

# Predictive multi-scale occupancy models at range-wide extents: Effects of habitat and human disturbance on distributions of wetland birds

Bryan S. Stevens<sup>1</sup>  | Courtney J. Conway<sup>2</sup>

<sup>1</sup>Idaho Cooperative Fish and Wildlife Research Unit, Department of Fish and Wildlife Sciences, University of Idaho, Moscow, ID, USA

<sup>2</sup>U.S. Geological Survey, Idaho Cooperative Fish and Wildlife Research Unit, University of Idaho, Moscow, ID, USA

## Correspondence

Bryan S. Stevens, Idaho Cooperative Fish and Wildlife Research Unit, Department of Fish and Wildlife Sciences, University of Idaho, Moscow, ID 83844, USA.  
Email: bstevens@uidaho.edu

## Funding information

US Geological Survey GAP Analysis Program; Nebraska Game and Parks Commission; U.S. Fish and Wildlife Service; US Department of Defense

Editor: Josep Serra-Diaz

## Abstract

**Aim:** Predicting distributions is fundamental to ecology, yet hindered by spatially restricted sampling, scale-dependent relationships and detection error associated with field surveys. Predictive species distribution models (SDMs) are nonetheless vital for conservation of many species. We developed a framework for building predictive SDMs with multi-scale data and used it to develop range-wide breeding-season SDMs for 14 marsh bird species of concern.

**Location:** USA.

**Methods:** We built SDMs using data from range-wide surveys conducted over 14 years, and habitat and disturbance covariates measured at multiple spatial scales. We built hierarchical occupancy models that included heterogeneity in detectability during sampling, and used Bayesian model selection to regulate model complexity (covariates and scales) based explicitly on spatial predictive abilities. We thus integrated model selection for optimizing out-of-sample prediction, range-wide sampling over broad conditions, multi-scale analyses and scale optimization, and species-specific detectability for a suite of wide-ranging species.

**Results:** Distributions of marsh birds were affected by local wetland conditions, but also by agricultural, urban and hydrologic disturbances operating from local scales (100–500 m) to the watershed level. Variables measuring human disturbances improved prediction for most species, and every species was affected by attributes at >1 scale. Five species showed evidence for continental-scale range contraction during the study.

**Main conclusions:** We demonstrate how hierarchical occupancy models can be optimized for prediction across a species' range at the extent of a continent while also accounting for imperfect detection, and thus describe a generalizable approach that can be used for any species. We provide the first data-driven, empirical SDMs built at the range-wide extent for most of our 14 study species and demonstrate that previous studies focused on local distributions and the effects of fine-scale wetland vegetation missed important broadscale drivers of occupancy for marsh birds.

## KEYWORDS

Bayesian model selection, habitat model, hierarchical occupancy model, marsh bird, scale dependence, species distribution model

## 1 | INTRODUCTION

Predicting the spatial distribution of species is a fundamental task in ecology and biogeography (Guisan & Thuiller, 2005; Guisan et al., 2013). Yet, accurate prediction from species distribution models (SDMs) is a difficult problem that can be derailed by challenges common to ecological studies. First, many SDMs are based on spatially restricted sampling that achieves limited coverage of environmental gradients, and thus, data do not represent the range of conditions inhabited by a species, resulting in poor transferability (i.e. models may extrapolate poorly in space and time; Wenger & Olden, 2012; Yates et al., 2018). Second, modelled relationships are often sensitive to the measurement scales of occurrence and environmental covariate data (Wheatley & Johnson, 2009; Wiens, 1989) and tools for identifying optimal scales are still being developed (Chandler & Hepinstall-Cymerman, 2016; McGarigal, Wan, Zeller, Timm, & Cushman, 2016; Miguet, Fahrig, & Lavigne, 2017; Stevens & Conway, 2019). Third, field observations of animals are often corrupted by measurement error due to the failed detection of cryptic organisms (Kéry, 2002; MacKenzie et al., 2002), resulting in biased predictions if heterogeneity in detectability is not accounted for (Gu & Swihart, 2004; Guillera-Arroita et al., 2015; Lahoz-Monfort, Guillera-Arroita, & Wintle, 2014). Despite these challenges, predictive SDMs remain a vital component of many conservation programs (Guisan et al., 2013; Scott et al., 2002) and tools for addressing these problems are thus necessary.

Prediction is often a fundamental objective of SDM applications (Guisan et al., 2013), yet there is no consensus methodology for building and selecting predictive SDMs from multi-scale environmental covariate data (Elith & Leathwick, 2009; McGarigal et al., 2016). Studies of species–habitat relationships are commonly local or regional in nature and assess the effects of environmental attributes at relatively fine scales, and thus, the interplay of fine-scale and broadscale habitat and human disturbance features affecting species distributions is often poorly understood and not accounted for in predictive SDMs (Hobbs, 2003; Jackson & Fahrig, 2015; Wheatley & Johnson, 2009). Multi-scale analyses that seek to identify the best spatial scale for modelling species–environment relationships are desirable and thus commonly advocated (McGarigal et al., 2016; Wiens, 1989). The process of identifying the most appropriate scales for measuring species–environment relationships is made even more challenging, however, by the large number of hypothesized covariates that are commonly available to researchers as a result of remote sensing and geographic information system technologies (Guisan & Thuiller, 2005; Rushton, Ormerod, & Kerby, 2004). Selection of SDM structures (both covariates and their scales) for prediction must therefore balance the fit of models to existing data with the generality needed to predict to unsampled locations (Heikkinen, Marmion,

& Luoto, 2012; Wenger & Olden, 2012; Yates et al., 2018). Modelling approaches that optimize model complexity for the explicit purpose of spatial prediction are becoming more common (e.g. Wenger & Olden, 2012), especially for SDMs that predict the apparent distribution of organisms (i.e. the mathematical product of occupancy and detection probabilities; Kéry, 2011; Lahoz-Monfort et al., 2014). Yet employment of model selection techniques that explicitly optimize SDM complexity for predicting new data while also accounting for failed detection remains rare (but see Broms, Hooten, & Fitzpatrick, 2016; Stevens & Conway, 2019).

Despite their technical challenges, predictive SDMs aid conservation decision-making for many species of management and conservation concern, which often includes wetland and riparian species. Wetlands and riparian areas are among the most biologically productive and diverse ecosystems, yet wetlands are highly degraded and threatened globally (Bedford, Leopold, & Gibbs, 2001; Brinson & Malvárez, 2002; Tockner & Stanford, 2002). Many species of wetland-dependent birds have suffered population declines and range contractions as a consequence of habitat modification and degradation (Conway & Sulzman, 2007; Naugle, Johnson, Estey, & Higgins, 2001; Quesnelle, Fahrig, & Lindsay, 2013). Secretive marsh birds (marsh birds) are a cryptic group of species that are emergent-vegetative wetland obligates, and for which data-driven SDMs (i.e. built using analysis of field data and not simply expert opinion) to predict their distributions over broad extents are lacking. Substantial concern over marsh bird conservation exists within the USA (Conway & Timmermans, 2005; U.S. Fish & Wildlife Service, 2005; U.S. Fish & Wildlife Service, 2008), and multiple species are listed as threatened or endangered at the state, national and international levels (COSEWIC, 2002; Diario Oficial de la Federacion, 2002). As such, predictive SDMs built using field records of marsh bird occupancy are needed to identify existing threats and inform strategic conservation of these birds over broad spatial extents.

Aside from the obvious reliance of marsh birds on emergent wetland vegetation at fine scales, biologists know surprisingly little about the key habitat components necessary for marsh bird persistence, or how their distributions are shaped by anthropogenic modification of landscapes and watersheds. Like many other species, studies of marsh bird habitat relationships have been mostly local or regional in scope and focused on the effects of fine-scale vegetation (i.e. within 100 m of sampling points) on space use (e.g. Budd & Kremenetz, 2010; Conway, Eddleman, Anderson, & Hanebury, 1993; Darrah & Kremenetz, 2008; Darrah & Kremenetz, 2010; Darrah & Kremenetz, 2011; Lor & Malecki, 2006; Winstead & King, 2006). These studies generally lack the contrast of data collected over broadscale disturbance gradients needed to predict distributions across a species range. As an example, studies have often been conducted at wetlands containing the necessary emergent vegetation,

but also within watersheds that were already heavily modified by agriculture, human development, modification of hydrologic regimes or all of these factors (e.g. Alexander & Hepp, 2014; Baschuk, Koper, Wrubleski, & Goldsborough, 2012; Bolenbaugh, Cooper, Brady, Willard, & Kremetz, 2012; Darrah & Kremetz, 2008; Darrah & Kremetz, 2010; Darrah & Kremetz, 2011; Glisson, Brady, Paulios, Jacobi, & Larkin, 2015; Kremetz, Willard, Carroll, & Dugger, 2016; Valente, King, & Wilson, 2011; Winstead & King, 2006). Wetland degradation over broad spatial scales has been hypothesized as a driver of population declines for marsh birds (Eddleman, Knopf, Meanley, Reid, & Zembal, 1988), but little work has quantified these effects. For example, we are unaware of studies that assessed impacts of hydrologic modification on distributions of marsh birds over broad spatial extents (e.g. across a species range), despite the well-established influence of hydrologic modification on the structure and function of wetland ecosystems (Bedford et al., 2001; Tockner & Stanford, 2002) and evidence that water levels can impact fine-scale use of habitat by marsh birds (Nadeau & Conway, 2015; Roach & Barrett, 2015).

