survey points at a single site, to predict warbler densities across its breeding range. Ex. A at 2. Reidy notes that Mueller et al. (2022) does not provide an adequate explanation for how this sampling bias is addressed. *Id.* at 1, 3. Reidy commented, further, that the study's failure to validate its model through a comparison with independently collected data of banded warblers renders it impossible to determine the model's accuracy. *Id.* at 1, 8.[2]

O'Donnell, similarly, commented that Mueller et al. (2022) utilized a density model with a known pattern of overprediction, particularly at low densities, and failed to acknowledge the multiple factors that would contribute to this overprediction. *See* Ex. B at 2. Because of this, validation of density models with known warbler densities is important, yet Mueller et al. (2022) failed to validate its model with publicly available warbler densities. *Id.* O'Donnell requested

---

[2] In addition to these points, Reidy provided comment on each section of Mueller et al. (2022), including study area, survey implementation, and development of predictor variables. *See generally* Ex. A.

predicted warbler densities from the authors and compared those mean predicted densities with known densities on 15 intensive monitoring plots on the Balcones Canyonlands Preserve (BCP), and found the same pattern of overprediction—a pattern which she expected to be even more significant outside the BCP. *Id.* at 2–3. In fact, she calculated that the Mueller et al. (2022) abundance estimate for Kickapoo Cavern State Park, for example, was *up to 18 times* too high, given known densities at that site from banding and DNA samples. *Id.* at 2–3, 10. This demonstrated pattern of overprediction calls into question the Service's apparent reliance on the Mueller et al. (2022) study to establish "current conditions" for the golden-cheeked warbler in the ongoing Species Status Assessment process, as described in the agency's May 17 email. O'Donnell also commented that Mueller et al. (2022) placed an undue focus on population size, when recent population viability models show that population size has little effect on this species' population trajectories. *Id.* at 2.

Dr. Mueller and his coauthors have never provided a written, science-based response to the Reidy and O'Donnell critiques. This omission is particularly significant in the case of the O'Donnell critique because it is our understanding that Dr. Mueller's 2018 research permit to conduct point count surveys on the Balcones Canyonlands Preserve required, among other things, (1) sharing a draft of his paper with Ms. O'Donnell and her colleagues at the City of Austin prior to journal submission, and (2) comparing warbler abundance estimates derived from the Mueller study with observed densities from territory mapping data held by the City of Austin. Our understanding is that the City of Austin was not provided with a draft of the Mueller et al. (2022) paper until January 2022—eight months after it was submitted to the Journal of Wildlife Management. And instead of validating their model with territory mapping data held by the City of Austin, the study authors compared their model outputs against a "non-public database compiled by the U.S. Fish and Wildlife Service of known warbler locations." Thus, the Mueller paper's conclusions were not informed by comment from the City of Austin, which was specially positioned to evaluate the accuracy of the Mueller et al. (2022) density model, given its intensive warbler monitoring on the Balcones Canyonlands Preserve. No other scientific peer review of Mueller et al. (2022) afforded an equivalent opportunity for evaluation of that study's conclusions.

Even apart from these critiques of its population estimate, Mueller et al. (2022) does not support a decision to change the listing status of the golden-cheeked warbler because this species continues to face significant and even accelerating threats, as outlined above. The Service's own Recovery Plan states that the warbler will not be considered for delisting until, among other things, "sufficient breeding habitat has been protected to ensure the continued existence of at least one viable, self-sustaining population in each of eight [recovery] regions," and that this and the other delisting criteria have been "maintained for at least 10 consecutive years." Recovery Plan at 40. As previously discussed, recent research shows that the warbler's breeding habitat is threatened by urbanization, wildfires, and drought—indeed, there is no indication that "sufficient" breeding habitat is "protected" to ensure a viable population in each of the eight recovery regions. The abundance estimates of Mueller et al. (2022)—even if accepted as accurate, for the sake of argument—do not change this ongoing reality of severe threats to the continued existence of this species.

III. Given Dr. Mueller's primary authorship of a study with clear listing-related implications, and his role in the ongoing status review, the Service must take special care to ensure the status review's objectivity.

Finally, given the information discussed above, it is especially important for the Service to take affirmative steps to ensure both the appearance and the reality of objectivity in the Species Status Assessment process for the golden-cheeked warbler.

As discussed, the Service's May 17, 2022, email suggests the apparent centrality of the Mueller et al. (2022) study in the Species Status Assessment process. In this regard, Dr. Mueller is the Service's own staff biologist. Yet the information discussed above also underscores the contested nature of Dr. Mueller's findings. In these circumstances, there is an obvious potential for the perception, if not indeed the reality, to develop that the Service might favor the conclusions of its own scientist, Dr. Mueller, over the critiques and countervailing information offered by scientists from outside the Service.

To address this potential, the Service should take affirmative steps to ensure that the Species Status Assessment process proceeds objectively based on the best available scientific information without undue influence from any interested party. Yet, based on our understanding at this time, it appears that Dr. Mueller is playing a central role in the ongoing Species Status Assessment for the golden-cheeked warbler. Any decision regarding the golden-cheeked warbler's ESA listing status must be based on the "best scientific . . . data available," 16 U.S.C. § 1533(b)(1)(A), (c)(2), which necessarily entails an impartial evaluation of Mueller et al. (2022) alongside work of other golden-cheeked warbler expert scientists. The Service should take affirmative steps to ensure that the Species Status Assessment, and attendant status review, are conducted in the objective and rigorous manner that the ESA requires.

Sincerely,

/s/ Sharmeen Morrison
Attorney
Earthjustice
48 Wall Street, 15th Floor
New York, NY 10033
Tel.: (212) 284-8034
Fax: (212) 918-1556
smorrison@earthjustice.org

CC:

Amy Leuders, amy_lueders@fws.gov
Regional Director, Southwest Region, U.S. Fish and Wildlife Service
Catherine Yeargan, catherine_yeargan@fws.gov
Acting Field Supervisor, Austin Ecological Services Field Office, U.S. Fish and Wildlife Service

Enclosures:

As exhibits:

Exhibit A: Correspondence from Jennifer Reidy to Jim Mueller, Steven Sesnie, Jade Florence, and Christina Williams (June 27, 2022)

Exhibit B: Email from Lisa O'Donnell to Jim Mueller (Feb. 10, 2022)

As .ZIP file:

Correspondence from Jennifer Reidy to Jim Mueller, Steven Sesnie, Jade Florence, and Christina Williams (June 27, 2022)

Email from Lisa O'Donnell to Jim Mueller (Feb. 10, 2022)

Giridhar Athrey et al., *Crumbling Diversity: Comparison of Historical Archived and Contemporary Natural Populations Indicate Reduced Genetic Diversity and Increasing Genetic Differentiation in the Golden-cheeked Warbler*, 12 Conserv. Genet. 1345 (2011)

Sarah E. Crouchet et al., *Tree Mortality After a Hot Drought: Distinguishing Density-Dependent and -Independent Drivers and Why it Matters*, 2 Front. For. Glob. Change 1 (2019)

Victoria M. Donovan et al., *Surging Wildfire Activity in a Grassland Biome*, 44 Geophys. Res. Lett. 5986 (2017)

Lindsay Dreiss et al., *Spatiotemporal Patterns in Golden-cheeked Warbler Breeding Habitat Quantity and Suitability*, 17 Avian Conservation & Ecology 1 (2022)

James M. Mueller et al., *Multi-Scale Species Density Model for Conserving an Endangered Songbird*, 86 J. Wildlife Management 1 (2022)

John Nielsen-Gammon et al., *Assessment of Historic and Future Trends of Extreme Weather in Texas, 1900-2036*, Texas A&M University, Office of the Texas State Climatologist (March 5, 2020)

H. Wayne Polley et al., *Projected Drought Effects on the Demography of Ashe Juniper Populations Inferred from Remote Measurements of Tree Canopies*, 219 Plant Ecol. 1259 (2018)

Charlotte M. Reemts et al., *Slow Recolonization of Burned Oak-Juniper Woodlands by Ashe Juniper (*Juniperus ashei*): Ten Years of Succession After Crown Fire*, 255 Forest Ecology & Management 1057, 1058 (2008)

