Copeia

founded in 1913 by John Treadwell Nichols

Breeding Chorus Indices Are Weakly Related to Estimated Abundance of Boreal Chorus Frogs

Paul Stephen Corn[1], Erin Muths[2], Amanda M. Kissel[2,3], and Rick D. Scherer[4]

Official Publication of the American Society of Ichthyologists and Herpetologists

Breeding Chorus Indices Are Weakly Related to Estimated Abundance of Boreal Chorus Frogs

Paul Stephen Corn[1], Erin Muths[2], Amanda M. Kissel[2,3], and Rick D. Scherer[4]

Call surveys used to monitor breeding choruses of anuran amphibians generate index values that are frequently used to represent the number of male frogs present, but few studies have quantified this relationship. We compared abundance of male Boreal Chorus Frogs (*Pseudacris maculata*), estimated using capture–recapture methods in two populations in Colorado, to call index values derived from automated recordings. Single index values, such as might result from large monitoring efforts, were unrelated to population size. A synthetic call saturation index (CSI), the daily proportion of the maximum possible sum of index values derived from multiple recordings, was greater in larger populations, but the relationship was not highly predictive.

AUDITORY surveys of breeding anurans are a common tool used to verify distributions, investigate ecological relationships, and monitor trends of populations at various geographic and temporal scales (Dorcas et al., 2009). Call survey data are recorded typically on an integer scale of one to three (Bishop et al., 1997; Weir and Mossman, 2005), although Brodman (2009) used a five-point scale proposed by Karns (1986). Qualitatively, these scales indicate the number of frogs heard calling (one or two, a few, many). However, based on the reasonable expectation that a larger number of frogs should generate a greater amount of sound, there has long been a desire to treat these numbers as quantitative indices to population size. Some authors have done that explicitly (Fahrig et al., 1995; Mossman et al., 1998; Mazerolle, 2005; Eigenbrod et al., 2008; Brodman, 2009), but the more common approach is to treat these data conservatively as indicating whether a species was present or not detected, and use occupancy analysis to estimate detection probability (Weir et al., 2005; Pellet and Schmidt, 2005; Jackson et al., 2006; Sung et al., 2006; Brander et al., 2007).

Early in the development of call surveys as a sampling method, Vogt and Hine (1982) stated that for estimating abundance, ''Call indices for frogs are of limited use . . . '', and concluded, ''Its value as a census technique is generally limited to data on species presence.'' Vogt and Hine objected that in some species not all males in a breeding congregation may call, but that other species may have dense choruses of hundreds of calling males, making individual calls impossible to count, and they described biotic and abiotic factors that may influence data collected with call surveys, all of which have been verified by subsequent studies. These include that timing and intensity of calling activity varies among species and season (Bishop et al., 1997; Crouch and Paton, 2002; de Solla et al., 2006; Saenz et al., 2006), time of day (Bridges and Dorcas, 2000; Oseen and Wassersug, 2002), and may be dependent on the weather, particularly temperature and precipitation (Oseen and Wassersug, 2002; Kirlin et al., 2006; Saenz et al., 2006; Tupper et al., 2007; Steelman and Dorcas, 2010). Research also introduced the additional problem of observer bias, i.e., different observers often assign different index scores to a given chorus of frogs (Bishop et al., 1997; Shirose et al., 1997; Hemesath, 1998; Burton et al., 2006; Pierce and Gutzwiller, 2007) and may disagree on which species are present (Genet and Sargent, 2003; Lotz and Allen, 2007; Pierce and Gutzwiller, 2007).

Most call surveys are conducted manually (i.e., by humans who go to a field site and listen to frog choruses), and in large monitoring programs a single index value would be used to represent the relative abundance of a species at a site. The standard protocol of the North American Amphibian Monitoring Program (NAAMP) specifies that survey routes should be completed at least three times in a breeding season (Weir and Mossman, 2005). However, this is intended more to capture inter-specific variation in calling activity than to provide multiple observations for species at individual sites. To avoid the problem of intra-specific seasonal variation in activity (Crouch and Paton, 2002) and the statistical complications of repeated measures (Weir et al., 2009), analyses of call data as indices of abundance should be restricted to the observation obtained during a species' maximum breeding activity. This was the approach used by Mossman et al. (1998), but restricting the data to a single observation likely exacerbates problems with other sources of variation (environmental influences and observer bias). One potential solution may lie with the use of automated recording systems (ARS; Peterson and Dorcas, 1994; Dorcas et al., 2009), which can generate large amounts of data. The use of ARS has enormous potential to expand our understanding of anuran breeding phenology by allowing larger scale investigations and comparisons across landscapes. However, to take advantage of ARS, the relationship between the information we derive from sound recordings and the information gathered from more intensive population level sampling must be quantified. We explore the use of ARS (sometimes termed ''frogloggers'') to evaluate how both manual surveys and parameters derived from multiple recordings compare to abundance of Boreal Chorus Frogs (*Pseudacris maculata*) estimated using capture–recapture (CR) methods.

