# ASSESSING GOLDEN-CHEEKED WARBLER PRODUCTIVITY WITH A REPRODUCTIVE INDEX: IS IT GOOD ENOUGH?

FINAL REPORT SUBMITTED TO

U.S. FISH AND WILDLIFE SERVICE

ON NOVEMBER 3, 2014

Prepared by:

JENNIFER L. REIDY, Research Specialist
University of Missouri, 302 Natural Resources Bldg.,
Columbia, MO 65211

LISA O'DONNELL, Senior Biologist
City of Austin, 3635 RR 620 South,
Austin, TX 78738

FRANK R. THOMPSON III, Research Biologist
U.S. Forest Service,
202 Natural Resources Bldg.,
Columbia, MO 65211

# EXECUTIVE SUMMARY

It is critically important for managers and researchers to determine and understand relationships between endangered species populations and landscape and habitat features to effectively manage and conserve populations and the habitats they rely on. The golden-cheeked warbler (*Setophaga chrysoparia*), an endangered songbird, breeds exclusively in the juniper-oak woodlands of central Texas. Despite being listed as endangered since 1990 and being the focus of numerous studies, there are still gaps in our knowledge of breeding ecology, including reliable estimates of breeding success and demographic rates across the breeding range (Groce et al. 2010). Several recent studies focused on linking productivity to various factors of interest by using a reproductive index, but thus far, no field test has been conducted to assess how well a reproductive index serves as a surrogate for more rigorous measures of productivity, such as nest success and seasonal productivity (fledglings/territory). We compared the performance of a reproductive index (RI) to results from intensive territory mapping and monitoring of color-banded adults. Data for each monitoring approach were collected independently by different observers. We were only able to match territories delineated by the RI with 29-65% of territories delineated by intensive monitoring on a plot, and there was no significant correlation between the reproductive rankings from the two approaches for the territories we could match. At the plot level, the RI method tended to underestimate the total number of territorial males present on plots, but there was a significant weak correlation between the numbers of territories delineated by each approach. We found no significant relationship for pairing success between the RI and intensive monitoring methods, nest survival from intensive monitoring and breeding success from the RI method, or breeding success from intensive monitoring and RI methods when controlling for observer. We conclude that the RI approach did not produce reliable estimates of

reproductive performance at either the territory or plot level. Observers varied greatly in their delineation of territories and reproductive success rankings. We acknowledge that not all investigations can implement intensive monitoring of color-banded individuals, but when absolute estimates of reproductive success rates or other demographic measures are required, such as in species viability or threat assessments, we recommend intensive monitoring of color-banded individuals.

## INTRODUCTION

Knowledge of the factors limiting reproductive success is critical for conservation and management of wildlife populations, particularly endangered species (Watson et al. 2006, Anich et al. 2013). Nest survival is an important component of reproductive success and can provide important insight into how habitat and landscape factors affect productivity of songbirds (Flaspohler et al. 2001, Ribic et al. 2012, Peak and Thompson 2014). Meaningful nest survival estimates depend on a large and representative sample of nests, and finding and monitoring nests can be logistically unfeasible, particularly for rare or cryptically-nesting species (DeSante and Geupel 1987, Vickery et al. 1992, Groce et al. 2010). In lieu of monitoring nests, Vickery et al. (1992) proposed the use of a reproductive index to determine reproductive success and output. The Vickery method is considered a reproductive index (hereafter "RI") because it relies on indirect behavioral clues from the adults to establish reproductive success by ranking reproductive activity (e.g., evidence of feeding young receives a higher rank than evidence of nest building or associating with a female). It evolved as a method to estimate grassland songbird breeding success (portion of sampled population producing at least one host young) without having to locate and monitor, and potentially disturb, cryptic nests for rare species and was originally intended as a tool for wildlife managers when obtaining reliable and unbiased nest success data was not feasible. However, there are several potential problems with using a RI as a surrogate for intensively monitoring individual pairs. Observers may not be able to accurately identify territorial males (Bell et al. 1973, Best 1975, Verner and Milne 1990, Rivers et al. 2003, Morgan et al. 2010), which affects calculation of territory density and can easily lead to interpretation errors when observers are assigning ranks to territories. Based on limited observations, observers may fail to interpret nesting behaviors correctly (Best 1975) or may

assign incorrect rankings (Morgan et al. 2010), regardless of experience level (Rivers et al. 2003). While the use of a RI is appealing because it requires much less effort than traditional nest monitoring, Vickery et al. (1992) cautioned it should not be relied upon until field testing was conducted demonstrating its usefulness at assessing reproductive output for the focal species.

Despite this, several studies have used the RI to determine reproductive success of grassland (Powell and Collier 1998, DeLisle and Savidge 1996, Nocera et al. 2007), forest (Rangen et al. 2000), and shrubland (Hannah et al. 2008) birds without testing its efficacy. Few studies have tested the accuracy of the RI for predicting reproductive success on songbird species, but those that have reported the RI did a poor job of correctly predicting nest fate or breeding success, in part because observers were not able to determine the correct number of territories (Rivers et al. 2003, Althoff et al. 2009, Morgan et al. 2010). Rivers et al. (2003) concluded that the RI did not correlate well with reproductive success for their focal system and recommended conducting pilot trials to test the correlation between nest fate and final index ranking for the target species before relying on it. Althoff et al. (2009) found low concordance between reproductive success and the RI at the plot level, but concluded that depending on the study objectives, use of a RI may be a time-efficient endeavor, particularly for species with cryptic nests that are easily disturbed through nest-searching endeavors such as grassland birds. Morgan et al. (2010) concluded the RI was inaccurate enough at the territory and plot level as to not be useful in estimating reproductive success of their focal species. Christoferson and Morrison (2001) reported higher concordance (80-92%) between final rankings from a RI and from nest monitoring for three species with different nesting strata; however, they did not conduct an independent test of the method but rather used a combination of the RI and nest

monitoring to arrive at reproductive success. A RI may be a useful method to assess management impacts or habitat quality in a relative sense, even if it is inaccurate as an absolute measure, if there is a consistent or predictable relationship between RI and actual reproductive performance. Managers setting conservation and management guidelines and goals for species of conservation concern need to understand factors limiting populations, including reproductive success, and must therefore know how well a RI performs against more rigorous measures of reproductive success.

The golden-cheeked warbler (*Setophaga chrysoparia*; GCWA) is a federally-listed endangered songbird whose breeding range lies entirely within the Ashe juniper (*Juniperus ashei*)-oak (*Quercus* spp.) woodlands of central Texas (Pulich 1976, USFWS 1992). Several recent short-term studies have used a modified RI based on Vickery et al. (1992) to address specific research needs outlined in the GCWA Recovery Plan (U.S. Fish and Wildlife Service 1992), including: identifying a minimum patch-size threshold for occupancy and pairing and territory success in rural (Butcher et al. 2010) and urban areas (Robinson 2013); evaluating the impact of road and construction noise on GCWA productivity (Lackey et al. 2011); examining effects of tree species composition and foraging effort on GCWA productivity (Marshall et al. 2013); and determining the effects of canopy closure and tree species composition on GCWA productivity at the fringe of their breeding range (Klassen et al. 2012). Two additional publications have combined range-wide data collected from other studies to relate breeding success to landscape patch size and composition (Campomizzi et al. 2012) and tree structure (Long 2014). However, no field test evaluating the reliability of the RI at predicting reproductive success of GCWA, or other forest songbirds, has been conducted, so it is difficult to interpret the conclusions.

Despite strong interest in conserving and managing this species, only one long-term monitoring program, at Fort Hood, Texas, has monitored nest success and seasonal productivity of color-banded birds to determine breeding success and productivity (Peak 2007, Peak and Thompson 2014). The monitoring program at the Balcones Canyonlands Preserve (BCP), Texas, estimated pairing and breeding success on the basis of cumulative observations of unbanded males across the breeding season at several plots between 1998 and 2008 (City of Austin 2008); however, in 2009, they began color-banding adults on intensive monitoring plots to quantify territory density, nest success, and productivity (City of Austin 2009, 2011).

Better knowledge of the reliability of the RI as an indicator of reproductive success is needed to inform its future use, especially for endangered species such as the GCWA. Intensive territory mapping and nest monitoring of color-banded birds is generally considered the most reliable approach to estimating territory density and seasonal productivity. Therefore, we evaluated the use of the RI approach for GCWA by comparing estimates from the RI to estimates from intensive monitoring of color-banded birds. Our objectives were to determine correspondence in estimates from these two approaches of reproductive success at the territory level and territory density and reproductive success at the plot level. In addition, we evaluated the effect of observers on the RI estimates. We implemented each approach with independent observers across multiple plots and years.

