

# The effects of survey protocol on detection probabilities and site occupancy estimates of summer breeding anurans

Michelle M. Gooch<sup>1</sup>, Aubrey M. Heupel<sup>1,2</sup>, Steven J. Price<sup>1,3</sup> and Michael E. Dorcas<sup>1</sup>

<sup>1</sup> Department of Biology, Davidson College, Davidson, North Carolina 28035-7118, U.S.A.

<sup>2</sup> Present address: Dept. Nat. Resource Ecology and Management, Iowa State University, Ames, Iowa 50011, U.S.A.

<sup>3</sup> Corresponding author; e-mail: sjprice@davidson.edu

**Abstract.** Recent declines in amphibian populations have created an urgent need for large-scale, long-term monitoring efforts and many anuran monitoring programs have been established that utilize calling surveys. Calling surveys can be effective monitoring tools; however, differences among survey protocols may bias survey results. Failure to take into account detection probabilities when monitoring anurans can lead to inaccurate inferences about site occupancy, since non-detections in survey data do not necessarily mean that a species is absent unless the probability of detection is 1. We used a likelihood-based method, in the form of the computer program PRESENCE, to estimate detection probabilities and site occupancy rates for summer-breeding anurans in the Western Piedmont of North Carolina. Using detection data from calling surveys, we evaluated how detectability and site occupancy for five anuran species were influenced by 1) time spent listening at each site, 2) number of surveys per site, and 3) sample- and site-specific covariates. We found considerable variation among species with regards to detection probability and site occupancy across survey duration and sampling occasion. Although 13% of all species detection occurred after 3 min, longer surveys did not significantly increase detectability of individual species. We found that detectability varied more with sampling occasion than with survey duration for each species. Covariates had differing effects on occupancy and detectability among individual species. Multiple surveys per site within a season are necessary to eliminate biased detection probabilities, but we found that 3- or 5-min surveys were adequate for detecting all species breeding at the time of the survey.

*Key words:* Anuran calling survey; frog; monitoring program; North Carolina; PRESENCE.

## Introduction

Anthropogenic disturbances have resulted in the widespread decline of amphibian populations (Delis et al., 1996; Alford and Richards, 1999; Knutson et al., 2000; Gibbs et al., 2005). Consequently, scientists have increased amphibian monitoring

efforts to help determine the extent of these declines. Amphibian monitoring efforts have focused largely on anurans because declines have been documented primarily in this group and because of the relative ease with which anurans can be detected and identified via calling surveys (Weir and Mossman, 2005). Extensive anuran monitoring programs have been implemented throughout the United States and Canada (Shirose et al., 1997; ARMI, 2004; Weir and Mossman, 2005). Anuran calling surveys have provided scientists with valuable data pertaining to amphibian population trends and have led to the implementation of management practices for amphibians (Mossman et al., 1998).

Most calling surveys involve observers listening to and documenting the species of anurans heard as well as recording an index representing the perceived abundance of each species. However, specific sampling protocols often vary considerably among programs. Variation in protocol may affect the ability to detect all species, leading to imprecise site occupancy estimates (Crouch and Paton, 2002; Bailey et al., 2004). Because some breeding anurans may not be detected even when present at a location, recent monitoring initiatives, including the United States Geological Survey's Amphibian Research and Monitoring Initiative, have begun to incorporate detection probabilities into site occupancy estimates, greatly improving monitoring results (MacKenzie et al., 2003; Bailey et al., 2004; Pierce and Gutzwiller, 2004). However, information is still needed to determine how sampling protocols influence anuran detection probabilities and site occupancy estimates (Shirose et al., 1997; Crouch and Paton, 2002; Bailey et al., 2004; Pierce and Gutzwiller, 2004).

Survey protocols that may influence monitoring results include time spent listening at each survey station (i.e., survey duration) and number of surveys conducted within a sampling period (i.e., sampling occasion). The number of minutes spent listening often varies among monitoring programs. For example, the Marsh Monitoring Program requires an observer to listen for 3 min (Crewe et al., 2005), while North American Amphibian Monitoring program (NAAMP) volunteers listen for 5 min (Weir and Mossman, 2005) and several state monitoring programs require up to 10 min of listening time (Mossman et al., 1998). Shirose et al. (1997) found that longer surveys rarely resulted in the detection of species not heard in the first 3 min, while others have found that calling surveys should be conducted for 10 min in order to have a high probability of detecting all species that are present (Crouch and Paton, 2002). Pierce and Gutzwiller (2004) found that species detectability was significantly higher for 15-min surveys than it was for 5-min surveys.

Anuran calling surveys also specify "sampling periods" — prescribed seasonal time periods that are intended to collectively cover the peak breeding seasons of all local anurans — during which calling surveys should be conducted (Weir and Mossman, 2005). Sampling periods last between 2 and 6 wk and are based on anuran breeding phenology (Weir and Mossman, 2005). Calling surveys are usually carried out once during each sampling period; however, a great deal of interspecific variation in calling behavior occurs within each sampling period even during peak breeding seasons (Todd et al., 2003). Large nightly fluctuations in sample-specific

environmental variables such as air and/or water temperature may result in a species being undetected on nights within their peak breeding season. Such nightly variation may cause the observer to fail to detect a species during a single sampling occasion, resulting in occupancy estimates (unadjusted for imperfect detection) that underestimate true site occupancy probabilities for the species.

