

Monitoring bobolink abundance and breeding success in response to stewardship

Andrew J. Campomizzi, Zoé M. Lebrun-Southcott
Bird Ecology and Conservation Ontario

22 February 2019

1 Introduction

Long-term population declines of songbirds (i.e., order Passeriformes) that breed in North American grasslands has led to conservation concern for some of these species (Brennan and Kuvlesky 2005, Sauer et al. 2013). Bobolink (*Dolichonyx oryzivorus*) are listed as threatened by the governments of Canada (Government of Canada 2017) and Ontario (MNR 2017) because of population declines. Bobolink abundance in Ontario has declined by an average of 2.6% per year based on Breeding Bird Survey data from 1968 to 2008 (COSEWIC 2010). Likely causes of population declines that are associated with breeding grounds include habitat loss, habitat fragmentation, and incidental mortality of eggs and young from agricultural operations (particularly hay harvest and cattle grazing while birds are nesting; COSEWIC 2010, McCracken et al. 2013, MNR 2015). Conservation efforts to recover the population in Ontario are ongoing and important because 13% of the global bobolink population breeds in the province, which is the largest percent for any province and is surpassed by the states of North Dakota (22%) and Minnesota (13%) only (Partners in Flight Science Committee 2013).

Information about the impacts of stewardship actions on bobolink is needed to support recovery efforts (MNR 2015). A high priority action for research and monitoring in the Government Response Statement is the need to “Research, assess and identify best management practices that can benefit nest productivity with minimal disruption to farming activities (e.g. delayed or modified haying, shifted grazing)” (MNR 2015). Avian ecologists use various field survey types (e.g., Thompson et al. 1998) that could be applicable to assessing the impacts of stewardship on bobolink. Of course, no one method is perfect depending on particular monitoring goals and study designs (Block et al. 2001, Morrison et al. 2008, Johnson 2012). Although spot mapping (i.e., delineating locations used by unmarked territorial birds) and nest monitoring provide detailed information about songbird abundance and breeding success, these methods are labour intensive (Thompson et al. 1998). Less intensive survey types (e.g., transect survey, point count; Thompson et al. 1998) are common, but their reliability for assessing bobolink abundance and breeding success is unknown. Although various statistical techniques are available for analyzing field survey data, we largely ignore discussion of statistics to focus on field survey types.

Estimating the abundance of territorial songbirds from non-intensive survey types can be challenging. Abundance indices have largely been replaced by sophisticated statistical analyses that enable incorporation of covariates and estimates of detection probability (Johnson 2008). However, specific field implementation of surveys often impacts results. For example, Newell et al. (2013) found that 100-m radius point counts underestimated density of abundant forest songbird species, compared to spot mapping. Jones et al. (2000) found that 50-m radius point counts overestimated density of cerulean warblers (*Setophaga cerulea*); whereas, 100-m radius and variable circular plot counts underestimated density. Buckland (2006) found that transect surveys gave unbiased density estimates of common songbirds in Scotland and were more efficient to implement than point counts. Diefenbach et al. (2003) found that up to 60% of grassland birds in Pennsylvania went undetected > 50 m from transects, underscoring the importance of estimating the probability of detecting songbirds on surveys.

Assessing the breeding success of songbirds while avoiding labour-intensive field surveys, such as nest monitoring, is also challenging. Vickery et al. (1992) introduced an index of breeding success based on behavioural observations collected across multiple visits to territories during the breeding season (i.e., from lowest rank of territorial male to highest rank of fledging). Rivers et al. (2003) and Morgan et al. (2010) found that the Vickery index did not reliably identify breeding success of dickcissel (*Spiza americana*) and Savannah sparrow (*Passerculus sandwichensis*), respectively, across the breeding season. However, Christoferson and Morrison (2001) found that the Vickery index accurately identified the breeding status of 3 riparian songbird species compared to nest monitoring for each visit to territories. To our knowledge, the applicability of the Vickery index to bobolink has not yet been evaluated.

In 2017, we used 3 survey types (transect survey, point count, roadside count) to assess bobolink abundance and breeding success in pastures. We compared the performance of these survey types to intensive monitoring (i.e., spot mapping and nest monitoring) to evaluate the survey effort needed to assess bobolink abundance and breeding success. We used the results from the 2017 field season to develop a draft monitoring scheme to assess the impacts of stewardship actions on bobolink. In 2018, we implemented the draft monitoring scheme in delayed-harvest hay fields, restored grasslands, and fallow fields. Although we conducted roadside counts in 2018, we excluded them from analysis in this report because few fields were adjacent to roads, resulting in an insufficient sample size. Herein, we report on data collected in 2018 only.

Our goal was to evaluate the efficacy of low-intensity survey types (i.e., transect survey, point count) to assess bobolink abundance and breeding success in fields where common grassland stewardship actions were implemented. For comparison, we considered spot mapping and nest monitoring to provide the best available information about abundance and breeding success in fields. Our objectives were to: (1) evaluate the efficacy of survey types to estimate bobolink abundance, (2) evaluate the efficacy of survey types to assess bobolink breeding success, (3) use ≥ 1 survey type(s) to evaluate the ecological impact of grassland stewardship on bobolink, and (4) develop a monitoring scheme that can be used by other organizations to evaluate the ecological impact of various stewardship actions on bobolink.

2 Methods

2.1 Study species

Bobolink have a large breeding range across southern Canada and the northern USA (Renfrew et al. 2015). In contrast, the species has a relatively small non-breeding range in the southern interior of South America, predominantly in Paraguay, eastern Bolivia, northeastern Argentina, and southwestern Brazil (COSEWIC 2010, Renfrew et al. 2015). The spring trans-equatorial migration begins in late March and early April; adults arrive on breeding grounds in May (Renfrew et al. 2015). The typical duration of nesting is ~28 days (1 egg laid per day [mean clutch size = ~5], incubation begins on the day the last egg is laid, 12 days of incubation, 11 days from hatch to fledge; Renfrew et al. 2015). In 2018, the earliest date bobolink laid eggs was 18 May and the latest date of nesting activity was 23 July. Bobolink rarely raise > 1 brood per year; however, individuals may attempt a second nest following an initial nest failure (Renfrew et al. 2015). Males and females are easily distinguished because the species is sexually dichromatic and males sing, but females do not (Renfrew et al. 2015).