MATERIALS AND METHODS

We recorded calling *P. maculata* at two ponds in the Rocky Mountains in Larimer County, Colorado (Lily Pond [2969 m

[1] U.S. Geological Survey, Northern Rocky Mountain Science Center, Aldo Leopold Wilderness Research Institute, 790 E. Beckwith Avenue, Missoula, Montana 59801; E-mail: scorn@usgs.gov. Send reprint requests to this address.
[2] U.S. Geological Survey, Fort Collins Science Center, 2150 Centre Avenue, Building C, Fort Collins, Colorado 80526; E-mail: (EM) muthse@usgs.gov.
[3] Present address: Department of Biological Sciences, Simon Fraser University, Burnaby, BC V5A 1S6, Canada; E-mail: akissel@sfu.ca.
[4] Colorado State University, Department of Fish, Wildlife, & Conservation, Fort Collins, Colorado 80523; E-mail: scherer@rams.colostate.edu.
Submitted: 8 December 2010. Accepted: 4 May 2011. Associate Editor: M. J. Lannoo.
© 2011 by the American Society of Ichthyologists and Herpetologists DOI: 10.1643/CH-10-190

elevation] and Matthews Pond [2803 m]), where we have observed breeding by this species annually since 1986 (Corn et al., 1989, 2000; Corn and Muths, 2002). Both ponds are shallow (<2 m deep), fishless vernal wetlands situated within a matrix of second-growth forest (Vertucci and Corn, 1996). They vary in size and depth depending on winter snow pack. Lily Pond averages about 0.5 ha and Matthews Pond about 0.07 ha during breeding, and they typically dry completely by late summer. Wood Frogs (*Lithobates sylvaticus*) also breed at both ponds.

We deployed ARS at Lily Pond beginning in 1994 and Matthews Pond in 1995. Through 2005, we used analog ARS constructed using a stereo cassette tape recorder controlled by a solid-state timer (Peterson and Dorcas, 1994; Corn et al., 2000; Saenz et al., 2006). One channel recorded frog calls from an external omni-directional microphone, while the other channel recorded the time from a voice clock activated at the beginning of each recording with a microphone inside the case of the ARS. Calls were recorded for 12 sec every 30 min in 1994–1995, and 18 sec every hour in 1996–2005. In 2006–2007, we used ARS constructed around a monaural digital voice recorder (DVR, Diasonic Technology Ltd., model DDR-5100, 512 megabyte capacity) using an external omni-directional microphone and powered using external D-cell batteries through the DC input. The DVR could be programmed for ten daily recordings; we made 54-sec recordings hourly from 1600 to 0100 (1-min recordings were programmed; this discrepancy appears to be characteristic of the DVR). In 2008–2009, we deployed ARS designed for acoustic monitoring (Songmeter SM-1, Wildlife Acoustics, Inc.). The Songmeter has flexible programming capabilities, records in stereo using two external omni-directional microphones, and stores digital files on removable SD cards. We made hourly 5-min recordings at both locations. As with the analog ARS, recordings with the SM-1 were made 24 hr per day.

All recordings (analog and digital) were scored by a human listener. Only the first 15 sec of the digital recordings were assessed to keep the data consistent with the earlier tape recordings. The listener scored recordings with calling according to the standard three-point NAAMP index (Weir and Mossman, 2005): 1 = individual calls distinguishable, no overlap; 2 = individual calls distinguishable, some overlap; 3 = full chorus with continuous, overlapping calling. Our only deviation from this scoring was for recordings with only two or three clearly identifiable frogs calling. In this case, if a frog began calling just before the end of another frog's call, this was still scored as a 1, despite a small amount of overlap between calls.

We estimated the number of male *P. maculata* present each year at each site, based on two to four (usually three) sampling occasions in May and early June. Sampling occasions lasted 2–4 hr, beginning after sunset. The portion of each pond accessible by wading was searched, and frogs were captured by hand, given a unique mark if newly captured, assessed for color phenotype (Matthews and Pettus, 1966), and released at the point of capture. Marks were toe clips according to the system of Martof (1953), except that we removed no more than one toe per foot and did not remove thumbs. Sampling occasions averaged about four days apart (maximum 14 days between session 1 and session 3) and occurred during the period of peak calling (Fig. 1). Therefore, we assumed that each population was closed to emigration and immigration during sampling and

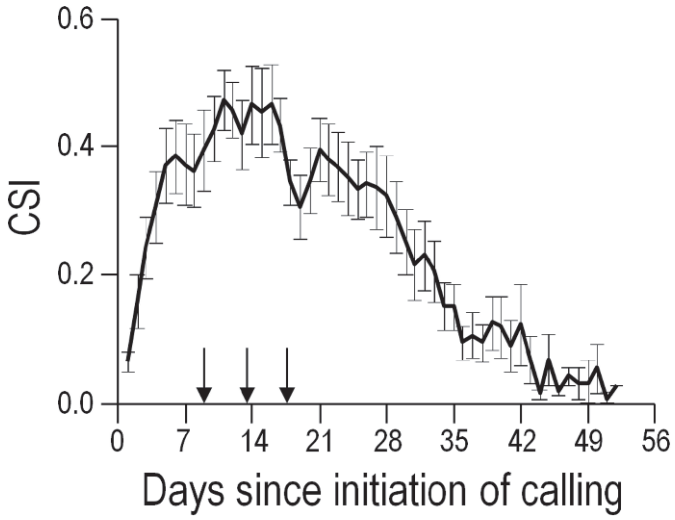

Fig. 1. Mean (±1 SE) daily call saturation index (for eight years at Lily Pond and ten years at Matthews Pond combined) relative to the initiation of calling activity. Data from years without recordings or when recordings failed to determine the initiation of breeding activity were not used in this figure. Arrows indicate the mean dates for the first three CR occasions.