## STUDY AREA

We conducted this study on seven 40-ha plots within the BCP. The BCP is a 12,300 ha network of preserves in western Travis County, Texas, that was established in 1996 under a federal permit to mitigate habitat loss for the GCWA and other endangered and rare species (Fig. 1; USFS 1996). Our study plots were located on City of Austin and Travis County properties

and were predominately mature closed-canopy woodland comprised of Ashe juniper, plateau live oak (*Quercus fusiformis*), shin oak (*Quercus sinuata* var. *breviloba*), Texas red oak (*Quercus buckleyi*), and cedar elm (*Ulmus crassifolia*). We surveyed two plots (CV, VI) in 2013 and five additional plots (BC, FR, HA, JJT, RR) in 2014. Our seven plots spanned the gradient of density observed on 18 monitoring plots spread across the BCP (City of Austin et al. 2013).

## METHODS

### Study species

Golden-cheeked warblers nest exclusively in the Ashe juniper-oak woodlands of central Texas, where they are locally common and abundant (Pulich 1976, USFWS 1992). They are socially monogamous, although polygyny has been documented in rare instances (Peak et al. 2010, J. Reidy, personal observation). They are typically single-brooded, but a portion of pairs that successfully fledge young early in the nesting season will attempt a second brood (Peak and Thompson 2014, J. Reidy and City of Austin, unpublished data). Pairs may attempt up to five nesting attempts if a nest fails (Peak and Thompson 2014), although only up to 4 have been observed on BCP (J. Reidy and City of Austin, unpublished data). Females build nests in the canopy of dense woodlands, often in Ashe junipers (Reidy and Thompson 2012). Only females incubate eggs, but both adults feed nestlings and fledglings. Predation is the main cause of nest failure and failure due to parasitism by brown-headed cowbirds (*Molothrus ater*) is not common (Reidy et al. 2008).

### Intensive monitoring

*Territory mapping*.— We banded adult GCWA from mid-March through early May 2009-2014 to establish number of territorial males, delineate territorial boundaries, and determine pairing and breeding success (City of Austin 2009). We played pre-recorded GCWA

songs beneath the middle of 6-m long mist-nets to capture adults and banded males and females with a unique combination of 2-3 colored leg bands and a U.S. Geological Survey numbered aluminum band. We already had a large number of banded adults on plots when we initiated this study at the beginning of the 2013 season, which contributed to the accuracy of our intensive monitoring.

Two experienced biologists surveyed plots with > 5 territorial males 2-3 times per week from 15 March to 15 June 2013-2014. Plots with ≤ 5 territorial males were surveyed by one biologist who visited the plot 2-3 times per week. Once a week, biologists surveyed the entire plot and attempted to resight all color-banded adults; recorded locations of adults in universal transverse mercator (UTM) coordinates; documented presence of females, nesting behaviors/nests, and fledglings; and focused on mapping territories. Each week, they rotated the starting point in a clockwise direction. Biologists returned to each plot 1-2 more times each week for targeted productivity surveys where effort was focused on searching for territorial birds who had not been previously detected that week, locate nests and record nesting behaviors, and monitor previously found nests. They attempted to collect ≥ 5 unique locations (> 30 m apart) during each visit with a goal of accumulating ≥ 33 locations across the breeding season to delineate territories season (Davis et al. 2010).

*Nest monitoring and fledgling surveys*.—Biologists searched for nests when they observed nesting behavior such as an adult carrying nesting material or food and used a combination of behavioral clues from adults and systematic searching to locate nests. If they failed to locate the nest during that visit, they attempted to locate it on subsequent visits and continued to visit each territory 2-3 times per week until the nest or fledglings were found. They monitored active nests every 3-5 days until nests fledged or failed, and more frequently near

predicted hatch and fledge dates. They placed unmarked flags ~5-10 m from nests and recorded directions to nests in their notebooks. Other random flags were distributed throughout the plot so only plot biologists knew nest locations. Biologists spent proportionately more time later in the season in territories without an active nest or fledglings. Biologists sharing a plot exchanged information on territory status and shared nest monitoring duties. Once a nest was determined to have fledged or adults were found attending fledglings, they spent less time in that territory; however, they continued to monitor all territories for double-brooding and additional fledglings at least once per week until the end of the season. They attempted to document all fledglings for all territories, regardless of whether a nest had been located for that pair, until mid-June. If no fledglings were confirmed by the end of the season and adults were still present, they considered that pair to be unsuccessful. At the end of the season, biologists assigned a reproductive rank to each territory according to the highest rank observed during the breeding season (described under "Ranking").

*Territory delineation*.— We created minimum convex polygons (MCP) in ArcMap 10.0 (ESRI, Redlands, California) by bounding observations assigned to unique males for cumulative observations contributed by biologists across the season. We included only observations of males who achieved an overall rank of 1 (males detected for ≥ 4 weeks) and excluded non-territorial observations (e.g., a silent male following a neighbor's female, adults feeding fledglings) and outlier observations (single observations > 150 m from other locations). We considered a territory as "full" if the male was detected outside the plot boundaries on no more than one survey and as "edge" if the male was detected within and outside the plot boundary on multiple surveys (City of Austin et al. 2013). All territories had ≥ 10 observations.

**Reproductive index monitoring**

*Territory mapping*.—In 2013, one observer surveyed CV and two observers surveyed VI weekly from 15 March to 10 June. In 2014, two independent observers surveyed BC, CV, and VI and each surveyed two additional plots (FR, JJT, HA, RR) once per week from 15 March 15 to 10 June. For plots with multiple observers, each observer visited the plot on a different day from each other and the intensive monitoring biologist, except for BC, which was surveyed on the same day with observers starting on opposite sides of the plot. One of the three observers in 2013 and both observers in 2014 had prior experience surveying for GCWA. We considered observers with prior experience conducting GCWA surveys as "experienced" and those without prior experience as "inexperienced".

Index observers initially established the approximate numbers and locations of GCWA by conducting two surveys along parallel transects spaced 100 m apart once a week from 15 March 15 to 29 March beginning at sunrise and finishing 4-6 hr post-sunrise (2 surveys per plot). Males were first detected in the Austin area 7 March 2013 and 7 March 2014 by BCP biologists, so this period coincided with territory establishment and pair-bonding. Observers recorded time, sex, detection type (song, chip, visual), location in UTM coordinates and the distance and bearing to each detected GCWA on a transect survey form, and mapped out the general area of each detection for future reference. Index observers began recording individual location data to define territories and collecting productivity data after the initial two transect surveys. They returned to areas where they had previously detected GCWA and also recorded locations for any newly identified territories. Observers spent up to 30 min searching for a previously identified male and up to 30 min following him once he was detected and attempted to record $\geq 3$ locations at least 15 m apart and any reproductive behaviors indicative of breeding stage (presence of a female, evidence of nesting or fledging). They recorded the highest rank (described below)

observed for each observation period and assigned a final rank at the end of the season based on the cumulative observations. They conducted 10 territory/productivity surveys and collected ≥ 10 locations for each territorial male during the breeding season. Index observers rotated their starting point each week in a clockwise direction and mapped territories by following males and noting changes in vocal behaviors, contemporaneous singing, and using past locations (Bibby et al. 2000). They did not attempt to determine whether GCWA were color-banded, and if they did detect bands, they made no record of it for future reference. Index observers were allowed to share information (such as nest locations or fledgling information) with intensive monitoring biologists, but were not allowed to share information with each other or to receive information from biologists. They acted as independent observers and did not have access to information regarding banding status or nest locations. Additionally, index observers and intensive monitoring biologists did not survey the same plot on the same day, except on rare occasions when weather forced surveyors to switch survey days. During those days, index observers and intensive monitoring biologists coordinated prior to the survey to ensure that they surveyed the plot in different directions to avoid interfering with each other.

*Territory delineation*.— We created MCP in ArcMap 10.0 by bounding observations assigned to unique males for each index observer. We included only observations of males who achieved an overall rank of 1 (males detected for ≥ 4 weeks) and excluded non-territorial observations (e.g., a silent male following a neighbor's female, adults feeding fledglings) and outlier observations (single observations > 150 m from other locations). We created MCP for territories with ≥ 10 observations. To compare rankings between territories, we overlaid MCP produced for index monitoring over those delineated from intensive monitoring. We attempted to match territories delineated by the RI method to territories delineated by intensive monitoring.

We paired a RI territory to an intensive monitoring territory if > 50% of its area fell within the intensive monitoring territory. If multiple RI territories met this criterion for a given intensive monitoring territory, we only paired the territory with the largest area in the intensive monitoring territory. We also evaluated clustering of points for both methods to aid in assigning matches.