2.2 Study area

In 2018, we studied bobolink in the Luther Marsh Wildlife Management Area (43°57'45'' N, -80°24'2'' W) in Dufferin and Wellington counties in southern Ontario and on 4 farms within 42 km. We monitored bobolink in 36 fields, totaling 254 ha: hay fields ($n = 9$) on 3 farms, fields seeded with native grasses and forbs for grassland restoration ($n = 18$) in the Luther Marsh WMA, and fallow fields (i.e., no longer farmed and left to naturalize; $n = 9$) in the Luther Marsh WMA and on 1 farm. Vegetation in hay fields was primarily grasses (e.g., timothy [*Phleum pratense*], brome species [*Bromus* spp.], redtop [*Agrostis gigantea*]) and secondarily forbs (e.g., alfalfa [*Medicago sativa*], red clover [*Trifolium pratense*]). Restored fields tended to have more forb cover (e.g., goldenrod species [*Solidago* spp.], vetch species [*Vicia* spp.]) than grasses (e.g., timothy, redtop). Whereas, the proportion of grass cover (e.g., Canada bluegrass [*Poa compressa*], redtop) and forbs (e.g., goldenrod species, bird's-foot trefoil [*Lotus corniculatus*]) was variable in fallow fields. Field size ranged from 2 to 23 ha. Cutting was delayed in hay fields to provide nesting habitat for bobolink. Restored fields were planted between 2010 and 2015 to support grasslands birds and other wildlife (L. Campbell, Grand River Conservation Authority, personal communication). All but 1 of the fallow fields

were agricultural fields that were intentionally left to naturalize to support grassland birds and other wildlife; 1 was an abandoned hay field. All fields were undisturbed by management (e.g., mowing) during May, June, and the first week of July (i.e., when surveys occurred). However, a small amount of mowing did occur in 2 fields before mid-July, but occurred in areas with no active bobolink nests.

2.3 Study design

We used spot mapping and nest monitoring to acquire the most accurate assessment possible of abundance and breeding success in each field. We used 2 additional less intensive survey types (i.e., transect surveys and point counts) and compared results to assess whether these surveys accurately assessed abundance and breeding success. We originally planned to include roadside surveys, but few fields were adjacent to roads, resulting in insufficient data. For analysis and results, we refer to combined data from nest monitoring and spot mapping as spot map.

2.4 Spot mapping and nest monitoring

We used a modified spot mapping method (*sensu* Wiens 1969) to delineate and monitor bobolink territories. From mid-May through mid-July, we visited each field approximately twice per week to search for bobolink. When we detected a bobolink, we followed the individual or pair for up to 30 min, if possible. For each individual or pair, we recorded 3 to 6 global positioning system (GPS) locations per visit and recorded bird behaviour at each location to document breeding activity. Repeated visits enabled us to delineate territories based on clusters of GPS locations and the number of individuals we detected on each visit. We classified behavioural observations using a modified Vickery index of breeding status (Table 1; Vickery et al. 1992), providing evidence of nesting (i.e., adult(s) incubating a nest or carrying nest material or food) and fledging (i.e., adult(s) delivering food to multiple locations after evidence of a nest observed, flightless dependent fledglings, or adults carrying food for ≥ 11 days) in each bobolink territory. We considered adults carrying food to 1 location for ≥ 11 days as evidence of fledging because bobolink young remain in the nest for 10 to 11 days (Renfrew et al. 2015). We also included a behavioural observation category to record an adult bobolink behaviour we called agitated alarm calling, indicating the presence of mature nestlings or young fledglings.

During spot mapping, we also searched for bobolink nests. We used behavioural cues and systematic searching to locate nests (Martin and Geupel 1993, Winter et al. 2003). We did not approach nests when females were building to minimize disturbance when the chance of nest abandonment can be high (Renfrew et al. 2015). Once bobolink were incubating eggs, we visited nests approximately once every 3 days, on the expected fledge date, and on subsequent days until a nest was no longer active. On each visit, we recorded the number of eggs, number of young, age of young, condition of the nest, and adult behaviour. We considered several factors to determine if a nest fledged, including if we observed ≥ 1 flightless dependent fledgling near a nest, adults alarm calling near a nest that had large nestlings on the previous visit, and adults delivering food near a nest that had large nestlings on the previous visit.

Table 1: Behavioural classes recorded for each bobolink detected on surveys in fields. Behavioural classes are ranked from lowest (i.e., loafing) to highest (i.e., dependent fledgling(s) observed) evidence of breeding activity.

Code ^a	Description	Breeding behaviour ^b
LO	Loafing	No
VO	Vocalizing	No
FO	Foraging	No
TE	Territorial or aggressive interactions toward conspecifics or heterospecifics	No
CO	Courtship interactions; female-male interactions, displays, copulations, male feeding female	No
NB	Nest building, carrying nest material	Yes
IN	Incubating	Yes
FN	Food carry to nest	Yes
AA	Agitated alarm calling	Yes
FF	Food carry to fledgling(s)	Yes
FL	Dependent fledgling(s) observed (incapable of sustained flight, fed by adults)	Yes

^a AA code was not recorded for individual bobolink detections in 2017

^b CO was insufficient to confirm breeding

2.5 Transect surveys and point counts

We used aerial photographs in QGIS (version 2.14.19, qgis.osgeo.org, accessed 23 Sep 2017) to determine survey locations in each field. Transect surveys were located so that each transect line bisected the length of a field, with transect lines separated into 100-m segments. Each transect included at least one 100-m survey segment. Fields usually contained 1 transect line, but included 2 transect lines if their area was large enough to separate transect lines by ≥ 250 m to minimize the chance of counting the same individuals on multiple surveys. We walked a pace of 1 step per second and stopped when we detected a bobolink to record its distance from the transect line when first detected; we recorded detections ≤ 75 m of the transect line. We visited transects during 2 discrete time periods: 17 through 25 May and 08 through 17 June. To provide additional data about breeding success, we visited fields that had ≥ 1 bobolink territory, based on spot mapping, again from 01 through 04 July using the same transect survey protocol.

Most point count surveys were located approximately in the centre of each field. Fields usually contained 1 point count, but included multiple point counts if their area was large enough to separate point count centres by ≥ 250 m to minimize the chance of counting the same individuals on multiple surveys. We conducted 5-min, 75-m radius circular point counts, recording bobolink detections within distance bands of 0 to 25, > 25 to 50, and > 50 to 75 m. To estimate abundance, we visited point counts during 4 discrete time periods: 17 through 25 May, 29 May through 06 June, 08 through 17 June, and 20 through 29 June. To provide additional data about breeding success, we visited fields that had ≥ 1 bobolink territory, based on spot mapping, again from 01 through 04 July using the same point count protocol.

On both survey types, we recorded sex, detection method (i.e., song, call, visual), and behaviour for each individual. For behaviour, we recorded the highest ranked evidence of breeding based on a modified Vickery index (Table 1; Vickery et al. 1992). We considered a detection of 1 or more of the following to be evidence of nesting: nest building, incubating, food carry to nest, agitated alarm calling, food carry to fledgling(s), dependent fledgling(s) (Table 1). Likewise, detections of food carry to fledgling(s) or dependent fledgling(s) provided evidence of fledging.

We scheduled surveys across dates while most bobolink were nesting. We surveyed from sunrise to 1000 and avoided surveying during rain or excessive wind, which could affect bobolink activity and ability of observers to detect birds. Wind during surveys was ≤ 3 on the Beaufort scale during May and June and ≤ 5 in July (NOAA 2018). To minimize differences among the 4 observers, we regularly practiced surveys together with a laser rangefinder to ensure consistency of survey techniques and detections of bobolink.