used the closed capture–recapture model in program MARK (White and Burnham, 1999) to estimate population size at each site for each year. The closed capture–recapture model comprises parameters that represent capture probability (p_i, the probability of first capture on occasion i, and c_i, the probability of recapture on occasion i) and estimated number of adult males ($\hat{N}$). We evaluated three of the standard models of capture probability (Otis et al., 1978): M_0—capture probability is constant across individuals and sampling occasions within each year (null model); M_b—capture probability varies between individuals that have or have not been captured previously (behavior effect) but does not vary across sampling occasions; and M_t—capture probability varies across sampling occasions but not across individuals. Akaike's Information Criterion for small sample sizes (QAICc) was used to determine which of the models in the set best represented the data.

To assess the capacity for manual call surveys to serve as an index to population size of *P. maculata*, we tabulated the frequencies of index values over 11 nights centered on the date of peak breeding activity and between 2100 and 0100 hr (the interval prescribed in NAAMP protocols; Weir and Mossman, 2005). We defined the date of peak calling activity annually as the mid-point of the three days with the highest mean daily call saturation index (CSI; Corn and Muths, 2002), which is the mean daily index value divided by three. The CSI is the proportion (0–1) of total call saturation (i.e., if all recordings in a day have an index of 3, CSI = 1). Recordings for a given date were those from 0600 to before 0600 the next morning, so that all nocturnal recordings in 24 hr had the same date.

To assess whether the greater amount of data available from ARS could serve as an index to abundance, we examined the effect of $\hat{N}$ and two weather variables, temperature and precipitation, on the maximum three-day mean CSI using an information-theoretic approach (Burnham and Anderson, 2002). Temperature and precipitation are the most common abiotic influences on amphibian

Table 1. Estimated Abundance of Calling Male *P. maculata*, Call Indices, and Weather at Lily Pond and Matthews Pond. CSI is the maximum three-day mean, call index frequencies are between 2100 and 0100 hr for the 11 days centered on the date of maximum calling activity, and weather variables are the mean minimum temperature and total precipitation for the three days that generated the CSI. Dashes indicate missing data.

Pond	Year	$\hat{N}$ (SE)	Maximum 3-day CSI	Call index frequency			T_{min} (°C)	Prec (mm)
				1	2	3		
Lily	1994	178 (41)	0.667	0	12	90	—	10.2
	1995	53 (12)	0.674	3	11	67	−0.70	2.5
	1996	—	—	—	—	—	−3.03	0
	1997	111 (11)	0.656	0	10	26	0.57	10.2
	1998	205 (36)	0.532	13	11	13	0.23	0
	1999	147 (77)	0.776	3	4	39	−1.80	12.7
	2000	716 (299)	—	—	—	—	—	—
	2001	764 (567)	—	—	—	—	—	—
	2002	344 (225)	0.728	6	15	27	−2.57	10.2
	2003	311 (84)	0.790	0	7	48	0.93	5.1
	2004	127 (20)	0.685	3	10	35	−0.10	0
	2005	206 (41)	0.560	2	20	18	−0.20	38.1
	2006	240 (60)	—	11	14	2	0.77	2.5
	2007	95 (31)	—	0	14	39	−0.03	0
	2008	555 (133)	0.583	8	16	24	0.70	2.5
	2009	1845 (792)	0.798	0	4	50	−0.93	7.6
Matthews	1995	18 (5)	0.525	9	13	76	−1.57	2.5
	1996	19 (8)	0.537	1	20	26	−2.60	0
	1997	140 (35)	0.773	5	7	38	−1.60	2.5
	1998	75 (6)	0.449	6	9	18	−2.07	5.1
	1999	75 (48)	0.606	1	16	28	−2.87	0
	2000	87 (18)	—	—	—	—	—	—
	2001	140 (22)	0.699	2	10	39	—	2.5
	2002	82 (16)	0.667	1	6	28	−2.27	0.0
	2003	82 (14)	0.752	0	10	40	0.93	5.1
	2004	55 (14)	—	—	—	—	—	—
	2005	107 (36)	0.477	6	4	15	−2.40	5.1
	2006	58 (26)	—	4	10	21	−2.87	0
	2007	267 (121)	—	—	—	—	−2.77	0
	2008	404 (127)	—	—	—	—	—	0
	2009	83 (36)	0.602	0	33	22	−0.53	12.7

breeding activity (Oseen and Wassersug, 2002; Saenz et al., 2006) and are frequently significant variables in analyses of detection in calling surveys (Gooch et al., 2006; Kirlin et al., 2006; Tupper et al., 2007; Steelman and Dorcas, 2010). We obtained weather data from the USDA Natural Resources and Conservation Service for the Joe Wright snowpack telemetry (SNOTEL) station (Natural Resources and Conservation Service, 2010) located 5.3 km west from Lily Pond and 7.3 km southwest from Matthews Pond at an elevation of 3084 m. We evaluated seven linear models with various combinations of a categorical variable, site (Lily Pond or Matthews Pond), and three continuous variables: $\hat{N}$, T_{min}, and Prec (mean minimum air temperature and total accumulated precipitation, respectively, for the three days that generated the maximum mean CSI). Models were ranked according to their differences in the Akaike Information Criterion, adjusted for small sample size (ΔAICc). Model residuals were computed using the general linear models module in SYSTAT ver.13 (Systat Software Inc.; http://www.systat.com), and ΔAICc were calculated using the formula in Burnham and Anderson (2002). We assessed the predictive ability of the top-ranked model using linear regression.