Steven E. Sesnie et al., *Airborne Laser Altimetry and Multispectral Imagery for Modeling Golden-cheeked Warbler* (Setophaga chrysoparia) *Density*, 7 Ecosphere 1 (2016)

U.S. Fish & Wildlife Serv., Species Status Assessment Framework (Aug. 2016), https://www.fws.gov/sites/default/files/documents/species-status-assesment-framework-2016-08-10.pdf

EXHIBIT A

Review of Mueller et al. 2022

I appreciate the opportunity to provide feedback and comments on the recent publication "Multi-scale species density model for conserving an endangered songbird", by Mueller et al. (2022). This paper expressed two main objectives: 1) to use continuous vegetation and climate data to evaluate density-habitat and climate relationships with the golden-cheeked warbler, and 2) to estimate the current population size of the warbler across the breeding range and compare this to a previous estimate by Mathewson et al. (2012). The study used novel and interesting covariates to predict relationships between warbler density and habitat and climate data. I commend the use of continuous habitat data across the breeding range and the exploration of novel covariates and I welcome the opportunity to do a deeper dive into these variables in the future. I consider the warbler density-habitat relationships as likely valid and reflective of the current knowledge of warbler habitat preferences. However, given the concerns I express in detail below, I do not regard the population estimate or the predicted range-wide densities as valid without more information. The results of this study should be used to highlight areas for protection of habitat important to the species and the resultant density map may best be viewed as an occupancy map.

I have three main concerns with the study, which will be discussed in further detail throughout my specific comments: 1) the sampling design is biased and there was no adequate explanation for how the authors dealt with this bias in the methods or results, 2) the model was not validated against known populations of banded warblers so we have no idea how accurate the model was at predicting warbler density, and 3) the authors did not provide biological reasons or predictions for the models selected and some of the modeling steps/models did not make sense as a result. The following comments will expand upon these concerns in greater detail. My comments follow the general flow of the paper.

Introduction: The introduction begins with reporting two wildly disparate population estimates, yet no explanation is provided as to how those population estimates were derived or what the potential problems with relying on them could be. They then refer to Klassen et al. (2012) as evidence that previous studies using patch-based measures of woodland to determine occupancy, density, or habitat loss likely excluded habitat that is used by warblers. Klassen et al. (2012) measured canopy cover in the field using a densiometer within individual warbler territories, and the reported canopy over was the average of the points within a territory. This is very different from the pixel-based classification of canopy cover used by the other studies. It would be an interesting (and worthwhile) comparison to calculate the pixel-based canopy cover within territories from the Klassen et al. (2012) study and determine correspondence with field-based measurements. My point being, this is comparing apples to oranges without the above comparison determining how congruent the two methods are.

Based on the concluding statement alluding to greater need for knowledge of warbler use in low canopy areas, I expected a greater emphasis on surveying across all levels of canopy cover throughout the range. The entire second paragraph makes the argument that previous habitat delineations for the breeding range are erroneous and falsely based on assumptions that warblers exclusively use woodland habitat, or avoid areas of low canopy cover. My concern here is that it is generally well established that the species is a woodland habitat specialist, but if the authors' wish to dispute this (and to suggest that the species uses a broader range of canopy conditions than otherwise reported), why did they follow up with a sampling design that was completely contrary to measuring this?

Study Area: The study area was cited as 67,861 km^2, but it was not stated how the study area was defined. Was it based on habitat delineation (stated that the area includes mosaic of habitats, some non-warbler habitat), total area of all 35 counties with a record for warblers, recovery regions, something else? Was this the area the population estimate was extrapolated to? What criteria was used to distinguish woodlands and shrublands – canopy height, canopy structure, canopy cover, ... – if that is relevant for defining warbler habitat in the future. The authors stated they expected warblers to occur in woodlands, but throughout the paper, it seemed a main point was to include non-woodland (or low canopy shrubland-type habitat) in future habitat delineations and population estimates (potentially inflating such estimates if they are erroneously included as habitat).

The authors reported permission to survey 32 private and 37 public properties, but they did not include the size they surveyed of the respective properties. I would like to know the percentage of points and surveyed area on public lands vs private lands. The majority of warbler habitat is unprotected (Groce et al. 2010), and therefore understanding what portion of the population was being estimated from (apparently mostly public) and then extrapolated to (mostly private) is of critical importance. Additionally, it was stated the actual sampled area spans 2,112 km^2; in other words, the sampled area was 3% of the area extrapolated to (if that was the area extrapolated to)!

Methods: How was live oak defined under the broadleaf and juniper classification? Was it included in broadleaf or in juniper? While live oak is an important component of the tree canopy throughout the warbler's range, it is important to identify it correctly, particularly because live oak savannas are common in the western portion of the range and do not constitute warbler habitat alone (i.e., in the absence of mature junipers).

Why did you use a model from Sesnie et al. (2016) to predict warbler density across the range? That model was validated and performed well for the area sampled (BCNWR), but those results do not mean the predictions automatically hold across the entire breeding range. For example, we used the most supported model from the Sesnie et al. (0216) study to predict detectability and warbler density on BCP and found different variables were supported for modeling detection and density was generally biased high, particularly at low density plots (City of Austin 2018, Appendix S3). We found the detection probability was much lower on BCP using the supported variables from Sesnie et al. (2016), and that contributed substantially to the bias.

I do not understand the stratified sampling design based on predicted densities across the arbitrary watersheds. If the objective was to determine relationships across the tree canopy spectrum (as seemingly stated in the introduction), why not calculate the distribution of tree canopy cover (based on the 10-m pixel value) and use breakpoints that allow for balanced sampling across the breeding range (or rather, the sampling area) that reflects the tree cover range? Instead, a model from a small number of points at a single site was used to predict warbler densities range-wide and then the sampling was stratified using this as a guide. The authors also stated they did not randomly sample points because it would have reflected the sizes of the properties that they had permission to study, an admission that the available sampling area may not reflect the available habitat, which points to unrepresentative sampling. The very fact that >40% of possible points to survey (from >27,000) were on one (public) property highlights the probability that habitat conditions (or densities) across the breeding range were not representatively sampled. This design seems like a sophisticated way to avoid admitting the sample was not random and representative of all available habitat.

Regarding the designation of survey points into density strata: was the survey point designated with the density predicted for the 10-m pixel if fell within? What if that pixel was different than the neighboring pixels and perhaps not representative of the predicted densities at a larger spatial scale? When discussing the density stratification cutoff values, what variation were they referring to within and among strata: Number of points? Number of birds expected within each level? Something else? It is not clear. Then, only 3 of 4 designated strata were even sampled. What percentage of the total available points (and available habitat, or breeding range area) did this eliminate? And why eliminate these when the objective was to determine the warbler's use of low canopy cover (and again, referencing Klassen et al.'s study, with a different measure of canopy cover, and Farrell's study, where warbler settlement patterns were manipulated as part of the study, for the reasoning)?

The sampling allocation scheme resulted in oversampling in areas of higher predicted warbler densities and not sampling at all in areas of low predicted densities. I am not sure how this sampling design achieves the stated desire to "redefine warbler habitat and refine estimates of habitat and abundance" (from introduction) because it failed to sample across the tree canopy spectrum. Ultimately, more sampling is needed in areas of low or medium tree canopy cover to better determine warbler densities within these landscapes, especially given that the majority of the range supports the lower density stratum of warblers (<0.25 males/ha) as the authors later report in the discussion. (And none of this touches on the other potentially important and predictive variables used herein or by others.)

I read the referenced papers Johnson et al. (2010), Conn et al. (2017), and Schmidt and Deacy (2021), along with related papers, to better understand the unbalanced sampling design. The general conclusion from these papers is preferential (or biased) sampling results in known and implicit bias. Each paper used sophisticated modeling to deal with the sampling bias and Conn et al. (2017) even stated that a design-based approach to random and representative sampling is preferred over what they termed "model-based triage". I did not see any evidence within the analysis that the sampling bias was remedied or dealt with in any way.

In summary, I cannot tell what modifications the authors made to the distance analysis that were in line with the referenced papers by Johnson et al. (2010) or Schmidt and Deacy (2021). I also do not see anything in the Johnson et al. (2010) paper that supports the statement referencing it ("This approach of optimizing survey effort results..."). I appreciate the efforts made to provide reasoning for the biased sampling design (such as Table S1, but I have no idea how that proves the sampling design is unbiased), but I am missing or misunderstanding the execution that dealt with the biased sampling.