The interplay of fine-scale and broadscale habitat and disturbance features that shape breeding distributions of marsh birds is also poorly understood. Studies have attempted to tease apart effects of habitat over a range of spatial scales, but they typically measured different attributes at each scale (e.g. Glisson, Brady, et al., 2015; Harms & Dinsmore, 2013; Naugle et al., 2001; Pickens & King, 2012; Roach & Barrett, 2015; Valente et al., 2011), thus confounding scalar effects with effects of the habitat features themselves (McGarigal et al., 2016; Wheatley & Johnson, 2009). As a consequence, prioritizing wetlands for marsh bird conservation over broad extents via extrapolation of site-specific and fine-scale results assumes that results are transferable to other regions and scale up linearly to broader spatial extents; these assumptions may not be valid and, hence, are unlikely to result in reliable prediction (Miller, Turner, Smithwick, Dent, & Stanley, 2004; Roach, Hunter, Nibblelink, & Barrett, 2017; Schneider, 2001).

Current information gaps mean that biologists are ill-equipped to identify the best remaining habitat for marsh birds across their ranges. Predictive SDMs built using multi-scale covariate data collected over existing environmental and disturbance gradients are therefore needed to predict marsh bird occupancy across large landscapes. Moreover, while marsh bird populations are of significant conservation concern within the USA, the status of many species is uncertain (i.e. range contracting or expanding). To address these needs, our objectives were to (a) develop predictive SDMs for each of 14 priority marsh bird species using data collected from across their breeding ranges within the continental USA, (b) assess the impacts of wetland features and anthropogenic disturbances measured over a variety of spatial extents on marsh bird distributions and (c) assess empirical evidence for contraction of range-wide breeding distributions for each marsh bird species during our 14-year study. To accomplish these objectives, we built optimally predictive multi-scale SDMs for 14 species of marsh birds and selected covariates and their spatial extents to optimize

out-of-sample prediction across the U.S. breeding range of each species. We also accounted for spatial and temporal heterogeneity in detectability during field sampling through use of hierarchical Bayesian occupancy models. Thus, we took a holistic, multispecies and multi-scale approach to building predictive SDMs and identifying threats to marsh birds. Moreover, our approach for developing optimally predictive occupancy models using multi-scale environmental covariate data is generalizable and can be applied to any species.

## 2 | METHODS

### 2.1 | Study area

We included survey data from 8,457 sites throughout the continental USA that were contained within the geographic range of 14 marsh bird species of conservation concern: pied-billed grebe (*Podilymbus podiceps*), American bittern (*Botaurus lentiginosus*), least bittern (*Ixobrychus exilis*), American coot (*Fulica americana*), common gallinule (*Gallinula galeata*), purple gallinule (*Porphyrio martinicus*), limpkin (*Aramus guarauna*), king rail (*Rallus elegans*), clapper rail (*Rallus crepitans*), Ridgway's rail (*Rallus longirostris*), sora (*Porzana carolina*), Virginia rail (*Rallus limicola*), black rail (*Laterallus jamaicensis*) and yellow rail (*Coturnicops noveboracensis*). We used ranges from the Gap Analysis Program (GAP) to identify the geographic range extent for each species (Gergely & McKerrow, 2013) and included survey points (sampling described below) that fell within the range identified by the GAP model for that species. Thus, study areas were species-specific and were effectively the entire U.S. portion of the breeding range for each of the 14 species. We also visually inspected survey sites where a species was detected outside of its GAP range in ArcGIS (ArcMap 10.4.1, ESRI, Redlands, CA), and included these extralimital sites if they were near the range boundary or fell within plausible gaps of the range for a given species (Figures S1.1–S1.14). Our approach for including extralimital sites kept us from excluding plausible breeding-season detections, given the imperfect accuracy of GAP range models. However, the number of extralimital sites relative to sample sizes used to build SDMs for each species was miniscule (<1% of data points for each species), and thus, inferences were primarily determined by surveys at sites within the GAP range for each species.

Our focus was on predicting the distribution of a species within their breeding range as a function of environmental covariates, which is distinctly different from identifying the factors that determine range boundaries (Brown, Stevens, & Kaufman, 1996). Conservation investments for marsh birds, like many other species, are limited. This reality provided the impetus for developing models using samples located within the breeding range for each species, as opposed to using all samples within the continental USA for developing species-specific SDMs (as would be appropriate if studying range boundaries). We focused on predicting occupancy within the breeding range because necessary components of strategic conservation for these birds include identifying important areas within

each species' range and restoration of breeding habitat via spatial targeting of management efforts over broad extents.

## 2.2 | Data collection

Detection–non-detection data from marsh bird surveys were collected by >50 government agencies and non-governmental organizations from 1999 to 2012 at 8,457 sites throughout the USA (see Appendix S1). Field surveys were completed during the breeding season (1 March–15 July) following the North American standardized marsh bird monitoring protocol (Conway, 2011). Most surveys included both a passive and a call-broadcast segment. Broadcasting marsh bird calls through a speaker increases detection probabilities by eliciting vocal responses from birds occupying the area that otherwise are likely to go undetected (Conway & Gibbs, 2005; Conway & Nadeau, 2010). Field surveyors typically visited sites repeatedly over a breeding season. While field sampling was spatially extensive and included a large number of locations for each species, most of these sites were sampled for  $\leq 2$  years during the study (Table S1.1) and the frequency of sampling varied because of the large number of organizations involved in data collection. Not all of the 8,457 sites were located inside the range of each species, and thus, the number of sites included in analyses varied among the 14 species (903–8,457; Table S1.1; Figures S1.1–S1.14). Also, analyses for six species used data collected over a shorter temporal extent (10–13 years) because some years did not include sampling locations within all 14 species' ranges (Table S1.1).

For each visit at a given site within each year, surveyors recorded detection for each species if it was detected during the visit and recorded non-detection if it was not detected during the visit. Replicate visits within each year allowed estimation of occupancy–habitat relationships while also explicitly accounting for failed detections that commonly occur during field surveys (Kéry, 2002; MacKenzie et al., 2006). Because marsh bird detection probabilities are influenced by a number of factors, both within and beyond the control of investigators, we assessed four survey-level attributes recorded during field sampling as hypothesized covariates in species-specific detection models (Conway & Gibbs, 2011; Conway, Sulzman, & Raulston, 2004; Nadeau, Conway, Smith, & Lewis, 2008): (a) time of day at the start of each survey, (b) Julian date, (c) duration of the call-broadcast survey and (d) a binary variable indicating whether or not calls for the species of interest were included in the call-broadcast segment of the survey.

We obtained spatial data to build SDMs for marsh birds as a function of wetland habitat and anthropogenic disturbances measured over a variety of spatial extents (using the focal-patch-landscape design described by Brennan, Bender, Contreras, & Fahrig, 2002; covariate data and spatial analyses described in Appendix S2). We used data from the National Wetland Inventory (NWI) to characterize attributes of wetlands, and data from the National GAP Land Cover Dataset, version 2 (GAP) to describe anthropogenic disturbances surrounding field sites. We considered 21 wetland covariates representing combinations of NWI wetland system and

subsystem (wetland system–subsystem), wetland classes, water regimes and special modifiers as variables to predict the distribution of marsh birds (Tables S2.1–S2.3). We also considered nine GAP land cover variables related to anthropogenic disturbance (Table S2.4): five variables that described human development (developed open-space, low-intensity development, medium-intensity development, high-intensity development and all development land cover types combined [development]), three variables that described modification by agriculture (cultivated-cropland cover, pasture-hay cover and cultivated-cropland and pasture-hay combined [agriculture]) and one variable that described cover of non-native vegetation (non-native cover). Importantly, all of the NWI and GAP variables (Appendix S2) measured attributes of wetlands (e.g. water levels and flooding regimes, vegetation coverage) and anthropogenic disturbance (e.g. human development and agriculture) that are known to directly affect the structure and function of wetland ecosystems, and hence the distribution of organisms within wetlands (Bedford et al., 2001; Tockner & Stanford, 2002; Zelder & Kercher, 2005). Thus, all covariates considered were ecologically plausible factors affecting the distribution of breeding marsh birds, despite the fact that the impacts of many of these variables on distribution of marsh birds over broad extents has received limited study.