Additionally, site-specific habitat characteristics may influence anuran site occupancy estimates (Bailey et al., 2004). For example, some species are more sensitive to the extent, condition, and type of upland habitat surrounding their breeding location. Including habitat/land-use variables within a predefined buffer region surrounding a pond may provide strength to models that predict species' presence/absence (Johnson et al., 2002; Price et al., 2005).

The objectives of this study were 1) to evaluate the effects of time spent listening on detection probabilities and site occupancy estimates for summer breeding anurans, 2) to determine if detection probabilities and site occupancy estimates change with regard to the number of surveys conducted within a sampling period at each site, and 3) to create models that incorporate time-specific sample covariates and site-specific habitat variables in order to assess the importance of covariates on detection probability and site occupancy estimates.

## **Methods**

### *Study sites*

We monitored anurans at 35 ponds in Mecklenburg County, located in the western Piedmont of North Carolina. Urbanization in this region is occurring rapidly, and only small patches of farmland and secondary-growth forest remain between areas of development (Griffith et al., 2003). We selected study ponds from digital aerial photographs using a geographical information system (GIS; ArcView 3.2, ESRI, Redlands, CA) and personal communication with the Mecklenburg County Natural Resources Division. All ponds were permanent and were isolated from other water bodies by at least 200 m. Ponds were selected to represent a wide variety of urbanization levels, from completely forested ponds to those surrounded entirely by urban development. We attempted to select ponds that equally represent the different levels of urbanization in the region.

### *Survey protocol*

We based our calling surveys on the NAAMP protocol (Weir and Mossman, 2005), although some variations to the protocol were made. All surveys were conducted within the sampling period (June 10-July 13, 2004) for summer-breeding anurans in the Piedmont of North Carolina when conditions were favorable (windspeed  $\leq 19$  km/h preferably after a rainfall). Surveys began at approximately 0.5 h after sunset and were completed no later than 01:00 a.m. the following morning. We

began listening for frog calls immediately after arriving at each pond and listened for 10 min. We recorded the species heard and the time that each species was first detected to the nearest second during the 10-min survey. We also recorded environmental conditions including a sky code (0 = few clouds, 1 = partly cloudy or variable sky, 2 = cloudy or overcast, 3 = fog or smoke, 4 = drizzle or light rain, 5 = snow, 6 = showers — did not conduct survey), wind speed estimate (0 = <1.6 km/h, 1 = 1.6-4.8 km/h, 2 = 6.4-11.3 km/h, 3 = 12.9-19.3 km/h, 4 = 20.9-29 km/h — did not conduct survey), relative humidity, air and shallow water temperature (C°) as well as the start time and end time of the survey. Surveys were conducted every 2 wk during the survey period, resulting in a total of three sampling occasions at each pond. The same two observers were present during each survey of all 35 ponds. The observers were very familiar with the survey protocol and the calls of potential species, and were coordinated in their estimations of abundance. Thus observer bias during calling surveys was minimized.

### *Landscape analysis*

Using digital aerial photographs of Mecklenburg County (2002) in the GIS, we created 200 m buffers around each pond. Although recent studies have focused on landscape scales larger than 200 m (Knutson et al., 1999; Johnson et al., 2002; Price et al., 2005), habitat conditions at smaller scales, such as 200 m, include critical habitat for amphibians (Semlitsch and Bodie, 2003). The 200 m buffer surrounding an aquatic site is within the range that Semlitsch and Bodie (2003) classify as core habitat needed by amphibians for activities such as foraging and overwintering. Within these 200 m buffer zones, we measured total area (to the nearest 0.1 ha) and percentage of that area representing developed areas (considered “urban”), forest, and pasture/grassland.

### *Data analysis*

We used the computer program PRESENCE (MacKenzie et al., 2002) for all calculations of site occupancy and detection probabilities. PRESENCE utilizes a likelihood-based method for estimating the proportion of sites occupied when species detection probabilities are <1. All estimation models assume 1) sites that are occupied by the species of interest remain occupied for the duration of the survey, 2) species are not detected when absent, and a species may or may not be detected when present, and 3) detecting a species at one site is independent of detecting a species at all other sites (MacKenzie et al., 2002). In addition, this method requires at least two sample occasions per sampling period where detection/nondetection data are recorded for each species. Estimable parameters given by PRESENCE include  $\psi_i$ , the probability that a species is present at site  $i$ , and  $p_{it}$ , the conditional probability that a species is detected at site  $i$  at time  $t$ , given it is present (MacKenzie et al., 2002). Both  $\psi_i$  and  $p_{it}$  can be expressed as a logit-function of site-specific

covariates, such as habitat variables, and  $p_{it}$  may also be expressed as a function of sample-specific variables such as air temperature and weather conditions.

#### *Detection probabilities ( $p$ ) and site occupancy estimates ( $\psi$ )*

We used PRESENCE to estimate  $p$  and  $\psi$  for five of the most common summer-breeding anuran species detected at our 35 ponds. We evaluated the effects of survey length on  $p$  using the results of our calling surveys. We sorted our data according to the time each species was first detected at each site and grouped data into three categories: 1) all species heard from 0-3 min, 2) all species heard from 0-5 min, and 3) all species heard from 0-10 min. We used  $\psi(\cdot)p(\cdot)$ , to determine the effects of survey length on  $p$  and  $\psi$ . This model assumed that  $\psi$  was the same for all sites and that  $p$  was constant across all three sampling occasions. Although the constant model does not represent the “best” model for each species because it excludes sample covariates, it allowed us to calculate detection probability for the three different time length intervals (3, 5 and 10-min surveys) without introducing biases that may confound our results due to different models (Boulinier et al., 1998).

