Habitat Relations

Distribution of Duck Broods Relative to Habitat Characteristics in the Prairie Pothole Region

JOHANN WALKER,1 Ducks Unlimited, Inc., Great Plains Regional Office, 2525 River Road, Bismarck, ND 58503, USA
JAY J. ROTELLA, Department of Ecology, Montana State University, Bozeman, MT 59717, USA
JOSHUA H. SCHMIDT, Central Alaska Network, National Park Service, Fairbanks, AK 99775, USA
CHARLES R. LOESCH, Habitat and Population Evaluation Team, United States Fish and Wildlife Service, Bismarck, ND 58501, USA
RONALD E. REYNOLDS,2 Habitat and Population Evaluation Team, United States Fish and Wildlife Service, Bismarck, ND 58501, USA
MARK S. LINDBERG, Department of Biology and Wildlife and Institute of Arctic Biology, University of Alaska Fairbanks, Fairbanks, AK 99775, USA
JAMES K. RINGELMAN, Ducks Unlimited, Inc., Great Plains Regional Office, 2525 River Road, Bismarck, ND 58503, USA
SCOTT E. STEPHENS,3 Ducks Unlimited, Inc., Great Plains Regional Office, 2525 River Road, Bismarck, ND 58503, USA

ABSTRACT Conservation programs for breeding ducks in the Prairie Pothole Region (PPR) of the United States and Canada require effective means of evaluating and characterizing breeding habitat across large landscapes. Extensive surveys of the distribution of duck broods in late-summer could help identify wetland basins with greater probabilities of occupancy. Broods are difficult to detect, however, rendering presence–absence data from single-visit surveys difficult to interpret, particularly when probability of detection is related to habitat features. Multiple-visit occupancy surveys offer a potential solution. From 20 July to 5 August 2007–2009, we conducted a 3-visit survey of wetland basins located on 167 10.4-km2 study plots in the PPR. Our survey focused on broods of the 5 most common breeding duck species (Anas spp.). Our main objectives were to investigate ecological relationships between occupancy of wetland basins by broods and habitat characteristics and to examine if habitat-specific detection was of enough concern to warrant multi-survey approaches in the future. We surveyed 3,226 wetland basins during the study. Probability of occupancy of a wetland basin by a brood was positively related to the log of wet area for all 5 study species and was greater on wetlands located on plots with a greater proportion of herbaceous perennial cover for 4 of 5 species. For example, the median probability of occupancy for gadwall (Anas strepera) increased from 0.08 (90% Credible Interval [CrI]: 0.07, 0.10) to 0.28 (90% CrI: 0.23, 0.33) as wet area increased from 0.19 ha to 2.12 ha, and increased from 0.12 (90% CrI: 0.09, 0.16) to 0.20 (90% CrI: 0.16, 0.25) as proportion of perennial grass cover on the study plot increased from 0.03 to 0.99. Because occupancy and detection were both related to attributes of wetland basins, we concluded that the multiple-visit survey was a useful approach for identifying habitat relationships of duck broods. Our results indicated that most broods of the study species were found in 10.4-km2 landscapes with greater densities of small- to mid-sized wetland basins and a greater proportion of herbaceous perennial vegetation. Our study provided new empirical support that could be used to help target conservation actions to the most productive landscapes for breeding ducks. © 2012 The Wildlife Society.

KEY WORDS Anas, brood survey, detection probability, ducks, habitat relationships, waterfowl conservation.

A rigorous understanding of wildlife-habitat relationships facilitates conservation of wildlife populations (Morrison et al. 2006). Habitat conservation for breeding ducks (Anatidae) in the 770,000-km2 Prairie Pothole Region (PPR) of north-central North America is a multi-million dollar effort that depends in part on identification of productive habitat (United States Fish and Wildlife Service and Canadian Wildlife Service 1986, Williams et al. 1999). Knowledge of habitat relationships derived from spatially and temporally extensive operational surveys of breeding pairs (e.g., Cowardin et al. 1995, Reynolds et al. 2006) and from detailed, but much less extensive, studies of reproductive parameters (e.g., nest survival probability; Greenwood et al. 1995, Reynolds et al. 2001, Stephens...
et al. 2005) provides much of the basis for habitat conservation efforts directed at breeding ducks in the PPR. The potential complementary value of information about habitat relationships of duck broods has been recognized by waterfowl ecologists and managers for some time (Hammond 1970, Cowardin and Blohm 1992). For example, occupancy of wetland basins by broods provides information about both reproduction and habitat use. Thus, investigation of presence of broods relative to habitat characteristics could lead to new knowledge about habitat relationships of breeding duck populations. Nevertheless, relatively little is currently known about the habitat relationships of broods at the broad extents most relevant to conservation and management of breeding ducks, in part because of long-standing challenges with interpretation of the results of brood surveys.

Because of the cryptic coloration, small size, and secretive behavior of ducklings, the probability of detecting a brood in a single visit to a wetland basin tends to be low (typically less than 0.5) and variable (reviewed in Pagano and Arnold 2009). Various sampling and analysis methods have been used to estimate detection probability of duck broods, but most of these methods have intensive sampling requirements, require that broods be individually marked, or both (e.g., Ringelman and Flake 1980, Rumble and Flake 1982, Cooper 1996, Giudice 2001). Furthermore, past approaches have largely focused on adjusting observed abundance for imperfect detection. As a result, these methods are more useful for detailed inventories of brood abundance than for extensive studies of habitat relationships.