**Ranking**

Observers assigned a rank of 1 to 5 to each territory following a modified Vickery RI (Campomizzi et al. 2012): 1 = male present ≥ 4 weeks, 2 = pair present for ≥ 4 weeks, 3 = nest material carry observed, 4 = food carry observed, 5 = host fledgling verified, and assigned a rank of 0 to a territory if the male was not detected during the weekly survey. We compiled rankings from index monitoring for each territory for each week and a final ranking based on the highest observed rank. We also assigned a final rank to territories from intensive monitoring based on the highest observed rank during the season. We considered a male territorial and unpaired if he was detected over a period exceeding 4 weeks but we did not verify evidence of a female or nesting activity (rank = 1). We considered a male territorial and paired if we detected him associating with a female or documented the presence of a female in his territory (rank = 2). If we observed evidence of nest-building, such as a female collecting or flying with nesting material, we ranked the territory as a 3. If either adult was detected collecting or carrying food, we assigned a rank of 4. We considered a territory successful, and ranked a 5, if we confirmed the adults feeding host fledglings or had other indirect evidence of fledging such as multiple food items carried in different directions or we heard begging chips.

**Data analysis**

Each observer interpreted their own data to derive rankings. We report summary statistics regarding the average survey and total survey effort, territory abundance, pairing

success, and breeding success from index and intensive monitoring. We report the number of territories delineated by each method and the number we considered a match. We calculated the rank correlation coefficients for matched territories and summarized the difference in rankings for matched territories.

We calculated plot-level territory density using Verner's method wherein each full territory was counted as one territory and each edge territory was counted as 0.5 (Verner 1985). We calculated plot-level pairing success (rank $\geq 2$) and breeding success (rank = 5) as the proportion of territories reaching these ranks. We calculated daily nest survival (DSR) and period nest survival (PSR) based on a 25-day nest cycle using the logistic exposure method (Shaffer 2004) and a model that included year, day of year, and plot.

We compared plot-level measures of territory density, pairing success, nest survival, and breeding success with a repeated measures generalized linear model (MIXED procedure, SAS 9.3, SAS Institute, Cary, North Carolina). We treated the RI-based estimated as the response variable and evaluated its relationship with the intensive monitoring estimate; we controlled for observer if the addition of observer as a fixed effect to the model was supported by a lower $AIC_c$ value for the model with observer. We used a repeated measures model with plot × year as the subject because plots were visited by multiple observers.

## RESULTS

### Intensive monitoring

We spent 730 and 1742 hours conducting intensive monitoring in 2013 and 2014, respectively (Table 1). Biologists ranked 33 (79% of territories with at least one banded adult) and 98 (72% of territories with at least one banded adult) territories as $\geq 1$ in 2013 and 2014 (Table 2), respectively. For both years combined, average territory density was 0.27 ± 0.10

(range: 0.10 to 0.40; Table 2). Biologists ranked 91 of 131 (69%) territories as a 5 (host young detected); they also monitored an additional 14 males that they ranked as 0 (male detected during breeding season but did not qualify as territorial). Plot-level pairing success averaged 96 ± 5% (range: 89-100%) and plot-level breeding success averaged 69 ± 11% (range: 53-89%; Table 2).

We monitored 102 nests (32 and 70 in 2013 and 2014, respectively) representing 90 territories (25 and 65 in 2013 and 2014, respectively), of which 65 (16 and 49 in 2013 and 2014, respectively) were considered successful (i.e., fledged at least one host young). Apparent nest success (64%; 65/102) was slightly lower than the observed breeding success (69%). We documented 3.3 ± 1.0 host young per successful nest (N = 65), and found 3.1 ± 1.1 fledglings per successful territory from the first brood for which successful nests were not monitored (N = 26 ). Average number of young per all nests monitored was 2.1 ± 1.8 (N = 102 nests) and for all territories was 2.2 ± 1.7 (excluding successful second broods; N = 131 territories). Average DSR was 0.98 (95% CI: 0.96-0.98) and PSR for the 25-day nest cycle was 0.54 (95% CI: 0.38-0.68) and did not vary by plot.

**Index monitoring**

We spent 29 and 107 total hours conducting transect surveys and 154 and 582 total hours conducting productivity surveys (Table 1) in 2013 and 2014, respectively. Average survey time (hours per survey) did not differ between intensive and index monitoring, but total survey time (total hours per plot) did (Table 1). Average territory density per plot across all plots and observers was 0.25 ± 0.4 and ranged from 0.09 to 0.43 (Table 3). Index observers ranked 77 of 151 territories as a 5. Overall pairing success was 81 ± 40% and breeding success was 51 ± 50% based on all index monitoring (Table 3). Experienced observers documented lower territory density (GLM, F = 4.96, df = 1, $p = 0.05$), but higher pairing (GLM, F = 37.11, df = 1, $p <$

0.0001) and breeding success (GLM, F = 28.26, df = 1, $p < 0.0001$) than inexperienced observers.

**Territory-level comparison**

We matched 84 territories delineated by the RI method with territories delineated by intensive monitoring, which represented an average of 45% (range: 21-80%) of territories delineated by intensive monitoring across all index observers and plots (Fig. 2). Ranks for 40 territories coincided (48%), 11 were overestimated (13%) and 33 were underestimated (39%; Fig. 3). Across all 84 territories we found a weak but non-significant correlation in reproductive ranks from the two methods ($r = 0.156$, $P = 0.157$). We identified 7 territories assigned a rank of 5 before intensive monitoring indicated the territories had fledged young.

**Plot-level comparison**

The distributions of reproductive success ranks from index and intensive monitoring were more similar in 2014 than 2013 when summarized by plot (Fig. 4); however, the number of territories reaching a rank of 5 was underestimated by the RI method in both years. We found strong support for observer effects in models comparing territory density, pairing success, nest survival, and breeding success, so we evaluated these relationships while controlling for observer. We found a significant linear relationship between territory density derived from the the RI and intensive monitoring methods, but density estimates from the RI method were generally lower than those from intensive monitoring (Table 4, Fig. 5). We found no significant relationship for pairing success between the RI and intensive monitoring methods, nest survival from intensive monitoring and breeding success from the RI method, or breeding success from intensive monitoring and RI methods when controlling for observer (Table 4; Figs. 6, 7).

## DISCUSSION

We found a general lack of consensus between intensive and index monitoring for measures of territory abundance, paring success, and breeding success. We have high confidence in the results from intensive monitoring because of biologists' substantial prior experience with the species coupled with the strong emphasis and time spent to locate females, nests, and fledglings. We found nests for ~69% of GCWA territories monitored, only slightly less than observed breeding success. Based on the observed period nest success of 54% and 65 successful nests, we would expect to have found ~120 active nests (65/0.54) and we monitored 102 active nests, suggesting we monitored a representative sample of nests. Additionally, we found high consensus between the total number of fledglings detected from successful and total nests and successful and total territories; this was just slightly less than that observed from video-observation of successful nests (Reidy et al. 2008) and similar to seasonal productivity reported for GCWA at Fort Hood (Peak and Thompson 2014). Hence, we believe that our intensive monitoring effort provided a reasonable estimate of reproductive success. We rarely documented behavior of adults indicative of disturbance (e.g., alarm calls, distraction displays) while nest monitoring and believe our protocol for nest searching and monitoring did not present a disturbance to this species. The RI was originally developed for grassland birds, which are easily disturbed by nest searching and monitoring methods.

Index observers tended to underestimate the total number of territorial males (average: 80%, range: 58-125%), although experienced observers were often closer than inexperienced observers in determining territory abundance and locating females and fledglings. Additionally, there was disparity in total abundance and density, pairing success, and breeding success between experienced observers surveying the same plots, with the greatest disparity on the plot

with highest abundance. Index observers, regardless of experience, had limited opportunities compared to intensive monitoring biologists to map territories and locate females and fledglings, and without the ability to uniquely identify each bird, they were unable to reliably discern territories or assign females and fledglings. Intensive monitoring revealed more movement than index monitoring was able to detect. We documented within-season dispersal and movement of color-banded adults that caused some males to ultimately be ranked as zero, and some of these males were even confirmed with females or active nests before disappearing from the plot before the 4-week mark required to be ranked. We also documented extra-territorial locations where males were skulking (silent or softly chipping) in close proximity to fertile females or family groups up to several hundred meters from their territory. Males often shifted territorial boundaries from the beginning of the nesting season to the end, sometimes so completely that other neighboring or late-arriving males took over portions of the male's territory. All plots had some, and often high, degree of territory overlap and some areas not occupied by territorial males. Often males were silent even when other males were singing nearby or within their territories. We were only able to detect these types of movements and interactions among individuals, and hence determine accurate territory locations and reproductive success, by observations of color-banded birds.