We evaluated possible survey-specific effects on detection probability using a predefined model with survey-specific  $p$  but constant  $\psi$  (i.e.,  $\psi(\cdot)p(t)$ ). This model assumed that  $\psi$  is the same for all sites, but  $p$  differs between the three sampling occasions. This model allowed us to calculate detection probabilities for each species during each of the three surveys within our sampling period. The equations we used to calculate survey-specific  $p$  (for our three surveys) were:

$$p_{t=1} = e^{int+\beta_1} / 1 + e^{int+\beta_1} \quad (1)$$

$$p_{t=2} = e^{int+\beta_2} / 1 + e^{int+\beta_2} \quad (2)$$

$$p_{t=3} = e^{int} / 1 + e^{int} \quad (3)$$

where  $int$  was the intercept for sampling covariates, and  $\beta_i$  was the sampling covariate coefficient for survey  $i$  (MacKenzie et al., 2002). For this analysis, we used results from our 10-min surveys.

We explored the importance of sample and site covariates on  $p$  and  $\psi$  by first modeling  $p$  as a logit function of survey effects and each sample covariate separately. Since our number of weather variables was relatively small (seven), we were able to model each variable rather than selecting variables *a priori*. However, each of our weather variables has been found to be potentially important in determining anuran calling activity on a given night (Oseen and Wassersug, 2002). We held the proportion of sites occupied constant,  $\psi(\cdot)$ , and allowed  $p$  to vary with time (survey effects) and each covariate separately,  $p(t)$  and  $p(\text{Cov})$ . We used a constant model,  $\psi(\cdot)p(t)$ , as a reference. Each model was ranked according to Akaike Information Criterion (AIC) values (Akaike, 1973; Burnham and Anderson, 1998) calculated by PRESENCE. The lowest ranked sample covariate model for each species was then combined with each of three site covariates. The lowest ranked model was considered to be the “best” model fit for that species. It should be

noted that the model selected “best” does not necessarily represent all environmental or biological processes that influence site occupancy or detection probabilities (Bailey et al., 2004).

## Results

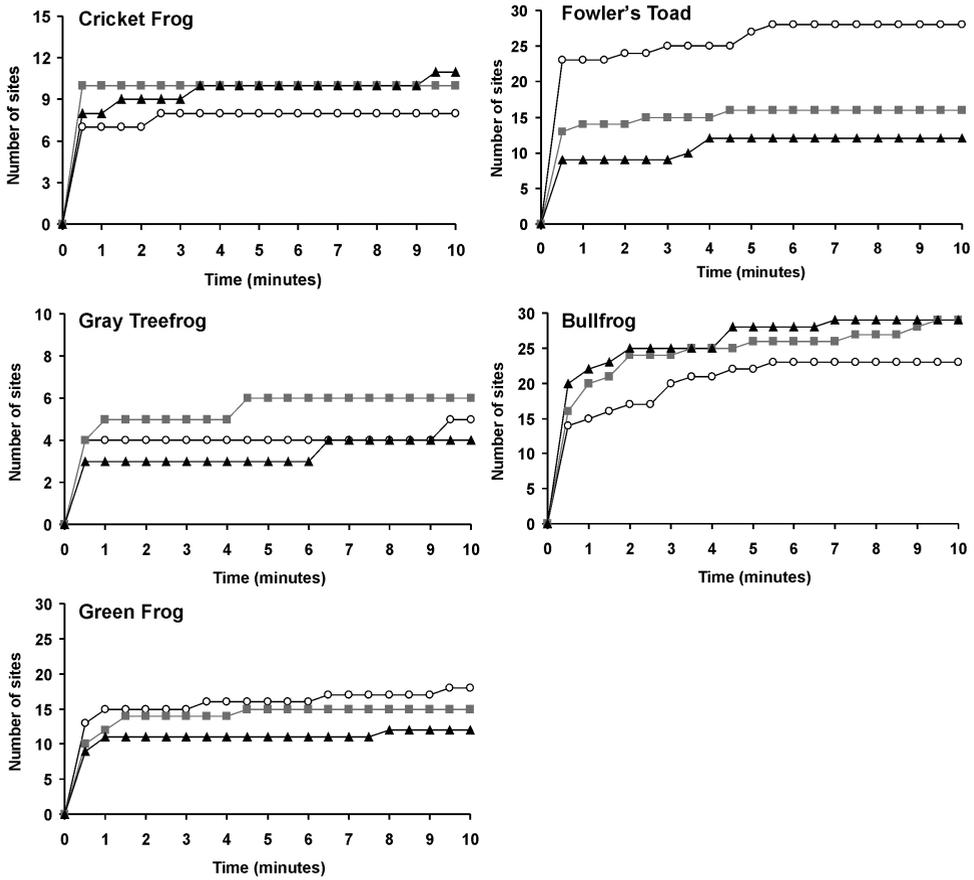
During our one-month sampling period we detected a total of seven vocalizing anuran species; however only northern cricket frog (*Acris crepitans*; detected at 12 sites), Fowler’s toad (*Bufo fowleri*; 28 sites), Cope’s gray treefrog (*Hyla chrysoscelis*; 11 sites), bullfrog (*Rana catesbeiana*; 33 sites), and green frog (*Rana clamitans*; 20 sites) were common enough to be used in our analyses. The eastern narrowmouth toad (*Gastrophryne carolinensis*) was detected at two sites, and green treefrog (*Hyla cinerea*) was detected at only one site. At least one species of vocalizing anuran was detected at all 35 sites on at least one occasion, although there were four different occasions where no calling anurans were heard at specific sites.