Side note: I am curious how 73 inaccessible points were available for sampling – it seems like the majority would have been eliminated by being in the lowest, unsampled stratum.

Below are some excerpts from some of the literature I read reinforcing my concerns about results from the sampling design being applied to a breeding range population estimate:

"We develop a hierarchical statistical modelling framework for detecting and alleviating the biasing effects of preferential sampling in spatial distribution models fitted to count data. The approach works by specifying a joint model for population density and the locations selected for sampling, and specifying a dependent correlation structure between the two processes. Using simulation, we show that moderate levels of preferential sampling can lead to large (e.g. 40%) bias in estimates of animal density and that our modelling approach can considerably reduce this bias. In contrast, preferential sampling did not

appear to bias inferences about parameters informing species–habitat relationships (i.e. slope parameters)." – Abstract from Conn et al. 2017

"We show that abundance is highly overestimated at low abundance locations when preferential sampling effects are not accounted for, in both a simulated example and a practical application using fishery data. This highlights that ecologists should be aware of the potential bias resulting from preferential sampling and account for it in a model when a survey is based on non-randomized and/or non-systematic sampling." - Abstract from Pennino et al. 2019

"Unbiased estimates are dependent upon either (i) distribution of sampling effort being random throughout the study area (for design-based inference) or (ii) model correctness (for model-based inference). It is easier to have confidence in the former rather than in the latter because our models are always wrong. Nevertheless, model-based inference will play an increasing role in population assessment as the availability of spatially referenced data increases."-Miller et al. 2013

Survey implementation: Just some general notes and observations- I suspect that training was not adequate (1 week) to sufficiently train a crew in distance estimation and song identification. Distance estimation is hard, and requires a lot of training with multiple observers to sync. Warblers have a large repertoire of songs, especially range-wide, beyond the typical recordings of A and B songs. For these reasons, I had my surveyors on BCP in the field for a month before they did point counts and that was restricted to one preserve, so there was less song variation. I also collected point count data on BCNWR but with less training and the data was so poor that I couldn't use it. Starting prior to sunrise can also impart some bias into the overall surveys for 2 reasons: 1) singing rates tend to be highest at this time (so detection rates may differ), 2) because singing rates are highest for most songbirds, the extra layer of noise can be distracting and cause some distant warblers to go undetected that would have been heard later in the morning (so detection distances may differ).

Development of predictor variables: There were several novel and interesting variables used in this study covering the range of habitat, land use, and climate data available. For vegetation variables, I am not clear how proportion of canopy that is juniper is calculated – e.g., is it the total percent of a pixel that is juniper or is it total percent of the canopied portion of a pixel that is juniper (and how is canopy defined)? How it is defined will make a difference. For example, think of the difference in a pixel that is 100% canopy, of which 20% of that is juniper, vs a pixel that is 20% canopy but 100% of that canopy is juniper. Both pixels are 20% juniper, but one is embedded in a pixel with high canopy cover while the other is in a pixel with low canopy cover. Similarly, I am concerned about the use of the % of canopy above 3m and how this variable is defined. Figure 2 in supplemental material (distribution of predictor variables) shows the majority of the entire study area supports a high percentage of canopy cover above 3m even in areas that otherwise show they support low canopy cover. Was there no threshold for how much of a pixel had to be >3 m to be considered canopy? As one example, in Figure 1, area 30b shows up as almost all <20% canopy, but in Figure 2, the same area shows up with >50% seemingly with a canopy >3m.

I am interested in the hypotheses supporting the inclusion of the various NDVI-derived variables. I think these are interesting to a species with a broader range, but was there enough variation across the range to warrant inclusion for modeling density of warblers? Same question regarding the topographic and climate variables – what were the hypotheses and what was the variation? My thought is the topographic variables are correlated with important habitat structural features, such as red oak occurrence (or other favorable trees), water availability, age/maturity of woodland, woodland complexity, etc. Why was the precipitation data used from 1970-2000 – why use a 30-year period and why is the end year ~20 years prior to sampling? We have strong evidence that the climate in central Texas is becoming warmer and drier, so it makes more sense to me to use data from the last 10 years as more representative of current conditions. I would appreciate more information regarding the data layers also. For example, forest loss was estimated from a global data set provided by Hansen et al. (2013), but no additional information was provided such as what spatial data layer(s) were used to define forest and what spatial resolution was used for analysis.

I am also concerned about the lack of presentation for biological hypotheses regarding the predictor variables and the scales used. For example, looking at Table S2, why did the authors choose to look at vegetative phenology at the site and landscape scale but not the territory scale? Was it because of the original resolution or a lack of variation within the small spatial scale or another reason? What were the hypotheses for using them at all? Same question regarding the climate variables and hypotheses. Why were the variables for anthropogenic disturbance considered at the territory scale, where I suspect values for all (except perhaps forest edge) were very low, and not considered at the landscape scale, where there was likely large variation? Why were these included in a range-wide study of this nature when they are probably not important within the context of this type of study.

Detection and density model selection: I believe it is customary to include the distance interval bins in distance-based studies. The authors indicated they used a 3-min duration survey to avoid overestimation and cited Peak (2011) and Sesnie et al. (2016). Peak (2011) indicated distance sampling overestimated by 29% using a 3-min interval, and indicated all durations resulted in overestimates compared to a known population of marked birds. In a follow-up study, Peak and Thompson (2013) reported that predicted density was 44% greater than known populations using a 2-min duration and attempting to meet all assumptions for distance sampling with extensive training. Sesnie et al. (2016) reported predicted densities were 83-110% of territory counts across 2- to 6-min surveys and found a slight underestimate using the 3-min duration. These disparate results may be the result of different predictor variables, the difference in range of predictor variables at the two sites, or because Sesnie et al. (2016) visited each point 4 times within a season rather than just once. The current study also only visited points once and therefore, the bias may be closer to the Peak study and warrants further investigation.

Paragraph 4 detailed evaluating different canopy height breaks and concluded with the statement that they used the canopy height variable with the most support, but it was not clearly reported what that was in the results section (it appears to be 3m). I personally found the description of what they did confusing also.

Why did the authors lump the site and territory scales in the analysis? There was no reporting on the total number of models evaluated or the results of the step-by-step model/variable selection process. Table S2 just provided the variable list and the scales they were calculated for. I am leery of this model selection procedure because it feels closer to data dredging rather than testing predictions. Also, we have ample strong evidence that density is landscape-dependent and relationships differ by spatial scale. For example, I would expect warblers to use areas of low canopy at a territory scale within an area of greater juniper-oak canopy at the site or landscape scale. Likewise, I would predict an area of high canopy cover at the territory scale to have lower occurrence or density within a landscape of low canopy cover. Therefore, I would have a model that reflected that prediction. The final model should have been included meaningful variables of interest from multiple spatial scales that resulted from well thought-out predictions and literature reviews.

I do not see any model-based adjustment to deal with the spatially-biased sampling design within the distsamp procedure. Did they perform an adjustment and if so, where was it reported?

Generation of raster: What was the minimum canopy cover – 20%? It was not clearly stated; from Figure 2B, it appears to be ~25%. Why was there a need to restrict coverage to forest or shrub-scrub if the density estimates were gathered across a continuum of the habitat variables? Wouldn't it happen naturally that those areas were estimated to have 0 warblers? What was the pixel-size of the raster? Did they predict mean abundance for every 10-m pixel within the forest and shrub-scrub layer? Was the total abundance the sum of predicted abundance from each pixel?

I am completely confused as to the pertinence of the non-public and E-bird warbler locations. Was this an attempt at validating the model? At best, this can be considered an attempt to validate an occupancy model. The majority of E-bird locations were from public land that is protecting juniper-oak woodland. And I assume the majority (all?) of the FWS locations were coming from FWS land or those with a permit to survey for warblers and required to submit a report (surveying private or public land?). The authors generated abundance estimates for individual counties where almost no sampling occurred, so they should have been able to generate them for Fort Hood and BCP, where ~40-50% of the points were located and compared those to expected populations extrapolated from plots with marked birds.