Wetland-based brood surveys provide a potential basis for identifying habitat features associated with greater probability of occupancy by broods across broad extents. Presence-absence surveys that do not separate occupancy from detection can produce useful information about habitat relationships when detection is unrelated to habitat features (see Johnson 2008), but this condition could be too restrictive for wetland-based surveys of duck broods. For example, the amount of emergent vegetation covering a wetland affects both occupancy (Bloom 2010) and detection (Giudice 2001) of mallard (Anas platyrhynchos) broods. When a habitat variable affects both occupancy and detection, relationships with detection probability can be erroneously attributed to occupancy. When habitat relationships differ in direction or functional form and occupancy and detection are not separated by design, underestimation of regression coefficients, failure to detect relationships, and misidentification of functional forms are all possible (Kéry 2008, Royle and Dorazio 2008). Sampling designs based on multiple surveys of a set of spatially distinct units allow separate and simultaneous estimation of the probability of occupancy and the probability of detection of unmarked animals relative to covariates (MacKenzie et al. 2006) thereby relaxing the assumptions of an unadjusted survey. Because duck broods occupy spatially distinct wetland basins that can be rapidly surveyed and multiple-visit surveys do not require individually marked broods, this approach potentially provides a path to better information about brood-habitat relationships.

We conducted an extensive, multiple-visit survey of wetlands in the PPR, counted broods of the 5 most common duck species, and applied hierarchical occupancy models that corrected for incomplete detection to the resulting data. Our main objective was to investigate predictions about relationships between occupancy of wetland basins by broods and habitat characteristics. We hypothesized broadly that the distribution of duck broods in late-summer in the PPR is a realization of the effects of environmental conditions and land use on 1) the availability of productive, shallow-water wetland habitat for breeding pairs, nesting females, and developing ducklings and 2) survival rates of nests, ducklings, and breeding females.

**STUDY AREA**

Our study was conducted during 20 July to 5 August 2007–2009 in the PPR of North and South Dakota, essentially the area east of the Missouri River in these 2 states (Fig. 1). The study area was characterized by its high density of wetland basins (van der Valk 1989), heterogeneous land cover and land use (Johnson et al. 1994), and large population of breeding ducks (Zimpfer et al. 2009). Abundant, glacially formed wetland basins were the characteristic ecological feature of the landscape and ponded water in these wetland basins attracted high densities of breeding waterfowl during years of adequate precipitation. Most land was in private ownership, and the most common land uses were related to production of small grains, row crops, and beef cattle. Land cover was a mosaic of annually cultivated cropland, stands of introduced perennial grasses and forbs used for forage, perennial grassland used for pasture, and restored, mostly idle perennial grassland retired from cultivation under the United States Department of Agriculture’s Conservation Reserve Program. Climate, physiography, dominant land uses, and ecology of the study area are described in detail in previously published work (e.g., Johnson et al. 1994, Cowardin et al. 1995, Reynolds et al. 2006).

**METHODS**

**Sampling Design and Data Collection**

Sampling units in our brood occupancy study were National Wetlands Inventory (NWI; United States Fish and Wildlife Service 2010) wetland basins (hereafter wetlands) located within a sample of 167 10.4-km² study plots taken from ongoing projects in the study area. We sampled seasonal, semipermanent, and temporary wetlands, which are the 3 most abundant and productive wetland classes in the PPR (Stewart and Kantrud 1971, Reynolds et al. 2006). We focused our effort on broods of the 5 most abundant breeding duck species in the PPR: blue-winged teal (Anas discors), gadwall, mallard, northern pintail (A. acuta), and northern shoveler (A. clypeata; Zimpfer et al. 2009). We chose the late-summer survey period as a compromise between the breeding chronologies of earlier- and later-nesting species (Klett et al. 1988, Cooper 1996).

We selected the majority (140) of the sample plots from the 384 plots that made up the United States Fish and Wildlife
Service (USFWS) Four-Square-Mile Survey (FSMS) sample frame in North and South Dakota. As part of an operational survey of distribution and abundance of breeding duck pairs, these plots were randomly selected in 1987 and 1990 from 3 strata that described areas with high, medium, and low proportion of land owned by the USFWS (Cowardin et al. 1995). This plot dimension was chosen for the FSMS because of its relationship to the observed breeding season home range of radio-marked female mallards (Dwyer et al. 1979). We selected plots from the FSMS sample with the objectives of 1) including plots that spanned a representative gradient of agricultural intensity (indexed by the proportion of herbaceous perennial cover) and number of wetlands on a plot and 2) minimizing travel time between plots. We included another group of 27 10.4-km² plots in the sample during 2008 and 2009 where Ducks Unlimited, Inc. (DU) was conducting an ongoing investigation of nest survival probability of ducks. The DU plots were selected randomly from 9 strata formed by combining 3 levels of proportion of perennial grass cover (<0.47, 0.48–0.74, and ≥0.75; Stephens et al. 2005) and 3 levels of number of NWI basins (<90, 91–180, and ≥180).

We required a survey design that could be implemented across a large region by field assistants with a broad range of experience, thus we developed a relatively simple survey protocol (Sewell et al. 2010). Because of the logistical difficulty of obtaining trespass permission from private landowners and our interest in covering a large geographic area, we conducted all surveys from public roadsides, and observers remained in the vehicle during the survey. We provided observers with geo-referenced maps consisting of an aerial photo of each plot (United States Department of Agriculture 2010) overlaid with uniquely numbered NWI wetland polygons and public roads. We buffered public roads on both sides to a distance of 200 m from the center of the road. During each visit, the observer surveyed only the portion of every wetland that contained water; was not completely obscured by vegetation, terrain, or other obstructions; and intersected the 200-m buffer. A plot-level survey comprised 3 visits to every wetland on a sample plot. We conducted all 3 visits within a 24- to 36-hour period. Visits began at sunrise and continued until sunset. A minimum of 4 hours elapsed between visits. The same observer conducted all 3 visits to a given plot in a given year. During each visit,
observers used binoculars and spotting scopes to observe broods and recorded species, age class, and number of ducklings of each observed brood. If no broods were observed, then observers recorded a zero. Observers remained at each wetland for at least 2 minutes and recorded date, time, and wind speed (Beaufort scale; Simpson 1926) at the beginning of each visit to a plot. Observers recorded an ocular estimate (±10%) of the proportion of the surveyed wetland area covered by emergent vegetation during the first visit. We did not differentiate types of vegetation.