We observed considerable interspecific variation in calling activity with respect to number of minutes spent surveying (fig. 1). Although most species were first detected within the first 3 min of the survey (87% of detections in all surveys combined), 9% and 4% of all species first detections occurred at or after 3 min and 5 min, respectively, and there were five occasions where a species did not start calling until 9 min into the listening period. Cricket frogs were usually detected within the first 3 min of the survey. Fowler’s toads showed a gradual increase in number of sites in relation to time spent listening, and were detected at some sites after 3 min in each survey period. We normally detected gray treefrogs and green frogs within the first 3 min of the survey; however there were instances in which they were detected at a site after 4 min. Detections of bullfrogs increased gradually with increasing time spent listening.

We also detected variation in calling activity among the three surveys within the sampling period (fig. 1). Cricket frogs were found at fewer sites during survey 1 than during surveys 2 and 3. Detections of Fowler’s toads and green frogs decreased throughout the survey period. Gray treefrogs were found at more sites during survey 2 than in survey 1, but were detected less during survey 3 than survey 1. Bullfrogs were found in increasing numbers of sites throughout the sampling period.

### *Detection probability (p)*

Although longer time spent listening resulted in slightly higher detection probabilities ( $p$ ) for all species except cricket frogs, overlapping error bars suggest differences are not significant (fig. 2). There was some interspecific variation in  $p$ ; based on 10-min surveys, we found bullfrogs to be the most detectable ( $0.81 \pm 0.04$ ) followed by cricket frogs ( $0.80 \pm 0.07$ ), green frogs ( $0.74 \pm 0.06$ ), Fowler’s toads ( $0.63 \pm 0.06$ ), and gray treefrogs ( $0.48 \pm 0.17$ ).

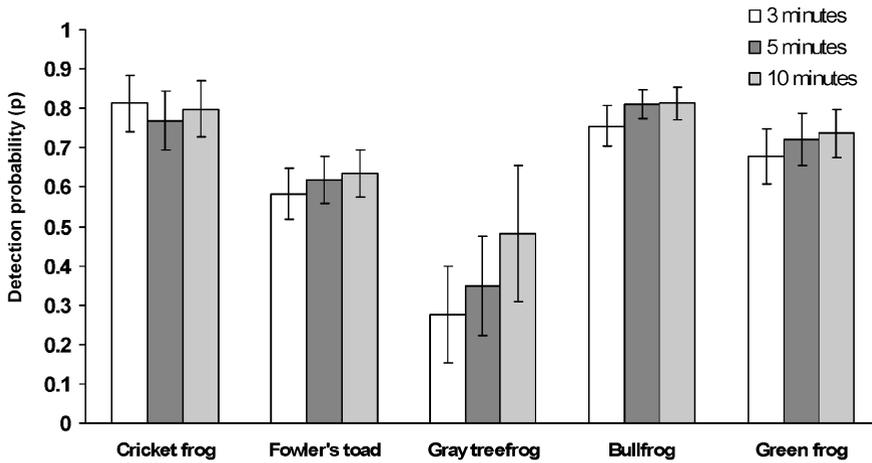


**Figure 1.** Number of sites where five anuran species were detected as influenced by survey length for each of three sampling occasions. Species include northern cricket frog (*Acris crepitans*), Fowler’s toad (*Bufo fowleri*), Cope’s gray treefrog (*Hyla chrysoscelis*), bullfrog (*Rana catesbeiana*), and green frog (*Rana clamitans*). Open circles represent survey 1, squares (medium gray) represent survey 2, and triangles (dark gray) represent survey 3.

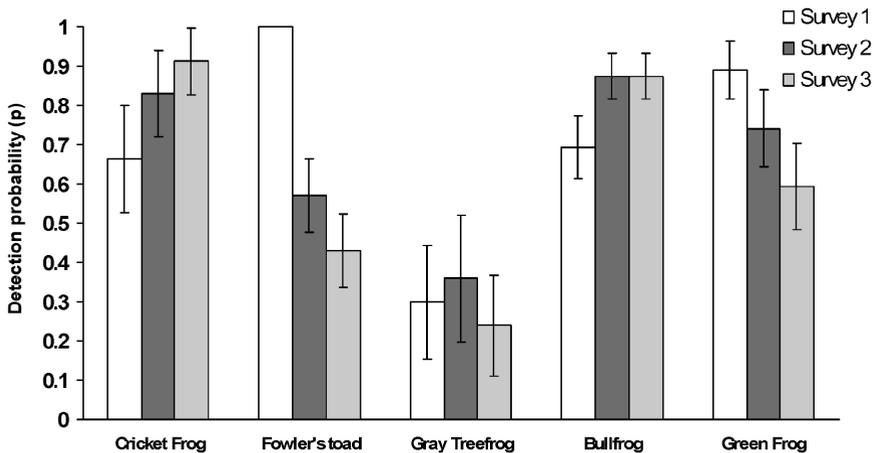
Sampling occasion strongly affected  $p$  (fig. 3). Cricket frogs had a consistently increasing  $p$  throughout the sampling period (survey 1 =  $0.66 \pm 0.14$ , survey 2 =  $0.82 \pm 0.11$ , survey 3 =  $0.90 \pm 0.08$ ). Both Fowler’s toads and green frogs had a decreasing  $p$  throughout the sampling period ( $1.00 \pm 0.00$ ,  $0.57 \pm 0.09$ ,  $0.42 \pm 0.09$  for Fowler’s toads and  $0.88 \pm 0.07$ ,  $0.74 \pm 0.10$ ,  $0.59 \pm 0.11$  for green frogs). Gray treefrogs had a higher  $p$  during survey 2 ( $0.36 \pm 0.16$ ) than survey 1 ( $0.30 \pm 0.15$ ) but a lower  $p$  for survey 3 ( $0.24 \pm 0.13$ ) than for surveys 1 or 2. Bullfrogs had the same  $p$  for surveys 2 and 3 ( $0.87 \pm 0.06$ ) and a lower  $p$  for survey 1 ( $0.69 \pm 0.08$ ).