2.6 Analysis

We used R (version 3.5.1, cran.r-project.org, accessed 23 Oct 2018) for all analyses. We first assessed the efficacy of transect surveys and point counts to capture bobolink abundance. We used the `cor.test` function to calculate the Spearman correlation (Zar 1999) between the number of territories in each field, based on spot mapping, compared to the maximum number of males detected in each field on any 1 visit on (1) transect surveys and (2) point counts. The maximum number of males detected in each field for transect surveys and point counts included all males detected across multiple transects lines or point count locations on any 1 visit to each field.

We used descriptive statistics to evaluate the efficacy of transect surveys and point counts to assess bobolink breeding success compared to spot mapping and nest monitoring. We calculated the percent of fields with (1) bobolink detections on surveys and territories based on spot mapping, (2) evidence of nesting, and (3) evidence of fledging.

We used distance sampling (Buckland et al. 1993a) to assess the overall number of bobolink territories impacted by grassland stewardship. We considered the delayed-harvest hay fields, restored grasslands, and fallow fields as 3 land cover types benefiting bobolink through stewardship. We treated this analysis as a demonstration of how to estimate the number of bobolink territories impacted by stewardship of fields enrolled in a stewardship program. We used the `Distance` package in R to conduct the distance sampling analysis. For analyses, we used detections of males only in May and June, excluding females and fledglings. We ran models for each key function (half normal, hazard, uniform) and compared relative model performance using Akaike's Information Criterion (AIC; Akaike 1974, Burnham and Anderson 2002). We included numerical covariates relevant to detection in models (i.e., date, min since sunrise). We z-transformed numerical covariates to ensure models converged. We did not include categorical covariates (e.g., observer) because models would not converge. We used the `gof_ds` function to apply the Cramer-von Mises test to evaluate goodness-of-fit and considered $\alpha < 0.05$ as evidence of poor model fit. We included transect length and

the sum of the area of all fields to enable estimates of abundance for surveyed fields. We compared the estimate of the abundance of males for all fields combined to the sum of territories for all fields, based on spot mapping. We excluded point counts from this analysis because models would not converge.

To further assess the ecological impact of stewardship, we used spot map data to estimate the number of bobolink territories in each field type (i.e., hay field, restored grassland, fallow field). Lastly, we quantified the observed percent of nests that fledged ≥ 1 young in each field type (i.e., uncorrected for nest survival).

3 Results

Bobolink territories occurred in 78% ($n = 36$) of fields, based on spot mapping (Table 2). Transect surveys correctly identified bobolink presence in all fields with territories and detected bobolink in 1 field that did not have a territory (Table 2). Point counts failed to detect bobolink in 2 fields with territories and detected bobolink in 1 field that did not have a territory (Table 2).

Table 2: The percent of fields with detections of bobolink, nesting, and fledging across 3 survey types (i.e., spot mapping and nest monitoring, transect survey, point count). We conducted surveys in hay fields, restored grasslands, and fallow fields in southern Ontario in 2018.

Survey type	Fields ($n = 36$)		
	% with detections ^a	% with nesting	% with fledging
Spot map	78	78	78
Transect	81	50	28
Point count	75	61	8

^a percent of fields with territories for spot map

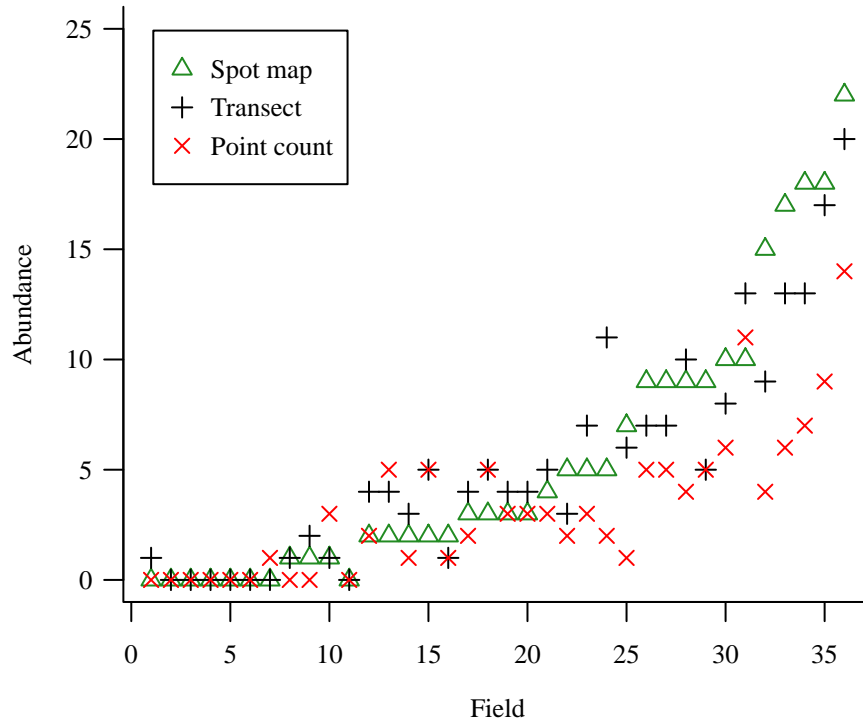


Figure 1: The number of bobolink territories in each field, based on spot mapping, compared to the maximum number of males detected in each field across 2 visits to transects and 4 visits to point counts in 36 fields in southern Ontario in 2018.

3.1 Abundance

The maximum number of male bobolink we detected on transect surveys in each field corresponded more closely to the number of territories in each field than did the maximum number of males detected on point counts (Figure 1). The maximum number of males detected in each field appeared to become increasingly biased low for point counts as the number of territories per field increased (Figure 2). The maximum number of males detected in each field on transect surveys was strongly correlated with the number of territories in each field, based on spot mapping ($S = 382.88$, $P < 0.001$, $\rho = 0.95$). Similarly, the maximum number of males detected in each field on point counts was also strongly correlated with the number of territories in each field, based on spot mapping ($S = 1117.2$, $P < 0.001$, $\rho = 0.86$).

