Wildlife Society Bulletin 39(4):721–731; 2015; DOI: 10.1002/wsb.576

Original Article

Evaluation of a Reproductive Index for Estimating Songbird Productivity: Case Study of the Golden-Cheeked Warbler

JENNIFER L. REIDY,[1] *Department of Fisheries and Wildlife Sciences, University of Missouri, 302 Natural Resources Building, Columbia, MO 65211, USA*

LISA O'DONNELL, *Wildland Conservation Division, City of Austin, 3635 RR 620 S, Austin, TX 78738, USA*

FRANK R. THOMPSON, III, *Northern Research Station, United States Department of Agriculture Forest Service, 202 Natural Resources Building, Columbia, MO 65211, USA*

ABSTRACT It is critically important 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. Several recent studies focused on the golden-cheeked warbler (*Setophaga chrysoparia*), an endangered songbird that breeds exclusively in central Texas, USA, have used a reproductive index to link pairing and breeding success to various factors of interest. However, no field test has evaluated the performance of a reproductive index at estimating productivity of golden-cheeked warblers or similar forest songbirds. We computed reproductive index ranks for 5 observers from multiple plots during 2013 and 2014 and compared these ranks with actual reproductive success determined from territory mapping and nest monitoring of color-banded adults. At the territory level, we found no significant correlation between 84 matched territories delineated by the reproductive index and territories delineated by intensive monitoring. At the plot level, index monitoring underestimated the total number of territorial males, but density from index monitoring was weakly correlated with actual density. We found no significant relationship for pairing success estimated from the reproductive index and actual pairing success, or with breeding success from the reproductive index with nest survival or actual breeding success, when controlling for observer. The reproductive index did not produce reliable estimates of reproductive performance at either the territory or plot level and we recommend intensive monitoring of color-banded individuals when absolute estimates of reproductive success rates or other demographic measures are required, such as in species viability or threat assessments.

KEY WORDS breeding success, color-banded birds, forest songbird, intensive monitoring, pairing success, territory mapping, Vickery index.

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 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; however, 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). Vickery et al. (1992) proposed the use of a reproductive index to determine reproductive success and output in lieu of monitoring nests. The Vickery method is a reproductive index (hereafter, "RI") because it relies on ranks based on indirect behavioral clues from the adults to establish reproductive success. It was developed to estimate grassland songbird breeding success (portion of sampled population producing $\geq$1 host young) without having to locate and monitor, and potentially disturb, cryptic nests for rare species and was originally intended as a tool for 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 monitoring individual pairs using more intensive methods such as color-banding and nest monitoring. Observers may not be able to accurately identify territorial males (Best 1975, Paul and Roth 1983, 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 RI ranks

Received: 29 December 2014; Accepted: 15 May 2015
Published: 27 August 2015

[1]*E-mail: jennifer.reidy@gmail.edu*

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). Although 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 that demonstrated its usefulness at assessing reproductive output for the focal species.

Several studies have used the RI to determine reproductive success of grassland (DeLisle and Savidge 1996, Powell and Collier 1998, Nocera et al. 2007), forest (Rangen et al. 2000, Campomizzi et al. 2012), and shrubland (Hannah et al. 2008) birds without testing its efficacy, despite the cautionary statement by Vickery et al. (1992). A few studies have evaluated the accuracy of the RI and found that it 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 RI 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 they concluded that, depending on the study objectives, use of an 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 level 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 3 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. Despite these conclusions, a RI may be a useful, less-intensive 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 the 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*) is a federally listed endangered songbird whose breeding range lies entirely within the Ashe juniper (*Juniperus ashei*)–oak (*Quercus* spp.) woodlands of central Texas, USA (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 Golden-cheeked Warbler Recovery Plan (USFWS 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 golden-cheeked warbler productivity (Lackey et al. 2011); examining effects of tree species composition and foraging effort on golden-cheeked warbler productivity (Marshall et al. 2013); and determining the effects of canopy closure and tree species composition on golden-cheeked warbler 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 golden-cheeked warblers, or other forest songbirds, has been conducted, so it is difficult to interpret the conclusions.