I would argue against reporting that this species has high detectability (after the initial settlement period) and that because a birdwatcher indicated a lack of detection at a location means there were no birds. It is much harder to prove absence than presence and after pairing the singing rates in warblers becomes highly variable. I have sat for HOURS waiting for a male to sing in a territory. I also did a study on song rates for which I recorded every minute I detected a song for up to an hour. The results demonstrated that warblers will stay silent for very long periods after pairing and in particular during pair-bonding and building.

Results:

Point count surveys: Table S4 showed the point distribution by the geographic region that resulted from the allocation of points based on preliminary predicted densities. It appears the % of points sampled within a region is on par with the % of the region's area compared to the total study area, but there is a disparity in the % of the region sampled of the total study area, hence the clustered point allocation. The sampled area of the Lower Brazos region was 13% of its total area, in comparison to 1-3% of the other regions, but it was 45% of the total cumulative sampled area. The Colorado region supported over 1/3 of the total surveyed points but the sampled area was <2% of the total area of the region. All of this raises red flags that can be lessened with the inclusion of descriptive statistics of the covariate data. Variables used in the final model should have been presented in the paper and the full covariate data should have been included in supplementary material. This covariate data should include the mean, SD, and range for all covariates for the points surveyed, the sampled area, and for the study area they extrapolated to. Refer to Reidy et al. (2016) Table 1, where covariate data is reported for the surveyed points and for the study area. Additional descriptive statistics such as number of points surveyed in public/private (along with area) and within the density strata identified previously would also be helpful.

Warblers were only detected at 37% of points. This does not make a good case for warblers being highly detectable.

Models: Buckland et al. (2001) suggests truncating the farthest 10% of observations (does Buckland et al. 2015 say something different?) and this truncation distance only eliminated <3% of detections. In the rugged terrain that most warblers are found in, a truncation distance of 180 m seems high and suggests overestimation of distance was possible (same as in Sesnie et al. 2016). Overall detection probability and number of observers were not reported (top of page 6). Both are of interest and can be compared to similar studies so it should have been reported.

I am a little confused about the road density in the detection models. Am I correct in thinking you mean that in the original step evaluating detection models, all three variables were significant, and therefore included in all subsequent density models, but was no longer significant in the final combined scales model, so you excluded it? Or did you re-run detection variables with density models?

It was not reported what canopy height was ultimately used, or how many models were ultimately considered. The models shown in Table 3 do not appear to be all candidate models considered, so Table 3 should include the criteria used for cutoffs, even if it was just wi>0.0. Regarding canopy height, in Table 3 and Figure 2, it is clear that canopy height is referring to the % of canopy cover that is above 3 m (correct?), and it appears that the authors chose this from a single preliminary step and then used the most supported canopy height % in all subsequent models, but it should be reported in text for clarity. I did not follow the final step where the authors combined results into an overall density model from all scales. I followed Table 3 and could see the step-wise approach to variable selection for the landscape and the territory/site scales, but then the final overall model had a completely new set of variables. In other words, I expected to see the landscape variables from the top landscape model combined with the territory and site variables from the top territory/site model. Were variables from across scales correlated or did important predictors become insignificant when all variables across scales were included? I do not agree with the approach taken (at least not without more information) for the final model. In particular, the only landscape scale variable in the final model is annual temperature range. What is the biological significance of this variable? What does this effect mean? Why was it even considered? Is it a surrogate for something more meaningful/manageable? What was the point of the step-wise approach if they were ultimately going to use all uncorrelated variables in all possible combinations in the end?

The density model results should begin with the stating the top model and model support for each spatial scale and then follow up with the variable effects (or refer results to supplemental material). Alternatively, at a minimum, the authors need to state in the first sentence what the combined final model was and then proceed with variable effects. By comparing the final model variables to the variable predictions (page10, second paragraph), I was able to determine these were generated from the combined (final) model, but it was not clear from text. I was expecting to see (and how I initially read it) was the most important variables at the territory scale were..., at the site scale were..., at the landscape scale were... and then a final model (more akin to Reidy et al. 2017). The results were especially confusing because the step-wise variable selection was very different from the final model variable set. Also, mean density was not reported.

In the final paragraph, it would be helpful to have similar information on density ranges for the site and landscape scale covariates as is reported for territory variables. The annual temperature range was very small, so I found it interesting that it had such a significant effect and it warrants more discussion as to what this effect means. My guess is that this effect is spurious and related to the greater availability of suitable habitat in the southern portion of the range (with less temperature variability) and less habitat in the north with a slighter higher range of temperatures. I also struggle to find the tight confidence intervals credible, given the relatively small number of points that should have been covering a broad range of conditions.

Distribution (and population size): Why was Figure 3 presented within ecoregions, as opposed to either the same regions specified in Figure 1, or perhaps the current recovery regions (included instead as Figure S2)? The estimated >3,000 males per recovery unit is meaningless to most readers and there was no reason given for that information. I understand that it was because that was the population goal for each of 8 recovery units in the FWS recovery plan, but the majority of readers will not know this. I am curious why a reviewer would not have asked for clarification on this information. What area was the population actually extrapolated to?

The population estimate reported by Mathewson et al. (2012) was not directly comparable to the current one when both employ the 5-min survey duration because the previous estimate was generated using an entirely different survey design: they conducted point counts under different field conditions and design considerations, including a two-step of occupancy surveys then abundance surveys using double observers per point to correct imperfect detection, and a maximum distance of 100 m, among other things.

The comparison with independent data should only be considered a comparison of potential occupancy, not density, and further evaluation of the data's representation of available habitat rather than public habitat is warranted.

In order to produce a credible population estimate: 1) the sampled area should be random and representative of the area to be extrapolated to, and 2) the model should be validated by comparing to independently collected data in areas of well-documented densities. This study does neither so I cannot have confidence in the population estimates.

Discussion: How was it determined that 46 points with warbler detections were considered non-habitat by Duarte et al. (2013)? The range map provided in Duarte et al. (2013) looks very similar to Figure 1. More information is needed to substantiate this statement. Additionally, I consider it misleading to state the threshold Duarte et al. (2013) used as 50% canopy cover. Duarte et al. (2013) did not define or use canopy cover at all but rather defined a 30-m pixel as woodland if >50% was covered by woody vegetation (which should include the continuum of shrub/scrub to forest and based on the figures, it does).

The "rules" used in this study may also exclude potential habitat, correct? The majority of the sampled points were within areas of high predicted densities and the entire stratum of low-density areas was not even surveyed.

What was the source of discrepancy in the study area (6.8 mil ha) compared to that of Mathewson et al. (2012) and Duarte et al. (2013), which were 1.7 mil ha and 1.6 mil ha, respectively? Are the authors suggesting that the potential habitat is >4x what previous delineations were (this goes back to my question in the study area section about how the study area was defined and what area the population was extrapolated to).

The authors point out, correctly, that areas of low canopy cover may be used when in proximity to adequate tree cover but they do not provide the information managers need, such as in what scenarios will a warbler use low canopy cover. I agree that warblers may (and do) use areas of low canopy cover, but only when the low canopy cover is at the small scale and the woodland cover is much greater at the site scale. In other words, warblers' use of areas of low canopy cover is landscape-driven and context-specific. If they use areas of low canopy cover at a broader spatial scale (such as a savanna, or a shrub/scrub site, not connected to an area of greater woodland cover), this is important and novel information that needs deeper study. Regarding the potential of this study to predict abundance to areas of low canopy cover requires an assumption that density is not density-dependent but rather habitat-driven (which may not be true, Farrell et al. 2012) and all the sampled points are representative of the available habitat at multiple spatial scales. In other words, there's an implicit assumption that all areas of low canopy cover at the site or territory scale are the same, whereas I would argue that the amount of canopy cover required for settlement is very dependent on the surrounding landscape and will change based on it. Yet, no vegetation variable was included at the landscape scale in the final model used to project the population.

"We allowed the warbler survey data to determine the habitat relationships" – I would argue that this is misleading and perhaps false because the authors seemingly used a preliminary model based on a small sample from BCNWR to predict warbler densities across the entire range, then divided the range into strata based on the predicted relationships before data was even gathered. A side note: the value of these areas with low canopy cover remains unclear because the current evidence suggests low densities and reproductive success (Klassen et al. 2012). These areas may represent sinks and warrant more intensive study to determine their value to conservation of the species.