Index observers conducted 10 weekly productivity surveys to each plot after the two initial transect surveys, similar to the design the City of Austin employed prior to 2009 and similar to other studies on GCWA (Campomizzi et al. 2012, Long 2014). Additional survey effort may have resulted in more concordance between territory abundance from intensive and index monitoring, particularly for plots with high abundance. However, we believe that it would still be difficult to determine the true number of territorial males even with considerably more

survey effort without uniquely marking birds. This is partly because of the high degree of territory overlap and movement established by intensive monitoring and because males often did not countersing. We saw a similar pattern in City of Austin's territory mapping before and after color banding was initiated in 2009. Territories overlapped very little prior to banding (City of Austin 2007), whereas there is substantial overlap documented after banding (City of Austin et al. 2013). Territory density also fluctuated greatly between observers prior to banding (City of Austin 2007).

The RI method failed to produce results concordant with intensive monitoring at the territory level. We were only able to match territories for 29-65% of territories delineated by index observers, and final rankings from the two methods were not significantly correlated for the territories we were able to match. We also had evidence that observers assigned fledglings to the wrong territory because the RI method assigned a rank of 5 to some territories before intensive monitoring indicated they fledged young. These territories ultimately did fledge young, making the final rankings seemingly align. By only considering intensive monitoring territories that matched with RI-derived territories, our approach should have produced the most favorable comparison possible for the RI method because intensive monitoring territories for which RI observers found no corresponding territory were not included as discordant. Other studies comparing these methods at the territory-level also found that the RI rankings were not a reliable indicator of territory-level reproductive success. Morgan et al. (2010) was much more successful at matching territories than we were, but found rankings were only weakly correlated to actual breeding success. Rivers et al. (2003) opted to not match territories because the number of territorial males estimated from each method were so different.

We did not find concordance between rankings from index and intensive monitoring at the plot level either. We found a stronger correlation between index observer and reproductive measures from intensive monitoring than between the RI and reproductive measures, suggesting strong observer effects on the RI. Given the difficulty of assigning females and fledglings to a color-banded population with the majority of nests found through intensive monitoring, the lack of concordance between index observers surveying the same plot, and the territories prematurely assigned a rank of 5, we suspect there is more error in the assignment of females and fledglings to territories by index observers than we were able to document. We would likely have documented even more observer disparity had we asked observers to interpret weekly surveys from other observers rather than just their own (see Verner and Mile 1990). Other studies have also concluded observers provide inconsistent results regarding the number of territories and final reproductive rankings (Verner and Milne 1990, Rivers et al. 2003, Morgan et al. 2010). The influence of observers is critical to interpreting the conclusions from studies utilizing a RI because the majority of studies on this species and other forest songbirds are of short duration, typically 2-3 years, and involve many observers, many of whom may lack intimate experience with this species needed to correctly assign breeding activity.

## MANAGEMENT IMPLICATIONS

Color banding adults and finding and monitoring nests are time- and skill- intensive endeavors, but we believe they provide the most accurate measures of reproductive output. We did not find support that the RI produced reliable results of actual breeding performance at either the territory or plot level. Index monitoring tended to underestimate breeding success at the territory level more than overestimate. However, we did find evidence of further overestimation at the weekly survey level for territories that were ultimately successful. Relationships between

pairing and breeding success from a RI and landscape or habitat attributes at the territory level assume correct assignment of females and fledglings to territories. If observers fail to detect females or fledglings, or incorrectly assign them (e.g., assign fledglings from the same territory to multiple territories), pairing and breeding success can be under- or over-estimated and we found that this error is biased by observer experience and interpretation. At the plot level, index observers were able to differentiate plots with low and high abundance, but pairing and breeding success were more correlated to observer than to actual success rates based on intensive monitoring. Since the RI was developed to document breeding success and not relative density, we consider the RI an inadequate method to determine breeding success for GCWA. Based on our results, we agree with Morgan et al. (2010) and Rivers et al. (2003) that the RI is a poor substitute for intensive monitoring to determine productivity. While our study focused on a single species, we believe our results are applicable to other forest songbirds, particularly those with large and overlapping territories such as GCWA. We acknowledge that not all investigations can implement intensive monitoring of color-banded individuals, but when absolute estimates of reproductive success rates or other demographic measures are required, such as in species viability or threat assessments, we recommend intensive monitoring of color-banded individuals.

## ACKNOWLEDGMENTS

We thank the numerous field biologists and volunteers who assisted with data collection. The U.S. Fish and Wildlife Service provided funding for the study. Additional funding, in-kind, or logistical support were provided by the City of Austin, Travis County, the USDA Forest Service Northern Research Station, and the University of Missouri.

## LITERATURE CITED

Althoff, D., P. Gipson, J. Pontius, and R. Japuntich. 2009. Evaluation of a reproductive index to estimate grasshopper sparrow and Eastern meadowlark reproductive success. Wildlife Biology in Practice 1:33–44.

Anich, N. M., M. Worland, and K. J. Martin. 2013. Nest-site selection, nest survival, productivity, and survival of spruce grouse in Wisconsin. Wilson Journal of Ornithology 125:570-582.

Bell, B. D., C. K. Catchpole, K. J. Corbett, and R. J. Hornby. 1973. The relationship between census results and breeding populations of some marshland passerines. Bird Study 20:127-140.

Best, L. B. 1975. Interpretational errors in the "mapping method" as a census technique. Auk 92:452-460.

Bibby, C. J., N. D. Burgess, D. A. Hill, S. Mustoe, and S. Lambton. 2000. Bird Census Techniques, 2nd edition. Academic Press, San Diego, California, USA.

Butcher, J. A., M. L. Morrison, D. Ransom, Jr., R. D. Slack, and R. N. Wilkins. 2010. Evidence of a minimum patch size threshold of reproductive success in an endangered songbird. Journal of Wildlife Management 74:133–139.

Campomizzi, A. J., S. L. Farrell, T. M. McFarland, H. A. Mathewson, M. L. Morrison, and R. N. Wilkins. 2012. Species conservation at a broad spatial scale: reproductive success of golden-cheeked warblers across their breeding range. Wildlife Society Bulletin 36: 440-449.

Christoferson, L. L., and M. L. Morrison. 2001. Integrating methods to determine breeding and nesting status of three western songbirds. Wildlife Society Bulletin 29:688–696.

City of Austin. 2007. Golden-cheeked warbler (*Dendroica chrysoparia*) and black-capped

vireo (*Vireo atricapilla*) monitoring program. Balcones Canyonlands Preserve Annual Report Fiscal Year 2006-07. Austin, Texas, USA.

City of Austin. 2009. Golden-cheeked warbler *(Dendroica chrysoparia)* and black-capped vireo *(Vireo atricapilla)* monitoring program: Balcones Canyonlands Preserve annual report FY 2008-09. City of Austin, Austin Water Utility, Wildland Conservation Division, Balcones Canyonlands Preserve, Austin, Texas, USA.

City of Austin. 2011. Golden-cheeked warbler *(Dendroica chrysoparia)* and black-capped vireo *(Vireo atricapilla)* monitoring program: Balcones Canyonlands Preserve annual report FY 2010-11. City of Austin, Austin Water Utility, Wildland Conservation Division, Balcones Canyonlands Preserve, Austin, Texas, USA.

City of Austin, Travis County, and U.S. Forest Service. 2013. 2013 annual report: Golden-cheeked warbler (*Setophaga chrysoparia*) monitoring program. Balcones Canyonlands Preserve. City of Austin, Austin Water Utility, Wildland Conservation Division, Balcones Canyonlands Preserve, Austin, Texas, USA.

Davis, C. A., D. M. Leslie, Jr., W. D. Walter, and A. E. Graber. 2010. Mountain biking trail use affects reproductive success of nesting golden-cheeked warblers. Wilson Journal of Ornithology 122:465-474.

DeLisle, J. M., and J. A. Savidge. 1996. Reproductive success of grasshopper sparrows in relation to edge. The Prairie Naturalist 28:107-113.

DeSante, D. F., and G. R. Geupel. 1987. Landbird productivity in central coastal California: the relationship to annual rainfall, and a reproductive failure in 1986. Condor 89:636-653.

Flaspohler, D. J. S. A. Temple, and R. N. Rosenfield. 2001. Species-specific edge effects on nest

success and breeding bird density in a forested landscape. Ecological Applications 11:32-46.

Groce, J. E., H. A. Mathewson, M. L. Morrison, and R. N. Wilkins. 2010. Scientific evaluation for the 5-year status review of the golden-cheeked warbler. Final report submitted to the Texas Parks and Wildlife Department, Austin, Texas, and to the U.S. Fish and Wildlife Service, Region 2, Albuquerque, New Mexico, USA.

Hannah, K. C., F. K. A. Schmiegelow, and K. E. H. Aitken. 2008. White-throated sparrow response to forest harvesting in north-central Alberta: results not so clear-cut? Avian Conservation and Ecology 3;6.