Observers had a wide range of experience with waterfowl surveys (from 0 yrs to >20 yrs), and the overall experience level of the observer pool varied among years. We therefore worked to mitigate this potential source of heterogeneity in detection by requiring observers with no prior brood survey experience to conduct training surveys for 3–5 weeks prior to collecting data for our study. Training surveys were conducted under the supervision of an observer with previous brood survey experience. We provided training on the protocol only to observers who self-identified as having experience with waterfowl surveys. All observers used the same survey protocol. Our observation protocol was approved in June 2009 by the University of Alaska Fairbanks Institutional Animal Care and Use Committee (protocol # 09-35).

**Data Analysis**

Our analysis focused on 2 parameters: probability of occupancy of a wetland by a brood ($\psi$) and the probability that a brood was available for detection and was detected during a survey of a wetland ($p$) given presence (Nichols et al. 2009). The distribution of broods among wetlands during late summer is most likely a complicated function of pair abundance, nesting chronology, nest survival rate, brood survival rate, and the use of available wetland habitat by female ducks during the brood-rearing period. We hypothesized that variation in $\psi$ would therefore be related to habitat characteristics that were related to the size and distribution of the breeding pair population and to the reproductive success and habitat use of breeding females. We hypothesized that variation in $p$ would be related to covariates that described variation in the detectability of broods associated with brood behavior, brood age, brood size, habitat characteristics, survey timing, and observer experience. We selected covariates based on their linkage to our hypotheses given existing knowledge of brood ecology and wildlife count surveys.

Our hypotheses and predictions about $\psi$ were based on knowledge of the breeding chronology, habitat-use, and reproductive success of the study species. Because the plot-level surveys were completed over an approximately 2-week survey season in each year, we hypothesized that broods of later-nesting species might increase within a season because of additions of new broods from late-hatched nests and that broods of earlier-nesting species might decrease because of losses of broods to mortality and fledging. An increasing within-season trend in nest hatching probability (e.g., Greenwood et al. 1995) could produce the same pattern. We therefore predicted greater average $\psi$ for broods of blue-winged teal, gadwall, and northern shoveler (later-nesting species) than mallard or northern pintail (earlier-nesting species). Further, we predicted an increasing within-year trend in probability of occupancy for later-nesting species and a decreasing within-year trend in probability of occupancy for earlier-nesting species.

Given our objective of learning about the utility of our survey design for extensive assessment of habitat relationships, we were particularly interested in the relationships between $\psi$ and characteristics of wetlands. We hypothesized that gadwall, northern shoveler, and northern pintail might be less likely to use wetlands with greater coverage of emergent vegetation (Murkin et al. 1997, Reynolds et al. 2006) and that blue-winged teal and mallard would be more likely to use wetlands with more emergent cover (Ringelman and Flake 1980, Bloom 2010). Abundance of breeding pairs of dabbling ducks is positively, but nonlinearly, related to the area of a wetland containing ponded water (hereafter wet area), and this relationship is thought to result from a nonlinear increase in productive shallow-water habitat that is proportional to wetland perimeter (Cowardin et al. 1995, Reynolds et al. 2006). We therefore predicted a positive relationship between $\psi$ and the logarithm of wet area. Hatching probability of duck nests is positively related to the amount of herbaceous perennial cover (or negatively related to the amount of cropland) on the landscape (Greenwood et al. 1995, Reynolds et al. 2001, Stephens et al. 2005). We predicted that $\psi$ would thus be greater on plots with a greater proportion of herbaceous perennial cover. We included 2 plot-level covariates describing the wetland conditions on each plot in each year as a potential way of indirectly controlling for variation in the size of the breeding pair population and the availability of wetland habitat for brood-rearing. We predicted that if the size of the plot-level pair population was positively related to the number of seasonal, semipermanent, and temporary wetlands containing ponded water in May (Cowardin et al. 1995, Reynolds et al. 2006) then $\psi$ would be positively related to the count of ponds in wetlands of these classes on the plot in May. Because female mallards, and presumably other ducks, select brood-rearing habitat non-randomly (Rotella and Ratti 1992, Raven et al. 2007, Bloom 2010), we predicted that $\psi$ would be less in landscapes with more available ponded water in seasonal, semipermanent, and temporary wetlands in July. That is, given a greater amount of available wetland habitat, we predicted reduced probability of occupancy at the level of individual wetlands.

We predicted that given our somewhat limited and indirect knowledge of brood distribution, variation in $\psi$ beyond that explained by the covariates was likely. We therefore included a plot-year-level distribution of random effects (i.e., a random effect for each plot in each year) in our model of $\psi$. These random effects provided additional flexibility to control for missing covariates and potential overdispersion (Link and Barker 2009) and thus rendered our results more generally applicable to the population of 10.4-km² landscapes in the study area.

Our predictions about probability of detection were based on a relatively large group of past assessments of probability
of detection of broods in the PPR. These surveys used diverse methods and were conducted on many different study areas during a roughly 30-year period, thus they provided a large body of information for constructing predictions about \( p \). Unlike past studies that used individually identifiable broods (e.g., Pagano and Arnold 2009), we could not use brood-level covariates in our wetland-based study of unmarked broods. We nonetheless predicted that part of the variation in \( p \) among wetlands would be related to differences in brood-rearing behavior, brood age, and brood size among species. We expected \( p \) to be greatest for gadwall because of their use of relatively open, semi-permanent wetlands (Murkin et al. 1997), and their likely larger average brood size in July (Klett et al. 1988). Further, we expected that older broods might be more detectable than younger broods (Ringelman and Flake 1980). Given that mallards and northern pintails nest earlier than other species, we expected mallard and pintail broods to be older than other broods and to therefore have greater \( p \) in our surveys.