*Proportion of sites occupied ( $\psi$ )*

Overlapping error bars indicate that site occupancy estimates ( $\psi$ ) for each species were not significantly affected by time spent listening when detection probability



**Figure 2.** Detection probabilities ( $\pm 1$  SE),  $p(\cdot)$ , for five summer-breeding anuran species influenced by amount of time spent listening at each sampling occasion.

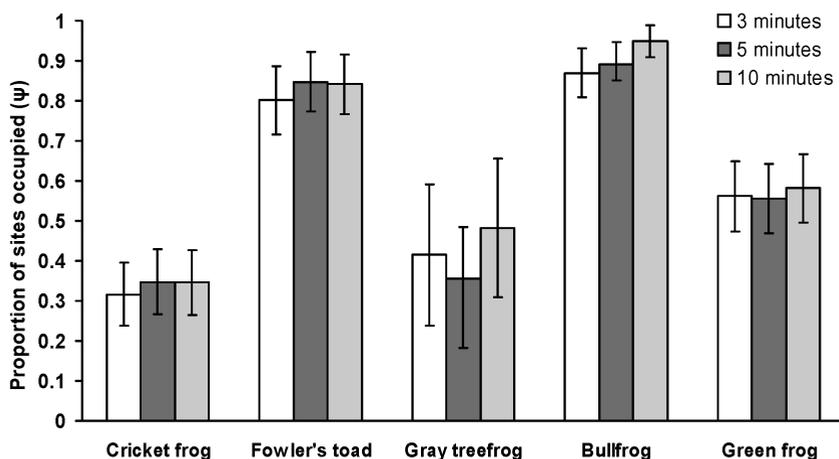


**Figure 3.** Detection probabilities ( $\pm 1$  SE),  $p(t)$ , for five summer-breeding anuran species influenced by sampling occasion.

differences (caused by differences in sampling effort) were corrected for (fig. 4). This was to be expected, because true  $\psi$  at the ponds had not changed, only our ability to detect the species. However, differences in  $\psi$  among species were evident. At 10 min, bullfrogs had the highest  $\psi$  ( $0.95 \pm 0.04$ ), followed by Fowler's toads ( $0.84 \pm 0.7$ ), green frogs ( $0.58 \pm 0.09$ ), gray treefrogs ( $0.48 \pm 0.17$ ) and cricket frogs ( $0.35 \pm 0.08$ ).

#### *Analysis of covariates*

Best-fit models were generated using sample and site covariates for each species (table 1). Cricket frogs were best predicted by water temperature and % forest



**Figure 4.** Proportion of sites occupied ( $\pm 1$  SE),  $\psi(\cdot)$ , for five summer-breeding anuran species influenced by amount of time spent listening at each sampling occasion. In calculating  $\psi$ , PRESENCE takes detection probability into account.

within 200 m radius of the pond ( $w_i = 0.59$ ). All other models for cricket frogs provided virtually zero weight. Although Fowler's toads appear to be positively associated with urban land surrounding a pond ( $w_i = 0.17$ ), all four models considered for this species had virtually identical weight, suggesting that all models provide a similar description of the data. Likewise, no conclusive predictor was identified for gray treefrogs, although there is some suggestion that pasture/grass and water temperature provide a slightly better description of the data ( $w_i = 0.15$ ). Bullfrogs also showed no strong association with any habitat variable, although pasture/grassland surrounding a pond provided the best descriptive model ( $w_i = 0.22$ ). Green frogs were best predicted by air temperature and % forest within 200 m radius of the pond ( $w_i = 0.89$ ).