RESULTS

We obtained estimates of the number of calling males and usable recordings at both sites in most years (Table 1). Gaps in the calling data are due to equipment failure. Because of the programming limitations of the DVR-based ARS used in 2006–2007 (ten recordings per day), these data could not be used to calculate daily CSI that were comparable to the values based on recordings over the entire 24 hr of each day obtained from the other ARS. However, we were able to analyze the call index frequencies during peak breeding for 2100–0100 hr from these data, except at Matthews Pond in 2007, when the volume level of the recordings was much lower than other years, and we distrust the index values we obtained from them. Altogether, tape-based ARS were successful in providing season-long usable recordings in 18 of 23 installations (78%), and digital ARS succeeded in six of eight installations (75%).

The timing of breeding activity varied annually and was mainly related to volume of the winter snow pack (Corn and Muths, 2002). Peak breeding occurred earlier at the lower elevation Matthews Pond (mean day of year = 139) than at Lily Pond (mean day of year = 150). Calling activity was lowest in the early morning, increased during the day, and

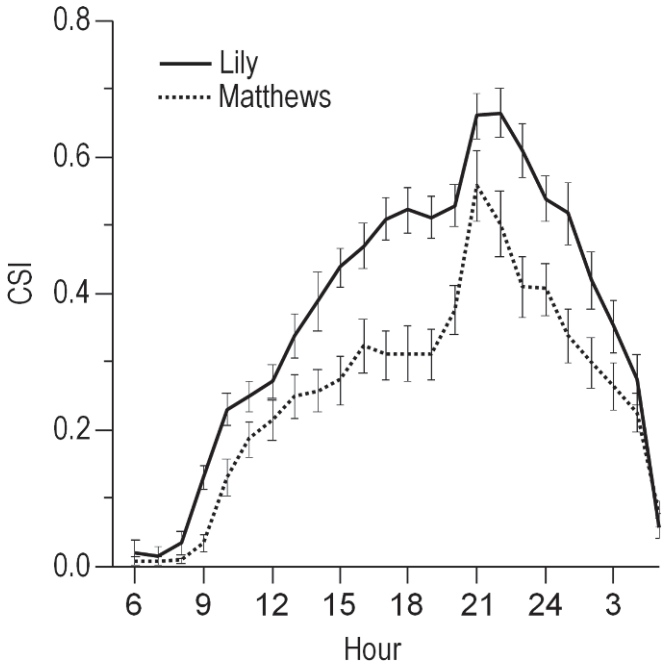

Fig. 2. Mean (±1 SE) hourly call saturation index (years combined, but excluding 2006–2007, when recording did not include all 24 hr) of *P. maculata* at Lily Pond and Matthews Pond.

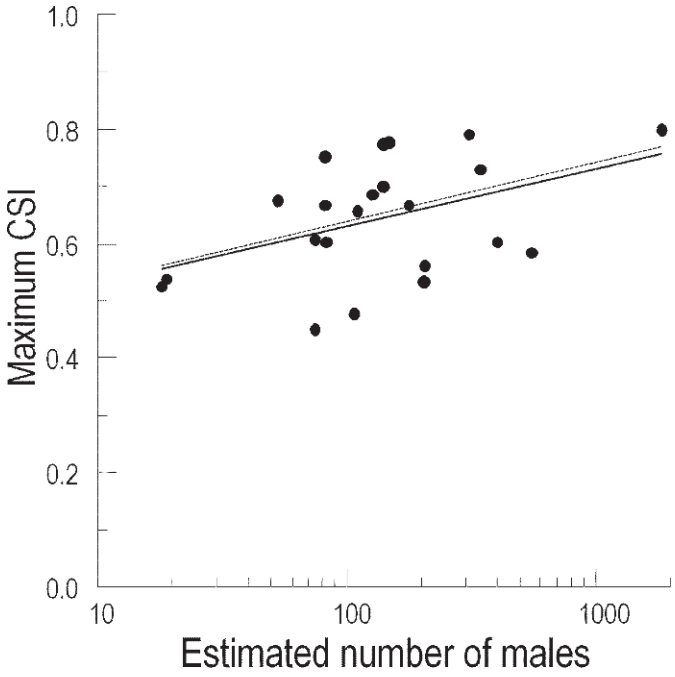

Fig. 3. The maximum three-day mean call saturation index versus the estimated number of male *P. maculata* in breeding choruses at Lily Pond and Matthews Pond ($R^2 = 0.18$; regression equation [solid line] $CSI = 0.432 + 0.099 \times \log_{10}(\hat{N})$, $P = 0.046$). The dashed line indicates the regression with the slope (0.102) corrected for error in estimating $\hat{N}$.

peaked shortly after sunset at both study sites (Fig. 2). During peak breeding activity and the window prescribed for NAAMP surveys, most index values were 3 (70%), followed by 2 (23%) and 1 (7%). The modal value for a year was 3 in 21 of 24 occasions (Table 1).