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 golden-cheeked warbler. Intensive monitoring, including territory mapping and nest monitoring of color-banded birds, is generally considered the most reliable approach to estimating territory density and seasonal productivity (Bell et al. 1973, Best 1975, Paul and Roth 1983, Verner 1985, Bibby et al. 2000). Therefore, we evaluated the reliability of the RI for golden-cheeked warblers by comparing reproductive ranks derived from the RI to "actual" reproductive status determined from intensive monitoring of color-banded birds. Our objectives were to determine correspondence in 1) reproductive success estimated from the RI to actual reproductive success determined from intensive monitoring at the territory level; and 2) territory density and reproductive success estimated from the RI to actual 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 at 7 40-ha plots within the Balcones Canyonlands Preserve. The Preserve is a ~12,300-ha network of preserves in western Travis County, Texas, that was established to mitigate habitat loss for the golden-cheeked warbler and other endangered and rare species (Fig. 1; USFWS 1996). Our study plots were located on City of Austin and Travis County properties and were predominately mature closed-canopy woodland consisted of Ashe juniper, plateau live oak (*Quercus fusiformis*), shin oak (*Q. sinuata* var. *breviloba*), Texas red oak (*Q. buckleyi*), and cedar elm (*Ulmus crassifolia*). In 2013, we surveyed 2 plots (CV, VI) of relatively high golden-cheeked warbler density (range = 0.36–0.41 males/ha) based on previous territory mapping of color-banded males (City of Austin 2011, City of Austin et al. 2012). In 2014, we surveyed 5 additional plots (BC, FR, HA, JJT, RR) of varying golden-cheeked warbler density (range = 0.10–0.41 males/ha; City of Austin 2011; City of Austin et al. 2012, 2013). Our 7 plots spanned the gradient of density observed on 18 monitoring plots spread

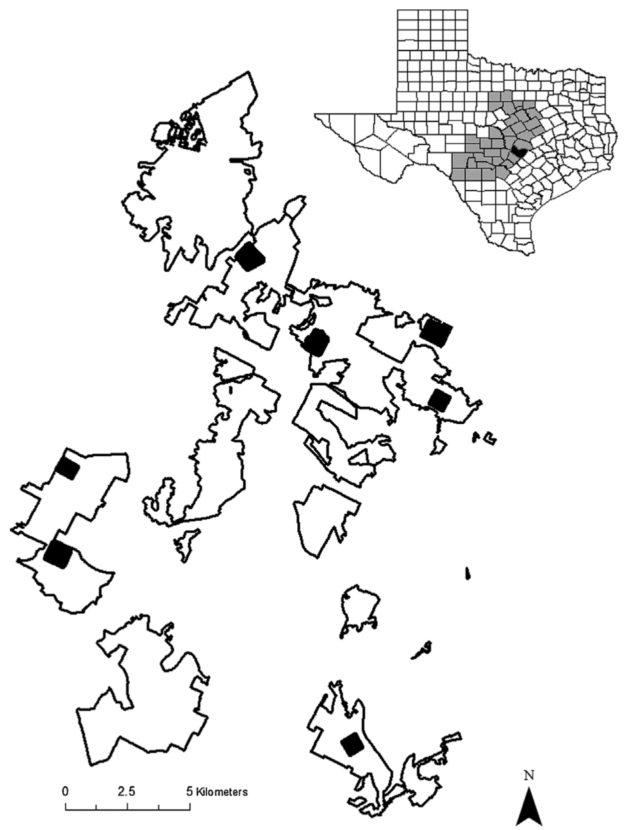

Figure 1. Location of 7 study plots (black) on Balcones Canyonlands Preserve (gray outline) in Travis County, Texas, USA, used to compare reproductive success of golden-cheeked warblers based on a reproductive index to intensive monitoring of color-banded birds, 2013 and 2014. Inset of Texas shows the breeding range shaded in gray and Travis County in black.

across the Balcones Canyonlands Preserve (City of Austin et al. 2013).

METHODS

Intensive Monitoring

Total territory mapping.—We banded adult golden-cheeked warblers from mid-March through early May 2009–2014 to establish the number of territorial males, delineate territorial boundaries, and determine pairing and breeding success (City of Austin 2009). We played prerecorded golden-cheeked warbler songs beneath the middle of 6-m-long mist nets to capture adults and banded adults with a unique combination of 2–3 colored leg bands and a U.S. Geological Survey numbered aluminum band (Animal Use Protocol 7414). Because the majority of the study population was color-banded, we consider our territory mapping described below to accurately reflect the number of territorial males present.