Relationships between an animal's use of an area and the attributes of that area are often sensitive to the observational scales at which data are collected (Hobbs, 2003; Levin, 1992; Mayor, Schneider, Schaefer, & Mahoney, 2009; Wiens, 1989). Thus, we evaluated occupancy–covariate relationships at multiple spatial extents surrounding each site. We measured the amount of cover for each wetland and disturbance variable at three spatial extents (100-m, 224-m and 500-m radii buffers) using the focal-patch-landscape design (Brennan et al., 2002) and circular moving-window analyses. Attributes of each site were described at a spatial grain of 30 m and assigned proportional coverage values for each covariate measured over each of the three extents. Multi-scale analyses such as these can result in overlap of spatial-extent buffers, resulting in broad-extent covariate measurements around neighbouring sites that are similar in value (Zuckerberg et al., 2012). This similarity may cause a reduction in covariate variance as the measurement scales increase, which some have suggested may result in analyses favouring small-scale covariate measurements as optimal (Lipse, Naugle, Nowak, & Lukacs, 2017). We did observe small reductions in covariate variance with increased measurement extent (i.e. from 100-m to 500-m buffers; Table S1.2), but this did not result in favouring of fine-scale habitat measurements for predictor variables in our analyses (see Section 2.3 and Section 3 below).

In addition to the local variables, we also measured 10 broad-scale disturbance variables at multiple levels within watersheds surrounding each site (Table S2.5). We used one variable describing the number of current and former mines and mineral processing plants to characterize the effects of mining on water quality, one variable describing the number of pollutant release and transfer reporting facilities to characterize the possibility for contamination, and six hydrologic-modification variables related to the storage capacity

of water by upstream dams, as well as the magnitude of restriction and interruption of natural flow regimes by dams for each site (see Appendix S2). We also measured the two composite GAP variables for development and agriculture (described above) to characterize anthropogenic disturbance within a watershed. Moreover, each of these variables was measured at two spatial levels: catchments and networks. Each catchment contained the drainage area for a single surface water reach. However, catchments receive water from upstream sources, and thus, wetlands within a catchment may be influenced by conditions upstream within a watershed. Sampling points were thus also contained within a broader watershed network that also included upstream wetlands (hereafter the network) to account for upstream dynamics. Disturbance variables measured at the catchment and network levels were described at a spatial grain of 30 m (i.e. similar to local variables but calculated over broader extents).

### 2.3 | Statistical analysis

We used a hierarchical formulation of the multiseason occupancy model (Mackenzie et al., 2006) to predict the probability of occupancy as a function of wetland and disturbance variables. We built models to predict net occupancy probability ( $\psi$ ) as a function of covariates (p. 312, Royle & Dorazio, 2008) and were not interested in modelling local extinction and colonization dynamics explicitly. Hence, we modelled occupancy and detection processes as:

$$z_{i,t} \sim \text{Bernoulli}(\psi_{i,t})$$

$$y_{ij,t} \sim \text{Bernoulli}(z_{i,t}p_{ij,t})$$

This parameterization assumed the true occupancy state ( $z$ ) for each site ( $i$ ) and year ( $t$ ) was a random variable arising from a stochastic Bernoulli process, where  $y_{ij,t}$  represents the detection (1) or non-detection (0) event at visit  $j$ , whose distribution is also Bernoulli but governed by the occupancy state and probability of detection on that survey occasion ( $p_{ij,t}$ ). As such, the temporal grain of occupancy was a single breeding season, whereas the temporal extent was the entire study duration for that particular species (10–14 years). We did not consider temporally dynamic occupancy models where  $z_{i,t+1}$  was dependent in a Markovian fashion to  $z_{i,t}$  because sites were sampled inconsistently across years, and most sites were sampled for only 1–2 years (Table S1.1). Thus, we did not believe our spatially extensive data were adequate for temporally dynamic models of marsh bird occupancy.

This statistical model treats the true occupancy at a site (i.e. present or absent) as closed and unchanging within a single breeding season, but assumes the status can change among seasons. In addition to modelling the relationship between occupancy and habitat, occupancy models explicitly model the process of detection during field surveys. Ignoring failed detection during presence-absence surveys results in mathematical confounding of detection and occupancy probabilities, and the resulting models

predict the joint probability of where a species both occurs and is likely to be found during sampling (i.e. apparent distribution; Guillaera-Arroita, Lahoz-Monfort, MacKenzie, Wintle, & McCarthy, 2014; Kéry, 2011; Lahoz-Monfort et al., 2014). This problem of confounding detection with occupancy is particularly serious for marsh birds because they have notoriously low detection probability (Conway & Gibbs, 2011). Thus, we built hierarchical occupancy models that allow for mathematical separation of occupancy and detection probabilities (Guillaera-Arroita et al., 2015) to predict the true distribution of marsh birds without making two restrictive but commonly made assumptions that are not valid for marsh birds and many other animals: (a) that occupancy status did not change over the study duration, and (b) that detection probability did not change among surveys.

We used a Bayesian statistical paradigm for all model fitting and inferences (model fitting details provided in Appendix S3) and used model selection to optimize the model structure and spatial extent of covariates for prediction. We randomly split sites into training and testing data for each species, where training data were used to fit models and estimate posterior distributions of parameters, and testing data were held out to assess predictive abilities for each model. We used 60% and 40% of sites for the training (542–5074 sites, depending on species) and testing (361–3385 sites, depending on species) data, respectively, where all data at a site were included in the set to which a given site was assigned (Table S1.1). All model ranking was conducted using the logarithmic scoring rule (hereafter log scoring rule; Broms et al., 2016; Gelman, Hwang, & Vehtari, 2014; Hooten & Hobbs, 2015), where log scores for each model were calculated from the posterior predictive density using the testing data set (Stevens & Conway, 2019; Appendix S3). This was equivalent to selecting a model based on its ability to maximize the joint probability of observing the testing data (i.e. log-likelihood evaluated at the testing data for every combination of model parameters contained within the posterior distributions; Gelman et al., 2014; Hooten & Hobbs, 2015), which consisted of samples from across the entire U.S. breeding range of each species.

Selecting models based on fit of their predictions to out-of-sample data is a key component of predictive model selection (Hooten & Hobbs, 2015; Houlahan, McKinney, Anderson, & McGill, 2017), and our approach to model selection optimized spatial prediction across the breeding range of each species. This approach regulated model complexity explicitly using out-of-sample predictive performance, thus naturally alleviating the problem of degraded predictive performance through overfitting of training data that is common to many SDMs (Heikkinen et al., 2012; Wenger & Olden, 2012). As such, this approach to model selection is conceptually different than selecting the appropriate level of complexity based only on within-sample fit to training data (i.e. selection via a model's explanatory ability), or by attempting to balance fit and complexity using information criteria, which can result in overparameterized models with degraded out-of-sample

predictive performance (Broms et al., 2016; Hooten & Hobbs, 2015; Stevens & Conway, 2019). Nonetheless, we also sought to make ecological inferences about effects of covariates identified for their ability to predict marsh bird occupancy, as is common in ecological studies when prediction is the goal (Broms et al., 2016). Such inferences are reasonable given the occupancy modelling framework (that separates detection and occupancy probabilities) and the suite of ecologically relevant covariates that we used. Moreover, prior to including occupancy covariates in models, we selected a detection model for each species based on its log scores from the intercept-only occupancy model (see Appendix S4) and then used the best detection model for examining covariates that affected occupancy.