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.

Lackey, M. A., M. L. Morrison, Z. G. Loman, N. Fisher, S. L. Farrell, B. A. Collier, and R. N. Wilkins. 2011. Effects of road construction noise on the endangered golden-cheeked warbler. Wildlife Society Bulletin 35:15–19.

Long, A. M. 2014. The influence of tree species composition on songbird abundance and productivity. Dissertation. Texas A&M University, College Station, Texas, USA.

Marshall, M. E., M. L. Morrison, and R. N. Wilkins. 2013. Tree species composition and food availability affect productivity of an endangered species: the golden-cheeked warbler. Condor 115: 882-892.

Morgan, M. R., C. Norment, and M. C. Runge. 2012. Evaluation of a reproductive index for estimating productivity of grassland breeding birds. Auk 127:86–93.

Nocera, J. J., T. M. Fitzgerald, A. R. Hanson, and G. R. Milton. 2007. Differential habitat use

by Acadian Nelson's sharp-tailed sparrows: implications for regional conservation. Journal of Field Ornithology 78:50-55.

Peak, R. G. 2007. Forest edges negatively affect golden-cheeked warbler nest survival. Condor 109:628−637.

Peak, R. G., and F. R. Thompson, III. 2014. Seasonal productivity and nest survival of golden-cheeked warblers vary with forest type and edge density. Condor 116:546-559.

Peak, R. G., D. J. Lusk, and J. D. Peppers. 2010. Polygyny in the golden-cheeked warbler. Bulletin of the Texas Ornithological Society 43:44–47.

Powell, A. N., and C. L. Collier. 1998. Reproductive success of Belding's savannah sparrows in a highly fragmented landscape. Auk 115:508-513.

Pulich, W. 1976. The golden-cheeked warbler, a bioecological study. Texas Parks and Wildlife Department, Austin, Texas, USA.

Reidy, J. L., M. M. Stake, and F. R. Thompson, III. 2008. Golden-cheeked warbler nest mortality and predators in urban and rural landscapes. Condor 110:458–466.

Reidy, J. L., F. R. Thompson, III, and R. G. Peak. 2009. Factors affecting golden-cheeked warbler nest survival in urban and rural landscapes. Journal of Wildlife Management 73:407-413.

Reidy, J. L., and F. R. Thompson, III. 2012. Predatory identity can explain nest predation patterns. Pages 135-148 *in* C. A. Ribic, F. R. Thompson III, and P. J. Pietz, editors. Video surveillance of nesting birds. Studies in Avian Biology 43. University of California Press, Berkeley, California, USA.

Ribic, C. A., M. J. Guzy, T. J. Anderson, D. W. Sample, and J. L. Nack. 2012. Bird productivity

and nest predation in agricultural grasslands. Pages 119-134 *in* C. A. Ribic, F. R. Thompson III, and P. J. Pietz, editors. Video surveillance of nesting birds. Studies in Avian Biology 43. University of California Press, Berkeley, California, USA.

Rivers, J. W., D. P. Althoff, P. S. Gipson, and J. S. Pontius. 2003. Evaluation of a reproductive index to estimate dickcissel reproductive success. Journal of Wildlife Management 67:136–143.

Robinson, D. H. 2013. Effects of habitat characteristics on occupancy and productivity of a forest-dependent songbird in an urban landscape. Thesis. Texas A&M University, College Station, Texas, USA.

Shaffer, T. L. 2004. A unified approach to analyzing nest success. Auk 121:526–540.

U.S. Fish and Wildlife Service. 1992. Golden-cheeked warbler (*Dendroica chrysoparia*) recovery plan. Albuquerque, New Mexico, USA.

Verner, J. 1985. Assessment of counting techniques. Current Ornithology 2:247-302.

Verner, J., and K. A. Milne. 1990. Analyst and observer variability in density estimates from spot mapping. Condor 92:313-325.

Vickery, P. D., M. L. Hunter, Jr., and J. V. Wells. 1992. Use of a new reproductive index to evaluate relationship between habitat quality and breeding success. Auk 109:697–705.

Watson, M., J. M. Wilson, M. Koshkin, B. Sherbakov, F. Karpov, A. Gavrilov, H. Schielzeth, M. Brombacher, N. J. Collar, and W. Cresswell. 2006. Nest survival and productivity of the critically endangered Sociable Lapwing *Vanellus gregarius*. Ibis 148:489-502.

**Figure 1.** Location of seven golden-cheeked warbler monitoring plots (black) used to compare intensive monitoring with weekly index monitoring across Balcones Canyonlands Preserve (gray outline), 2013 and 2014.

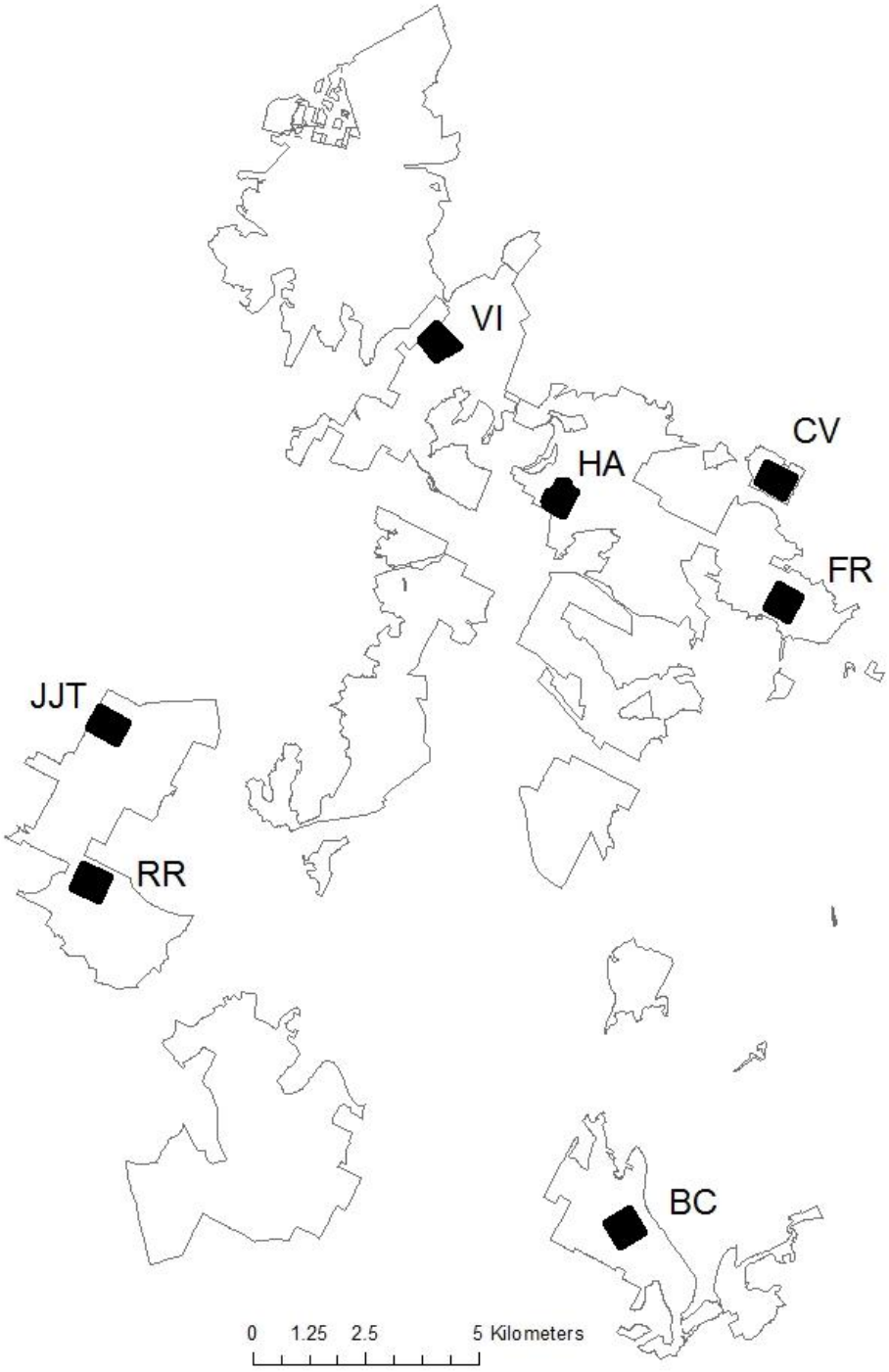

**Figure 2.** Golden-cheeked warbler territories delineated from intensive monitoring (gray outline polygons) and index monitoring (filled light and dark gray polygons) on seven plots across Balcones Canyonlands Preserve, March-June 2013-2014. Top row shows two plots surveyed by three index observers (one on CV, two on VI) in 2013. Second row shows three plots surveyed by two index observers in 2014. Bottom two rows are four plots surveyed by one of two index observers in 2014.