## Discussion

We found that 10-min surveys resulted in more detections than 3-min surveys for every species, and that 13% of all species detections occurred after 3 min. However, unlike Pierce and Gutzwiller (2004) who found that 77% of all species were detected in the first 5 min of their calling surveys and that 15 min were required to detect >90% of all species known to be present and calling at least once during 30-min surveys, we found that 94% of all species known to be present were detected in the first 5 min of our surveys. In our study, detection probabilities generally increased with increasing time spent listening, although changes in  $p$  may be considered negligible. Our findings agree with those of Shirose et al. (1997), who found that 3-min surveys may be adequate to sample presence/absence of most species, and that while increasing the length of surveys will decrease the probability of overlooking species, the weakness of the relationship between number of species

**Table 1.** Relative differences in AIC from the best model ( $\Delta\text{AIC}$ ), AIC model weights ( $w_i$ ), detection probabilities ( $p$ ) for surveys 1, 2 and 3, overall estimates of the fraction of sites occupied ( $\psi$ ) with associated standard error ( $\text{SE}(\psi)$ ) for five summer-breeding anuran species. The best-fit sample and site covariate model,  $\psi(\text{Cov})p(\text{Cov})$ , is shown, along with the model including only the best-fit sample covariate,  $\psi(\cdot)p(\text{Cov})$ , only the best-fit site covariate,  $\psi(\text{Cov})p(t)$ , and the constant model,  $\psi(\cdot)p(t)$ . Overall best-fit models for each species are shown in bold.

Model, by species	$\Delta\text{AIC}$	$w_i$	$p_1$	$p_2$	$p_3$	$\psi$	$\text{SE}(\psi)$
Cricket frog							
<b><math>\psi(\text{forest})p(\text{watertemp})</math></b>	<b>0.00</b>	<b>0.59</b>	<b>0.66</b>	<b>0.83</b>	<b>0.91</b>	<b>0.35</b>	<b>0.07</b>
$\psi(\cdot)p(\text{watertemp})$	8.42	0.01	0.54	0.82	0.85	0.35	0.08
$\psi(\text{forest})p(t)$	8.90	0.01	0.66	0.83	0.91	0.34	0.07
$\psi(\cdot)p(t)$	17.43	0.00	0.66	0.83	0.91	0.35	0.08
Fowler's toad							
$\psi(\text{urban})p(\text{watertemp})$	0.63	0.13	1.00	0.60	0.45	0.80	0.36
$\psi(\cdot)p(\text{watertemp})$	1.21	0.09	1.00	0.60	0.45	0.80	0.07
<b><math>\psi(\text{urban})p(t)</math></b>	<b>0.00</b>	<b>0.17</b>	<b>1.00</b>	<b>0.57</b>	<b>0.43</b>	<b>0.80</b>	<b>0.07</b>
$\psi(\cdot)p(t)$	0.58	0.13	1.00	0.57	0.43	0.80	0.07
Gray treefrog							
<b><math>\psi(\text{pasture/grass})p(\text{watertemp})</math></b>	<b>0.00</b>	<b>0.15</b>	<b>0.31</b>	<b>0.37</b>	<b>0.25</b>	<b>0.46</b>	<b>0.14</b>
$\psi(\cdot)p(\text{watertemp})$	0.15	0.13	0.28	0.34	0.22	0.52	0.17
$\psi(\text{pasture/grass})p(t)$	0.59	0.11	0.32	0.38	0.26	0.45	0.14
$\psi(\cdot)p(t)$	1.43	0.07	0.30	0.36	0.24	0.48	0.13
Bullfrog							
$\psi(\text{pasture/grass})p(\text{airtemp})$	0.94	0.14	0.69	0.88	0.88	0.95	0.04
$\psi(\cdot)p(\text{airtemp})$	1.75	0.09	0.69	0.87	0.87	0.95	0.04
<b><math>\psi(\text{pasture/grass})p(t)</math></b>	<b>0.00</b>	<b>0.22</b>	<b>0.70</b>	<b>0.88</b>	<b>0.88</b>	<b>0.94</b>	<b>0.04</b>
$\psi(\cdot)p(t)$	0.84	0.14	0.69	0.87	0.87	0.94	0.05
Green frog							
<b><math>\psi(\text{forest})p(\text{airtemp})</math></b>	<b>0.00</b>	<b>0.89</b>	<b>0.90</b>	<b>0.73</b>	<b>0.61</b>	<b>0.57</b>	<b>0.07</b>
$\psi(\cdot)p(\text{airtemp})$	8.87	0.01	0.89	0.73	0.60	0.58	0.09
$\psi(\text{forest})p(t)$	5.19	0.07	0.89	0.74	0.60	0.58	0.07
$\psi(\cdot)p(t)$	14.35	0.00	0.89	0.74	0.59	0.58	0.09

heard after the initial 2 min (3 or 5 min in our case) causes the return for monitoring effort to drop off after these specified lengths of time. Our results were also in agreement with the results of Pierce and Gutzwiller (2004), who found that detectability of two species of anurans (*Acris crepitans* and *Rana sphenoccephala*) did not differ among 5- and 10-min surveys. We found that 3- or 5-min surveys provided ample time for detecting all species known to be present.

Local density and variation in calling behavior likely affects the detection probability and site occupancy estimates of most species (Royle and Nichols, 2003). As the density of a population of a certain species increases, so does the probability of detecting a single individual of that species. When employing 3-min surveys, species with small populations at some sites may go undetected, while large populations might be detected every time. The high level of detection for most species in our study may reflect the fact that these species are abundant at our study sites; however, more intensive and localized searches would be needed to validate

this assumption. Gray treefrogs had the lowest detection probability in our study. This might be explained by the fact that gray treefrogs typically call from roadside ditches and small streams rather than at permanent farm ponds (Wright and Wright, 1949).