Two experienced biologists surveyed plots with >5 territorial males 2–3 times/week from 15 March to 15 June 2013–2014. Plots with $\leq$5 territorial males were surveyed by one biologist who visited the plot 2–3 times/week. Once per 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 and nests, and fledglings; and focused on mapping territories. Each week, they rotated the starting point in a clockwise direction. Both biologists returned to each plot 1–2 more times each week for targeted productivity surveys where effort was focused on searching for territorial birds that had not been previously detected that week, locate nests and record nesting behaviors, and monitor previously found nests. They attempted to collect $\geq$5 unique locations (>30 m apart) during each visit, with a goal of accumulating $\geq$33 locations across the breeding season to delineate territories (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/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 approximately 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, the biologists 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 ranked 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 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 $\geq$4 weeks) and excluded nonterritorial 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 boundary 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).

Reproductive Index Monitoring

Spot-mapping.—In 2013, one observer surveyed CV and 2 observers surveyed VI weekly from March 15 to June 10. In

2014, 2 independent observers surveyed BC, CV, and VI and each surveyed 2 additional plots (FR, JJT, HA, RR) once per week from 15 March 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 3 observers in 2013 and both observers in 2014 had prior experience surveying for golden-cheeked warblers. We considered observers with prior experience conducting golden-cheeked warbler surveys as "experienced" and those without prior experience as "inexperienced."

We generally followed methods used in studies of golden-cheeked warblers in Texas woodlands (Lackey et al. 2011, Campomizzi et al. 2012, Marshall et al. 2013) that were adapted by these investigators from Vickery et al. (1992). This approach uses a combination of territory, or spot, mapping (Verner 1985, Bibby et al. 2000) and the RI to arrive at territory density and reproductive success. Spot-mapping differs from total territory mapping in that observers rely on counter-singing and high site fidelity to differentiate territorial males and territories are delineated by examining clusters of observations made across the season rather than observations of uniquely identifiable adults (Bibby et al. 2000). Index observers initially established the approximate numbers and locations of golden-cheeked warblers by conducting 2 surveys along parallel transects spaced 100 m apart once per week from 15 March to 29 March beginning at sunrise and finishing 4–6 hours post-sunrise (2 surveys/plot). Males were first detected in the Austin area 7 March 2013 and 7 March 2014 by Balcones Canyonlands Preserve 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 golden-cheeked warbler on a transect survey form, and mapped out the general area of each detection for future reference. Index observers began spot-mapping territories and collecting productivity data after the initial 2 transect surveys. They returned to areas where they had previously detected golden-cheeked warblers and also recorded locations for any newly identified territories. Observers spent up to 30 minutes searching for a previously identified male and up to 30 minutes following him once he was detected, and they attempted to record $\geq$3 locations $\geq$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 a final rank at the end of the season based on the cumulative observations. We conducted 10 territory/productivity surveys and collected $\geq$10 locations for each territorial male during the breeding season. Index observers rotated their starting point each week in a clockwise direction and spot-mapped territories by following males and noting changes in vocal behaviors, contemporaneous singing, and using past locations (Bibby et al. 2000) and did not attempt to determine whether golden-cheeked warblers were banded. 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 $\geq$4 weeks) and excluded nonterritorial 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 $\geq$10 locations. We defined full and edge territories in the same way as for intensive monitoring.

Ranking

Observers assigned a rank of 1 to 5 to each territory following a modified Vickery RI (Campomizzi et al. 2012): 1 = male present $\geq$4 weeks, 2 = pair present for $\geq$4 weeks, 3 = nest material carry observed, 4 = food carry observed, 5 = host fledgling verified; and we assigned a rank of 0 to a territory if the male was not detected during the weekly survey. We compiled rankings from RI monitoring for each territory for each week and a final ranking based on the greatest observed rank. 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 observing multiple food items carried in different directions or hearing begging chips. We also assigned a final rank to territories from intensive monitoring based on the greatest observed rank during the breeding season.

Data Analysis

Each observer interpreted their own data to derive rankings. We report summary statistics (mean $\pm$ SD) 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 portion of the known males detected by index observers each week and the number we considered a match. 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 territorial observations for both methods to aid in assigning matches. We calculated the Spearman rank correlation coefficients for matched territories by index monitoring observer 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 (RI rank ≥ 2) and breeding success (RI rank $= 5$) as the proportion of territories reaching these ranks. We calculated daily nest survival and period nest survival 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 general linear model (MIXED and GLM procedures; SAS 9.3, SAS Institute, Cary, NC). We treated the RI-based estimates as the response variable and evaluated their relationship with the known measure based on intensive monitoring. We controlled for observer if the addition of observer as a fixed effect to the model was supported by a lower Akaike Information Criterion (AIC_c) value for the model with observer (Burnham and Anderson 2002). We used a repeated measures model with plot $\times$ year as the subject because plots were visited by multiple observers.