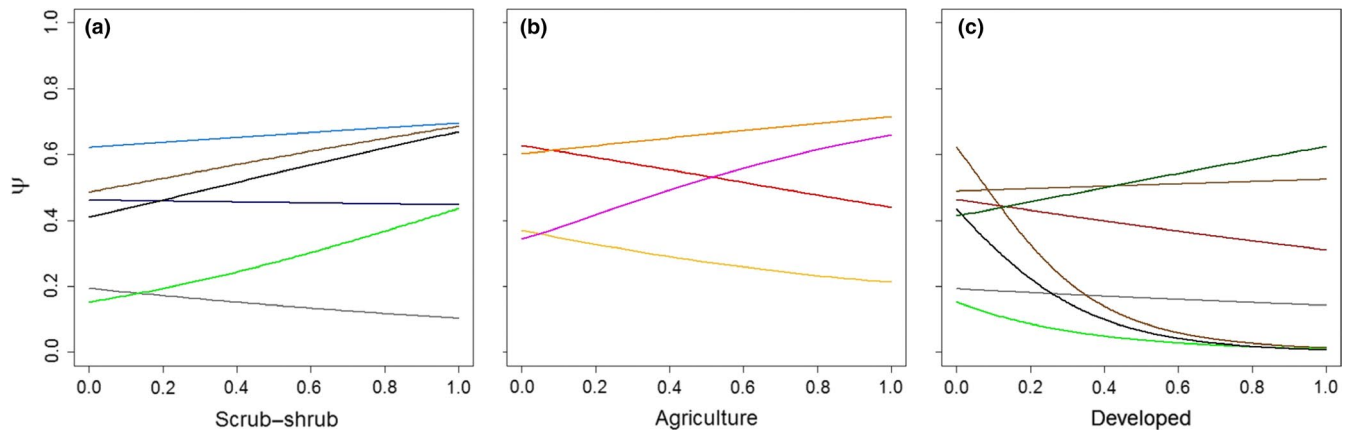
We identified optimal multi-scale SDMs for predicting range-wide occupancy for each marsh bird species. We first conducted scale optimizations for each covariate by fitting univariate occupancy models for each local variable at all three extents (100, 224 and 500 m) and then selecting the optimal extent for each variable (see Appendix S5). Next, we identified the best scale-optimized

habitat covariate within each of the four groups of variables representing wetland types and disturbance (wetland system-subsystem, class, water regime and special modifiers and GAP disturbance variables). This resulted in four scale-optimized covariates that were used to create a local-habitat model set for each species. Specifically, the model set consisted of all additive combinations of each scale-optimized covariate, for a total of 15 local-habitat models, and the best habitat model was selected from this set for each species. In addition, we identified the best broadscale disturbance variable by fitting univariate models for each covariate at both the catchment and network levels and identifying the best variable-level combination for each species. We then created a final model set that included the best local-habitat model from the prior step plus the best catchment or network level disturbance variable, as well as hypothesized temporal trends in occupancy over the study duration that considered linear, quadratic and cubic trends (Appendix S5), and selected the best model for each species. Importantly, habitat and disturbance covariate data sets were not changing over the study duration (i.e.

**TABLE 1** The most supported hierarchical occupancy model predicting range-wide breeding distribution within the continental USA for each of the 14 marsh bird species

| Species           | Top model   | Bayes-P |
|-------------------|---|---------|
| Pied-billed grebe | - riv.lp(224) + sc.shrub(500) + art.flood(500) + dev.hi(500) - ag(c) + trend - trend <sup>2</sup>       | 0.48    |
| American bittern  | - riv.lp(500) - shore(500) + seas.flood(500) + dev.os(100) + r.storfluc(c) + trend - trend <sup>2</sup> | 0.87    |
| Least bittern     | + palus(224) - exc(100) - ag(100) - maxstor(c) + trend - trend <sup>2</sup>                             | 0.82    |
| American coot     | + palus(500) - sc.shrub(500) + semi.flood(224) - dev(500) - dev(n) - trend + trend <sup>2</sup>         | 0.37    |
| Common gallinule  | + palus(100) - art.flood(224) + cult.crop(500) - ag(n) + trend - trend <sup>2</sup>                     | 0.62    |
| Purple gallinule  | - sat(100) + dev(c) - trend - trend <sup>2</sup>  | 0.32    |
| Limpkin           | - l.litt(100) + exc(100) + ag(100) + dev(n)   | 0.50    |
| King rail         | + sc.shrub(224) + per.flood(500) - dev.os(224) - ag(n)  | 0.44    |
| Clapper rail      | + exc(100) - dev.li(500)  | 0.50    |
| Ridgway's rail    | + riv.up(100) + sc.shrub(224) - r.storfluc(n) + trend - trend <sup>2</sup>                              | 0.42    |
| Sora              | - sc.shrub(500) + di(100) - nflowrest(c) + trend - trend <sup>2</sup>                                   | 0.02    |
| Virginia rail     | + l.litt(500) + exc(500) - dev.hi(100) - flowint(n)   | 0.47    |
| Black rail        | + sc.shrub(500) - art.flood(500) - dev.li(100) - trend  | 0.34    |
| Yellow rail       | + shore(224) - di(500) - ag(500) + ag(c)  | 0.41    |

Occupancy model covariates are artificial flooding (art.flood), diked-imponded (di), excavated (exc), lacustrine-littoral (l.litt), palustrine (palus), permanently flooded (per.flood), riverine lower-perennial (riv.lp), riverine upper-perennial (riv.up), saturated (sat), scrub-shrub (sc.shrub), seasonal flooding (seas.flood), semi-permanently flooded (semi.flood), shoreline (shore), agriculture (ag), cultivated croplands (cult.crop), developed (dev), high-intensity development (dev.hi), low-intensity development (dev.li), developed open-space (dev.os), watershed flow interruption (flowint), maximum water storage capacity in watershed (maxstor), normal water flow restriction (nflowrest), relative storage fluctuation (r.storfluc) and time trend across the study duration (trend). Direction of effects is indicated by sign (±), and spatial extents are indicated parenthetically and include 100, 224 and 500 m for local-habitat variables, and catchment (c) and network (n) for watershed-level disturbance variables. Goodness-of-fit of each model to the marsh bird field-sampling data is indicated by the Bayesian *p*-value scores (Bayes-P), where scores close to zero or one indicate lack of fit.



**FIGURE 1** Effects of natural wetland vegetation (a), agriculture (b) and human development (c) on breeding-season occupancy ( $\psi$ ) within the continental USA for marsh birds. Solid lines represent expected occupancy probabilities when all additional covariates were held at their median values, and colours represent individual species. Species affected by natural vegetation (scrub-shrub wetland; a) were pied-billed grebe (light brown), American coot (grey), king rail (light green), Ridgway's rail (light blue), sora (dark blue) and black rail (black). Species affected by agriculture (b) were least bittern (red), common gallinule (orange), limpkin (magenta) and yellow rail (yellow). Species affected by human development (c) were pied-billed grebe (light brown, linear decreasing), American bittern (dark green), American coot (grey), king rail (light green), clapper rail (dark brown, nonlinear decreasing), Virginia rail (dark red) and black rail (black). Specific covariates and scales of effect for each (100, 224 or 500 m) are provided in Table 1

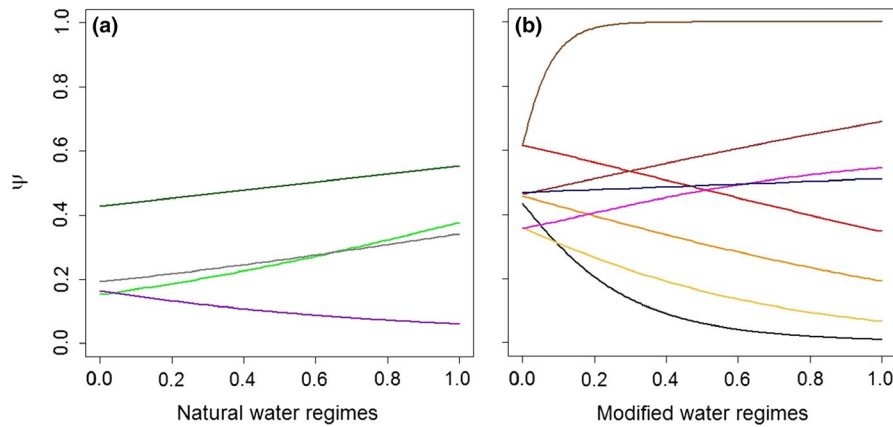
they were temporally static), and thus serve to aide spatial prediction of occupancy rather than temporal assessment of species persistence at a site. This limitation is due to the lack of available time-specific measurements of wetland and disturbance conditions over the study duration at the extent of the continental USA, which is a common limitation of SDMs developed over very broad spatial extents (Elith & Franklin, 2013). In summary, we used a multistage Bayesian model selection procedure to identify which scale-optimized covariates were included in the final model set for each species, and selected the top model from this set as the model that was best at predicting out-of-sample data (see Figure S1.15 for visual depiction of process). Lastly, we refit top models to the entire data set for each species (combining training and testing data; Hooten & Hobbs, 2015) to generate final parameter estimates and to evaluate goodness-of-fit by generating Bayesian  $P$ -values based on the sum of absolute residuals (pgs. 246–249 in Kéry, 2010).