To account for variation in \( p \) due to survey timing and habitat characteristics, we included several covariates in each species-specific model: date and time of the survey, wind speed during the survey, percent emergent vegetation cover and wet area of the surveyed wetland, and a binary variable describing whether a brood had been observed on that wetland on the previous visit. We predicted that \( p \) would increase during the 15-day survey period if observers became more proficient or if broods became more detectable because of changes in age distribution (Ringelman and Flake 1980, Giudice 2001, Pagano and Arnold 2009). Broods are generally more active and thus more easily detected early and late in the day (Ringelman and Flake 1980, Pagano and Arnold 2009). Therefore, we modeled detection as a function of time of day using a quadratic curve, which we predicted would result in a concave-up functional relationship with the smallest detection probability associated with visits conducted during mid-day. Emergent vegetation can obscure broods from view (Giudice 2001). Thus, we predicted that \( p \) would be negatively related to the area of the wetland covered by emergent vegetation. We predicted that \( p \) would be negatively related to wind speed if broods increasingly took cover in emergent vegetation or on shorelines as wind speed increased (Ringelman and Flake 1980, Giudice 2001, Pagano and Arnold 2009). We predicted that \( p \) would be negatively related to wet area because the observer’s effectiveness might decrease as area surveyed increased (Pagano and Arnold 2009, A. Royle, United States Geological Survey, unpublished data). Other researchers have found strong evidence that \( p \) is greater on sampling units where the study organism was detected on a previous visit (Riddle et al. 2010), thus we predicted that detection probability would be greater when broods had been detected in a previous visit.

Observer experience can also affect \( p \) (Diefenbach et al. 2003, MacKenzie et al. 2006), and our surveys were conducted by a large group of observers whose experience level varied within and among years. For efficiency, the same observer conducted all visits to a given plot in a given year. Thus, observer effects were largely confounded with unspecified plot-year-level variation that may have been present in \( p \). We consequently used random-effects at the plot-year level (i.e., the data collected on a given plot in a given year) to account for variation in \( p \) associated with variation in observer experience and variation not captured by the covariates.

We obtained values of covariates for input to the model from 3 sources: 1) information recorded by observers; 2) georeferenced, 1-m aerial videography of study plots (Cowardin et al. 1995, Reynolds et al. 2006); and 3) USFWS land cover maps (M. E. Estey, USFWS, unpublished data). Using the 3-visit encounter history for each wetland, we derived a binary covariate describing the expectation of detecting a brood, given a brood was detected on the previous visit (Previous detection). We used aerial videography of each surveyed plot, captured in May and July of each year, to calculate wetland variables at a 1-m resolution (Cowardin et al. 1995, Reynolds et al. 2006). We georeferenced aerial videography, digitized wetlands, and calculated the year-specific area of ponded water associated with each surveyed wetland basin (Wet area) with ArcView 9.3 (Environmental Systems Research Institute, Redlands, CA), and we used wet area to characterize variability in area of wetlands. We derived the number of wetlands containing ponded water in May (May ponds) and the total area of ponded water on the plot during the survey (July wet area) from the May and July videography, respectively. We calculated the proportion of upland area on the plot composed of herbaceous perennial cover (Perennial cover) from USFWS land cover (M. E. Estey, USFWS, unpublished data).

Our sampling design included a broad gradient of wetland basin size and perennial cover, but we could not control for potential correlation among habitat covariates due to changing wetland conditions among years. We consequently calculated summary statistics for observed covariates so that we could better assess our inference space. Multicollinearity can cause difficulty with estimation of regression coefficients and sampling variances in statistical models (Graham 2003). We assessed potential multicollinearity by estimating pairwise correlations and variance inflation factors (VIFs) using R (R Development Core Team 2010) and the AED package (Zuur et al. 2007). We removed any variables with VIF > 3 from the analysis.

We used a version of the binomial mixture model developed and described by MacKenzie et al. (2006) to estimate \( p \) and \( \psi \) given the observation and habitat covariates (Royle and Dorazio 2008). The basic structure of this hierarchical, generalized linear mixed model was described by the following statements:

\[
\begin{align*}
y_{ij} \mid z_i & \sim \text{Bin}(J, z_i p_{ij}) \\
\log \left( \frac{p_{ij}}{1 - p_{ij}} \right) &= \alpha_0 + \alpha_1 z_{i1} + \cdots + \alpha_s z_{i2} + \epsilon_{i\text{plot-year}} \\
z_i & \sim \text{Bin}(1, \psi_i) \\
\log \left( \frac{\psi_i}{1 - \psi_i} \right) &= \beta_0 + \beta_1 x_{i1} + \cdots + \beta_k x_{ik} + \epsilon_{i\text{plot-year}}
\end{align*}
\]
where \( y_{ij} \) was the observed presence of broods on wetland \( i = 1, 2, \ldots, I \) during visit \( j = 1, 2, \ldots, J \). \( \psi_j \) was the visit-level probability of detection, \( \psi_j \) was the wetland-level probability of occupancy, the \( \alpha_k \) and \( \beta_k \) were logit-scale regression coefficients associated with the predictors of detection and occupancy respectively, and the \( \epsilon_{\text{plot-year}} \) and \( \epsilon_{\text{site-year}} \) were random effects associated with detection and occupancy specific to a plot and year. Random effects were modeled as normally distributed deviations from the overall mean (i.e., \( \alpha_k \) and \( \beta_k \) with mean 0 and standard deviation \( \sigma_{\text{plot-year}} \). We did not model multiple random effects in a hierarchical structure (e.g., plots within years) because our study only captured 3 years, thus our models assumed independence of plots within and among years. The model was valid given the following assumptions: 1) occupancy status of the wetland remained constant during the survey, 2) variation in occupancy and detection among wetlands and visits was adequately described by the predictors and the random-effects, 3) detections of broods were independent among wetlands and visits, and 4) false detections were rare or non-existent (MacKenzie et al. 2006:104). We addressed the first assumption in 2 ways. At the wetland level, we sought to minimize the potential effect of changes in occupancy status of wetlands by conducting all 3 visits in a 24-hour period. At the year level, we incorporated the date covariate to account for a potential within-year trend in \( \psi \) due to hatching and immigration of new broods and mortality and emigration of existing broods over the 15-day survey season. We addressed the second and third assumptions by 1) using a group of covariates that were consistent with existing knowledge, 2) including the previous detection covariate to explicitly account for potential dependence in \( \rho \) among visits due to previous detections, 3) incorporating the more flexible random-effects structure, and 4) testing the model for major lack-of-fit before making inferences. False positive detections of broods could have occurred if observers mistook a group of fully feathered adult ducks for a brood. However, we believe that such mistakes were rare because 1) broods can be distinguished from groups of adult ducks by plumage characteristics and behavior (Gollop and Marshall 1954) and 2) this potential violation of the fourth assumption was addressed in the sampling protocol and during observer training. Specifically, observers were instructed to carefully evaluate fully feathered broods for bright plumage, a tendency to remain together, and flightlessness.