Due to high model selection uncertainty in most years, we derived model-averaged estimates of population size (Burnham and Anderson, 2002) for each year (Table 1). We did not use the population estimate from Lily Pond in 1996. Sampling occurred that year after the peak of breeding activity, and males may have begun leaving the pond, which would produce a biased estimate. The estimated population of male *P. maculata* was larger at Lily Pond (median = 206) than at Matthews Pond (median = 82).

Mirroring estimated abundance, the median CSI was higher at Lily Pond (0.674) than at Matthews Pond (0.602). The linear model with the greatest support was a simple regression of abundance (log-transformed $\hat{N}$) on CSI (Table 2). Models including minimum temperature, precipitation, or the interaction between temperature and precipitation as additional factors received lesser support, and there was no support for differences between sites affecting the relationship between abundance and CSI (Table 2). Estimated abundance explained only 18% of the variation in CSI (Fig. 3), but because the independent variable, $\hat{N}$, was subject to error, the slope of the regression is biased low (Snedecor and Cochran, 1989). We estimated the magnitude of the bias (κ) by dividing the variance in $\hat{N}$ (1.46×10^5) by the variance in $\hat{N}$ plus the variance in the precision of $\hat{N}$ (the mean standard error squared [6889]). The result indicates that bias was less than 5% ($\kappa = 0.955$). The corrected regression line is represented by the dashed line in Figure 3 (corrected slope = $\kappa^{-1} \times$ slope of CSI versus $\log_{10}[\hat{N}]$).

Table 2. Comparison of Linear Models Relating Abundance, Temperature, Precipitation, and Study Site to Annual Maximum Three-Day Mean Call Saturation Index (CSI).

Model	Parameters	ΔAICc	Weight
$\log(\hat{N})$	3	0	0.542
$\log(\hat{N})$ + Prec	4	2.18	0.182
$\log(\hat{N})$ + T_{min}*Prec	4	2.83	0.132
$\log(\hat{N})$ + T_{min}	4	3.03	0.119
$\log(\hat{N})$ + T_{min} + T_{min}*Prec	5	6.57	0.020
$\log(\hat{N})$ + T_{min} + Prec + T_{min}*Prec			
$\log(\hat{N})$ + Site + T_{min} + Prec + T_{min}*Prec	6 7	9.82 14.92	0.004 0.000

DISCUSSION

The use of indices as surrogates for population size in wildlife monitoring has a long, but controversial history (Johnson, 2008). Most of the discussion has focused on whether counts in the absence of some measure of detectability have any meaning (Anderson, 2001), but reliable estimates of detection probability can be difficult and expensive to obtain (Johnson, 2008). Indices from anuran call surveys, because they are abbreviated categorical estimates of relative abundance, are one step removed from indices based on counts, but they are still subject to many of the factors that affect detection in other taxonomic groups, including abiotic influences on activity and observer bias (Dorcas et al., 2009). Nevertheless, methods have been proposed for using manual call data as indices of abundance (Royle, 2004; Royle and Link, 2005), and research is needed on the relationships between call indices and abundance (Royle and Link, 2005; Weir et al., 2009).

Such research remains relatively rare. Some authors have assigned numbers of frogs to index values (Lepage et al.,

1997; Crouch and Paton, 2002) without explanation for how these relationships were derived. A few studies have compared the estimated number of calling frogs to abundance. Shirose et al. (1997) found significant regressions between numbers of Fowler's Toads (*Anaxyrus fowleri*) and Bullfrogs (*Lithobates catesbeianus*) estimated to be calling and numbers (uncorrected for detection probability) caught by hand the same night. Fogarty and Vilella (2001) found a significant correlation between point counts of calling *Eleutherodactylus coqui* and densities estimated using line-distance sampling. However, Goldberg and Schwalbe (2004) did not observe agreement between the number of Barking Frogs (*Craugastor augusti*) estimated to be calling and population size estimated by capture–recapture methods. Very few studies have compared abundance to the standard NAAMP index values. Nelson and Graves (2004) observed that CR estimates of numbers of male Green Frogs (*Lithobates clamitans*) were significantly larger in 12 ponds that had been assessed with a call index of 2, compared to 25 ponds with an index of 1. Stevens and Paszkowski (2004) assigned index values to individual choruses of *L. sylvaticus* within ponds and found that the maximum rank was a good predictor of the number of egg masses in a pond. They also estimated the number of calling males for each index value by counting the numbers of calling males in index 1 (mean = 1.3) and index 2 (mean = 3.7), and used twice the number of egg masses (assuming a breeding chorus sex ratio of 2:1 male:female) for index 3 choruses (mean = 118.1). The validity of these estimates is questionable, because Stevens and Paszkowski (2004) used two types of indices (counts of calling males and egg masses) to quantify the call survey index. They did not make an independent CR estimate of the number of males.