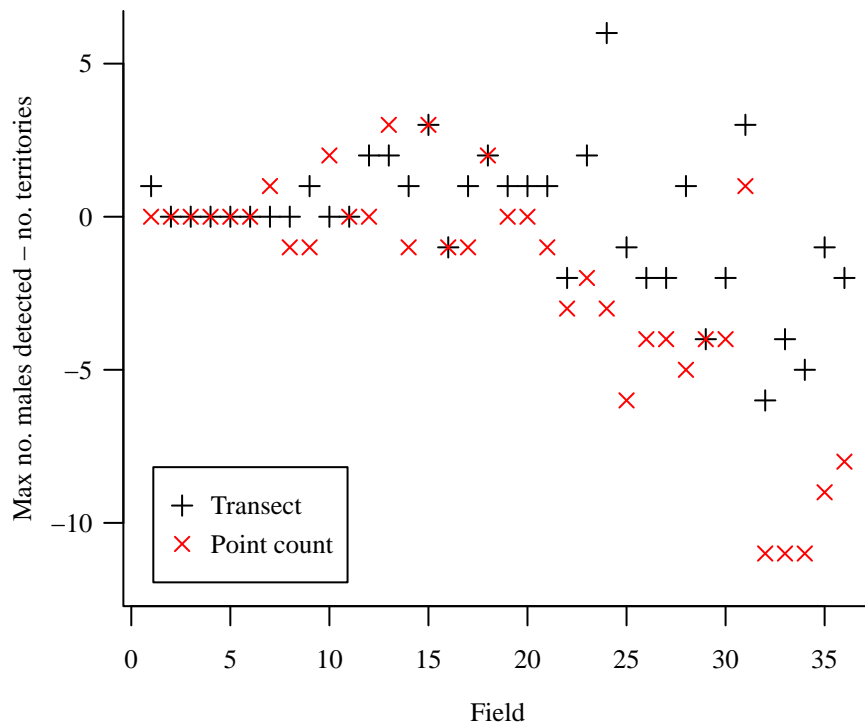


Figure 2: Difference in bobolink abundance per field, based on (1) the maximum number of males detected across visits to transect surveys minus the number of territories observed through spot mapping and (2) the maximum number of males detected across visits to point counts minus the number of territories observed through spot mapping.

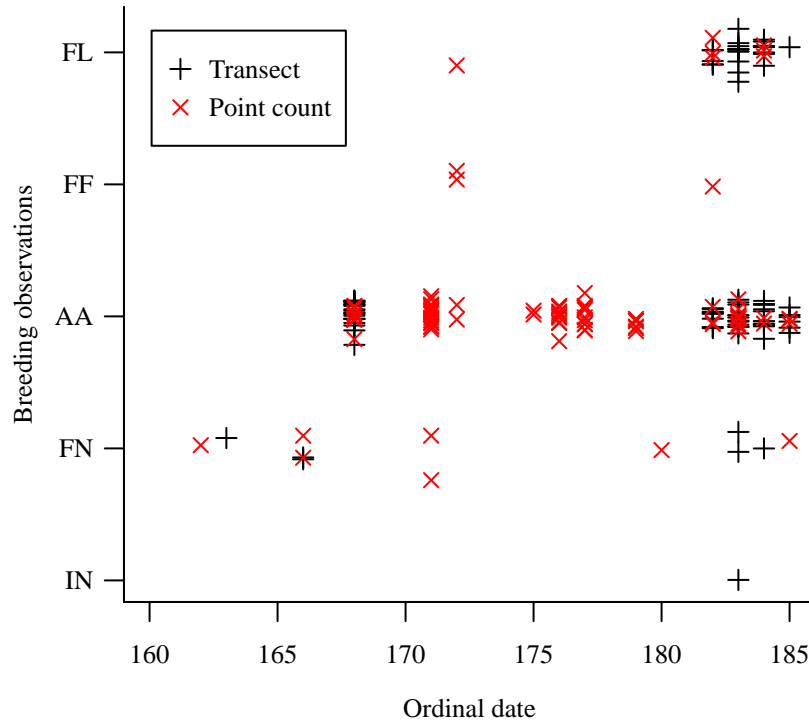


Figure 3: Detections of bobolink breeding behaviours across the breeding season on transect surveys and point counts in 36 fields in southern Ontario in 2018. Breeding behaviours were: IN = incubating; FN = food carry to nest; AA = agitated alarm calling; FF = food carry to fledgling(s); FL = dependent fledgling(s) observed (incapable of sustained flight, fed by adults). Ordinal date 160 = 09 June. Data points were jittered on the y-axis to make points visible.

3.2 Breeding success

Bobolink nested in and fledged young from all fields with territories (i.e., 78% [$n = 36$]), based on spot mapping and nest monitoring. We detected evidence of nesting in 61% of fields on point counts, compared to 50% of fields on transect surveys (Table 2). In contrast, we detected evidence of fledging in more fields on transect surveys (28%) than on point counts (8%; Table 2). We did not have false positive detections of nesting or fledging in fields on transect surveys and point counts. On transect surveys, we did not detect evidence of fledging until our third visit to fields from 01 to 04 July, although we did detect evidence of nesting in June during our second visit to fields (Figure 3). Agitated alarm calling by adults was the breeding behaviour we detected most frequently on transect surveys (65%, $n = 92$) and point counts (82%, $n = 97$).

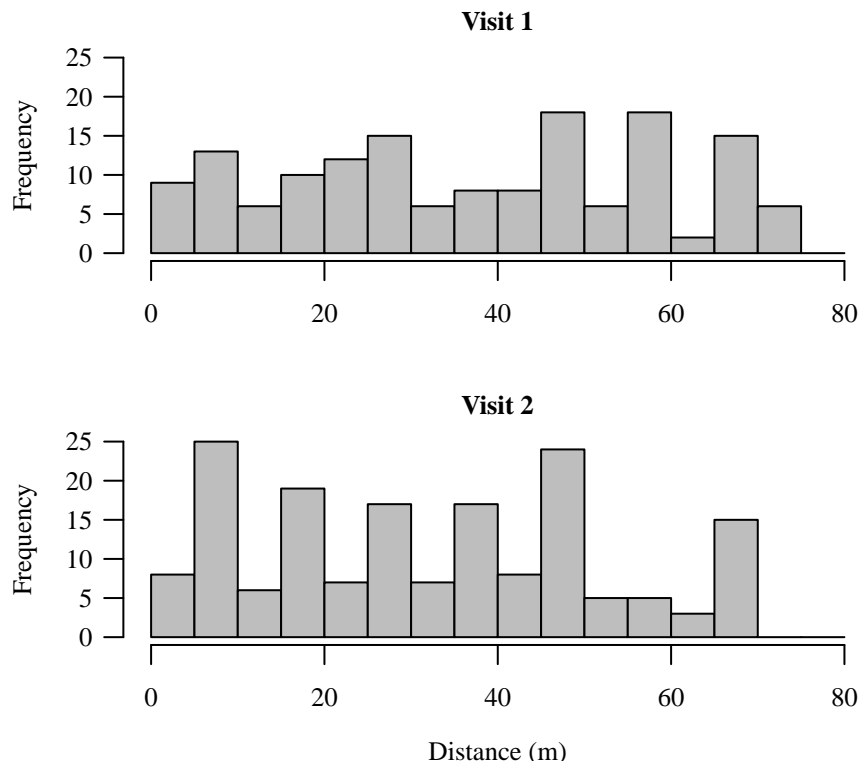


Figure 4: Histograms showing the number of male bobolink detected in 5-m distance bins on transect surveys in hay fields, restored grasslands, and fallow fields in southern Ontario in 2018.

3.3 Impact of stewardship

On transect surveys, we detected 318 male bobolink, 152 on visit 1 and 166 on visit 2 (Figure 4). We primarily detected bobolink visually (85%, $n = 318$).