Our decision to use multiple-visit surveys and a hierarchical modeling approach helped to alleviate potentially serious problems associated with non-detection, missing covariates, and overdispersion but created a challenge for selecting an approximating model. Hierarchical mixed-effects models are realistic and flexible, but they take much longer to program and run than their non-hierarchical counterparts. They also present significant, unresolved challenges for model selection (Gelman et al. 2004, Royle and Dorazio 2008, Link and Barker 2009). The number of covariates for \( \psi \) (6) and \( \rho \) (7) defined a large potential set of nested models for each species (i.e., \( \geq 2^{13} - 1 \) models \( \times 5 \) species). All additive combinations of the covariates were plausible given the a priori ecological justification for the inclusion of each covariate, but we could not run such a large set of models in a reasonable amount of time. On the other hand, data-based identification of a parsimonious model for each species was desirable both for further protection against spurious effects and to avoid unnecessary loss of precision due to estimation of unsupported parameters. Consequently, we took a 2-stage approach to data analysis.

In the first stage of analysis, we used the contributed R (R Development Core Team 2010) package unmarked (Fiske and Chandler 2010) to select a reduced model by eliminating unsupported parameters from the full 15-parameter regression model for each of the 5 species. We used a drop 1 approach analogous to model reduction based on Type II sums-of-squares in analysis of variance (ANOVA; Chambers 1992) to identify unsupported parameters. Specifically, we removed individual parameters from the full model and estimated the Akaike’s Information Criterion (AIC; Burnham and Anderson 2002) value of the resulting reduced model. We repeated this step for each of the parameters in the full model. This process created a set of 13 models with either 13 or 14 parameters per model (a 13-parameter model was created when both Time and \( T^2 \) were removed). We then examined the AIC difference between each reduced model and the full model. We defined unsupported parameters as those that were associated with a decrease in AIC relative to the full model when they were removed from the full model. We then removed the entire subset of unsupported parameters from the full model to create a reduced model for each species. Reduced models thus included the subset of parameters associated with increases in AIC relative to the full model when held out separately. Finally, we ran the resulting reduced model and compared its AIC value to the AIC value of the full model. We used the reduced model for inference when its AIC value indicated that it was a substantive improvement over the full model or when the full and reduced models differed by the inclusion of 1 or more uninformative parameters (sensu Arnold 2010).

In the second stage of analysis, we used a Bayesian approach and Markov Chain Monte Carlo (MCMC) simulation to obtain an estimate of the joint posterior distribution of the logit-scale coefficients and standard deviations of the random effects distributions for each of the species-specific, reduced models selected in the first stage. We conducted the MCMC simulation in WinBUGS 1.4 (Spiegelhalter et al. 2003) through the R package R2WinBugs (Sturtz et al. 2005). The Bayesian analysis followed established conventions, as described in Gelman et al. (2004). We used minimally informative, compact prior distributions for model parameters (Royle and Dorazio 2008). To facilitate convergence, we standardized values of each covariate by subtracting the mean and dividing by the standard deviation of the observed values. We squared time after standardization and standardized log(wet area) after taking the log. We ran 2 Markov chains with random initial values until we had 200,000 simulations for each chain, and we discarded the first 100,000 simulations from each chain to assure minimal
influence of initial values on the estimates (see Supplementary Material for details of prior distributions, model structure, and model diagnostics, available online at www.onlinelibrary.wiley.com).

Using the 200,000 post-convergence simulations, we constructed estimates of the posterior distributions of 1) logit-scale regression coefficients, 2) standard deviations of random effects distributions, and 3) back-transformed estimates of occupancy and detection probability at selected covariate values for each of the 5 study species. We evaluated relationships between \( \psi \) (or \( \rho \)) and the covariates based on the magnitude and direction of estimated coefficients, the number of species showing a similar pattern, and the medians and 90% credible intervals (90% Credible Interval [CrIs]) of the estimated posterior distributions of logit-scale regression coefficients and back-transformed probabilities. To evaluate the magnitude of covariate effects, we varied a single covariate while holding other variables constant at their mean values and evaluated the change in the estimated probability of occupancy or detection. When the same coefficient was included in models for multiple species and 90% credible intervals did not include zero, we presented results for the 2 species with the smallest and largest proportional response to the covariate.

RESULTS

Sample Characteristics

Our sample comprised 3,226 wetlands on 167 plots. We conducted 4,356 wetland surveys during the study; 2,213 were surveyed in a single year, 896 were surveyed in 2 of the 3 years, and 117 were surveyed in all 3 years. We visited each surveyed wetland 3 times, thus the data consisted of \( 3 \times 4,356 = 13,068 \) visits. We sampled 77 plots in 2007, 146 plots in 2008, and 148 plots in 2009; 26 plots were surveyed in a single year, 78 plots were surveyed in 2 of the 3 years, and 63 plots were surveyed in all 3 years. Among-year variation in the number of sample plots and wetlands resulted largely consistent with predictions. At mean values of the covariates, posterior median \( \psi \) (or \( \rho \)) and VIFs ranged from 1.01 to 1.13. Correlations among covariates were not large enough to hamper estimation or interpretation. Correlations among detection predictors ranged from \(-0.33\) to 0.20 and VIFs ranged from 1.00 to 1.13. Observers with some prior experience with waterfowl surveys conducted 100%, 36%, and 15% of the surveys in 2007, 2008, and 2009, respectively. Overall, observers with some prior experience collected 38% of the data.