Corn et al. (2000) found that manual call surveys distinguished between the small population at Matthews Pond in 1995 and larger populations at Lily Pond and another site, but examination of the additional data in this study indicates that single observations from manual call surveys would be entirely unreliable as indices of population size for *P. maculata*. The modal call index each year (for the 11 days of peak breeding and the times when manual surveys would comply with NAAMP guidelines) was less than 3 only three times in 24 opportunities, despite estimates of adult males that varied over two orders of magnitude. Moreover, the smallest populations did not generate the smallest indices. At Matthews Pond, the two very small populations in 1995 and 1996 (both <20 males) produced modal call indices of 3. At Lily Pond, the two years with modal indices of 2 had 206 and 240 males, but five years with smaller estimated populations (53–147 males) all had modal indices of 3. For *P. maculata* in particular, and perhaps chorus frogs in general, a single index value from a manual call survey contains little information about the actual abundance of frogs at a site. Royle (2004) describes a method for inferring the maximum index value when more than one value is collected at a site, to account for variation in calling activity due to external factors. However, although NAAMP protocols mandate three surveys of a route each year (Weir and Mossman, 2005), this is designed more to account for seasonal changes in species composition at a site than to provide multiple index values for individual species.

In contrast to single index values, the larger amount of data available from ARS and use of the CSI correctly identified Lily Pond as having a larger population of *P. maculata*, similar to Nelson and Graves' (2004) finding that ponds with a call index of 2 had larger populations of *L. clamitans* than ponds with a call index of 1. The predictive power of the relationship between abundance and CSI was low, and correction for the bias due to variation in estimating abundance does not change this. For example, the uncorrected regression in Figure 3 predicts CSI values of 0.630 and 0.699 for population sizes of 100 and 500 male frogs, respectively. Correcting for the bias due to imprecision in $\hat{N}$ increases the CSI estimates to 0.639 and 0.711. This increases the predicted difference in CSI between the two populations by 0.003. In contrast, the 11 population estimates with 95% confidence intervals that include 100 had CSI values that varied between 0.477 and 0.776 (Table 1), and the 95% confidence interval around the mean CSI for these 11 observations is 0.613–0.732. Because CSI was so variable, it would not be advisable to use it as a direct surrogate for abundance, for example, by inferring population status by analyzing the trend of annual changes in CSI, because there would be very little power to detect changes in abundance. However, if a large number of ARS were deployed in a monitoring program so that the data could be subjected to an occupancy analysis, use of CSI as a covariate related to abundance would be more justified than use of single index values from manual surveys.

Use of ARS for large monitoring projects has always been constrained by the cost and reliability of the hardware, the volume of data generated (>20 gigabytes per year for our two sites), and the time necessary to process the data. Although older tape-based analog ARS and new digital ARS failed at the same rate in our study, the failures of tape-based ARS were largely due to mechanical breakdowns of aging equipment, while the failures with digital ARS were likely attributable to battery failure and operator error. Over time, we expect digital ARS to be much more reliable than the older, analog version. This has been the experience of researchers in Florida who have deployed larger numbers of Songmeters (>50) and have encountered no instances of equipment failure (W. J. Barichivich, USGS, Gainesville, FL, pers. comm.). The new generation of digital ARS and software that can analyze the recordings and identify calls shows some promise for reducing the problem of too much data. For example, the program Song Scope (Wildlife Acoustics, Inc.; http://www.wildlifeacoustics.com) uses sound files (recognizers) that contain sample vocalizations and digital signal processing algorithms to identify calls in a recording (Agranat, 2009). Output from this software includes the duration of each call identified, and it is simple to sum these observations to obtain the total amount of time in a recording that species' calls were detected. The proportion of this time relative to the total length of the recording is analogous to the CSI we used in this study and could be compared to abundance in the same manner. However, computer analysis of sound files may prove difficult in practice. Waddle et al. (2009) evaluated the performance of Song Scope for three anuran species in Louisiana and found significant error rates for both false positives (3–16%) and false negatives (45–51%). Error rates this high, particularly due to false negatives, would call into question any statistics derived from such analyses. Building successful recognizer files is somewhat of an art, particularly in deciding on the trade-offs between types of errors (Waddle et al., 2009). Another consideration for long-term monitoring is that the output from the analysis is specific to the recognizers used. Therefore, a study must either stick

with an initial set of recognizers, or, if improved recognizers are developed, reanalyze all of the data that have been collected. Computer analysis of sound files from ARS is a topic that needs considerable additional research.

ACKNOWLEDGMENTS

Trade names are used for informational purposes only and such use does not imply any endorsement by the United States government. We thank J. Barichivich, B. Glorioso, and H. Waddle for comments and corrections. Research was conducted under animal care and use protocols approved by the USGS Fort Collins Science Center. This is contribution 378 of the U.S. Geological Survey Amphibian Research and Monitoring Initiative.

LITERATURE CITED

Agranat, I. 2009. Automatically identifying animal species from their vocalizations. Wildlife Acoustics, Concord, Massachusetts. http://www.wildlifeacoustics.com/sscope/aiasftv/

Anderson, D. R. 2001. The need to get the basics right in wildlife field studies. Wildlife Society Bulletin 29:1294–1297.

Bishop, C. A., K. E. Pettit, M. E. Gartshore, and D. A. Macleod. 1997. Extensive monitoring of anuran populations using call counts and road transects in Ontario (1992 to 1993), p. 149–160. *In*: Amphibians in Decline: Canadian Studies of a Global Problem. D. M. Green (ed.). Herpetological Conservation 1, Society for the Study of Amphibians and Reptiles, St. Louis, Missouri.