For transects, the top distance sampling model had date and min since sunrise as covariates and a hazard-rate key function (Tables 3, 4). There was minimal model uncertainty because the second model had ΔAIC 3.62; however, this model was similar to the top model, lacking the min since sunrise covariate only (Table 3). Detection probability was high, 0.95 (SE = 0.04), based on the top model. Estimated abundance from the top model was 230 (95% CI: 187, 282) male bobolink in the 254 ha of surveyed fields, compared to 197 territories based on spot mapping. For additional comparison, the sum of the maximum number of males we detected on transect surveys in each field was 193.

Table 3: Model results for distance sampling from transect surveys for male bobolink in hay fields, restored grasslands, and fallow fields in southern Ontario in 2018.

Model ^a	Key function	k ^b	Delta AIC ^c	P ^d	Detection ^e	Abundance ^f
date + time	Hazard-rate	4	0.00	0.19	0.95	230
date	Hazard-rate	3	3.62	0.18	0.95	229
date + time	Half-normal	3	15.53	0.11	0.89	244
date	Half-normal	2	15.94	0.14	0.91	239
intercept	Half-normal	1	22.66	0.11	0.95	230
intercept	Uniform, cosine adjustment	1	23.10	0.11	0.96	228
intercept	Hazard-rate	2	23.91	0.12	0.96	227
time	Half-normal	2	24.00	0.20	0.94	232
time	Hazard-rate	3	24.14	0.27	0.96	227

^a date = z-transformed ordinal date, time = z-transformed min since sunrise

^b No. parameters in model

^c AIC for top model = 2,724.51

^d Cramer-von Mises goodness-of-fit test

^e Estimated probability of detection for male bobolink on transect surveys

^f Estimated abundance of male bobolink in 254 ha of surveyed fields

Table 4: Parameters for the top distance sampling model for transect surveys for male bobolink in hay fields, restored grasslands, and fallow fields in southern Ontario in 2018.

	Parameter ^a	Estimate	SE
Scale coefficients	intercept	4.946	0.318
	date	-0.706	0.292
	time	0.056	0.037
Shape coefficient	intercept	4.721	8.093

^a date = z-transformed ordinal date, time = z-transformed min since sunrise

Based on spot mapping data, there were 54 bobolink territories in the 9 hay fields (57 ha). Eight of the 18 restored fields (125 ha) did not have bobolink territories. There were 78 territories in the 10 restored fields (98 ha) where territories occurred. There were 65 bobolink territories in the 9 fallow fields (72 ha). Based on nest monitoring, 63% ($n = 59$) of nests fledged young in hay fields, 70% ($n = 100$) fledged in restored grasslands, and 70% ($n = 66$) fledged in fallow fields.

4 Discussion

Overall, we found that transect surveys provided useful data for assessing bobolink abundance and estimating the number of territories impacted by stewardship actions. Point counts were less effective than transect surveys for assessing bobolink abundance. Transect surveys and point counts were unreliable for assessing breeding success, underestimating the number of fields with nesting and fledging, compared to spot mapping and nest monitoring. Our

results suggest transect surveys are a promising field survey type for estimating the number of bobolink territories in fields during the breeding season; however, assessing breeding success likely requires more intensive methods.

We found detection probability of male bobolink on transect surveys was high (i.e., 0.95), which is largely consistent with previous estimates in the literature. Lueders et al. (2006) estimated detection probability of bobolink females and males as 1.0, based on distance sampling for point counts. Shustack et al. (2010) found that detection probability of male bobolink was > 0.99 within 50 m of point counts, based on Huggin's closed capture removal models. Rotella et al. (1999) found male bobolink detection probability was 0.91 at 50 m and 0.80 at 75 m, based on distance sampling for point counts. However, lower detection probabilities have been reported in the literature. Thompson et al. (2014) found detection probability for bobolink was 0.68 early in the breeding season and 0.34 during late-season point counts, based on multinomial-negative binomial mixture models. We suspect that detection probability is higher for male bobolink than females because (1) females are often inconspicuous during incubation (whereas, males do not incubate) and (2) males sing and females do not.

Because detection probability of males on transect surveys was high, the sum of the maximum number of males we detected in each field across 2 visits to transect surveys (193) was similar to the sum of the number of territories across all fields (197). Although mean abundance across multiple visits is often used as an index of abundance (Betts et al. 2005), we found that the maximum number of males detected across visits was an accurate estimate of the number of territories. Although indices of abundance have their uses for monitoring (Johnson 2008), not accounting for imperfect detection of birds on surveys can lead to biased abundance estimates (Nichols et al. 2000, Farnsworth et al. 2002, Buckland 2006). Surveying for index of abundance for bobolink may be acceptable in some cases, but is unlikely to be received well by peer reviewers. Additionally, biases could arise because of visual obstructions in fields, particularly shrub cover, because we detected most male bobolink visually. Impacts of detection probability could bias estimates of abundance across a gradient of percent shrub cover or temporal trends as shrub cover increases across years or shrub cover is cleared for grassland management. Probability of detecting bobolink could also vary by date and time, as we found, in addition to observers. We recommend that estimates of bobolink abundance include estimates of detection probability to provide rigorous results that are likely to be widely accepted by researchers, can be compared across time and space, and can be used to support conservation (Kellner and Swihart 2014, Hayward et al. 2015).

Surveys late in the breeding season may give higher estimates of abundance than early-season surveys for 2 reasons. We suspect that we detected more male bobolink on visit 2 than visit 1 because adults often gave conspicuous alarm calls and approached observers when caring for mature nestlings or young fledglings. Additionally, bobolink disperse from hay fields following mowing and pasture following heavy grazing during the breeding season. Dispersal during the breeding season is problematic for some methods of estimating occurrence or abundance, (e.g., the geographic closure assumption in occupancy modeling [MacKenzie et al. 2002]). We did not observe substantial immigration of adults to surveyed fields in June after hay harvest began in the study area. Conducting surveys after bobolink begin nesting, but before dispersal from fields disturbed by agricultural activity, will likely provide the best estimates of male abundance. The best timing for surveys in our study area would likely be during the last week of May and first week of June.

Using distance sampling, we obtained a reasonable estimate of male bobolink abundance across all of the fields we surveyed. Although the mean estimate was higher than the number of territories, the number of territories was within the 95% confidence interval. Distance sampling appears promising for estimating the number of bobolink territories impacted by future stewardship programs. As with any method, there are challenges to overcome and assumptions to meet. For example, a minimum of 60 to 80 detections of the species of interest are recommended to analyze data using distance sampling (Buckland et al. 1993b). Additionally, substantial effort is needed to design surveys and meet the assumptions of distance sampling (Buckland et al. 1993b). Given various analysis options and criticisms thereof (e.g., Johnson 2008), distance sampling is a viable option that is well accepted in the literature and fairly straightforward to implement in the field. Tools are also available for planning study design and sampling for research and monitoring (Thomas et al. 2010).