RESULTS

Intensive Monitoring

We spent 730 and 1,742 hours conducting intensive monitoring in 2013 and 2014, respectively (Table 1). Biologists ranked 33 (79% of territories with ≥ 1 banded adult) and 98 (72% of territories with ≥ 1 banded adult) territories as ≥ 1 in 2013 and 2014, respectively (Table 2). For both years combined, average territory density was 0.27 ± 0.10 males/ha (range = 0.10–0.40; 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 that did not qualify as territorial). Plot-level pairing success averaged $96\% \pm 5\%$ (range = 89–100%) and plot-level breeding success averaged $69\% \pm 11\%$ (range = 53–89%; Table 2).

We monitored 102 nests (32 and 70 in 2013 and 2014, respectively) representing 84 territories (25 and 59 in 2013 and 2014, respectively), of which 65 (16 and 49 in 2013 and 2014, respectively) were considered successful (i.e., fledged ≥ 1 host young). Apparent nest success (64%; 65/102) was slightly lower than the observed breeding success (69%). We confirmed 3.3 ± 1.0 host young/successful nest ($N = 65$), and found 3.1 ± 1.1 fledglings/successful territory from the first brood for which successful nests were not monitored ($N = 31$). 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 daily and period nest survival were 0.976 (95% CI = 0.962–0.984) and 0.54 (95% CI = 0.38–0.68), respectively, 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 males/ha and ranged from 0.09 to 0.43 (Table 3). Index observers detected $46\% \pm 20\%$ (range = 29–68%) of known territorial males during each week. Index observers ranked 77 of 151 (51%) territories as a 5. Overall pairing success was $81\% \pm 40\%$ and breeding success was $51\% \pm 50\%$ based on all index monitoring (Table 3). Experienced observers documented lower territory density (GLM, $F = 4.96$, df = 1, $P = 0.050$), but higher pairing (GLM, $F = 37.11$, df = 1, $P < 0.001$) and breeding

Table 1. Summary of survey effort by year and plot for intensive monitoring and index monitoring (transect and productivity) surveys on Balcones Canyonlands Preserve, Texas, USA, during March–June 2013–2014. We report number of plot visits (n), the total hours surveyed (sum), and the mean and standard deviation (SD) of number of hours per survey conducted to assess golden-cheeked warbler reproductive success. Total survey effort differed by method ($P < 0.001$), but average per survey time did not ($P = 0.750$).

		Intensive monitoring				Transect surveys				Productivity surveys			
Year	Plot	n	Sum	$\bar{x}$	SD	n	Sum	$\bar{x}$	SD	n	Sum	$\bar{x}$	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 on Balcones Canyonlands Preserve, Texas, USA, 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*), density (Verner's *N*/area), and the number and mean and standard deviation of paired males (rank ≥ 2), successful males (rank = 5), banded males, and number of nests monitored.

					Pairing success			Breeding success			No. banded			No. nests		
Year	Plot	*N*	Verner's *N*[a]	Density[b]	*N*	%	SD	*N*	%	SD	*N*	%	SD	*N*	%	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	82	39
2014	BC	11	8.0	0.20	11	100	0	7	64	50	7	64	50	8	73	47
2014	CV	23	16.0	0.40	21	91	29	17	74	45	14	61	50	26	83	39
2014	FR	18	13.5	0.34	18	100	0	16	89	32	10[c]	56	51	5	33	49
2014	HA	12	8.5	0.21	12	100	0	8	67	49	8[c]	67	49	8	75	45
2014	JJT	5	4.0	0.10	5	100	0	3	60	55	3	60	55	3	40	55
2014	RR	10	7.0	0.18	9	90	32	8	80	42	8	80	42	5	40	52
2014	VI	19	14.0	0.35	17	89	32	12	63	50	13[c]	68	48	15	74	45

[a] Verner's *N* sums the no. of territories by considering territories that fall completely within the plot as 1 and territories that are inside and outside the plot boundaries as 0.5.
[b] All plots were 40 ha and density was calculated by dividing Verner's *N* by the area.
[c] FR, HA, and VI had 4, 2, and 1 additional territories with banded females.

success (GLM, $F = 28.26$, df = 1, $P < 0.001$) than did 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). We were able to match a greater proportion of territories at low-density plots. Ranks for 40 of the 84 (48%) territories coincided, 11 (13%) were overestimated, and 33 (39%) were underestimated (Fig. 3). Across all 84 territories we found no significant correlation in reproductive ranks from the 2 methods ($r = 0.156$, $P = 0.157$). We identified 7 territories where RI observers assigned a rank of 5 before intensive monitoring indicated the territories had fledged young, suggesting that RI observers incorrectly assigned fledglings to these territories.