### 3 | RESULTS

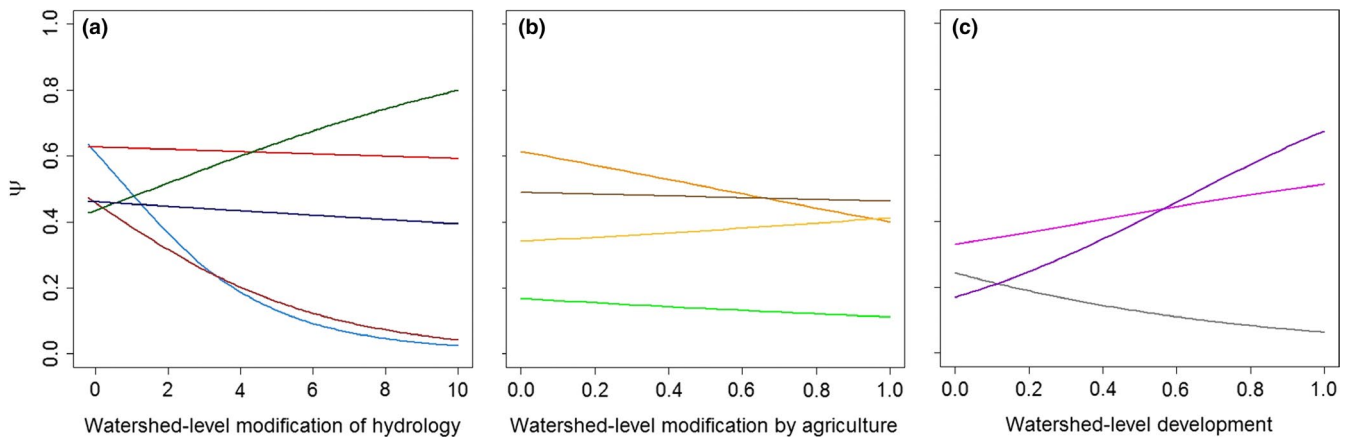
Optimally predictive models included wetland and human disturbance variables measured over all spatial extents (100, 224, 500 m) and levels (catchment, network), and the variables differed among species (Table 1). Local disturbance variables (measured at 100-m, 224-m or 500-m extents) improved predictive performance for all species, and watershed-level disturbances (measured over catchments or networks) further improved performance for 12 species (Tables 1 and S5.1–S5.28). Moreover, top models demonstrated that marsh birds had species-specific detection processes during surveys, where the relationships between detection and the detection

covariates varied among species (Tables S4.1–S4.16). The relationships between detection and both Julian date and time of day were heterogeneous among species and often nonlinear, where quadratic effects for  $\geq 1$  of these were found for 10 species (Table S4.2). In addition, models for eight species included interactions between Julian date and time of day, and thus, joint, nonlinear effects of covariates on detection were common but species-specific. Top models fit the data well for 13 species (all except sora; Table 1).

Factors affecting occupancy varied among species, but some commonalities emerged (Table 1, Figures 1–4). Natural wetland vegetation typically had positive effects on occupancy that were manifested at intermediate extents (224 m or 500 m; Figure 1a). Natural water regimes (Figure 2a) and palustrine wetlands (Table 1) also had positive effects at intermediate extents (224 m or 500 m) for most species. In contrast, human disturbance metrics (development, agriculture, artificial flooding, diking and hydrologic modification) measured over a range of extents often had negative or neutral effects on occupancy (95% credible interval overlapped zero) when present in top models (Tables 1 and S5.1–S5.14; Figures 1–3). Watershed-level hydrologic disturbance variables also typically had negative effects (Figure 3a), and agriculture within a watershed had either negative (common gallinule and king rail) or neutral effects (pied-billed grebe and yellow rail; Figure 3b). However, some exceptions to these patterns did emerge (Figures 1–3). For example, the effect of a natural water regime (saturated wetlands) on purple gallinule was negative, and some water regimes created by human modification (e.g. excavated) had positive and negative effects on occupancy, depending on the species (Figure 2b). Similarly, agriculture had positive effects on occupancy for limpkin and common gallinule within 100 m and 500 m, and development within 100 m had a positive effect on American bittern occupancy.



**FIGURE 2** Effects of natural water regimes (a) and human-modified water regimes (b) on breeding-season occupancy ( $\psi$ ) within the continental USA for marsh birds. Solid lines represent expected occupancy probabilities when all additional covariates were held at their median values, and colours represent individual species. Species affected by natural water regimes (a) were American bittern (dark green), American coot (grey), purple gallinule (purple) and king rail (light green). Species affected by human-modified water regimes (b) were least bittern (red, linear decreasing), common gallinule (orange), limpkin (magenta), clapper rail (dark brown), sora (dark blue), Virginia rail (dark red, linear increasing), black rail (black) and yellow rail (yellow). Specific covariates and scales of effect for each (100, 224 or 500 m) are provided in Table 1



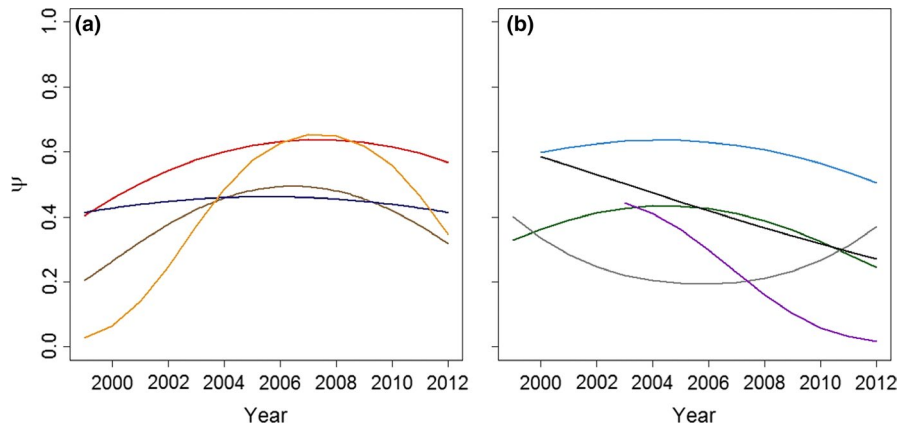
**FIGURE 3** Effects of watershed-level disturbance associated with modification of hydrology (a), agriculture (b) and human development (c) on breeding-season occupancy ( $\psi$ ) within the continental USA for marsh birds. Solid lines represent expected occupancy probabilities given all additional covariates were held at their median values, and colours represent individual species. Species affected by modification of hydrology (a) were American bittern (dark green), least bittern (red, linear decreasing), Ridgway's rail (light blue, nonlinear decreasing), sora (dark blue, linear decreasing) and Virginia rail (dark red, nonlinear decreasing). Species affected by agriculture (b) were pied-billed grebe (light brown), common gallinule (orange), king rail (light green) and yellow rail (yellow). Species affected by human development (c) were American coot (grey), limpkin (magenta) and purple gallinule (purple). Values along the x-axis represent standardized units for each hydrologic-modification variable (a). Specific covariates and level of effects for each (catchment or network) are provided in Table 1

In addition, watershed-level hydrologic modification had positive effects on American bittern occupancy (relative storage fluctuation at the catchment level; Figure 3a). Human development at the watershed level had positive, negative or neutral effects depending on the species (Figure 3c): American coot (-), purple gallinule (+) and limpkin (neutral). Moreover, inclusion of local-habitat, local-disturbance and watershed-level disturbance variables resulted in clear predictive gains relative to null, intercept-only occupancy models (Figure S5.1), despite small gains for some species (limpkin, yellow rail). Inclusion of local-disturbance variables (i.e. GAP variables at 100, 224 or 500 m) and watershed-level disturbance variables also

resulted in predictive gains over wetland habitat variables alone (Figure S5.1), but small predictive gains were observed for some species (pied-billed grebe, sora, limpkin, Ridgway's rail, yellow rail). Disturbance variables added only modest predictive gains for four of these five species (excluding Ridgway's rail) as the disturbance coefficients included in the top model were estimated imprecisely (95% credible intervals overlapped zero; Tables S5.1–S5.14). Lastly, partial effects plots for all covariates in top models for each species are provided in Appendix S5 (Figures S5.2–S5.15).