Brander, S. M., J. A. Royle, and M. Eames. 2007. Evaluation of the status of anurans on a refuge in suburban Maryland. Journal of Herpetology 41:52–60.

Bridges, A. S., and M. E. Dorcas. 2000. Temporal variation in anuran calling behavior: implications for surveys and monitoring programs. Copeia 2000:587–592.

Brodman, R. 2009. A 14-year study of amphibian populations and metacommunities. Herpetological Conservation and Biology 4:106–119.

Burnham, K. P., and D. R. Anderson. 2002. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. Second edition. Springer, New York.

Burton, E. C., M. J. Gray, and A. C. Schmutzer. 2006. Comparison of anuran call survey durations in Tennessee wetlands. Proceedings of the Annual Conference of the Southeastern Association of Fish and Wildlife Agencies 60:15–18.

Corn, P. S., and E. Muths. 2002. Variable breeding phenology affects the exposure of amphibian embryos to ultraviolet radiation. Ecology 83:2958–2963.

Corn, P. S., E. Muths, and W. M. Iko. 2000. A comparison in Colorado of three methods to monitor breeding amphibians. Northwestern Naturalist 81:22–30.

Corn, P. S., W. Stolzenburg, and R. B. Bury. 1989. Acid precipitation studies in Colorado and Wyoming: interim report of surveys of montane amphibians and water chemistry. U.S. Fish and Wildlife Service Biological Report 80 (40.26), Washington, D.C.

Crouch, W. B., III, and P. W. C. Paton. 2002. Assessing the use of call surveys to monitor breeding anurans in Rhode Island. Journal of Herpetology 36:185–192.

de Solla, S. R., K. J. Fernie, G. C. Barrett, and C. A. Bishop. 2006. Population trends and calling phenology of anuran populations surveyed in Ontario estimated using acoustic surveys. Biodiversity and Conservation 15:3481–3497.

Dorcas, M. E., S. J. Price, S. C. Walls, and W. J. Barichivich. 2009. Auditory monitoring of amphibian populations, p. 281–298. *In*: Amphibian Ecology and Conservation: A Handbook of Techniques. C. K. Dodd, Jr. (ed.). Oxford University Press, New York.

Eigenbrod, F., S. J. Hecnar, and L. Fahrig. 2008. The relative effects of road traffic and forest cover on amphibian populations. Biological Conservation 141:35–46.

Fahrig, L., J. H. Pedlar, S. E. Pope, P. D. Taylor, and J. F. Wegner. 1995. Effect of road traffic on amphibian density. Biological Conservation 73:177–182.

Fogarty, J. H., and F. J. Vilella. 2001. Evaluating methodologies to survey *Eleutherodactylus* frogs in montane forests of Puerto Rico. Wildlife Society Bulletin 29:948–955.

Genet, K. S., and L. G. Sargent. 2003. Evaluation of methods and data quality from a volunteer-based amphibian call survey. Wildlife Society Bulletin 31:703–714.

Goldberg, C. S., and C. R. Schwalbe. 2004. Considerations for monitoring a rare anuran (*Eleutherodactylus augusti*). Southwestern Naturalist 49:442–448.

Gooch, M. M., A. M. Heupel, S. J. Price, and M. E. Dorcas. 2006. The effects of survey protocol on detection probabilities and site occupancy estimates of summer breeding anurans. Applied Herpetology 3:129–142.

Hemesath, L. M. 1998. Iowa's frog and toad survey, 1991–1994, p. 206–216. *In*: Status and Conservation of Midwestern Amphibians. M. J. Lannoo (ed.). University of Iowa Press, Iowa City, Iowa.

Jackson, J. T., F. W. Weckerly, T. M. Swannack, and M. R. J. Forstner. 2006. Inferring absence of Houston Toads given imperfect detection probabilities. Journal of Wildlife Management 70:1461–1463.

Johnson, D. H. 2008. In defense of indices: the case of bird surveys. Journal of Wildlife Management 72:857–868.

Karns, D. R. 1986. Field herpetology: methods for the study of amphibians and reptiles in Minnesota. James Bell Ford Museum of Natural History Occasional Paper 18, Minneapolis, Minnesota.

Kirlin, M. S., M. M. Gooch, S. J. Price, and M. E. Dorcas. 2006. Predictors of winter anuran calling activity in the North Carolina piedmont. Journal of the North Carolina Academy of Science 122:10–18.

Lepage, M. L., R. Courtois, C. Daigle, and S. Matte. 1997. Surveying calling anurans in Quebec using volunteers, p. 128–140. *In*: Amphibians in Decline: Canadian Studies of a Global Problem. D. M. Green (ed.). Herpetological Conservation 1, Society for the Study of Amphibians and Reptiles, St. Louis, Missouri.

Lotz, A., and C. R. Allen. 2007. Observer bias in anuran call surveys. Journal of Wildlife Management 71:675–679.

Martof, B. S. 1953. Territoriality in the Green Frog, *Rana clamitans*. Ecology 34:165–174.

Matthews, T. C., and D. Pettus. 1966. Color inheritance in *Pseudacris triseriata*. Herpetologica 22:269–275.

Mazerolle, M. J. 2005. Peatlands and Green Frogs: a relationship regulated by acidity? Ecoscience 12:60–67.