Plot-Level Comparison

The distributions of reproductive success ranks from the RI 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 RI and actual territory density; however, index observers tended to underestimate the total number of territorial males ($\bar{x} = 80\%$, range = 58–125%; Table 4; Fig. 5). We found no significant relationship for pairing and breeding success estimated from the RI and actual pairing and breeding success, or between nest survival and breeding success from the RI method when controlling for observer, with observer accounting for a large percentage of the variation in RI ranks (Figs. 6 and 7; Table 4).

DISCUSSION

We found a general lack of agreement between estimates of territory density, pairing success, and breeding success from the reproductive index and from intensive monitoring. We

Table 3. Summary of plot-level golden-cheeked warbler territory abundance, density, and productivity determined by index observers on Balcones Canyonlands Preserve, Texas, USA, 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*), density (Verner's *N*/area), and the number and mean and standard deviation of paired males (rank ≥ 2) and successful males (rank = 5).

						Pairing success			Breeding success		
Year	Plot	Observer	*N*	Verner's *N*	Density	*N*	%	SD	*N*	%	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

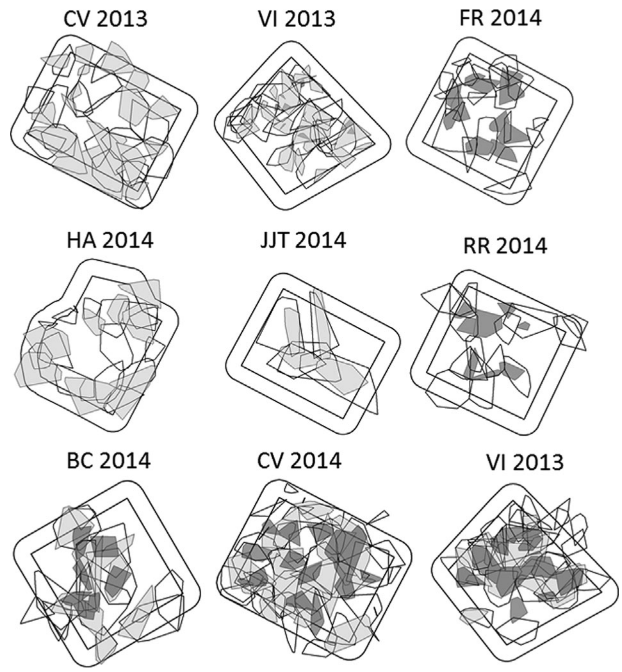

Figure 2. Golden-cheeked warbler territories delineated from intensive monitoring (gray outline polygons) and index monitoring from each observer (filled light and dark gray polygons) on 7 plots across Balcones Canyonlands Preserve, Texas, USA, March–June 2013–2014.

have high confidence in the results from intensive monitoring because of biologists' extensive prior experience with the species coupled with the use of color-banded birds and substantial effort to locate females, nests, and fledglings. We found nests for approximately 64% of golden-cheeked warbler territories monitored. Based on the estimated period nest success of 54% and 65 successful nests, we would expect to have found approximately 120 active nests and we monitored 102 active nests, suggesting we monitored most,

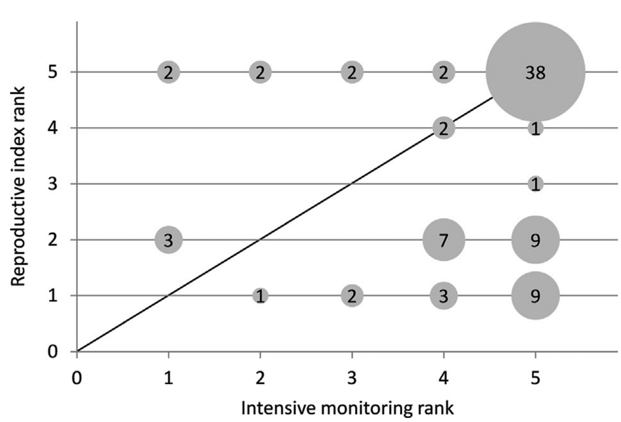

Figure 3. Correspondence in ranks of reproductive success (1–5) determined from reproductive index monitoring and intensive monitoring conducted during March–June 2013–2014 for 84 matched golden-cheeked warbler territories on Balcones Canyonlands Preserve, Texas, USA. Numbers in bubbles are the number of territories; bubbles above and below the diagonal line represent territories where the reproductive index method overestimated or underestimated the rank, respectively, compared with intensive monitoring.