I disagree with the statement that "warblers were abundant" for 3 reasons: 1) the term is not quantifiable so what qualifies as "abundant", 2) sampled points were purposely placed within areas expected to be of higher abundance, and 3) despite the biased point sampling, only ~1/3 of points had a warbler detected.

Why the switch to discussing ecoregions, when these areas were not previously specified, not used for the sampling allocation, were not used for recovery purposes, and were not defined elsewhere regarding sampling coverage? Is the 26% of study area referring to the 6.8 mil ha or to the 2.6 mil ha where they report warblers from at the end of page 17 (or something else)? I appreciate the efforts of this study and Mathewson et al. (2012) to provide more robust and defensible population estimates than the cursory efforts from the 1990s, but I do not consider them more or less valid than those earlier ones because of the flaws in the sampling design.

I enjoy exploring and explaining patterns, so the use of new covariates intrigued me. But I also appreciate the need to predicate these on biological reasons or predicted relationships and I am not sure the use of some variables was necessary or appropriate. For example, what was the reason annual temperature range was used to begin with? Was it an alternative to the geographic variability reported in Collier et al. (2012) and modeled using latitude and longitude? Was there an attempt to investigate other landscape features that were more pertinent, such as e.g., a topographic*canopy cover interaction? Same questions regarding the canopy cover*CTI variable. The relationship shown likely relates to woodland integrity, prior land use, woodland age, woodland composition, etc. Areas with high CTI are flatter (based on the description), so the effect on warbler densities with canopy cover makes sense as this probably represents more closed-canopy, older, taller woodlands with mixed oaks and the lower canopy areas on flat sites are often grading into more xeric conditions or live oak savannas, where we don't expect many warblers. In summary, I do not see a compelling argument that warbler density is driven by minute differences in annual temperature, but I do think there are other habitat-related differences that also follow that same spatial pattern, such as degree of fragmentation, effects of current and prior land use, terrain roughness (which also plays into anthropogenic disturbance and into habitat effects such as red oak occurrence), etc. I think exploring those relationships is also more meaningful for land managers.

I agree that the most important information from this study was identifying important spatially-explicit predictors across the range that are readily accessible and temporally updated. What is "prime" habitat? Similar to the earlier statement that warblers were abundant, saying density was greater in "prime habitat" doesn't provide the reader, or the land manager, with much information. Most of the remaining information in the paragraph should have been reported in results. The reported habitat totals ~2.6 mil ha, only 39% of the study area - was the rest of the study area within the unsurveyed stratum? What area was extrapolated to for the population estimate? Why were 2 density categories (Figure 3: 0.06-0.13, 0.13-0.25) lumped together in the reporting of how much area supported the category?

I believe the greatest value in this study can be in identifying areas for future research and for conservation. For example, this study, along with previous studies, have identified portions of the southern range as supporting the most available potential habitat and therefore likely the most warblers, and this area is therefore valuable to conservation and protection of this species and should be used in future land-protection decisions. To take this one step forward, identify the 116K ha with a high density of predicted warblers, overlay with the boundaries of protected lands, and see what areas should be targeted for protection. My calculations result in 80% of the predicted population across just 1.2 mil ha (a fraction of the original study area), and the additional 1.4 mil ha with low densities is questionable – <0.06 males/ha is very low. Additionally, given previous work showing distance sampling tends to overpredict by a much larger bias on low-density sites (e.g., O'Donnell et al. 2019), it is really difficult to assess the validity of the estimates. The results also highlight areas of additional monitoring needs – such as the areas of 30a and 30b and how much these woodland areas support warblers. The entire northern half of the range is patchy and low density, suggesting further study is warranted there also to determine conservation potential.

I do not consider the information regarding juniper growth patterns or fire effects pertinent to this study. It is not studied herein or otherwise related to the predictor variables. Additionally, it is misleading to state that grassland can become mature juniper-oak woodlands within 75 years. It is highly site-dependent and past land use (and abuse), along with soil types, sun exposure, precipitation patterns, browsing pressure, and topography will all interact to affect how a woodland stand develops. Growing a juniper is a lot simpler than growing warbler habitat unfortunately.

Abundance was estimated for every pixel and totals were provided for individual counties, even those with very little sampling and small populations. Why was it unreasonable to produce abundance estimates for large properties such as BCP or Ft Hood, especially when the majority of points were on these properties, and thus enable comparison to known marked populations? The use of fine-scale and continuous data should make it fairly easy to scale down the population estimates.

Distance sampling is known to overestimate density of this species (see Peak 201, Peak and Thompson 2013, Reidy et al. 2016, Fig. 2); other analytical designs have proven problematic as well (O'Donnell et al. 2019). Simply using a shorter survey duration does not eliminate overestimation bias and this needs to be explicitly stated in the discussion rather than ignored. Upward bias tends to increase inversely with density (, Reidy et al. 2016, O'Donnell et al. 2019), such that the predicted densities of 84% of the population (<0.25 males/ha) may be overestimated by 1-30 times actual densities.

Figure 3: Are the areas of white assumed to be non-habitat, areas that were excluded as part of <0.02 warbler density criteria, areas that were surveyed but had 0 warbler detections, part of the 6,786,100 ha of the study area? I am curious too about the stratification design of not surveying areas with predicted densities <0.02 males/ha but then extrapolating to areas with estimated densities of 0.01-0.2 males/ha. I am also really struggling with the extrapolation of the 30a and 30b areas, where almost no sampling occurred and where the woodland patches were small, highly fragmented, and mainly appear associated with canyons. At least in places I have surveyed, there do not tend to be many warblers in these juniper-dominated shrub-scrub habitats and I would argue that much more sampling needs to be done in this area to get good estimates on densities. Why were the density categories different than Figure S2 (inclusion of 0.01-0.2 males/ha in Figure 3)?

Additional notes: Virtually the entire BCP and much of BCNWR appear in red, indicating mean density was predicted to be >0.25 (from the predicted densities raster). Unfortunately, very few areas support densities this high on either of these properties, confirming my suspicion that the model is overpredicting densities. If the model overpredicted in an oversampled area that supported a large percentage of warblers (and of protected land), then how accurate was it at predicting densities where habitat conditions were under sampled, including most areas of low predicted densities or canopy cover?

Literature Cited

Conn, P. B., J. T. Thorson, and D. S. Johnson. 2017. Confronting preferential sampling when analysing population distributions: diagnosis and model-based triage. Methods in Ecology and Evolution 8:1535–1546.

Duarte, A., J. L. R. Jensen, J. S. Hatfield, and F. W. Weckerly. 2013. Spatiotemporal variation in range wide golden cheeked warbler breeding habitat. Ecosphere 4:1–12.

Farrell, S. L., M. L. Morrison, A. J. Campomizzi, and R. N. Wilkins. 2012. Conspecific cues and breeding habitat selection in an endangered woodland warbler. Journal of Animal Ecology 81:1056–1064.

Klassen, J. A., M. L. Morrison, H. A. Mathewson, G. G. Rosenthal, and R. N. Wilkins. 2012. Canopy characteristics affect reproductive success of golden-cheeked warblers. Wildlife Society Bulletin 36:54–60.

O'Donnell, L., C. C. Farquhar, J. W. Hunt, K. Nesvacil, J. L. Reidy, W. Reiner Jr., J. L. Scalise, and C. C. Warren. 2019. Density influences accuracy of model-based estimates for a forest songbird. Journal of Field Ornithology 90:80–90.

Peak, R. G. 2011. A field test of the distance sampling method using golden-cheeked warblers. Journal of Field ornithology 82:311–319.

Peak, R. G., and F. R. Thompson III. 2013. Amount and type of forest cover and edge are important predictors of golden-cheeked warbler density. Condor 115:659–668.

Pennino, M. G., I. Paradinas, J. B. Illian, F. Muñoz, J. M. Bellido, A. López-Quílez, and D. Conesa. 2019. Accounting for preferential sampling in species distribution models. Ecology and Evolution 9:653–663.

Reidy, J. L., F. R. Thompson III, C. Amundson, and L. O'Donnell. 2016. Landscape and local effects on occupancy and densities of an endangered wood-warbler in an urbanizing landscape. Landscape Ecology

Reidy, J. L., F. R. Thompson III, and L. O'Donnell. 2017. Density and nest survival of golden-cheeked warblers. Journal of Wildlife Management 81:678–689.