We found that delayed-harvest hay fields, restored grasslands, and fallow fields provided nesting habitat for bobolink. If we make some assumptions about land use in these fields in the absence of stewardship, then we can evaluate the likely ecological impacts on bobolink. For hay fields, we assumed that the same number of bobolink territories would occur. For restored grasslands and fallow fields, we assumed that previous agricultural use of fields for row crops (e.g., soy) would not have provided nesting habitat for bobolink, except for 1 fallow field which was previously a hay field. The 54 bobolink territories in hay fields were spared the impact of mowing that would have likely occurred during the breeding season, inadvertently destroying some nests or leaving them without vegetative cover

and vulnerable to predation. The 78 territories in restored fields would not have existed if fields were still farmed with row crops. Bobolink may not have established territories in 8 of the 18 restored fields because vegetation had not yet developed enough since planting to provide sufficient cover; some of these fields had patches of bare ground. Unlike restored fields, all fallow fields had bobolink territories. Fifty-eight of the 65 territories in fallow fields would not have existed if fields were still farmed with row crops. In the fallow field that was previously hay, there were likely bobolink. However, these nests would have been susceptible to hay harvest; whereas, in the fallow fields nests were undisturbed throughout the breeding season. More than half of nests fledged young in each field type. However, observed nest success we report here is preliminary; we are analyzing these data elsewhere to correct for biases in observed nest success. These 3 stewardship practices provided nesting habitat for bobolink through different means that have their own trade-offs for landowners.

Transect surveys and point counts did not effectively detect evidence of bobolink nesting and fledging in fields. Although we did not find previous assessments of detecting evidence of breeding on low-intensity surveys for bobolink, other studies have documented weaknesses in the Vickery index (Vickery et al. 1992) for other species of grassland birds (Rivers et al. 2003, Morgan et al. 2010). We think low-intensity surveys (i.e., transect surveys) are best suited to document male bobolink abundance, rather than breeding success. If breeding success is of interest to managers of a conservation program, then perhaps spot mapping (*sensu* Wiens 1969) and nest monitoring (Martin and Geupel 1993, Winter et al. 2003) could be implemented at a representative subset of fields where stewardship practices are implemented.

The time and effort needed to implement field monitoring is often a concern for conservation programs. Anecdotally, transect surveys and point counts took roughly the same effort to plan and implement in the field per visit. Bollinger et al. (1988) quantified survey effort for bobolink and found that transect surveys were slightly more efficient than point counts. We suspect that transect surveys may be a more efficient use of effort because we were able to collect data useful for estimating abundance from 2 visits to transects; whereas, 4 visits to point counts were of limited use for analysis. However, it is possible that point counts would have been more useful for analysis using distance sampling if we estimated exact distances to birds, rather than recording detections in distance bins.

5 Monitoring scheme

The purpose of this monitoring scheme is to assess the abundance and breeding success of bobolink in fields (e.g., hay field, pasture, restored grassland) that are being managed to provide breeding areas for the species. Assessing abundance and breeding success is important for understanding the impact of stewardship programs. Assessments of stewardship actions should provide a feedback loop into stewardship programs for program evaluation and adaptation of program actions to benefit bobolink. We focus on field methods to provide reliable information about bobolink use of fields during the breeding season and ideas for study designs to address particular stewardship actions.

5.1 Monitoring bobolink abundance

We recommend using transect surveys to estimate bobolink abundance in fields managed for conservation. Transect surveys should be located so that each transect line bisects the length of a field, with transect lines separated into 100-m segments. Each transect should contain the number of 100-m segments needed to bisect the entire field. Each field contains at least one transect line, but should include multiple transect lines if the area of a field is large enough to separate transect lines by ≥ 250 m to minimize the chance of counting the same individuals on multiple surveys. Use GIS and aerial photographs to determine locations of transects in each field before conducting surveys. Free open-source GIS software is available (e.g., QGIS: <https://www.qgis.org/en/site/>). Additionally, aerial photos are available from Land Information Ontario (<https://www.javacoeapp.lrc.gov.on.ca/geonetwork/srv/en/main.home>).

A surveyor (i.e., field biologist) visits each transect twice from 21 May (i.e., coinciding with first-egg date) to 30 June (i.e., before bobolink form flocks). On each survey, a surveyor walks a pace of one step per second. The surveyor stops walking to record each bobolink detected by sight or sound within 75 m of either side of the transect line. The surveyor notes the detection method (i.e., detected by sight, call, or song), sex, and perpendicular distance from the transect line. Surveyors require practice before and occasionally during data collection to ensure consistency in data collection, correct identification of bobolink, and accurate distance estimates (practicing with a laser rangefinder is recommended). Surveyors must also be familiar with best practices to ensure minimal disturbance to all nesting birds in fields.

By estimating distance to bobolink detected on transect surveys, detection probability and the sum of abundance across fields can be estimated using distance sampling (Buckland et al. 1993a). Although various analysis methods are available for estimating songbird abundance, we found that distance sampling was appropriate for bobolink surveyed from transects. Effort is needed to plan study design and sampling (Buckland et al. 1993, Thomas et al. 2010). Free analysis software is available (Thomas et al. 2010, <http://distancesampling.org/>). Biologists or conservation practitioners should consult statisticians or others familiar with distance sampling and study design before planning monitoring to ensure program goals can be achieved.

5.2 Monitoring bobolink breeding success

If information about breeding success is needed for fields managed for conservation, then we recommend nest monitoring. Nest monitoring is labour intensive and thus, it may be necessary to select a representative subset of fields to monitor.

Spot mapping is useful for delineating and monitoring bobolink territories (*sensu* Wiens 1969). From mid-May through mid-July, surveyors visit each field approximately twice per week to search for bobolink. When a bobolink is detected, a surveyor follows the individual or pair for up to 30 min, if possible. For each individual or pair, a surveyor records 3 to 6 GPS locations per visit. GPS locations are downloaded to GIS software. Repeated visits enable delineation of territories based on clusters of GPS locations and the number of individuals detected on each visit to each field.

During spot mapping, surveyors also search for bobolink nests. Surveyors use behavioural cues and systematic searching to locate nests (Martin and Geupel 1993, Winter et al. 2003). Surveyors should not approach nests when females are building to minimize disturbance when the chance of nest abandonment can be high (Renfrew et al. 2015). Once bobolink are incubating eggs, surveyors visit nests approximately once every 3 days, on the expected fledge date, and on subsequent days until a nest is no longer active. On each visit, surveyors record the number of eggs, number of young, age of young, condition of the nest, and adult behaviour. Expected fledge dates for nests can be estimated based on breeding phenology: typical duration of nesting is ~28 days (1 egg laid per day [mean clutch size = ~5], incubation begins on the day the last egg is laid, 12 days of incubation, 11 days from hatch to fledge; Renfrew et al. 2015). Several factors need to be considered to determine if a nest fledges young. Observations such as ≥ 1 flightless dependent fledgling near a nest, adults alarm calling near a nest that had large nestlings on the previous visit, and adults delivering food near a nest that had large nestlings on the previous visit may indicate fledging, depending on particulars of each nest (e.g., proximity of nearby nests). Surveyors familiar with monitoring songbird nests are needed to ensure minimal disturbance to nesting birds, appropriate data collection, and interpretation of observations and bird behaviour. Additionally, permits are needed from various jurisdictions to monitor nests of migratory birds and species at risk.