Distributions of marsh birds were also affected by wetland and disturbance variables at all five spatial extents (Tables 1 and





**FIGURE 4** Time trends for breeding-season occupancy ( $\psi$ ) within the continental USA for each of the nine species of marsh birds whose optimal model contained a time trend. Lines represent expected occupancy probabilities given all additional covariates were held at their median values, and colours indicate species with increased (a) or decreased (b)  $\psi$  over the study duration (1999–2012). Species with increased  $\psi$  were pied-billed grebe (light brown), least bittern (red), common gallinule (orange) and sora (dark blue). Species with decreased  $\psi$  were American bittern (dark green), American coot (grey), purple gallinule (purple), Ridgway's rail (light blue) and black rail (black)

S5.1–S5.56), and all species were affected by features at >1 extent. Occupancy of pied-billed grebe, American bittern, American coot, Virginia rail, black rail and yellow rail was most affected by local-habitat features at the broadest spatial extent (500 m), whereas occupancy of least bittern, limpkin and purple gallinule was most often affected by habitat features at fine extents (100 m). However, most species responded to one or more attributes at all three local scales, and the patterns varied among species (Table 1). Importantly, temporal trends in occupancy were included in top models for nine species, and often included nonlinear trends over the study duration (Table 1). Occupancy probability decreased over the 14-year study duration for American bitterns, American coots, purple gallinules, Ridgway's rails and black rails, whereas occupancy probability increased over the study for common gallinules, least bitterns, pied-billed grebes and soras (Figure 4; see Appendix S5). Many of these species also had quadratic occupancy trends over time, with peaks during the mid-2000s and declines towards the end of the study, whereas purple gallinules showed a nonlinear decreasing occupancy probability over time (i.e. rate of decline became stronger; Figure 4). Temporal trends were included in competitive models for the other five species (Tables S5.15–S5.28), but occupancy changed little over time for these species, except for slight negative trends for king rails and Virginia rails (Figures S5.16–S5.20).

## 4 | DISCUSSION

We built SDMs for 14 species of marsh birds that incorporate wetland features and anthropogenic disturbances while also addressing challenges that hinder predictive performance of many SDMs: (1) regulation of model complexity based on out-of-sample predictive ability, (2) range-wide sampling over a broad range of conditions, (3) multi-scale analyses with spatial scales optimized for prediction and (4) direct modelling of heterogeneity in species-specific detectability.

These four issues have often been addressed in isolation or in pairs (e.g. Broms et al., 2016, integrated #1 and #4 above in a multispecies model, but not #2 or #3), yet we are unaware of another study that integrated all of these elements for a wide-ranging animal at a continental scale, in particular using occupancy models that include detection error. Our process can be used with Bayesian SDMs for any species to optimize model complexity and the ecological neighbourhood size (Addicott et al., 1987) of species–environment relationships for prediction. Moreover, this approach is conceptually and empirically different from the common tactic of selecting models via within-sample data followed by post hoc predictive validation (Hooten & Hobbs, 2015; Houlahan et al., 2017; Stevens & Conway, 2019).

Our models are also among the first SDMs built for marsh birds at the extent of their entire breeding ranges using data collected in the field (as opposed to expert opinion; see also Glisson, Conway, Nadeau, Borgmann, & Laxson, 2015). We provide evidence that marsh bird occupancy is affected by landscape modification, both locally and within watersheds. Not all localities within a geographic range are equally suitable for a given species, and our models can be used to deduce predictions of occupancy probability as a function of habitat and disturbance variables that improves upon binary range models that currently exist to support conservation of these birds over large extents (e.g. GAP models). As such, our models can assist decision-makers in focusing limited resources to inform conservation and monitoring efforts for marsh birds. We also demonstrate the complexity of species-specific detection error that occurs during field sampling, which included nonlinear effects and covariate interactions. Accounting for failed detection and its contributing factors is thus essential to avoid biased predictions when building SDMs for wetland-dependent birds.

We created SDMs to maximize spatial predictive ability across the range of each marsh bird species. Inadequate spatial transferability is a common problem for SDMs that limits a model's predictive ability in new areas where data were not collected; a problem

that occurs when models are generated from sampling a subset of the environmental conditions over which a species is found (Elith & Leathwick, 2009; Elith & Franklin, 2013). When this occurs, prediction involves extrapolation in both geographic and environmental space and commonly results in poor performance despite metrics that may imply good fit of the model to training data (Heikkinen et al., 2012; Sequeira et al., 2016; Yates et al., 2018). We did not test spatial transferability by using regionally distinct training and testing data (Wenger & Olden, 2012), as our goal was not to use data from one subregion to predict occupancy in another. Instead, we selected models based explicitly on their predictive performance using out-of-sample data collected from across each species' range within the USA. The range-wide extent of field sampling meant that these analyses were conducted, by design, to minimize problems of spatial transferability within the breeding range by selecting models to optimize fit of predictions to out-of-sample data collected from across that range. Consequently, resulting models have the best spatial predictive capacity within the constraints of the covariates available over such broad extents, and predictions that are expected to result in interpolation when applied broadly, rather than extrapolation to novel conditions.

In addition, we identified the optimal spatial extents for a suite of species–environment relationships used to predict marsh bird occupancy. Ecological theory and empirical studies demonstrate that animal responses to environmental conditions are sensitive to scales of observation (Wheatley & Johnson, 2009; Wiens, 1989) and that local space use is often affected by conditions operating over broader spatial extents (Addicott et al., 1987; McGarigal et al., 2016). We selected optimally predictive spatial extents for wetland and disturbance covariates, and then multi-scale models that included covariates at their optimal extents. Species responded individually to wetland conditions and disturbances across numerous scales, including conditions within focal patches (100, 224 or 500 m) and entire watersheds. Each species also responded to conditions at more than one scale, and thus, arbitrary selection of a single scale would have resulted in suboptimal prediction. These results highlight the benefits of a multi-scale approach to characterize habitat suitability, as opposed to the common approaches of measuring habitat relationships only at fine spatial scales or confounding changes in measurement extent with changes to the variables measured. More specifically, our results imply that the emphasis on effects of fine-scale vegetation on breeding marsh birds has likely missed other important drivers of occupancy across broad spatial domains.

Our models provide an important step towards prediction, and ultimately conservation, of areas with suitable breeding habitat for marsh birds across the continental USA. Existing predictions of marsh bird habitat are primarily based on expert opinion and collections of fine-scale and regionally focused studies (Bolenbaugh et al., 2012; Darrah & Krementz, 2008; Roach & Barrett, 2015; Roach et al., 2017). Prediction of suitable habitat over broad extents has thus commonly relied on spatial–temporal extrapolation and assumed that fine-scale habitat relationships scale up linearly across a landscape and that region-specific models are

transferable. Not surprisingly, such predictions of suitable habitat have proven difficult to validate with field data (Bolenbaugh et al., 2012; Roach et al., 2017). In contrast, we used detection–non-detection data collected throughout each species' breeding range to build predictive SDMs with their structure explicitly selected to optimize the range-wide spatial predictive capabilities for each species. These models provide a rigorous foundation for predicting habitat suitability using existing spatial covariate data measured at their appropriate extents, avoiding cross-scale extrapolation and providing more reliable predictions than models based solely on wetland attributes. Moreover, the covariates we used for modelling occupancy represent attributes known to affect the structure and function of wetland ecosystems. Thus, the interpretation of ecological effects on breeding-season occupancy for marsh birds is appropriate even though our primary emphasis was spatial prediction.

We provide evidence that anthropogenic disturbance, in a variety of forms and over a variety of extents, affects habitat suitability for marsh birds. Agriculture and urban development are among the biggest threats to wetland ecosystems globally, with direct effects from conversion of wetlands for other uses, but also degrading effects on remaining wetlands via fragmentation, changes to hydrology, increased nutrient loads and changes to water chemistry (Bedford et al., 2001; Tockner & Stanford, 2002; Zelder & Kercher, 2005). These changes have resulted in substantial effects on the distribution of plant and animal species that reside within wetlands. Agriculture may benefit wetland birds in some regions or circumstances (e.g. flooded rice fields in the southeastern USA; Czech & Parsons, 2002; Pierluissi & King, 2008), yet loss of wetlands to agriculture and development have been implicated as a threat to marsh birds (Glisson, Brady, et al., 2015; Naugle et al., 2001). Although our results imply the effects of individual disturbances were species-specific, they also provide direct evidence of anthropogenic impacts on the breeding distributions of marsh birds that have been speculated for 30 years (Eddleman et al., 1988).

Modification of hydrology, both locally and within the larger watersheds where birds reside, affects the breeding distributions of marsh birds. Dams, diking, excavation and other physical modifications alter water flows, water depths and seasonality of fluctuating water levels within a watershed and, in turn, affect soil saturation, water chemistry, water temperatures and sediment dynamics that drive distributions of plants and animals (Bednarek, 2001; Tockner & Stanford, 2002; Zelder & Kercher, 2005). Water depth and its fluctuation can directly affect wetland birds in a variety of ways, including through accessibility to food, nesting habitat and escape cover (Ma, Cai, Li, & Chen, 2010; Nadeau & Conway, 2015), and consequently, water depth is a reported driver of marsh bird habitat use (Flores & Eddleman, 1995; Harms & Dinsmore, 2013; Krementz et al., 2016; Lor & Malecki, 2006). Lack of natural water fluctuations can facilitate wetland succession by eliminating seasonal disturbances (e.g. high flows), which in turn can affect the composition and structure of wetlands to the detriment of marsh

birds (Conway, Nadeau, & Piest, 2010; Winstead & King, 2006). Thus, the extent of effects caused by hydrologic modifications may reach beyond those sampled by many field studies, extending far downstream of the source(s). Our findings are novel because we demonstrate that hydrologic changes can affect marsh birds via processes that operate at a variety of spatial scales, including local effects as well as effects that extend to the level of watersheds. As such, we demonstrate that hydrologic modification has multi-scale, watershed-level and species-specific effects on marsh bird occupancy.