EXHIBIT B

City of Austin | Austin Water
AustinWater.org

February 10, 2022

Dr. Jim Mueller, Zone Biologist
U.S. Fish and Wildlife Service
National Wildlife Refuge System
Jim_Mueller@fws.gov

Hello Jim,

Thank you for providing your draft manuscript, “Multi-scale Species Density Model for Conserving an Endangered Songbird” for our review. With only three weeks to review the draft, our comments are somewhat cursory but extensive. We had hoped to avoid this by working collaboratively over the past four years, which is why we included special conditions in your 2018 BCP research permit. In addition to providing the draft prior to submission, we mutually agreed to the following:

- BCP and WQPL staff request that reporting and/or publication of predicted Golden-cheeked Warbler density and abundance estimates derived from this study include comparison with observed density and abundance derived from territory mapping data. BCP and WQPL staff will provide these data to the permittee.

- The annual report will provide detailed results that allow for comparison with the City of Austin's territory mapping data for the Golden-cheeked Warbler, including model outputs, map of habitat strata within the BCP/WQPL, predicted density/abundance estimates with 95% confidence intervals for each habitat strata, and sufficient information to allow for independent replication and evaluation of the results.

We are unsure why the above information was never provided. While we did use the link in the draft manuscript of rasters with predicted Golden-cheeked Warbler densities, we were only able to do a quick comparison of the mean predicted densities (derived from 3-minute surveys) to known densities from the same year (2018) for our 15 intensive monitoring plots within the BCP. Given that the consistent overprediction of density models remains unresolved, especially at low densities, we are unable to support many of the findings as currently presented in the draft. Most of these concerns could be addressed if the paper were to focus on using new insights from your habitat models to promote habitat protection efforts instead of unreliable predictions of abundance. We provide general and more specific comments below.

Sincerely,

Lisa O’Donnell, Senior Biologist
Austin Water (AW), Wildland Conservation Division (WCD), Balcones Canyonlands Preserve (BCP)
Lisa.Odonnell@austintexas.gov

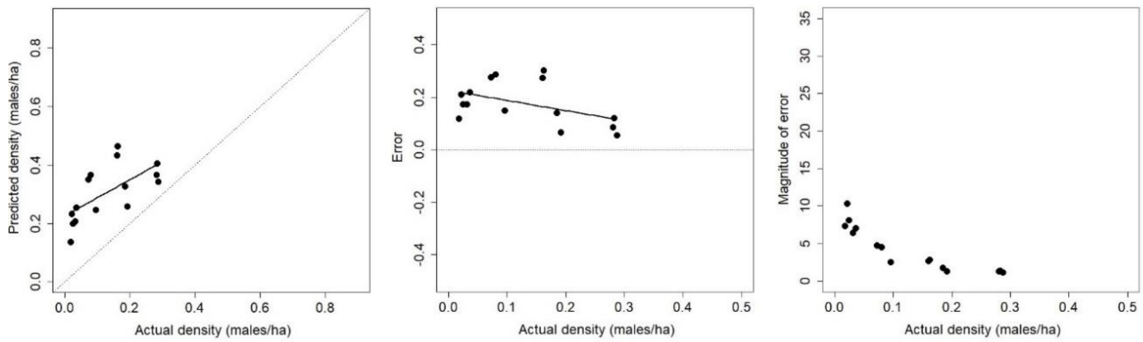
Cc: Nico Hauwert, BCP Program Manager; Kevin Thuesen, Water Quality Protection Lands (WQPL) Program Manager; Matt Hollon, WCD Division Manager; Sherri Kuhl, AW Environmental Resource Officer; Kimberlee Harvey, Balcones Canyonlands Conservation Plan Secretary

General comments: The emphasis on population size obscures the more salient points of the paper. We recommend focusing on new insights from your habitat model that have not previously been published. For example, this is the first model we are aware of that includes canopy height as a variable across the warbler's breeding range. Identifying areas with prime habitat (high tree canopy cover, 60-80% Ashe juniper, and >3 m trees), including prime habitat on public land, would be a major improvement over previous models. At the other end of the spectrum, the distribution of habitat variables associated with zero counts, low predicted occupancy, and the multitude of imminent threats (urban development, tree mortality, conversion to non-forest, climate change, etc.) would help identify areas and actions that are detrimental to this species. The temperature variable as a predictor of warbler density is new, and we are also interested in the compound topographic index, which we had not heard of before.

Changing the focus in such a way would address the majority of our comments. Given the important implications of many of your findings, we are unsure why population size seems to be the main focus. Based on population viability models (most recently Duarte et al. 2016 and Reidy et al. 2020), population size has little effect on population trajectories, which are driven by productivity, juvenile survival, and dispersal (all of which depend on habitat quality). Further, the warbler was listed based on habitat loss (not population size), and Duarte et al. found that habitat changes are unidirectional, with losses more likely than gains. This supports the need to protect prime habitat and identify where threats are occurring so they can be minimized or avoided, which are essential to the warbler's continued survival and recovery. With your new models, your paper has the potential to address these conservation needs.

While models fit to point count data can be useful in predicting occupancy, distribution, and habitat relationships, they remain problematic for predicting abundance due to the large overestimation error, particularly at low densities (Reidy et al. 2018, O'Donnell et al. 2019). This includes density models used to predict warbler abundance on the BCP and range-wide (Mathewson et al. 2012), and the model used in Sesnie et al. (2016) when applied to the BCP. Given the known pattern of overprediction, we are curious why the draft manuscript does not mention this and assumes the major factor leading to density/abundance overestimation is survey duration (3-min vs. 5-min). As noted in Reidy et al. (2018) and O'Donnell et al. (2019), this is likely only one of multiple factors, including biological ones (i.e., territory size increases with decreasing density/habitat quality). We are unaware of any density model that has resolved this issue, which is why Reidy et al. (2016, 2020) relied on validation using known densities, and why we included model validation as one of your permit conditions. For model validation, your study used a "non-public database compiled by the U.S. Fish and Wildlife Service of known warbler locations" to assess how well the model predicts density, so warbler territory data from BCP and other sites with known densities (BCNWR, Camp Bullis, Fort Hood) should be appropriate.

Following our request for the predicted warbler densities on BCP and WQPL, you referred us to links in the draft manuscript to obtain this information. Using the rasters of predicted warbler density, we only had time to compare the mean predicted densities (derived from the 3-minute surveys) to known densities for the 15 intensive monitoring plots on BCP that we surveyed in 2018:



Even with the 3-min survey duration, we see the same pattern of overestimation (on BCP, these are areas of low canopy cover, short trees, and/or small patches). Given that the BCP supports high densities compared to many other areas within the warbler's breeding range, we suspect these overpredictions are even more problematic beyond the BCP. For example, we banded birds and collected DNA samples at Kickapoo Cavern State Park in 2020-2021 and believe the abundance estimate for this site (355.7 males) overestimates 12 to 18 times (see specific comments for Figure 4).

Below are some additional questions and comments on the density model:

1) Given that the sample was intentionally selected to have unequal probabilities, this needs to be taken into account in the analyses by applying sample weights that are the reciprocal of the selection probabilities. We were unable to determine if this was done when the regression models were built to predict abundance from site characteristics.

2) The selection of the 1,804 survey points, distributed among 69 sites, are treated as a random sample rather than a cluster sample. This will make the estimated variances too large—in other words, they will have less information than presumed, since there are many correlated sites.

3) Areas with zero or low numbers of birds were not surveyed, implying that the only data that matters is nonzero counts. However, relating habitat features to counts requires the ability to predict when those counts are low as well as high. In other words, zero is just as important as nonzero values. While the methods state that you didn't project to unsuitable areas, more predictors than just the ones deemed unsuitable from the satellite imagery were used. One suggestion would be to add in multiple zero counts, along with the habitat variables for those zero counts, to see if you get the same prediction equations.

4) We are not clear how the habitat classification model was validated using the 25% holdout sample. How did you know the true classification? We understand that the study included a field survey of some sites and were checked to see how the predictor worked on those. But it appears there are error rates from the holdout sample, and are wondering how you knew the correct classification for those.