By monitoring an adequate number of nests, the daily survival rate of nests can be estimated to correct for biases in the observed percent of nests that fledge young (Dinsmore et al. 2002, Shaffer 2004). This analysis enables estimates that can be compared to other research and quantification of relationships with predictor variables (e.g., shrub cover, field size, distance to field edge) that may be of interest to conservation practitioners. Again, biologists or conservation practitioners should consult statisticians or others familiar with nest survival analyses and study design before planning monitoring.

5.3 Study designs to assess stewardship actions

Stewardship actions currently implemented to benefit nesting bobolink are identified in the 2018 Grassland Stewardship Program from the Ontario Soil and Crop Improvement Association (OSCIA 2018). The program identified 4 actions: (1) mowing to remove encroaching trees and shrubs from fields, (2) grassland restoration, (3) delayed grazing in rotational-grazing systems, and (4) delayed haying. Monitoring the ecological response of bobolink to each practice will require particular study designs. The field methods outlined above would need to be integrated into a study design specific to each stewardship action.

Impact assessment designs would likely be useful to monitor the effects of tree and shrub removal from fields on nesting bobolink (Morrison et al. 2008). Ideally, a field scheduled to receive the treatment (i.e., tree and shrub removal) would be sampled for bobolink during year 1. The treatment would occur after the breeding season in year 1 (e.g., in the fall). During year 2, bobolink would be monitored in the field again to enable comparisons with year 1.

To evaluate impacts on bobolink across a number of fields involved in a conservation program, treatments should occur in multiple years to reduce confounding the effect of year and treatment. For example, if all fields receive the treatment in year 1, then differences found in year 2 could be attributed to inter-annual variation as well as the treatment. Additionally, it may take ≥ 1 year for bobolink to respond to this stewardship action; therefore, monitoring in subsequent years (i.e., after year 2) may be necessary. Lastly, it may also be helpful to monitor control fields, which could have received the tree and shrub removal treatment, but were not selected for the program (Morrison et al. 2008).

Grassland restorations require years to establish; thus, multiple-time designs (Morrison et al. 2008) may be useful to monitor the response of bobolink using the field methods above. Fields restored to grassland would need to be monitored for several years to assess use by bobolink. Again, to reduce year and treatment being confounded, treatment of multiple fields should occur across multiple years to enable an assessment of the stewardship action in general.

Delayed grazing and haying primarily impacts bobolink in the same year the practice is implemented. The field methods above could be used to assess the abundance and breeding success of bobolink in a particular field to quantify the impact of the stewardship action. To assess the impact of the practice across a stewardship program, particular metrics of bobolink in fields receiving delayed grazing or haying could be compared with conditions in other fields (e.g., farmer-directed grazing based on production needs). Such a comparison could estimate the benefits for bobolink (e.g., the number of male bobolink in fields with delayed grazing compared to the number in fields disturbed by grazing). For fields where grazing or haying is delayed in multiple sequential years, a multi-year monitoring plan could provide information about long-term effects on bobolink abundance in such a field (e.g., if abundance increases across years where haying is delayed in multiple sequential years).

6 Acknowledgments

We thank participating farmers and the Grand River Conservation Authority for the opportunity to study bobolink on their land. We are grateful for assistance in the field from M. Bateman, N. Conroy, M. Fromberger, J. Horvat, G. Morris, J. Put, and D. Stonley. R. Dobson, D. Ethier, C. Lituma, C. Risley, and L. Van Vliet provided productive discussion about and review of this work. Funding was provided by (1) the Government of Ontario, (2) the Ontario Soil and Crop Improvement Association through the Species at Risk Partnerships on Agricultural Lands (SARPAL) program, an Environment and Climate Change Canada initiative, (3) Mitacs, (4) Echo Foundation, (5) Colleges and Institutes Canada, Career-Launcher Internship program, and (6) individual donors.

Assistance for this project was provided by the Government of Ontario.

This work was supported by Mitacs through the Mitacs Accelerate program.

This project was undertaken with the financial support of the Government of Canada through the federal Department of the Environment and Climate Change. Ce projet a été réalisé avec l'appui financier du gouvernement du Canada agissant par l'entremise du ministère fédéral de l'Environnement et du Changement climatique.

Funding for this initiative is provided by the Government of Canada through Natural Resources Canada's Green Jobs - Science and Technology Internship Program, as part of the Youth Employment Strategy. Le financement de cette initiative est fourni par le gouvernement du Canada par l'entremise du Programme de stages en sciences et technologie - Emplois verts de Ressources naturelles Canada dans le cadre de la Stratégie emploi jeunesse.

7 Literature cited

- Akaike, H. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19:716-723.
- Betts, M. G., N. P. P. Simon, and J. J. Nocera. 2005. Point count summary statistics differentially predict reproductive activity in bird-habitat relationship studies. *Journal of Ornithology* 146:151-159.
- Block, W. M., A. B. Franklin, J. P. Ward, J. L. Ganey, and G. C. White. 2001. Design and implementation of monitoring studies to evaluate the success of ecological restoration on wildlife. *Restoration Ecology* 9:293-303.