and likely a representative sample, of nests. Additionally, we found 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 golden-cheeked warblers at Fort Hood (Peak and Thompson 2014). Hence, we believe that our intensive monitoring effort provided an accurate measure of reproductive success. We rarely documented behavior of adults indicative of disturbance (e.g., alarm calls, distraction displays) while nest monitoring and when we did it was usually immediately prior to, during, or immediately after fledging. Hence we do not believe our protocol for nest searching and monitoring represented a disturbance to this species.

Index observers tended to underestimate the total number of territorial males, 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 greatest abundance. Index observers, regardless of experience, had limited opportunities compared with 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 mapped 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, because most of our adults were banded and because we were surveying plots multiple times per week.

Index observers conducted 10 weekly productivity surveys on each plot after the 2 initial transect surveys, similar to the design the City of Austin employed prior to 2009 (City of Austin 2007) and similar to other studies on golden-cheeked warblers (Campomizzi et al. 2012, Long 2014). Additional survey effort may have resulted in more concordance between territory abundance from intensive and RI monitoring, particularly for plots with high abundance. However, we

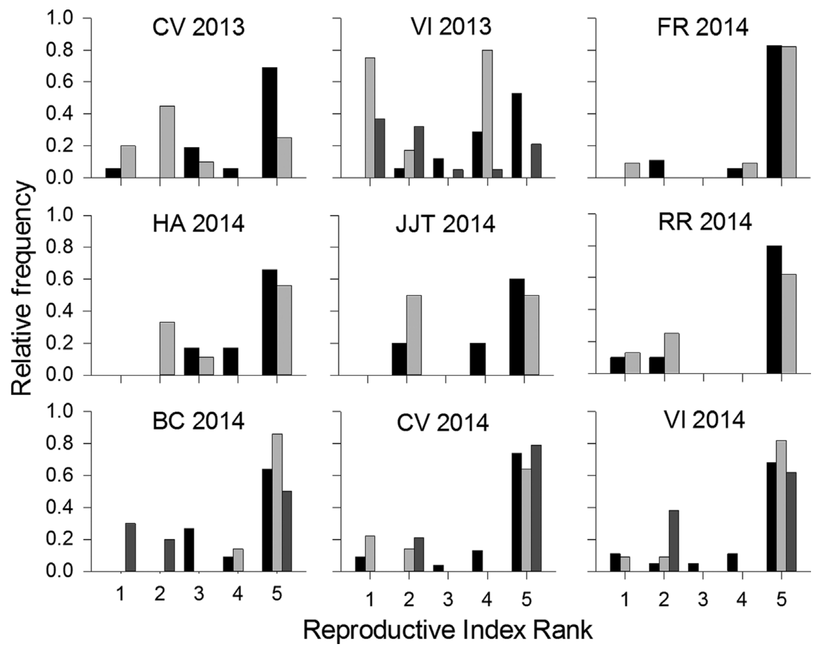

Figure 4. Distribution of relative reproductive index ranks of golden-cheeked warbler territories determined from intensive monitoring (black) and a reproductive index from each observer (light and dark gray) on Balcones Canyonlands Preserve, Texas, USA, March–June 2013–2014.

believe that it would still be difficult to determine the true number of territorial males even with considerably more survey effort without uniquely color-banding birds. This is partly because of the high degree of territory overlap and movement documented by intensive monitoring and because males often did not counter-sing. We saw a similar pattern in the City of Austin's territory mapping before and after color banding was initiated in 2009. Territories overlapped very little prior to color 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 color banding (City of Austin 2007). These issues have been documented in other studies comparing territory delineation from spot-mapping with and without color-banded birds (Bell et al. 1973, Best 1975, Paul and Roth 1983).