5) There is so much complication in the model fitting section, we couldn't keep straight how many different models were fit and for what reason. We understand that there is a detectability model for the distance sampling. From that, were counts adjusted upward for the sites that were sampled?

And then those adjusted (inputed) counts were used as the predicted value in a regression where predictors are the habitat variables. There appear to be several other models discussed, and model averaging as well, so it is not clear which model was used to estimate the final population numbers and how.

Citations for the above reports and publications that address overprediction of warbler density models:

O'Donnell, L., C.C. Farquhar, J.W. Hunt, K. Nesvacil, J.L. Reidy, W. Reiner Jr., J.L. Scalise, and C.C. Warren. 2019. Density influences accuracy of model-based estimates for a forest songbird. Journal of Field Ornithology 90(1):80-90.

Reidy, J.L., F.R. Thompson, C. Amundson, and L. O'Donnell. 2016. Landscape and local effects on occupancy and densities of an endangered wood-warbler in an urbanizing landscape. Landscape Ecology 31:365–382.

Reidy, J.L., F.R. Thompson, and L. O'Donnell. 2018. Evaluation of point counts and prediction of Golden-cheeked Warbler density on the Balcones Canyonlands Preserve, Travis County, Texas. Preliminary report. https://www.traviscountytx.gov/images/tnr/Docs/bccp/2018/appendix-s3.pdf

Reidy, J.L., F.R. Thompson, and L. O'Donnell. 2020. Population viability of Golden-cheeked Warblers in an urbanizing landscape. Wildlife Society Bulletin 44(3):502-511.

<u>Specific comments</u>:

Lines 22-23: We find it interesting that over half of the survey points had no detections and would like to know habitat conditions associated with those locations.

Lines 27-29: The draft abstract mentions “2 variables” and “5 of the 6 terms” -- it would be helpful to state what those are in the abstract.

Lines 31-33: As stated in the general comments, this assumes the major factor leading to density/abundance overestimation is survey duration, but this is likely only one of multiple factors, including biological ones (i.e., territory size increases with decreasing density/habitat quality).

Lines 36-38: It is unclear here what insights the model provides, especially for “the outcome of alternative management practices…that can be obscured using broad land cover categories and anthropogenically-defined habitat patches.” We are not sure what this means.

Lines 39-40: The benefits of this model for conservation and management decisions are not clear. It would be helpful to include a couple of examples. For example, the habitat model may be beneficial in identifying sites with prime habitat needed for long-term survival and recovery and providing management goals.

Lines 46-47: <u>This is incorrect</u>. The warbler was listed based on habitat loss and degradation. While population estimates were included in the final rule listing the warbler as endangered, these were not the basis for the listing decision.

Lines 49-50: The sentence “Today, the population is estimated to be much larger” should either be omitted or reworded so it is not stated as fact. If population estimates remain a focus of this paper, then the differences in methodology and inherent biases, including the consistent pattern of overestimation from multiple models fit to point count data, should be discussed.

Lines 52-55: As stated in the general comments, all density models evaluated to date have been shown to overpredict, particularly at low densities (Reidy et al. 2018, O’Donnell et al. 2019). This includes density models used to predict warbler abundance on the BCP, Mathewson et al. (2012), and the model used in Sesnie et al. (2016) when applied to the BCP. Because field methods have not resolved this issue, Reidy et al. (2016, 2020) validated their models using known densities.

Lines 57-62: This is misleading, since the density model used in this study does not provide any information on effects of more open canopy on species viability. For example, are the birds detected and incorporated into the predicted densities reproducing? We note that Klassen et al. found reproductive success rates to be much lower at Kickapoo Cavern State Park (39.5% and 59.4% in 2009 and 2010, respectively) than on the BCP (78% both years). What percentage are transient/non-breeders? Presence may be an artifact of birds running out of options elsewhere and settling on poor habitat because they have nowhere else to go.

Lines 62-64: “…an evaluation is needed of how warbler use of areas with lower tree canopy cover affect population estimates and our understanding of habitat relationships” – this would be interesting, but it is not clear that this was accomplished by this study.

Lines 69-71: Since we already know warblers prefer large patches of mature (>3 m tall trees provides a proxy for stand age), closed canopy, and mixed forests, focusing on where these areas are across the breeding range would be more useful, as well as areas of low quality habitat and threats (urban development, tree mortality, conversion to non-forest, climate change, etc.).

Line 74: Because the issues with overestimation from point counts extend beyond survey duration and have not been resolved, we suggest focusing on occupancy and habitat relationships rather than questionable estimates of population size.

Line 80: This mentions soils, but we did not see soils elsewhere in the paper.

Lines 84-86: This does not address the known issues of point counts overestimating, especially in low density areas. These need to be addressed before continuing to use point counts to predict abundance, especially across the warbler's breeding range.

Lines 104-106: We note that the list of oak species does not include lacey oak. Is there a citation(s) for the species list?

Lines 108-110: It is important to note here that habitat is being lost due to climate in addition to what is being lost due to urban expansion.

Lines 115-118: There is no comparison of densities on public vs. private lands, despite a roughly 50/50 mix of the two, or consideration of the likelihood that densities would vary by land use. Was land use considered as a variable?

Lines 159-166: Sampling in this clustered way needs some kind of model correction.

Line 164-166: Why were political boundaries used rather than natural boundaries?

Lines 171-172: What were the density ranges and predictors based on?

Lines 173-176: Surveying in the low density stratum is important to confirm absence and remove bias. Since this study was not designed to assess reproductive success, Farrell et al. (2012) doesn't seem relevant here. The model was developed to predict densities, not population viability. If viability was to be measured, more intensive studies involving reproductive success should have been deployed. We also note that although only 3 of the 4 strata were surveyed, warbler detections were found at <50% of the 1804 survey points; this would have been much lower had all 4 strata been surveyed.

Lines 177-179: The underlying assumption appears to be that all low density sites are comparable. However, a low density site next to a high density site is more likely to be occupied than an isolated site. For example, low density sites in Kimble County are more likely to be unoccupied than sites within the BCP. This is why equal sampling across all strata is needed. As with Mathewson et al. (2012), the surveys in this study were biased toward high density areas, which will continue to bias the results. Given that overestimation is most pronounced at low density sites, the survey design should have included an equal number of survey points in all 4 strata.

Lines 181-183: Using this approach might make the model more precise, but not more accurate if the model is overpredicting density. To remove bias, low density sites should have received the same level of survey effort.

Line 185-187: While this approach may be more cost-effective, it is not clear how the model will produce unbiased predictions given that no density model based on point count surveys has accomplished this for the warbler.

Lines 189: If the 1,877 survey points were selected based upon habitat, how could some of the 73 that were discarded be situated in unsuitable habitat? This suggests problems with the predictive ability of the habitat model and/or recent habitat loss, and the need for ground-truthing of the habitat classifications.

Lines 190-191: It would be interesting to note areas where habitat had been significantly altered since the imagery acquisition date.

Lines 195-197: This does not seem like a lot of training, especially if the observers had no prior experience conducting warbler surveys or point counts.

Lines 214-218: Distance estimation can lead to significant biases, especially if they were predominantly over-estimates or predominantly under-estimates. Routinely underestimating distances to singing birds would inflate population estimates. Range-finder estimates are relatively accurate, but are rarely possible in dense foliage. Even in areas with low canopy cover, a reading is rarely taken on the bird, but rather on the point where the observer believes the bird is situated (which may or may not be correct).

Lines 228-229: Note that a 100-m radius (3.14 ha) is approximately the size of a warbler territory only in high quality habitat. Territories in medium and low quality habitat (smaller patches, shorter trees, and/or lower canopy cover) may be 2-8 times larger (based on data from color-banded birds on BCP and from one site on WQPL).

Lines 239-242: Since canopy height has not been incorporated in habitat models that encompass the warbler's breeding range, the results and conservation/management implications would be a better focus of this paper.

Lines 281-282: While survey duration may minimize overestimation, there are other factors that need to be addressed to avoid overestimation, particularly at low densities. This includes differences in territory sizes from low to high quality habitat.

Lines 311-313: We are wondering if this is a statistically valid approach, since it appears to be building the models a posteriori after fishing for which ones that are significant in a linear model and then combining those into a multivariate model.

Lines 338-340: If the non-public data can be used to evaluate the density model, then the data should include BCP and other sites with known densities from intensive monitoring.