- Bollinger, E. K., T. A. Gavin, and D. C. McIntyre. 1988. Comparison of transects and circular-plots for estimating bobolink densities. *Journal of Wildlife Management* 52:777-786.
- Brennan, L. A., and W. P. Kuvlesky Jr. 2005. North American grassland birds: an unfolding conservation crisis. *Journal of Wildlife Management* 69:1-13.
- Buckland, S. T. 2006. Point-transect surveys for songbirds: robust methodologies. *Auk* 123:345-357.
- Buckland, S. T., D. R. Anderson, K. P. Burnham, and J. L. Laake. 1993*a*. Distance sampling: estimating abundance of biological populations. Springer.
- Buckland, S. T., D. R. Anderson, K. P. Burnham, and J. L. Laake. 1993*b*. Study design and field methods. Pages 295-349 *in* Distance sampling: estimating abundance of biological populations. Springer.
- Burnham, K. P., and D. R. Anderson. 2002. Model selection and multimodel inference. Springer-Verlag, New York, USA.
- Christoferson, L. L., and M. L. Morrison. 2001. Integrating methods to determine breeding and nesting status of 3 western songbirds. *Wildlife Society Bulletin* 29:688-696.
- Committee on the Status of Endangered Wildlife in Canada (COSEWIC). 2010. COSEWIC assessment and status report on the bobolink *Dolichonyx oryzivorus* in Canada. Ottawa, Canada.
- Diefenbach, D. R., D. W. Brauning, and J. A. Mattice. 2003. Variability in grassland bird counts related to observer differences and species detection rates. *Auk* 120:1168-1179.
- Dinsmore, S. J., G. C. White, and F. L. Knopf. 2002. Advanced techniques for modeling avian nest survival. *Ecology* 83:3476-3488.
- Farnsworth, G. L., K. H. Pollock, J. D. Nichols, T. R. Simons, J. E. Hines, and J. R. Sauer. 2002. A removal model for estimating detection probabilities from point-count surveys. *Auk* 119:414-425.
- Government of Canada. 2017. Species at risk public registry. https://www.registrelep-sararegistry.gc.ca/species/schedules_e.cfm?id=1. Accessed 13 Dec 2017.
- Hayward, M. W., L. Boitani, N. D. Burrows, P. J. Funston, K. U. Karanth, D. I. MacKenzie, K. H. Pollock, and R. W. Yarnell. 2015. Ecologists need robust survey designs, sampling and analytical methods. *Journal of Applied Ecology* 52:286-290.
- Johnson, D. H. 2008. In defense of indices: the case of bird surveys. *Journal of Wildlife Management* 72:857-868.
- Johnson, D. H. 2012. Monitoring that matters. Pages 54-73 *in* R. A. Gitzen, J. J. Millspaugh, A. B. Cooper, and D. S. Licht, editors. Design and analysis of long-term ecological monitoring studies. Cambridge University Press, Cambridge, UK.
- Jones, J., W. J. McLeish, and R. J. Robertson. 2000. Density influences census technique accuracy for cerulean warblers in eastern Ontario. *Journal of Field Ornithology* 71:46-56.
- Kellner, K. F., and R. K. Swihart. 2014. Accounting for imperfect detection in ecology: a quantitative review. *PLoS ONE* 9:e111436.
- Lueders, A. S., P. L. Kennedy, and D. H. Johnson. 2006. Influences of management regimes on breeding bird densities and habitat in mixed-grass prairie: an example from North Dakota. *Journal of Wildlife Management* 70:600-606.
- MacKenzie, D. I., J. D. Nichols, B. L. Gideon, S. Droege, J. A. Royle, and C. A. Langtimm. 2002. Estimating site occupancy rates when detection probabilities are less than one. *Ecology* 83:2248-2255.
- Martin, T. E., and G. R. Geupel. 1993. Nest-monitoring plots: methods for locating nests and monitoring success. *Journal of Field Ornithology* 64:507-519.
- McCracken, J. D., R. A. Reid, R. B. Renfrew, B. Frei, J. V. Jalava, A. Cowie, and A. R. Couturier. 2013. Recovery Strategy for the bobolink (*Dolichonyx oryzivorus*) and eastern meadowlark (*Sturnella magna*) in Ontario. Ontario Recovery Strategy Series. Prepared for the Ontario Ministry of Natural Resources, Peterborough, Canada.

- Ministry of Natural Resources and Forestry (MNR). 2015. Bobolink and eastern meadowlark Government Response Statement. <https://www.ontario.ca/page/bobolink-and-eastern-meadowlark-government-response-statement>. Accessed 10 Nov 2017.
- Ministry of Natural Resources and Forestry (MNR). 2017. Species at risk in Ontario. <https://www.ontario.ca/environment-and-energy/species-risk-ontario-list>. Accessed 13 Dec 2017.
- 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.
- Morrison, M. L., W. M. Block, M. D. Strickland, B. A. Collier, and M. J. Peterson. 2008. *Wildlife study design*. Second edition. Springer, New York, USA.
- Newell, F. L., J. Sheehan, P. B. Wood, A. D. Rodewald, D. A. Buehler, P. D. Keyser, J. L. Larkin, T. A. Beachy, M. H. Bakermans, T. J. Boves, A. Evans, G. A. George, M. E. McDermott, K. A. Perkins, M. B. White, and T. B. Wigley. 2013. Comparison of point counts and territory mapping for detecting effects of forest management on songbirds. *Journal of Field Ornithology* 84:270-286.
- National Oceanic and Atmospheric Administration (NOAA). 2018. Beaufort wind scale. <https://www.spc.noaa.gov/faq/tornado/beaufort.html>. Accessed 30 Nov 2018.
- Nichols, J. D., J. E. Hines, J. R. Sauer, F. W. Fallon, J. E. Fallon, and P. J. Heglund. 2000. A double-observer approach for estimating detection probability and abundance from point counts. *Auk* 117:393-408.
- Ontario Soil and Crop Improvement Association (OSCIA). 2018. Grassland Stewardship Program. <https://www.ontariosoilcrop.org/oscia-programs/sarpal/gsp/>. Accessed 14 February 2018.
- Partners in Flight Science Committee. 2013. Population Estimates Database, version 2013. <http://pif.birdconservancy.org/PopEstimates/>. Accessed 04 Oct 2018.
- Renfrew, R., A. M. Strong, N. G. Perlut, S. G. Martin, and T. A. Gavin. 2015. Bobolink (*Dolichonyx oryzivorus*), *The Birds of North America Online* in A. Poole, editor. Cornell Lab of Ornithology, Ithaca, 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.
- Rotella, J. J., E. M. Madden, and A. J. Hansen. 1999. Sampling considerations for estimating density of passerines in grasslands. *Studies in Avian Biology* 19:237-243.
- Sauer, J. R., W. A. Link, J. E. Fallon, K. L. Pardieck, and D. J. Ziolkowski Jr. 2013. The North American Breeding Bird Survey 1966-2011: summary analysis and species accounts. *North American Fauna* 79:1-32.
- Shaffer, T. L. 2004. A unified approach to analyzing nest success. *Auk* 121:526-540.
- Shustack, D. P., A. M. Strong, and T. M. Donovan. 2010. Habitat use patterns of bobolinks and Savannah sparrows in the northeastern United States. *Avian Conservation and Ecology* 5(2):11.
- Thomas, L., S. T. Buckland, E. A. Rexstad, J. L. Laake, S. Strindberg, S. L. Hedley, J. R. B. Bishop, T. A. Marques, and K. P. Burnham. 2010. Distance software: design and analysis of distance sampling surveys for estimating population size. *Journal of Applied Ecology* 47:5-14.
- Thompson, S. J., T. W. Arnold, and C. L. Amundson. 2014. A multiscale assessment of tree avoidance by prairie birds. *Condor: Ornithological Applications* 116:303-315.
- Thompson, W. L., G. C. White, and C. Gowan. 1998. *Birds*. Pages 261-300. *Monitoring vertebrate populations*. Academic Press, San Diego, USA.
- Vickery, P. D., M. L. Hunter Jr., and J. V. Wells. 1992. The use of a new reproductive index to evaluate relationships between habitat quality and breeding success. *Auk* 109:697-705.
- Wiens, J. A. 1969. An approach to the study of ecological relationships among grassland birds. *Ornithological Monographs* 8:1-93.
- Winter, M., S. E. Hawks, J. A. Shaffer, and D. H. Johnson. 2003. Guidelines for finding nests of passerine birds in tallgrass prairie. *Prairie Naturalist* 35:197-211.

Zar, J. H. 1999. Simple linear correlation. Pages 377-412. Biostatistical analysis. Fourth edition. Prentice-Hall, Upper Saddle River, USA.