Model Selection and Parameter Estimation

For blue-winged teal, mallard, northern pintail, and northern shoveler, the reduced model represented a substantive reduction in AIC relative to the full model. Number of parameters in reduced models ranged from 10 for mallard to 14 for gadwall. For gadwall, the AIC difference between the full and reduced models was less than 2 (Table 2). We therefore carried reduced models forward into the second stage of analysis for all 5 species. In the second-stage Bayesian analysis, all 5 species-specific, reduced models converged to the posterior distribution of fixed-effects coefficients and random-effects standard deviations and none of the models displayed evidence of lack-of-fit (see Supplementary Material available online at www.onlinelibrary.wiley.com).

Table 1. Descriptive statistics for observed covariates used in a hierarchical occupancy model describing variation in the distribution of duck broods in North and South Dakota, USA during 2007–2009.

<table>
<thead>
<tr>
<th>Probability</th>
<th>Covariate</th>
<th>Level</th>
<th>Min.</th>
<th>Q1</th>
<th>Median</th>
<th>Q3</th>
<th>Max.</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Detection (( \rho ))</td>
<td>Time</td>
<td>Visit</td>
<td>360</td>
<td>487</td>
<td>664</td>
<td>852</td>
<td>1,213</td>
<td>685.89</td>
<td>213.43</td>
</tr>
<tr>
<td>Wind speed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Occupancy (( \psi ))</td>
<td>July wet area</td>
<td>Plot</td>
<td>0.13</td>
<td>33.73</td>
<td>61.96</td>
<td>107.82</td>
<td>406.35</td>
<td>76.58</td>
<td>61.62</td>
</tr>
<tr>
<td>May ponds</td>
<td>Plot</td>
<td>1.00</td>
<td>27.00</td>
<td>45.00</td>
<td>72.00</td>
<td>442.00</td>
<td>64.22</td>
<td>65.24</td>
<td></td>
</tr>
<tr>
<td>Both (( \rho ) and ( \psi ))</td>
<td>Perennial cover</td>
<td>Plot</td>
<td>0.03</td>
<td>0.23</td>
<td>0.54</td>
<td>0.82</td>
<td>0.99</td>
<td>0.53</td>
<td>0.29</td>
</tr>
<tr>
<td>Emergent cover</td>
<td>Wetland</td>
<td>0.00</td>
<td>0.19</td>
<td>0.63</td>
<td>2.12</td>
<td>106.35</td>
<td>2.68</td>
<td>7.58</td>
<td></td>
</tr>
<tr>
<td>Wet area</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Date</td>
<td>Plot</td>
<td>1</td>
<td>2</td>
<td>6</td>
<td>8</td>
<td>17</td>
<td>5.87</td>
<td>4.04</td>
<td></td>
</tr>
</tbody>
</table>

\( a \) Time was the time in minutes since 0000 CDT that the count took place, wind speed was the wind level during the visit according to the ordinal Beaufort scale, July wet area was the summed hectares of ponded water on the 10.4-km\(^2\) study plot in July, May ponds was the count of wetland basins holding ponded water on the study plot in May, perennial cover was the proportion of the upland area of the study plot composed of perennial vegetation, emergent cover was the proportion of the surveyed wetland covered by emergent vegetation, wet area was the number of hectares of ponded water associated with the surveyed wetland basin, and date was the date the survey was conducted (Day 1 = July 20).

\( b \) First quartile (25th percentile).

\( c \) Third quartile (75th percentile).
0.06, 0.14) for blue-winged teal, mallard, gadwall, northern shoveler, and northern pintail, respectively.

When 90% credible intervals did not overlap zero, the direction of coefficient estimates describing relationships between $\psi$ and the covariates was consistent with our predictions for date (no. of species: 2), emergent vegetation (3), perennial cover (4), July wet area (5). For date (2) and May ponds (2), the direction of coefficient estimates was opposite of predictions (Fig. 2). Relationships between $\psi$ and Perennial cover and between $\psi$ and log(Wet area) were associated with large enough changes in $\psi$ to be of ecological and management interest (Fig. 3; Table S1 available online at www.onlinelibrary.wiley.com).

For example, the median probability of occupancy for gadwall increased from 0.08 (90% CrI: 0.07, 0.10) to 0.28 (90% CrI: 0.23, 0.33) as wet area increased from 0.19 ha to 2.12 ha, and increased from 0.12 (90% CrI: 0.09, 0.16) to 0.20 (90% CrI: 0.16, 0.25) as proportion of perennial grass cover on the study plot increased from 0.03 to 0.99.

We found substantial unexplained variation in occupancy for all 5 species. The logit-scale standard deviation of the plot–year random-effects for $\psi$ was 0.84 (90% CrI: 0.66, 1.04), 0.79 (90% CrI: 0.58, 1.00), 0.85 (90% CrI: 0.65, 1.05), 0.56 (90% CrI: 0.10, 1.04), and 1.10 (90% CrI: 0.87, 1.37) for blue-winged teal, gadwall, mallard, northern pintail, and northern shoveler, respectively. For northern pintail, the species with the smallest standard deviation, median probability of occupancy with all covariates held constant at their mean values ranged from 0.03 (90% CrI: 0.01, 0.10) to 0.24 (90% CrI: 0.14, 0.39) over a range of 2 $\sigma_{\text{year}}$. For northern shoveler, the species with the largest standard deviation, median probability of occupancy with all covariates held constant at their mean values ranged from 0.01 (90% CrI: 0.01, 0.03) to 0.55 (90% CrI: 0.44, 0.67) over a range of 2 $\sigma_{\text{year}}$.

Our predictions about variation in $\psi$ among species were not supported. At mean covariate values, median estimated detection probability ($\hat{p}$) was 0.27 (90% CrI: 0.27, 0.31) for blue-winged teal, 0.30 (90% CrI: 0.24, 0.35) for gadwall, 0.27 (90% CrI: 0.23, 0.31) for mallard, 0.12 (90% CrI: 0.07, 0.20) for northern pintail, and 0.23 (90% CrI: 0.18, 0.29) for northern shoveler. Estimated probability of detection varied with survey timing, survey conditions, and habitat characteristics (Fig. 4). Relationships between $\hat{p}$ and the covariates were consistent with predictions for Time and Time$^2$ (no. of species: 2), Wind speed (2), Emergent vegetation (3), July wet area (2), and the log of wetland area (5).