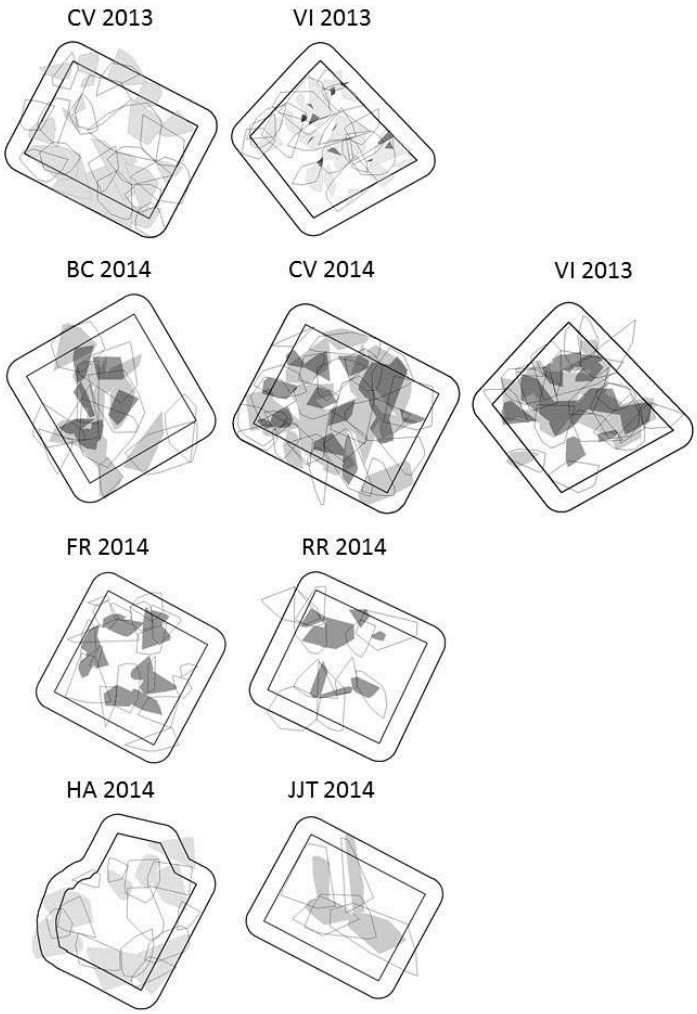

**Figure 3.** Correspondence in ranks of reproductive success (1-5) determined from reproductive index monitoring and intensive monitoring for 84 matched golden-cheeked warbler territories on Balcones Canyonlands Preserve. Numbers in bubbles are the number of territories; bubbles above and below the diagonal line represent territories where the reproductive index method overestimated or underestimated the rank, respectively, compared to intensive monitoring.

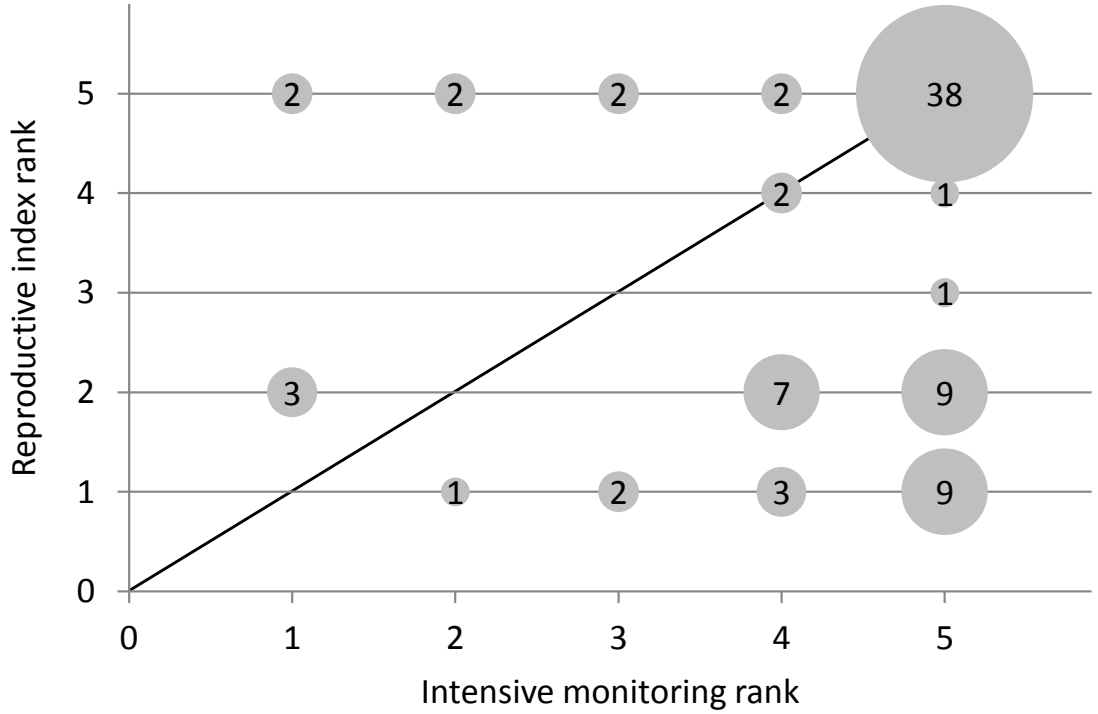

**Figure 4.** Distribution of relative reproductive index ranks of golden-cheeked warbler territories determined from intensive monitoring (black) and index monitoring (light and dark gray) for two plots in 2013 (top row), and seven plots in 2014 (bottom three rows).

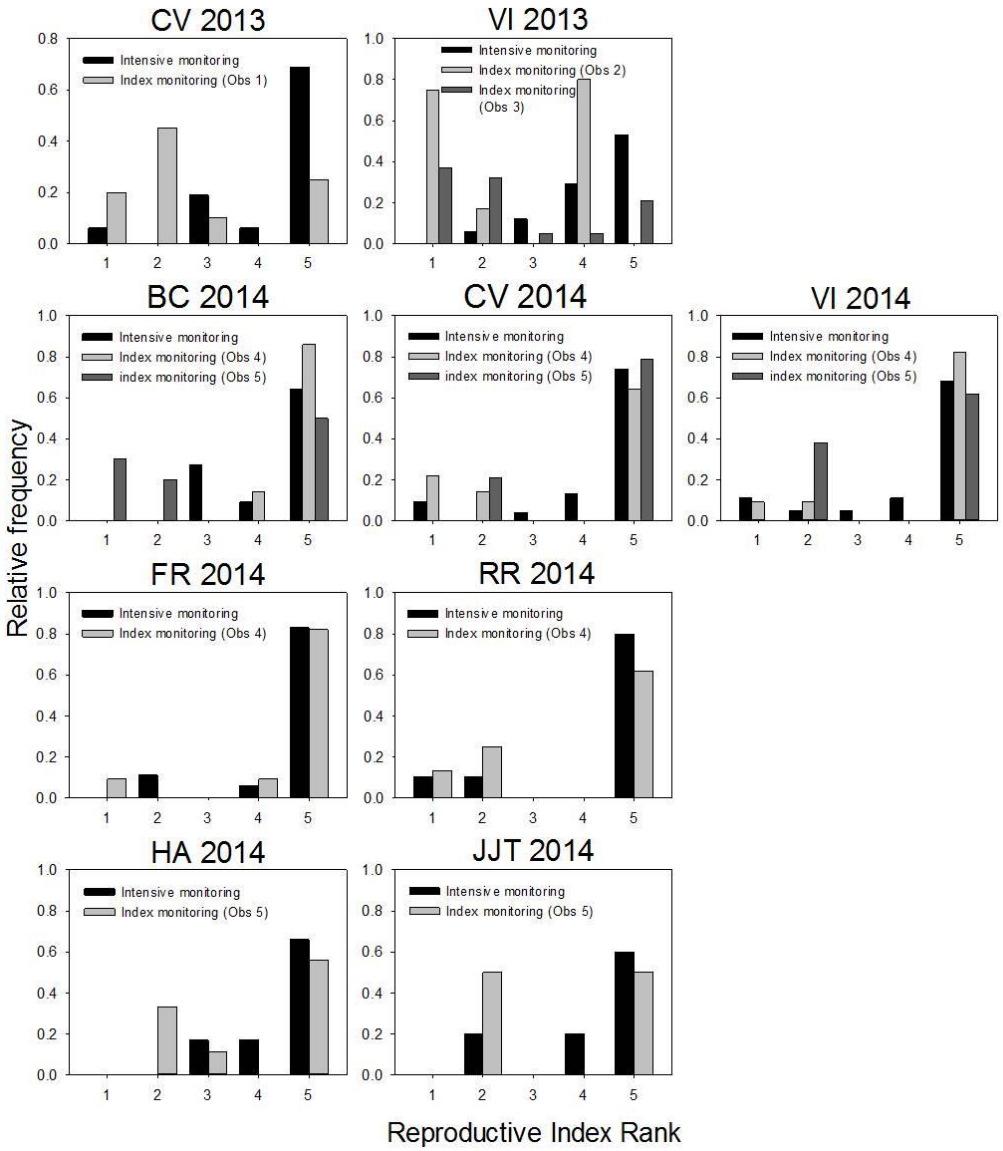

**Figure 5.** Number of golden-cheeked warbler territories per plot estimated from intensive monitoring versus index monitoring on Balcones Canyonlands Preserve, March-June 2013-2014. The total number of territorial males tended to be underestimated from index monitoring, whether observers were inexperienced (triangles) or experienced (circles).