The reproductive index 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 RI observers, and final rankings from the 2 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

Table 4. Significance of 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 based on Type-3 tests of fixed effects on Balcones Canyonlands Preserve, Texas, USA, 2013–2014. Nest survival from intensive monitoring was compared with breeding success for index monitoring because nests were not located during 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
Breeding success				
Breeding success	1, 7	0.05	0.836	<0.01
Observer	1, 4	7.5	0.011	0.81

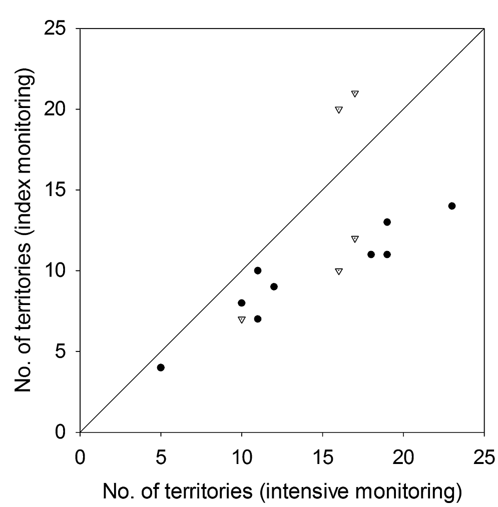

Figure 5. Number of golden-cheeked warbler territories per plot estimated from intensive monitoring versus index monitoring on Balcones Canyonlands Preserve, Texas, USA, March–June 2013–2014. Triangles represent inexperienced observers and circles represent experienced observers. The straight line is for a perfect correlation.

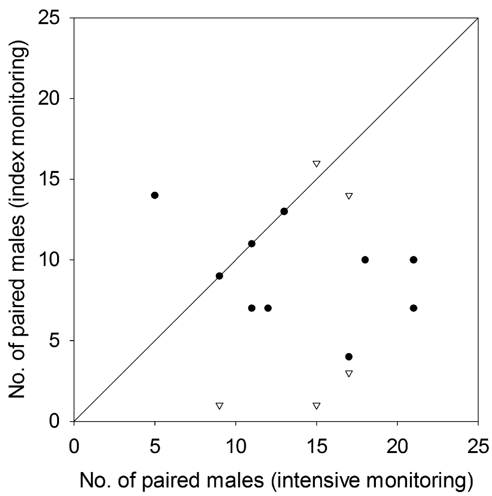

Figure 6. Number of paired male golden-cheeked warblers per plot estimated from intensive monitoring versus index monitoring on Balcones Canyonlands Preserve, Texas, USA, during March–June 2013–2014. Triangles represent inexperienced observers and circles represent experienced observers. The straight line is for a perfect correlation.

ultimately did fledge young, making the final rankings seemingly align. Furthermore, 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 reproductive index rankings were not a reliable indicator of territory-level reproductive success. Morgan et al. (2010) was more successful at matching territories than we were, but found rankings were only weakly correlated to actual breeding success. Rivers et al.

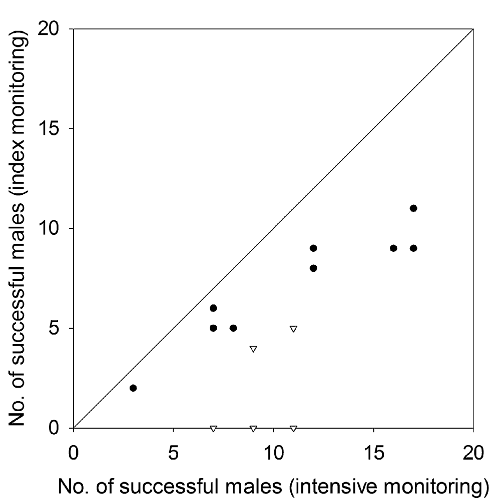

Figure 7. Number of golden-cheeked warbler territories that successfully fledged ≥1 host young/plot estimated from intensive monitoring versus index monitoring on Balcones Canyonlands Preserve, Texas, USA, March–June 2013–2014. Triangles represent inexperienced observers and circles represent experienced observers. The straight line is for a perfect correlation.

(2003) opted to not even attempt to match territories because the number of territorial males estimated from each method was 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 reproductive index and those reproductive measures, suggesting strong observer effects on the reproductive index. 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 error had we asked observers to interpret surveys from other observers rather than just their own (see Best 1975, Verner and Milne 1990). Other studies have also concluded observers provide inconsistent results regarding the number of territories and final reproductive rankings (Paul and Roth 1983, 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 reproductive index because the majority of studies on golden-cheeked warblers and other forest songbirds are of short duration, typically 2–3 years, and involve many observers, most of whom lack intimate knowledge of the species needed to correctly assign breeding activity.