### Table 2.

Models of occupancy and detection of duck broods in the Prairie Pothole Region (2007–2009) reduced by using changes in Akaïke’s Information Criterion (AIC) to identify and remove parameters that made little contribution to fit of the full model. Differences in AIC between the full and reduced models (i.e., $\text{AIC}_{\text{full}} – \text{AIC}_{\text{reduced}}$) were 5.7, 1.2, 8.2, 4.7, and 7.1 for blue-winged teal (BWTE), gadwall (GADW), mallard (MALL), northern pintail (NOPI), and northern shoveler (NSHO), respectively.

<table>
<thead>
<tr>
<th>Species</th>
<th>Date</th>
<th>Emergent cover</th>
<th>Perennial cover</th>
<th>May ponds</th>
<th>July wet area</th>
<th>log(Wet area)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BWTE</td>
<td>0$^a$</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>GADW</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>MALL</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>NOPI</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>NSHO</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Species</th>
<th>Date</th>
<th>Time</th>
<th>Time$^2$</th>
<th>Wind speed</th>
<th>Emergent cover</th>
<th>Previous detection</th>
<th>Wet area</th>
</tr>
</thead>
<tbody>
<tr>
<td>BWTE</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>GADW</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>MALL</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>NOPI</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>NSHO</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

$^a$ 1 indicates that the parameter was retained in the reduced model, and 0 indicates that it was not included.
Previous detection (4), and Wet area (5). Detection of mallard broods decreased linearly throughout the day. Other estimates were consistent with predictions, but 90% credible intervals overlapped zero (Fig. 4). The largest changes in $\hat{p}$ were associated with emergent vegetation and previous detection (Fig. 5 and Table S1 available online at www.onlinelibrary.wiley.com).

We observed substantial unexplained variation in $\hat{p}$ among plot-years. Estimated, logit-scale, median standard deviations of random-effects distributions (i.e., $\hat{\sigma}_{\text{plot-year}}$) were 0.78 (90% CrI: 0.62, 0.93), 1.07 (90% CrI: 0.84, 1.31), 0.65 (90% CrI: 0.45, 0.85), 1.17 (90% CrI: 0.77, 1.55), and 0.81 (90% CrI: 0.58, 1.06) for blue-winged teal, gadwall, mallard, northern pintail, and northern shoveler, respectively. For mallard, the species with the smallest standard deviation, median detection probability with all covariates held constant at their mean values ranged from 0.08 (90% CrI: 0.09, 0.11) to 0.57 (90% CrI: 0.52, 0.62) over a range of 2 $\hat{\sigma}_{\text{plot-year}}$. For northern pintail, the species with the largest standard deviation, median detection probability with all covariates held constant at their mean values ranged from 0.01 (90% CrI: 0.00, 0.05) to 0.58 (90% CrI: 0.47, 0.70) over a range of 2 $\hat{\sigma}_{\text{plot-year}}$.

**DISCUSSION**

Our research built on past studies of brood ecology and past brood surveys to develop a more extensive understanding of habitat relationships of duck broods. Although some of our predictions were not supported, we were able to identify habitat associations of duck broods that were consistent across a large geographic extent despite considerable variation in the data. The multiple-visit, wetland-based sampling design that we used provided new, useful information for conservation planners seeking to identify wetlands and landscapes associated with greater probability of occupancy by broods. Our study demonstrates that multiple-visit wetland-based brood surveys are a feasible, repeatable, and informative approach for broad-scale investigation of habitat relationships of duck broods.

Probability of occupancy increased with the log of wetland area for all study species. This wetland-level association was the strongest and most consistently observed relationship between $\psi$ and habitat characteristics. This pattern was broadly consistent with the repeatedly observed pattern of use of larger seasonal and semi-permanent wetlands by broods (Talent et al. 1982, Rotella and Ratti 1992, Krapu et al. 2006, Raven et al. 2007, Bloom 2010), but our research extended previous work by revealing that, given equal wet area, many small- to mid-sized wetlands would provide more brood habitat than fewer large wetlands. The number of breeding pairs of the study species occupying a wetland also tends to increase with wet area at a decreasing rate, possibly as a result of a nonlinear increase in the availability of productive shallow-water foraging habitat associated with
increasing wet area (Reynolds et al. 2006). We think that this explanation may also apply to broods. Probability of occupancy was positively related to the percent of herbaceous perennial cover on the study plot for 4 of the 5 study species. For example, the estimated rate of wetland occupancy for northern pintail increased from 1 brood per 20 wetlands (90% CrI: 13, 39) to 1 brood per 6 wetlands (90% CrI: 4, 9) as perennial cover increased from 3% to 99%. Multiple, non-mutually exclusive mechanisms could explain this pattern. Greater nest survival probabilities have been repeatedly observed in landscapes with greater amounts of perennial cover (e.g., Greenwood et al. 1995, Reynolds et al. 2001, Stephens et al. 2005). Perhaps greater hatching rates on plots with more perennial cover resulted in a larger number of broods. Alternatively, wetlands in less intensively cultivated landscapes might provide higher-quality food resources for breeding ducks (Gleason et al. 2003) leading to larger breeding populations, greater recruitment rates, or both on plots with greater amounts of perennial cover. Our results are inconsistent with recent observations that duckling survival is negatively related to the total amount of perennial vegetation (including trees; Amundson and Arnold 2011) and to the amount of managed hayland in the surrounding landscape (Bloom et al. 2012). However, our metric of perennial cover did not include trees, and probably included little hayland managed for duck nesting. In contrast to the other species, was not associated with perennial cover for mallard broods. This pattern might have been related to the distribution of mallard pairs in the study region. Breeding mallards tend to be more widespread than some other common duck species in the PPR, and mallards reach similar densities in both crop- and grass-dominated landscapes (Reynolds et al. 2006).