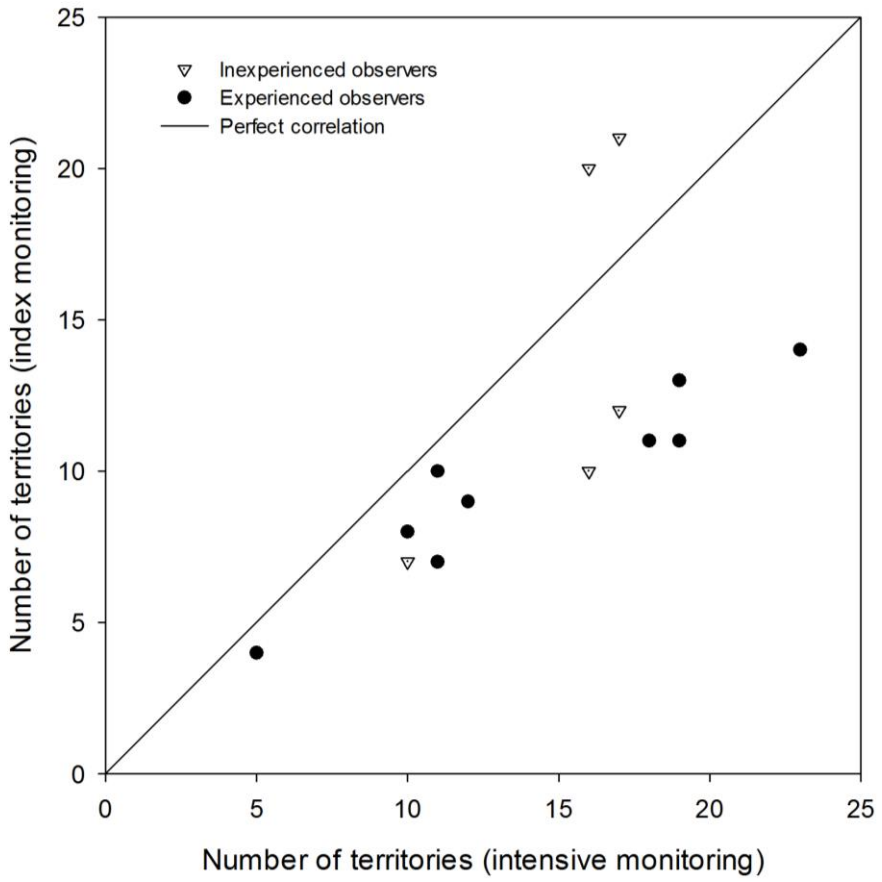

**Figure 6.** Number of paired male golden-cheeked warblers per plot estimated from intensive monitoring versus index monitoring on Balcones Canyonlands Preserve during March-June 2013-2014. Index observers tended to underestimate the total number of paired males, whether they were inexperienced (triangles) or experienced (circles).

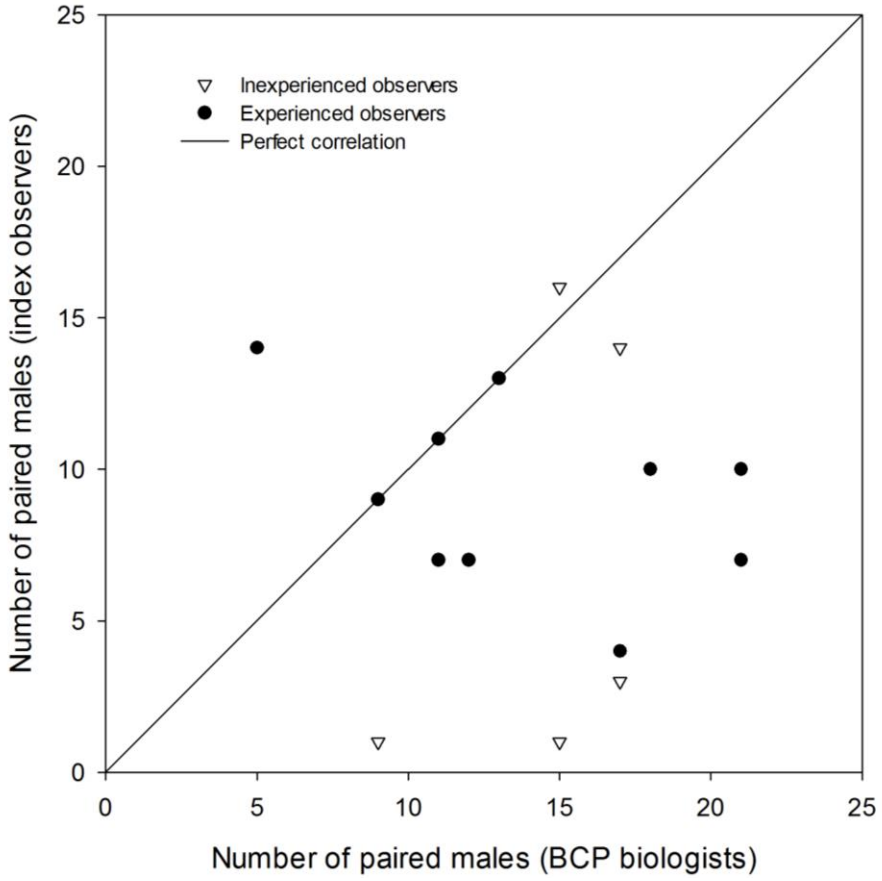

**Figure 7.** Number of golden-cheeked warbler territories that successfully fledged at least one host young per plot estimated from intensive monitoring versus index monitoring on Balcones Canyonlands Preserve, March-June 2013-2014. Index monitoring tended to underestimate the total number of territorial males, and inexperienced (triangles) observers underestimated more than experienced observers (circles).

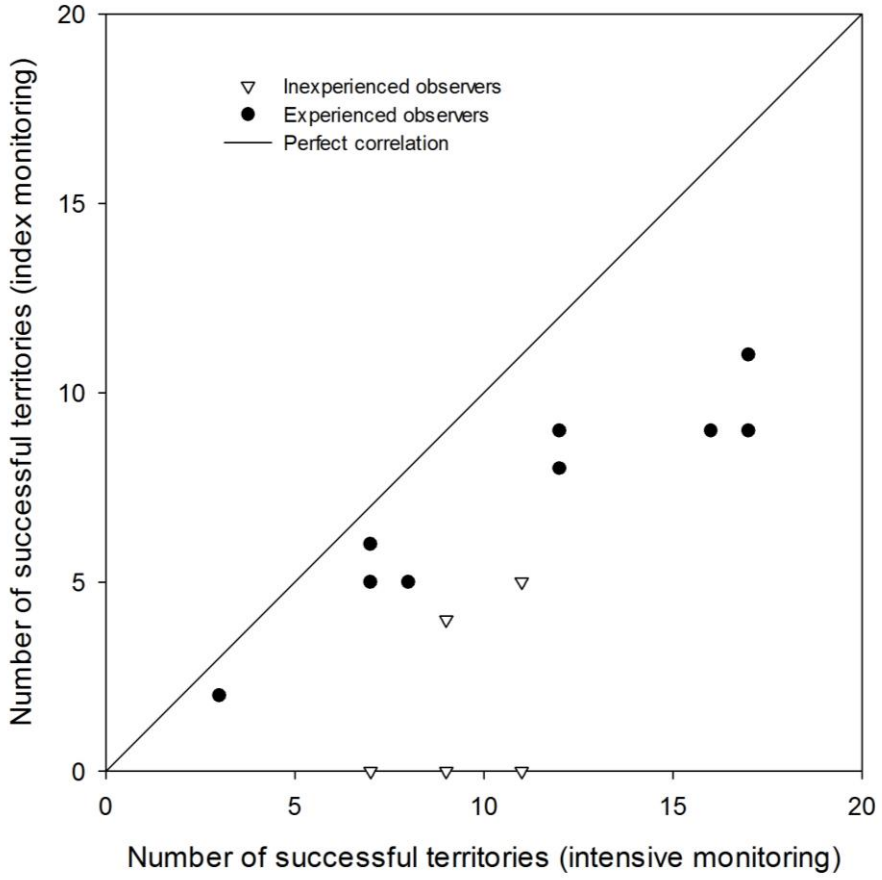

**Table 1.** Summary of survey effort by year and plot for intensive monitoring and index monitoring (transect and productivity) surveys on Balcones Canyonlands Preserve, March-June 2013-2014. We report number of plot visits (N) and the total hours surveyed (sum) and average length (mean) and standard deviation (SD) of surveys conducted to assess golden-cheeked warbler reproductive success. Total survey effort differed by method ($P < 0.001$), but average per survey time did not ($P = 0.75$).