Habitat quality and impacts of various habitat management actions are often assessed by measuring basic demographic measures such as pairing and breeding success using the reproductive index for populations of golden-cheeked warblers (Klassen et al. 2012, Marshall et al. 2013, Stewart et al. 2013). Our ability to evaluate habitat quality and recommend management and conservation strategies is reliant on accurate estimates of these measures in available habitat types or under various management scenarios. Our results suggest the use of a reproductive index to be an unreliable method for determining pairing or breeding success for golden-cheeked warblers. Number of territories and subsequent ranking based on observations of reproductive behaviors were observer-dependent. Our observers tended to underestimate total number of territories, and consequently pairing and breeding success, likely because many territories were overlapping and pairs on our plots experienced high pairing success and moderately high breeding success. Observers may be more prone to overestimate for populations of golden-cheeked warblers or other species that experience lower pairing or breeding success, and we consider this an avenue for future investigation. We did not evaluate whether multiple observers working together on a plot would reduce observer-related bias; we doubt there would be substantial improvement under the same survey effort; however, this is another issue worth exploring. Monitoring methods that rely on infrequent surveys are likely to miss most reproductive activity because paired males sing less than unpaired males (J. L. Reidy, unpublished data) and are therefore less likely to be detected during surveys. In the absence of nest monitoring and monitoring of color-banded birds, determining the correct number of territories and potentially mismatching fledglings and territories is likely a pervasive problem, and the outcome is erroneous conclusions about habitat quality or responses to impacts. Future investigations should focus on ways to improve the accuracy of the reproductive index. We suggest comparing breeding success determined from nest monitoring to reproductive index rankings for color-banded birds from independent observers. This protocol would still require the time and skill investment of banding adults, but may alleviate the need to also locate and monitor nests to verify productivity, which may be a necessary alternative for long-term monitoring where the effort to monitor nests is too great.

MANAGEMENT IMPLICATIONS

Color-banding adults and finding and monitoring nests are time- and skill-intensive endeavors, but they provide the most accurate measures of reproductive output. We did not find support that the reproductive index produced reliable results of actual breeding performance at either the territory or plot level. Our results indicate the reproductive index is a poor substitute for intensive monitoring for assessing productivity of golden-cheeked warblers. Though 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 golden-cheeked warblers. Where absolute estimates of reproductive success rates or other demographic measures are required, such as in species viability or threat assessments, we believe intensive monitoring of color-banded individuals is required. We believe additional research is needed to improve and validate approaches to assess reproductive performance for species such as golden-cheeked warblers when intensive monitoring is not possible.

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, and logistical support were provided by the City of Austin, Travis County, the U.S. Department of Agriculture Forest Service Northern Research Station, and the University of Missouri. O. Bocanegra, T. Benson, and 2 anonymous reviewers provided constructive feedback on an earlier version of this manuscript.

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. Second edition. Academic Press, San Diego, California, USA.

Burnham, K. P., and D. R. Anderson. 2002. Model selection and multimodel inference: a practical information-theoretic approach. Second edition. Springer, New York, New York, 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. City of Austin, Austin Water Utility, Wildland Conservation Division, Balcones Canyonlands Preserve, 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 fiscal year 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 fiscal year 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. 2012. 2012 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.

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 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, 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. 2010. 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.

Paul, J. T., Jr., and R. R. Roth. 1983. Accuracy of a version of the spot-mapping census method. Journal of Field Ornithology 54:42–49.

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.

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, USA.

Rangen, S. A., K. A. Hobson, and R. G. Clark. 2000. A comparison of density and reproductive indices of songbirds in young and old boreal forest. Wildlife Society Bulletin 28:110–118.

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.

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, 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, USA.

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

Stewart, L. R., M. L. Morrison, M. R. Hutchison, D. N. Appel, and R. N. Wilkins. 2013. Effects of a forest pathogen on habitat selection and quality for the endangered golden-cheeked warbler. Wildlife Society Bulletin 38:279–287.

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

U.S. Fish and Wildlife Service [USFWS]. 1996. Final environmental impact statement/habitat conservation plan for proposed issuance of a permit to allow incidental take of the golden-cheeked warbler, black-capped vireo, and six karst invertebrates in Travis County, Texas. Prepared by Regional Environmental Consultants (RECON) and U.S. Fish and Wildlife Service. Albuquerque, New Mexico.

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.

Associate Editor: Benson.