Mossman, M. J., L. M. Hartman, R. Hay, J. R. Sauer, and B. J. Dhuey. 1998. Monitoring long-term trends in Wisconsin frog and toad populations, p. 169–198. *In*: Status and Conservation of Midwestern Amphibians. M. J. Lannoo (ed.). University of Iowa Press, Iowa City, Iowa.

Natural Resources and Conservation Service. 2010. SNOTEL site: Joe Wright. http://www.wcc.nrcs.usda.gov/nwcc/site?sitenum=551&state=co (Accessed 10 May 2010).

Nelson, G. L., and B. M. Graves. 2004. Anuran population monitoring: comparison of the North American Amphibian Monitoring Program's calling index with mark–recapture estimates for *Rana clamitans*. Journal of Herpetology 38:355–359.

Oseen, K. L., and R. J. Wassersug. 2002. Environmental factors influencing calling in sympatric anurans. Oecologia 133:616–625.

Otis, D. L., K. P. Burnham, G. C. White, and D. R. Anderson. 1978. Statistical inference from capture data on closed animal populations. Wildlife Monographs 62:3–135.

Pellet, J. A., and B. R. Schmidt. 2005. Monitoring distributions using call surveys: estimating site occupancy, detection probabilities and inferring absence. Biological Conservation 123:27–35.

Peterson, C. R., and M. E. Dorcas. 1994. Automated data acquisition, p. 47–57. *In*: Measuring and Monitoring Biological Diversity: Standard Methods for Amphibians. W. R. Heyer, M. A. Donnelly, R. W. McDiarmid, L. C. Hayek, and M. S. Foster (eds.). Smithsonian Institution Press, Washington, D.C.

Pierce, B. A., and K. J. Gutzwiller. 2007. Interobserver variation in frog call surveys. Journal of Herpetology 41:424–429.

Royle, J. A. 2004. Modeling abundance index data from anuran calling surveys. Conservation Biology 18:1378–1385.

Royle, J. A., and W. A. Link. 2005. A general class of multimodal mixture models for anuran calling survey data. Ecology 86:2505–2512.

Saenz, D., L. A. Fitzgerald, K. A. Baum, and R. N. Conner. 2006. Abiotic correlates of anuran calling phenology: the importance of rain, temperature, and season. Herpetological Monographs 20:64–82.

Shirose, L. J., C. A. Bishop, D. M. Green, C. J. MacDonald, R. J. Brooks, and N. J. Helferty. 1997. Validation tests of an amphibian call count survey technique in Ontario, Canada. Herpetologica 53:312–320.

Snedecor, G. W., and W. G. Cochran. 1989. Statistical Methods. Eighth edition. Iowa State University Press, Ames, Iowa.

Steelman, C. K., and M. E. Dorcas. 2010. Anuran calling survey optimization: developing and testing predictive models of anuran calling activity. Journal of Herpetology 44:61–68.

Stevens, C. E., and C. A. Paszkowski. 2004. Using chorus-size ranks from call surveys to estimate reproductive activity of the Wood Frog (*Rana sylvatica*). Journal of Herpetology 38:404–410.

Sung, H. C., S. K. Kim, S. W. Cheong, S. R. Park, D. C. Roh, K. W. Baek, J. H. Lee, and D. S. Park. 2006. Estimating detection probabilities and site occupancy rates of three anuran species using call surveys in Haenam Gun, Korea. Journal of Ecology and Field Biology 29:331–335.

Tupper, T. A., R. P. Cook, B. C. Timm, and A. Goodstine. 2007. Improving calling surveys for detecting Fowler's Toad, *Bufo fowleri*, in southern New England, USA. Applied Herpetology 4:245–259.

Vertucci, F. A., and P. S. Corn. 1996. Evaluation of episodic acidification and amphibian declines in the Rocky Mountains. Ecological Applications 6:449–457.

Vogt, R. C., and R. L. Hine. 1982. Evaluation of techniques for assessment of amphibian and reptile populations in Wisconsin, p. 201–217. *In*: Herpetological Communities. N. J. Scott, Jr. (ed.). U.S. Fish and Wildlife Service Wildlife Research Report 13, Washington, D.C.

Waddle, J. H., T. F. Thigpen, and B. M. Glorioso. 2009. Efficacy of automatic vocalization recognition software for anuran monitoring. Herpetological Conservation and Biology 4:384–388.

Weir, L., I. J. Fiske, and J. A. Royle. 2009. Trends in anuran occupancy from northeastern states of the North American Amphibian Monitoring Program. Herpetological Conservation and Biology 4:389–402.

Weir, L., and M. J. Mossman. 2005. North American Amphibian Monitoring Program (NAAMP), p. 307–313. *In*: Status and Conservation of Midwestern Amphibians. M. J. Lannoo (ed.). University of Iowa Press, Iowa City, Iowa.

Weir, L. A., J. A. Royle, P. Nanjappa, and R. E. Jung. 2005. Modeling anuran detection and site occupancy on North American Amphibian Monitoring Program (NAAMP) routes in Maryland. Journal of Herpetology 39:627–639.

White, G. C., and K. P. Burnham. 1999. Program MARK: survival estimation from populations of marked animals. Bird Study 46(Supplement):120–138.