Most of our results also corroborate those of previous studies on the distribution and abundance of marsh birds. For example, we expected natural vegetation and water regimes to be beneficial because vegetation characteristics, water depth and flow affect the behaviour and distribution of breeding marsh birds (Alexander & Hepp, 2014; Baschuk et al., 2012; Conway & Sulzman, 2007; Darrah & Kremetz, 2008, 2011). For example, the interplay of vegetation and water depth can affect the distribution and abundance of prey species (Baschuk et al., 2012), and consequently marsh bird foraging behaviour. Conway and Sulzman (2007) reported that black rails avoided deeper and open water but were commonly located in shallow water ( $\leq 2$  cm) near transitional plant communities between wetlands and uplands. Natural flow regimes can maintain early successional wetlands beneficial to marsh birds, whereas artificial stabilization of flows can change the structure and composition of important vegetation communities (Conway et al., 2010; Winstead & King, 2006). Eddleman et al. (1988) suggested that marsh bird conservation should be prioritized at natural wetlands with a high degree of elevation diversity (i.e. a variety of water depths). Glisson, Brady, et al. (2015) provided direct evidence that marsh bird occupancy in Wisconsin may be greater in natural than restored wetlands, and our results suggest that both natural vegetation and flow regimes are often beneficial for breeding marsh birds.

Many of the effects estimated from our models are congruent with existing knowledge of marsh bird ecology, yet some results were not anticipated and warrant further investigation. For example, we found positive effects of scrub-shrub wetlands measured over a variety of spatial extents on five species, with strong effects on occupancy for three of those species (pied-billed grebe, king rail and black rail). Effects of scrub-shrub wetlands on marsh birds reported by previous studies are highly inconsistent. Many studies have aggregated scrub-shrub and forested wetland into one variable representing woody vegetation, making it impossible to isolate the effects of scrub-shrub wetlands (forest is poor quality habitat for marsh birds; Budd & Kremetz, 2010). Studies have variously reported negative effects (Darrah & Kremetz, 2010), no effects (Budd & Kremetz, 2010) and positive effects (Kremetz et al., 2016) of this composite woody vegetation variable on marsh birds. Moreover, Conway and Sulzman (2007) reported black rail occurrence was positively correlated with mixed shrub vegetation, and Eddleman et al. (1988) reported that multiple rail species utilize shrubby wetlands during the breeding season. At a minimum, our results call into question the practice of aggregating scrub-shrub and forested wetlands

when assessing multi-scale habitat relationships for marsh birds. Our findings also suggest a more refined understanding of the relationships between marsh bird use of wetlands and shrub composition measured over a range of spatial scales is needed.

Perhaps more surprising are the positive relationships we documented between human modification of habitats and occupancy for several marsh bird species. For example, areas with modified water regimes, such as diked and excavated wetlands, were beneficial for four species (clapper rail, limpkin, sora, Virginia rail). Others have suggested these practices may be detrimental to marsh birds, particularly in coastal areas (Ma et al., 2010). We interpret these relationships as providing evidence that man-made and managed wetlands may provide important habitat for these species, and many wetlands managed specifically for wetland birds can be categorized as diked and excavated (e.g. wetlands on National Wildlife Refuge System lands). We also suspect these variables may be acting as proxies for wetland attributes related to the depth and distribution of water during the breeding season in created wetlands, and its subsequent effects on foraging behaviour, nest site selection and habitat use (Eddleman et al., 1988; Legare & Eddleman, 2001; Ma et al., 2010). Less intuitive are the positive fine-scale relationships between occupancy and agriculture (limpkin, common gallinule), and the fine-scale (American bittern) and broadscale (purple gallinule) relationships with human development. Other studies have also reported positive associations between marsh bird occupancy and agriculture (Quesnelle et al., 2013; Valente et al., 2011). For example, gallinules, bitterns, coots and rails have all been observed foraging in agricultural wetlands (e.g. flooded rice fields; Czech & Parsons, 2002) that are common in some parts of the southeastern USA (Pierluissi & King, 2008). Thus, we hypothesize that flooded agricultural wetlands, wet meadows that serve as pasture or both may provide foraging opportunities for some species. We are not aware of other studies documenting positive associations between marsh birds and human development. We suspect this may reflect use of artificially created wetlands at fine scales near human development (American bittern) and covariation of broadscale development with important wetlands whose features were not adequately captured by our covariate set (i.e. important wetlands happen to be juxtaposed near development; purple gallinule). Clearly, however, these hypotheses warrant further investigation.

Many studies have also demonstrated the importance of fine-scale cover of emergent vegetation to marsh birds (Alexander & Hepp, 2014; Budd & Kremetz, 2010; Darrah & Kremetz, 2010; Flores & Eddleman, 1995; Winstead & King, 2006). Marsh birds use emergent vegetation for a variety of purposes, including nesting and brood rearing, escape cover and foraging (Darrah & Kremetz, 2011; Legare & Eddleman, 2001; Lor & Malecki, 2006; Rush, Woodrey, & Cooper, 2010). Presence of emergent vegetation is clearly important for marsh birds, but our results imply that emergent vegetation that is locally available may be a necessary prerequisite but insufficient predictor of occupancy over broad extents and that a suite of other habitat features can greatly improve prediction of marsh bird occupancy (i.e. can help differentiate optimal emergent wetlands from

suboptimal emergent wetlands). Area of emergent wetland vegetation was a suboptimal predictor for all 14 species, and other studies have similarly found the amount of emergent vegetation to be an unsatisfactory predictor of marsh bird occurrence over broad extents (Bolenbaugh, Lehnen, & Kremenz, 2011; Darrah & Kremenz, 2008). In contrast, other wetland types were commonly included in top models and increased occupancy probability (scrub-shrub, palustrine, etc.), which our results suggest is affected by the interplay of wetland attributes and anthropogenic disturbances over a variety of scales.

Presence of emergent vegetation is likely a necessary but insufficient component of marsh bird habitat, where quality is also determined by other attributes and disturbances within the surrounding landscape. Our data support this conclusion, as emergent vegetation was present within 500 m of >90% of our sampling sites, whereas marsh bird occupancy varied among sites. We suspect that if field sampling had occurred in areas lacking emergent wetland vegetation, then the amount of emergent wetland would likely have been a strong predictor of occupancy. Given that field sampling occurred at wetlands containing some emergent vegetation (i.e. areas considered to be breeding habitat for marsh birds; Eddleman et al., 1988) our scope of inference and prediction is thus limited to emergent-vegetated wetlands. This is appropriate given that non-wetland areas and wetlands lacking emergent vegetation are not considered habitat for breeding marsh birds. Thus, presence of emergent vegetation may enable differentiation of wetlands as potential habitat for marsh birds or not, but inclusion of other covariates allows distinction between high- and low-quality breeding habitat.

We also assessed empirical evidence for recent changes to breeding ranges for marsh birds and, to our knowledge, are the first to do so at the extent of the entire range of each species within the continental USA. Concern over population declines and range contractions for marsh birds has been prevalent for at least three decades, yet range-wide status assessments for this group are rare. Our results indicate recent range contraction for five species (American bittern, American coot, black rail, purple gallinule, Ridgway's rail). However, we found little evidence of changes to breeding occupancy for several species, and evidence of increased occupancy for four species (common gallinule, least bittern, pied-billed grebe and sora). Population declines and range contraction likely occurred prior to our study, and temporal changes in occupancy may fail to detect abundance declines if the area occupied remains unchanged. Nonetheless, five species had declining breeding-season occupancy over the 14-year study, and those are species of conservation concern. Importantly, these results provide a baseline for which future studies of the status and distributions of marsh birds can be compared, and thus fill a vital information gap for monitoring, conservation and recovery efforts for these birds.