| | | Intensive monitoring | | | | Transect surveys | | | | Productivity surveys | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Year | Plot | N | Sum | Mean | SD | N | Sum | Mean | SD | N | Sum | Mean | SD |
| 2013 | CV | 65 | 387.5 | 6.0 | 2.3 | 2 | 8 | 4.2 | 0.3 | 10 | 53 | 5.4 | 1.2 |
| 2013 | VI | 66 | 342.3 | 5.2 | 2.0 | 4 | 20 | 5.1 | 0.7 | 20 | 96 | 4.8 | 1.2 |
| 2014 | BC | 42 | 263.3 | 6.3 | 2.7 | 4 | 22 | 5.6 | 0.6 | 20 | 111 | 5.5 | 1.6 |
| 2014 | CV | 58 | 369.1 | 6.4 | 2.0 | 4 | 23 | 5.7 | 1.0 | 20 | 130 | 6.5 | 1.3 |
| 2014 | FR | 45 | 266.5 | 5.9 | 1.6 | 2 | 9 | 4.6 | 1.8 | 10 | 60 | 6.0 | 1.7 |
| 2014 | HA | 33 | 204.5 | 6.2 | 1.7 | 2 | 13 | 6.5 | 1.4 | 10 | 59 | 5.9 | 0.6 |
| 2014 | JJT | 19 | 99.5 | 5.2 | 1.7 | 2 | 9 | 4.5 | 0.7 | 10 | 49 | 4.9 | 0.7 |
| 2014 | RR | 34 | 210.0 | 6.2 | 1.7 | 2 | 11 | 5.3 | 0.4 | 10 | 56 | 5.6 | 1.8 |
| 2014 | VI | 58 | 329.5 | 5.7 | 2.5 | 4 | 20 | 5.0 | 1.5 | 20 | 117 | 5.9 | 1.4 |

**Table 2.** Summary of plot-level golden-cheeked warbler territory abundance, density and productivity determined by intensive monitoring by BCP biologists during 2013 and 2014. We report the total number of full and partial territories (N), the total number of full and one-half partial territories (Verner's N), and the number and mean and SD of paired males (rank ≥ 2), successful males (rank = 5), banded males, and number of nests monitored and proportion that were successful.

| | | | | | Pairing success | | | Breeding success | | | No. Banded | | | No. Nests | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Year | Plot | N | Verner's N[1] | Density[2] | N | Mean | SD | N | Mean | SD | N | Mean | SD | N | Mean | SD |
| 2013 | CV | 16 | 11.5 | 0.29 | 15 | 94 | 25 | 11 | 69 | 48 | 12 | 75 | 45 | 12 | 63 | 50 |
| 2013 | VI | 17 | 13.5 | 0.34 | 17 | 100 | 0 | 9 | 53 | 51 | 14 | 82 | 39 | 20 | 42 | 51 |
| 2014 | BC | 11 | 8 | 0.20 | 11 | 100 | 0 | 7 | 64 | 50 | 7 | 64 | 50 | 8 | 67 | 50 |
| 2014 | CV | 23 | 16 | 0.40 | 21 | 91 | 29 | 17 | 74 | 45 | 14 | 61 | 50 | 26 | 62 | 50 |
| 2014 | FR | 18 | 13.5 | 0.34 | 18 | 100 | 0 | 16 | 89 | 32 | 10[3] | 56 | 51 | 5 | 75 | 46 |
| 2014 | HA | 12 | 8.5 | 0.21 | 12 | 100 | 0 | 8 | 67 | 49 | 8[3] | 67 | 49 | 8 | 50 | 53 |
| 2014 | JJT | 5 | 4 | 0.10 | 5 | 100 | 0 | 3 | 60 | 55 | 3 | 60 | 55 | 3 | 67 | 58 |
| 2014 | RR | 10 | 7 | 0.18 | 9 | 90 | 32 | 8 | 80 | 42 | 8 | 80 | 42 | 5 | 80 | 44 |
| 2014 | VI | 19 | 14 | 0.35 | 17 | 89 | 32 | 12 | 63 | 50 | 13[3] | 68 | 48 | 15 | 81 | 40 |

[1] Verner's N sums the number of territories by considering territories which fall completely within the plot as 1 and territories that are inside and outside the plot boundaries as 0.5.

[2] All plots were 40 ha and density was calculated by dividing Verner's N by the area.

[3] FR, HA, and VI had 4, 2 and 1 additional territories with banded females

**Table 3.** Summary of plot-level golden-cheeked warbler territory abundance, density, and productivity determined by index monitoring on Balcones Canyonlands Preserve, 2013-2014. We report the total number of full and partial territories (N), the total number of full and one-half partial territories (Verner's N), and the number and mean (%) and SD of paired males (rank ≥ 2) and successful males (rank = 5) for each observer.

| | | | | | | Pairing Success | | | Breeding Success | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Year | Plot | Observer | N | Verner's N | Density | Sum | Mean | SD | Sum | Mean | SD |
| 2013 | CV | 1 | 18 | 12.5 | 0.31 | 15 | 83 | 38 | 4 | 22 | 43 |
| 2013 | VI | 2 | 12 | 11.5 | 0.29 | 3 | 25 | 45 | 0 | 0 | 0 |
| 2013 | VI | 3 | 20 | 17 | 0.43 | 12 | 60 | 50 | 4 | 20 | 41 |
| 2014 | BC | 4 | 7 | 7 | 0.18 | 7 | 100 | 0 | 6 | 86 | 38 |
| 2014 | BC | 5 | 10 | 8 | 0.20 | 7 | 70 | 48 | 5 | 50 | 53 |
| 2014 | CV | 4 | 14 | 12.5 | 0.31 | 11 | 79 | 43 | 9 | 64 | 50 |
| 2014 | CV | 5 | 14 | 10 | 0.25 | 14 | 100 | 0 | 11 | 79 | 43 |
| 2014 | FR | 4 | 11 | 10.5 | 0.26 | 10 | 91 | 30 | 9 | 82 | 40 |
| 2014 | HA | 5 | 9 | 6 | 0.15 | 9 | 100 | 0 | 5 | 56 | 53 |
| 2014 | JJT | 5 | 4 | 3.5 | 0.09 | 4 | 100 | 0 | 2 | 50 | 58 |
| 2014 | RR | 4 | 8 | 8 | 0.20 | 7 | 88 | 35 | 5 | 63 | 52 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2014 | VI | 4 | 11 | 11 | 0.28 | 10 | 91 | 30 | 9 | 82 | 40 |
| 2014 | VI | 5 | 13 | 11.5 | 0.29 | 13 | 100 | 0 | 8 | 62 | 51 |

**Table 4.** Significance of the relationship between plot-level golden-cheeked warbler territory density, pairing success, nest survival, and breeding success from intensive monitoring and estimates from index monitoring while accounting for observer for the RI. Nest survival from intensive monitoring was compared to territory success for index monitoring since nests are not located in index monitoring.

| Model and effect | DF | F-value | *P*-value | Partial $R^2$ |
|---|---|---|---|---|
| Territory density | | | | |
| Territory density | 1, 7 | 43.65 | <0.001 | 0.37 |
| Observer | 1, 4 | 10.03 | 0.005 | 0.34 |
| Pairing success | | | | |
| Pairing success | 1, 7 | 0.01 | 0.944 | <0.01 |
| Observer | 1, 4 | 7.64 | 0.011 | 0.74 |
| Nest survival | | | | |
| Nest survival | 1, 7 | 0.10 | 0.766 | <0.01 |
| Observer | 4, 7 | 5.63 | 0.024 | 0.39 |
| Territory success | | | | |
| Territory success | 1, 7 | 0.05 | 0.836 | <0.01 |
| Observer | 1, 4 | 7.5 | 0.011 | 0.81 |