Calling behavior and variation in calling patterns may also affect detection probabilities. For example, bullfrogs and green frogs called sporadically during our surveys, reducing the likelihood that they would be detected early in a survey. Other species, such as cricket frogs called more continuously, meaning they would have a higher probability of being detected within the first 3 min of a survey (Pierce and Gutzwiller, 2004).

In our study, detection probabilities varied more with sampling occasion than with survey duration, with some species becoming more detectable as the sampling period progressed (from survey 1 to survey 3) while others became less detectable. The fact that detection probabilities varied according to sampling occasion may be explained by interspecific differences in breeding phenology, nightly variation in calling behavior, and sampling covariates. A species nearing the end of its breeding season would likely exhibit reduced calling activity, which would cause it to be detected at fewer sites, thus decreasing  $p$  across surveys (i.e., Fowler's toads and green frogs). Species that were in the beginning of their breeding season during survey 1 would increase in calling activity from surveys 1 to 3 (i.e., cricket frogs and bullfrogs) and be more detectable. These results suggest that although it may be tempting to assign each species a detection probability, our results indicate that detection probabilities do change during a sampling period. Therefore, monitoring protocols should include more than one sampling occasion per sampling period in order to buffer against year-to-year fluctuations in timing of breeding of individual species due to sample covariates, as well as capture interspecific differences in breeding phenology.

According to Shirose et al. (1997), increasing survey length dramatically (even to 60 min) does not guarantee that all species present and breeding at a site will be heard because anuran calling activity is affected by environmental factors, and breeding activity is likely to be interrupted if weather conditions become unfavorable. In our study, detection probabilities were not strongly influenced by covariate models as best-fit models exhibited only minor changes in  $p$  despite changes in AIC. Since weather conditions remained relatively constant throughout our summer sampling period, time-specific sample covariates probably have much larger effects on detectability estimates during seasons when weather changes more drastically from night to night, such as late winter/early spring (Kirlin et al., 2006).

Site covariates had only slight effects on site occupancy estimates. Fowler's toads, gray treefrogs, and bullfrogs were not strongly associated with any habitat variables, indicating that landscape variables may be of little importance to these ubiquitous species. Although toads depend extensively on upland habitat, they are considered habitat generalists and can use most aquatic habitats for reproduction (Wright and Wright, 1949). Cricket frog site occupancy was positively associated with forest

in our study in North Carolina and another study in Iowa and Wisconsin (Knutson et al., 2000). Green frogs were also positively associated with forest; other studies have found this same association, although most have also found green frogs to be generalist and thus difficult to classify into a single habitat category (Knutson et al., 2000; Johnson et al., 2002; Price et al., 2005). Further studies should be done to examine species-specific responses to sample and site covariates in order to validate our results.

Our results indicate that calling survey protocol may bias detection probabilities of summer-breeding anurans. Time spent listening did not affect detectability of the species we studied, but number of surveys conducted did. Multiple surveys at each site within a sampling period can reduce some biases associated with anuran calling surveys (MacKenzie et al., 2002; MacKenzie et al., 2003; Bailey et al., 2004). It should be noted, however, that our conclusions are based on data collected from fairly ubiquitous, habitat generalist species in a heavily altered landscape. Our results may not be applicable to more pristine ecosystems where there may be more species present, each with their own breeding phenology and different detectabilities. In addition to calling surveys, alternate survey methods, such as automated recording systems or drift fences, may be needed to effectively monitor rare or cryptic species (Crouch and Paton, 2002). Choice of calling survey protocols ultimately depends on the degree of accuracy that is needed for the desired monitoring effort.

**Acknowledgements.** We thank all of the land owners of the ponds used in our study. We also thank Studie Daves and the Mecklenburg County DNR for their advice and assistance regarding study site selection. Darryl MacKenzie provided valuable help with PRESENCE, and Robert W. Klaver, Robin Jung Brown, Michael J. Tyler and one anonymous reviewer provided comments on the manuscript. We thank Joy Hester, Bill Johnson, Sarah Budischak, Pierson Hill, Yurii Kornilev and Betsie Rothermel for their advice and assistance with the project. Larissa Bailey was especially helpful providing statistical advice and help with PRESENCE as well as providing comments on the manuscript. Manuscript preparation was aided by the Environmental Remediation Sciences Division of the Office of Biological and Environmental Research, U.S. Department of Energy through Financial Assistance Award number DE-FC09-96SR18546 to the University of Georgia Research Foundation. This research was supported by Duke Power and National Science Foundation grants (DEB-0347326 and DBI-1039153) to MED.

## References

- Akaike, H. (1973): Information theory as an extension of the maximum likelihood principle. In: 2nd International Symposium on Information Theory, p. 267-281. Petrov, B.N., Csaki, F., Eds, Akademiai Kiado, Budapest.