Lines 345-347: Sampling locations a minimum of 100 m apart will not assure that the same bird is not sampled twice. Territories frequently span ≥500 m, especially in low-density areas (we have had observations of the same male up to 700 m apart in woodlands with young/short trees and lower canopy cover), so this criteria will not avoid observations of the same bird. Where intensive monitoring provide known densities (BCP, BCNWR, Camp Bullis, and Ft. Hood), these data should be used to validate the predicted densities, especially in low density areas.

Lines 347-350: We question the reliability of absence data from eBird. For example, many birders are not attuned to aural cues (songs, calls), are beginning birders not familiar with warblers, or may have poor hearing. Was an effort made to ensure confidence of the absence data (i.e., sites where several birders found no warblers during the appropriate season)?

Lines 352-356: Is this the only model validation that was done? How does it determine the accuracy of the density estimates? As previously stated, there needs to be a comparison of predicted densities with known densities, including data from BCP, BCNWR, Camp Bullis, and Fort Hood.

Line 363: We note here that this is <44% occupancy even though the locations were chosen specifically because they should be warbler habitat. The paper should explain why there is so much absence.

Lines 370: See comment for lines 345-347; with a sampling area of a 180-m radius, a bird perched in the same location could easily be recorded from two sampling locations only 100 m apart.

Lines 381-385: Instead of numbers of variables and terms, it would be more informative to identify what these are.

Line 395: Negative effect size indicates lower percent juniper has higher warbler density, is this correct?

Line 412-415: This method appears to show probability of detection rather than validating the predicted densities. The independently collected data should include known densities from intensive territory monitoring. This should also specify whether the mean or median was used for generating the AUC. Using the mean when densities vary will skew the number slightly higher for true positive detections.

Lines 417-422: While the classification of vegetation gradients has been improved, this model still does not address the underlying issues with overpredicting density (aside from survey duration, the paper doesn't even mention them). Thus, it is not entirely clear how this model is an improvement over previous models, although we believe it could be if the surveys had been random with an equal level of effort across all 4 strata, if the predicted densities had been validated against known densities, and if the focus was on the distribution of habitat features.

Lines 424-425: Given the importance of the Balcones Canyonlands ecoregion, discussion of the threats are warranted, specifically the urbanizing corridor from Georgetown to San Antonio. Habitat loss due to misinformation about Ashe juniper is also a major threat.

Lines 425-427: Without validation against known densities, this statement is not supported.

Lines 428-429: We would like to know what the 2-min and 4-min estimates were.

Lines 430-431: Given the consistent overestimation of warbler density models compared to known densities, it is not surprising that this new model is comparable to the predictions in Mathewson et al. (2012).

Lines 432-434: This is an unsupported assumption, especially given the known declining trends and ongoing habitat loss. Averaging the density of 10 intensive monitoring plots on the BCP from 2011-2019, we observed an overall decline of ~30%, with an apparent drop after 2015. Since we know the density models overestimate at low densities, we question whether the models would be able to detect these declines. Instead of focusing on questionable abundance estimates, information related to the distribution of prime warbler habitat would be more useful for conservation and management.

Lines 458-461: The discussion of compound topographic index is interesting and an example of new information that would be a better focus of this paper.

Lines 475-477: How much of this is public land?

Lines 479-482: The low-density habitat accounting for 22% of warbler abundance across the range is presented as being equivalent in value to the species to areas of higher density – with no consideration that these sites may be sinks with little to no reproduction.

Lines 480-482: We suggest changing this sentence to: "This suggests that the [remove "large"] expanse of low density areas across the breeding range may harbor a large [change from "an important"] fraction of the population", since "important" suggests conservation value, which is unknown without data on reproductive success/viability.

Lines 485-486: Add elimination due to urban expansion.

Line 490-497: Recommending management strategies seems to be beyond the scope of the paper. Given the focus on the breeding range, it would be helpful to prioritize areas needed to protect prime breeding habitat from anthropogenic losses. For example, urban expansion is now encroaching on BCNWR, which we understand is a little over 50% of its acquisition goal. So, at the very least, completing acquisition to minimize the edge to area ratio is a priority for that and other focal areas. For management, it may be sufficient to state the goals (i.e., high tree canopy cover, 60-80% Ashe juniper, >3 m trees) rather than try to identify specific methods to achieve those goals.

Line 490: As stated above, Duarte et al. (2016) has found that habitat loss tends to be unidirectional. Seventy-five years to grow from grassland to dense-canopy woodlands may be possible for some sites that are near high quality habitat and have old-growth trees remaining (e.g., https://www.nature.com/articles/s41477-021-01088-5), but could take longer, since the diameter growth rate of Ashe juniper is very slow (~0.5-1.0 inch/decade; see p. 31 in https://www.traviscountytx.gov/images/tnr/Docs/bccp/2018/appendix-s9.pdf).

Lines 490-498: No mention is made, when considering threats to remaining habitat, of anthropogenic destruction of habitat, the primary source of loss.

Lines 492-497: Results of the Andruk et al. and Reidy et al. studies are misinterpreted with a biased view that fires are beneficial for warbler habitat. The need for shade tends to be a more important factor than the need for light for oak regeneration (from seed) in Central Texas, along with nutrient-rich soils (e.g., https://esajournals.onlinelibrary.wiley.com/doi/full/10.1002/ecs2.3017).

Lines 499-502: This appears to be an effort to avoid validating the model using known densities. If this model is unable to predict local population sizes, why are territory-level and site-level models included?

Line 501: Precise predictions of local population sizes are not useful if they repeatedly overestimate. Accurate predictions are needed.

Lines 502-506: The methods state comparison of predicted densities with “non-public” data collected at the local scale, so comparison with known densities should be appropriate. If the accuracy of the predicted densities can’t be reliably determined, then the paper should focus on habitat metrics only.

Lines 507-520: Our understanding is that Peak had a much larger sample size of known densities from intensive territory monitoring to use for validation, so this is a better reference than the limited sample size in Sesnie et al. (2016).

Lines 524-526: As previously stated, a false equivalence is given to warbler habitat that is suboptimal. Relying upon these patches for the survival of the species is a poor decision, because they are ultimately dependent upon source populations for their continued existence. We also note that climate change could quickly eliminate subprime habitat (and convert prime to subprime).

Lines 526-528: At the scale of the map, it’s difficult to identify where improvement and restoration of habitat may have the greatest benefits, can this be more specific?

Lines 528-530: Add that point counts tend to overestimate, especially at low densities, and so should be validated against known densities.

Table 7: Recovery Region boundaries have been ‘shifted to nearby county boundaries to facilitate summary of county-scale data’, as per Groce et al. 2010. We question whether this shift extends boundaries of Regions containing low abundance to capture warblers from regions of higher abundance, thereby bolstering the minimum number of warblers per region required for recovery.

Figure 4: Using the rasters of predicted warbler density derived from the 3-minute surveys, we zoomed in on an area outside of the BCP that we are familiar with from our collections of warbler DNA. The model predicts abundance at Kickapoo Cavern State Park to be 355.7 males. However, we struggled to find birds at this site and covered all the viable warbler habitat over three days in 2021. We banded eight males, observed three other males that we were unable to capture, and resighted a male that we banded in 2020. Based on our experience, we would estimate 20 to 50 males, with 50 being a very generous overestimation.

New predicted layer (zoomed in to a portion of western Travis County): We noticed that virtually all of the BCP is shown as red (>0.25 males/ha) or orange (>0.125 – 0.25 males/ha), but some of this is low density or unoccupied due to habitat conditions and/or urban development. As an example of habitat conditions, the Cortaña tract has large areas of non-old growth that are unoccupied, with a few territories in wooded canyons. The non-old-growth areas are shaded orange almost uniformly. This suggests the model overestimates in this type of habitat, at least in the eastern part of the range. There also needs to be a consideration when predicting occupancy or density that habitat occupied in one part of the range may not be occupied in other parts of the range. For the urban influence, it appears that a lot of the medium and high density sites were manually removed/clipped from these areas. However, some of the red polygons did not change and still appear to be overestimating. The model seems to predict some of the high density habitat correctly but also assigns high density to small patches that are completely surrounded by development. Will you be setting new model parameters for urban density or edge to address these issues?