The ability to use data from extensive, wetland-based surveys of unmarked broods to separately and simultaneously estimate occupancy and detection relative to habitat covariates was a key strength of our study. We observed complicated relationships between occupancy, detection, and wetland attributes that could not easily have been resolved with a single-visit approach. We observed negative relationships between and percent emergent cover for blue-winged teal, gadwall, and mallard. Analysis of data from a single-visit would have supported a negative relationships between the product (i.e., ) of these parameters and emergent cover for all 5 species. We observed positive relationships between and the log of wet area and negative relationships between and wet area for all 5 species. The most likely consequence of combining these opposing relationships in a single-visit analysis would have been both misidentification of shape and direction and over- or under-estimation of the parameter. We therefore concluded that a multiple-visit, wetland-based sampling design was preferable to an unadjusted single-visit survey for understanding habitat relationships of duck broods at a regional scale.

The spatial and temporal extent of our data had important implications for the scope of our results. In the PPR, variation in wetland conditions and ecosystem productivity associated with episodes of wetness and drought is substantial (Johnson et al. 2004). Wet–dry episodes operate at more extensive scales than those captured by our study and produce substantial variation in abundance and reproductive success of breeding ducks (Reynolds et al. 2006, Walker 2011). Our study did not continue for a sufficient number of years to capture the full range of variation in wetness and drought that characterizes the PPR, nor did it include landscapes from the entire PPR. Because our sample included data from both wet and dry years taken from a large sample of landscapes, we speculate that the relationships between brood occupancy and habitat characteristics that we observed may be generalizable to the broader PPR. Nonetheless, we suggest that collection of similar data on multiple sites across the PPR for 10–15 years would help to identify wetlands and landscapes with consistently greater probabilities of supporting broods and further test our conclusions.

Our survey relied on data collected from a single survey period to characterize occupancy of wetlands by broods of 5 species of ducks, which was probably somewhat late for earlier-nesting species and somewhat early for later-nesting species. Our ability to fully understand species-specific use of brood habitat was therefore most likely somewhat reduced. As a compromise between fully species-specific survey timing and a single survey period, we recommend that 2 sets of
Multiple-visit occupancy surveys be conducted each year with an early survey period in late June–early July directed at broods of earlier-nesting species and a later survey period in late July–early August directed at broods of later-nesting species. This approach would be consistent with the approach used by the USFWS on the Four-Square-Mile Survey of breeding duck pairs (Reynolds et al. 2006).

We regard our study not only as a contribution to the understanding of broad-scale habitat relationships of duck broods in the PPR, but also as an example of the well-known trade-off between extent and detail in waterfowl studies (Cowardin and Blohm 1992). Much like other researchers conducting broad-scale studies in the PPR (e.g., Reynolds et al. 2006, Drever et al. 2007), we observed considerable unexplained variation at the scale of individual survey units. Inferences from our study are therefore probably most appropriate when made at a landscape or regional scale. Taken as a whole, our study nonetheless provided rigorous, quantitative information about broad-scale habitat relationships of broods with clear relevance to current habitat conservation efforts.

**MANAGEMENT IMPLICATIONS**

Habitat conservation efforts for breeding ducks in the PPR are largely focused on using limited budgets to perpetually protect grassland and wetland habitat from conversion to cropland. Our results provided information about habitat relationships of broods that could be used to help guide conservation planning efforts for breeding ducks in the PPR. For example, when combined with information about the distribution and abundance of breeding pairs and hatching probability of nests, our results could provide additional information to conservation planners seeking to understand potential trade-offs between protecting complexes of small wetlands in intensively farmed landscapes and protecting grassland in landscapes with fewer wetlands.

Presence of broods provides evidence of wetland use by broods and evidence of successful nesting. Our study showed that larger wetlands located in landscapes with a high proportion of perennial grassland cover were most likely to be occupied by broods of the study species. Nevertheless, the logarithmic relationship between $\Psi$ and wetland area indicated that the largest increases in $\Psi$ occurred across a range of small- to mid-sized wetlands. For example, estimated occupancy of wetlands by gadwall broods increased by a factor of 3.41 (90% CrI: 2.87, 4.08) when wetland area increased from 0.19 ha to 2.77 ha (a 15-fold increase) and increased by a factor of 2.89 (90% CrI: 2.57, 3.26) when wetland area increased from 2.77 ha to 106.35 ha (a 38-fold increase).

Therefore, we suggest that targeting of protection to complexes of small- to mid-sized wetlands in grass-dominated landscapes is critical because these habitats appear to be particular assets to breeding waterfowl. This recommendation is broadly consistent with current approaches, but our
study lends new empirical support to PPR conservation programs by quantifying complex brood–habitat relationships for multiple species in a single framework across a large spatial extent.

ACKNOWLEDGMENTS

Our study received funding, personnel, and logistical support from several cooperating partners that came together through the Prairie Pothole Joint Venture (PPJV). The PPJV, Ducks Unlimited, Inc., and the United States Fish and Wildlife Service provided funding for data collection and data analysis and in-kind contributions of vehicles and personnel-hours. The North Dakota Game and Fish Department, The South Dakota Department of Game, Fish, and Parks, and The Delta Waterfowl Foundation provided personnel time, vehicles and equipment for data collection. We appreciate the efforts of the numerous seasonal and permanent personnel who collected the field data. A. Royle provided advice on statistical analysis. A. Janus and A. Smith provided assistance with GIS analysis. This manuscript benefited from revisions based on reviews provided by C. Hunter, Associate Editor J. Schmutz, and 2 anonymous peer-reviewers.

LITERATURE CITED


Bloom, P. M. 2010. Mallard duckling survival and habitat selection in the Canadian Prairie Pothole Region. Thesis, University of Saskatchewan, Saskatoon, Saskatchewan, Canada.


Fiske, L., and R. Chandler. 2010. Unmarked 0.8-3: models for data from unmarked animals. R Development Core Team, Vienna, Austria.


Associate Editor: Joel Schmutz.