- Alford, R.A., Richards, S.J. (1999): Global amphibian declines: A problem in applied ecology. *Ann. Rev. Ecol. Syst.* **30**: 133-165.
- ARMI (2004): Amphibian Research and Monitoring Initiative, USGS. (<http://armi.usgs.gov>).
- Bailey, L.L., Simons, T.R., Pollock, K.H. (2004): Estimating detection probability parameters for *Plethodon* salamanders using the robust capture-recapture design. *J. Wild. Manag.* **68**: 1-13.
- Boulinier, T., Nichols, J.D., Sauer, J.R., Hines, J.E., Pollock, K.H. (1998): Estimating species richness: The importance of heterogeneity in species detectability. *Ecology* **79**: 1018-1028.
- Burnham, K.P., Anderson, D.R. (1998): Model selection and inference: A practical information-theoretic approach. New York, Springer-Verlag.
- Crewe, T.L., Timmermans, S.T.A., Jones, K.E. (2005): The Marsh Monitoring Program Annual Report, 1995-2003: Annual indices and trends in bird abundance and amphibian occurrence in the Great Lakes basin. Published by Bird Studies Canada in cooperation with Environment Canada and the U.S. Environmental Protection Agency. ([www.bsc-eoc.org/library.html](http://www.bsc-eoc.org/library.html)).
- Crouch, W.B., Paton, P.W.C. (2002): Assessing the use of call surveys to monitor breeding anurans in Rhode Island. *J. Herpetol.* **36**: 185-192.
- Delis, P.R., Mushinsky, H.R., McCoy, E.D. (1996): Decline of some west-central Florida anuran populations in response to habitat degradation. *Biodiver. and Conser.* **5**: 1579-1595.
- Gibbs, J.P., Whiteleather, K.K., Schueler, F.W. (2005): Changes in frog and toad populations over 30 years in New York state. *Ecol. App.* **15**: 1148-1157.
- Griffith, J.A., Stehman, S.V., Loveland, T.R. (2003): Landscape trends in Mid-Atlantic and Southeastern United States ecoregions. *Environ. Manage.* **32**: 572-588.
- Johnson, C.M., Johnson, L.B., Richards, C., Beasley, V. (2002): Predicting the occurrence of amphibians: An assessment of multiple-scale models. In: *Predicting Species Occurrences: Issues of Accuracy and Scale*, p. 157-170. Scott, J.M., Heglund, P.J., Morrison, M.L., Haufler, J.B., Raphael, M.G., Wall, W.A., Samson, F.B, Eds, Washington D.C., Island Press.
- Kirlin, M.S., Gooch, M.M., Price, S.J., Dorcas, M.E. (2006): Predictors of winter anuran calling activity in the North Carolina Piedmont. *J. North Carolina Acad. of Sci.* **122**: 10-18.
- Knutson, M.G., Sauer, J.R., Olsen, D.A., Mossman, M.J., Hemesath, L.M., Lannoo, M.J. (1999): Effects of landscape composition and wetland fragmentation on frog and toad abundance and species richness in Iowa and Wisconsin, USA. *Conserv. Biol.* **13**: 1437-1446.
- Knutson, M.G., Sauer, J.R., Olsen, D.A., Mossman, M.J., Hemesath, L.M., Lannoo, M.J. (2000): Landscape associations of frog and toad species in Iowa and Wisconsin, USA. *J. Iowa Acad. Sci.* **107**: 134-145.
- MacKenzie, D.I., Nichols, J.D., Lachman, G.B., Droege, S., Royle, J.A., Langtimm, C.A. (2002): Estimating site occupancy rates when detection probabilities are less than one. *Ecology* **83**: 2248-2255.
- MacKenzie, D.I., Nichols, J.D., Hines, J.E., Knutson, M.G., Franklin, A.B. (2003): Estimating site occupancy, colonization, and local extinction when a species is detected imperfectly. *Ecology* **84**: 2200-2207.
- Mossman, M.J., Hartman, L.M., Hay, R., Sauer, J.R., Dhuey, B.J. (1998): Monitoring long-term trends in Wisconsin frog and toad populations. In: *Status and Conservation of Midwest Amphibians*, p. 169-198. Lannoo, M.J., Ed., Aimes, IA, University Iowa Press.
- Oseen, K.L., Wassersug, R.J. (2002): Environmental factors influencing calling in sympatric anurans. *Oecologia* **133**: 616-625.
- Pierce, B.A., Gutzwiller, K.J. (2004): Auditory sampling of frogs: detection efficiency in relation to survey duration. *J. Herpetol.* **38**: 495-500.
- Price, S.J., Marks, D.R., Howe, R.W., Hanowski, J., Niemi, G.J. (2005): The importance of spatial scale for conservation and assessment of anuran populations in coastal wetlands of the western Great Lakes. *Land. Ecol.* **20**: 441-454.
- Royle, J.A., Nichols, J.D. (2003): Estimating abundance from repeated presence-absence data or point counts. *Ecology* **84**: 777-790.

- Semlitsch, R.D., Bodie, J.R. (2003): Biological criteria for buffer zones around wetland and riparian habitats for amphibians and reptiles. *Conserv. Biol.* **17**: 1219-1228.
- Shirose, L.J., Bishop, C.A., Green, D.M, MacDonald, C.J., Brooks, R.J., Helferty, N.J. (1997): Validation tests of an amphibian call count survey technique in Ontario, Canada. *Herpetologica* **53**: 312-320.
- Todd, M.J., Cocklin, R.R., Dorcas, M.E. (2003): Temporal and spatial variation in anuran calling activity in the Western Piedmont of North Carolina. *J. North Carolina Acad. Sci.* **199**: 103-110.
- Weir, L.A., Mossman, M.J. (2005): North American Amphibian Monitoring Program (NAAMP). In: *Amphibian Declines: Conservation Status of United States Species*, p. 307-313. Lannoo, M., Ed., University of California Press.
- Wright, A.H., Wright, A.A. (1949): *The Handbook of Frogs and Toads*. Ithaca, New York, Comstock Publishing Inc.

*Accepted: December 15, 2005.*