Lastly, our analyses reduced a large set of plausible predictor variables to a smaller set of variables that were used to construct final model sets for predictive comparison. This approach produced models with few covariates relative to size of the initial covariate set, yet it remains possible that other tactics to reduce the number of covariates (or their individual contributions towards prediction)

would result in different models, possibly with better predictive abilities. Alternative approaches to variable selection that still focus on out-of-sample prediction are also worthy of consideration for future studies. In particular, formal Bayesian variable selection methods can be employed to reduce large sets of predictor variables, or to use all variables simultaneously while optimizing the shrinkage of uninformative parameters for prediction. For example, continuous model selection techniques like the Bayesian Lasso or ridge regression can include all variables simultaneously and optimize the prior distribution of regression coefficients for prediction in order to regulate the magnitude of each coefficient and therefore its contribution towards prediction (Hooten & Hobbs, 2015; O'Hara & Sillanpää, 2009). These techniques have been successfully demonstrated in ecological applications for developing predictive models of the distribution and abundance of organisms (Gerber, Kendall, Hooten, Dubovsky, & Drewien, 2015; Stevens & Conway, 2019).

While our models provide the first data-driven SDMs built at the extent of the breeding range for most of our 14 species, they also provide a foundation for further refinement to improve understanding of multi-scale habitat relationships and predictions of habitat suitability for marsh birds. Our spatial covariate data sets were limited in their temporal resolution and were only available as time-invariant snapshots to aide spatial prediction. These data are updated periodically over time, however, as are the field-sampling databases that store marsh bird detection-non-detection records. Our models provide predictions that can be refined as additional data become available, and their Bayesian implementation provides a rigorous foundation for adaptation over time using Bayesian learning (i.e. using posterior distributions from our models as priors to regulate updated parameter estimates; Hobbs & Hooten, 2015). This approach could also be used for spatial adaptation, effectively enabling regional heterogeneity in relationships (Nice et al., 2019) to be modelled directly by updating our models with localized data. Further refinement could also include spatially hierarchical analyses (e.g. Lipsey et al., 2017) that define occupancy at multiple nested levels (e.g. pixel, catchment, watershed) and model conditional occupancy processes directly. This could facilitate modelling of cross-scale interactions among variables, reduce overlap of buffers used to measure covariate data, and enable study of range boundaries, furthering our understanding of why breeding marsh birds occur in some areas but not in others. Nonetheless, our models provide rigorous predictions of occupancy that were optimized for spatial prediction to take advantage of existing broadscale data on wetland attributes and human disturbances, and consequently will facilitate spatial prioritization of habitat conservation for this sensitive group of wetland birds.

#### ACKNOWLEDGEMENTS

Funding was provided by the Department of Defense Legacy Program, U.S. Fish and Wildlife Service, Nebraska Game and Parks Commission and the GAP Analysis Program of the U.S. Geological Survey. We thank C. Nadeau, K. Borgmann, and W. Glisson for

compiling data sets, M. Rice for discussions about marsh bird conservation, and J.J. Fontaine and K. Decker for assistance securing funding. Any use of trade, firm or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

## DATA AVAILABILITY STATEMENT

Data used for building species distribution models for marsh birds are archived on Dryad (<https://doi.org/10.5061/dryad.2z34tmpgk>).

## ORCID

Bryan S. Stevens  <https://orcid.org/0000-0001-9463-5426>

## REFERENCES

- Addicott, J. F., Aho, J. M., Antolin, M. F., Padilla, D. K., Richardson, J. S., & Soluk, D. A. (1987). Ecological neighborhoods: Scaling environmental patterns. *Oikos*, 49, 340–346. <https://doi.org/10.2307/3565770>
- Alexander, B. W., & Hepp, G. R. (2014). Estimating effects of habitat characteristics on abundances of three species of secretive marsh birds in central Florida. *Waterbirds*, 37, 274–285. <https://doi.org/10.1675/063.037.0306>
- Baschuk, M. S., Koper, N., Wrubleski, D. A., & Goldsborough, G. (2012). Effects of water depth, cover, and food resources on habitat use of marsh birds and waterfowl in boreal wetlands of Manitoba, Canada. *Waterbirds*, 35, 44–55. <https://doi.org/10.1675/063.035.0105>
- Bedford, B. L., Leopold, D. J., & Gibbs, J. P. (2001). Wetland ecosystems. In S. A. Levin (Ed.), *Encyclopedia of biodiversity* (Vol. 7, pp. 384–402). Amsterdam, The Netherlands: Academic Press.
- Bednarek, A. T. (2001). Undamming rivers: A review of the ecological impacts of dam removal. *Environmental Management*, 27, 803–814. <https://doi.org/10.1007/s002670010189>
- Bolenbaugh, J. R., Cooper, T., Brady, R. S., Willard, K. L., & Kremetz, D. G. (2012). Population status and habitat associations of the king rail in the midwestern United States. *Waterbirds*, 35, 535–545. <https://doi.org/10.1675/063.035.0404>
- Bolenbaugh, J. R., Lehnen, S. E., & Kremetz, D. G. (2011). Secretive marsh bird species co-occurrences and habitat associations across the Midwest, USA. *Journal of Fish and Wildlife Management*, 2, 49–60. <https://doi.org/10.3996/012011-JFWM-001>
- Brennan, J. M., Bender, D. J., Contreras, T. A., & Fahrig, L. (2002). Focal patch landscape studies for wildlife management: Optimizing sampling effort across scales. In J. Lui, & W. W. Taylor (Eds.), *Integrating landscape ecology into natural resource management* (pp. 68–91). Cambridge, UK: Cambridge University Press.
- Brinson, M. M., & Malvárez, A. I. (2002). Temperate freshwater wetlands: Types, status, and threats. *Environmental Conservation*, 29, 115–133. <https://doi.org/10.1017/S0376892902000085>
- Broms, K. M., Hooten, M. B., & Fitzpatrick, R. M. (2016). Model selection and assessment for multi-species occupancy models. *Ecology*, 97, 1759–1770. <https://doi.org/10.1890/15-1471.1>
- Brown, J. H., Stevens, G. C., & Kaufman, D. M. (1996). The geographic range: Size, shape, boundaries, and internal structure. *Annual Review of Ecology and Systematics*, 27, 567–623.
- Budd, M. J., & Kremetz, D. G. (2010). Habitat use by least bitterns in the Arkansas delta. *Waterbirds*, 33, 140–147.
- Chandler, R., & Hepinstall-Cymerman, J. (2016). Estimating the spatial scales of landscape effects on abundance. *Landscape Ecology*, 31, 1383–1394.
- Conway, C. J. (2011). Standardized North American marsh bird monitoring protocol. *Waterbirds*, 34, 319–346. <https://doi.org/10.1675/063.034.0307>
- Conway, C. J., Eddleman, W. R., Anderson, S. H., & Hanebury, L. R. (1993). Seasonal changes in Yuma clapper rail vocalization rate and habitat use. *Journal of Wildlife Management*, 57, 282–290. <https://doi.org/10.2307/3809425>
- Conway, C. J., & Gibbs, J. P. (2005). Effectiveness of call-broadcast surveys for monitoring marsh birds. *The Auk*, 122, 26–35. <https://doi.org/10.1093/auk/122.1.26>
- Conway, C. J., & Gibbs, J. P. (2011). Summary of intrinsic and extrinsic factors affecting detection probability of marsh birds. *Wetlands*, 31, 403–411. <https://doi.org/10.1007/s13157-011-0155-x>
- Conway, C. J., & Nadeau, C. P. (2010). Effects of broadcasting conspecific and heterospecific calls on detection of marsh birds in North America. *Wetlands*, 30, 358–368. <https://doi.org/10.1007/s13157-010-0030-1>
- Conway, C. J., Nadeau, C. P., & Piest, L. (2010). Fire helps restore natural disturbance regime to benefit rate and endangered marsh birds endemic to the Colorado River. *Ecological Applications*, 20, 2024–2035.
- Conway, C. J., & Sulzman, C. (2007). Status and habitat use of the California black rail in the southwestern USA. *Wetlands*, 27, 987–998. [https://doi.org/10.1672/0277-5212\(2007\)27\[987:SAHUOT\]2.0.CO;2](https://doi.org/10.1672/0277-5212(2007)27[987:SAHUOT]2.0.CO;2)
- Conway, C. J., Sulzman, C., & Raulston, B. A. (2004). Factors affecting detection probability of California black rails. *Journal of Wildlife Management*, 68, 360–370. [https://doi.org/10.2193/0022-541X\(2004\)068\[0360:FADPOC\]2.0.CO;2](https://doi.org/10.2193/0022-541X(2004)068[0360:FADPOC]2.0.CO;2)
- Conway, C. J., & Timmermans, S. T. (2005). Progress toward developing field protocols for a North American marsh bird monitoring program. In C. J. Ralph & T. D. Rich (Eds.), *Bird Conservation Implementation and Integration in the Americas: Proceedings of the 3rd International Partners in Flight Conference Asilomar, California* (pp. 997–1005). Albany, CA: US Dept. of Agriculture, Forest Service, Pacific Southwest Research Station.
- COSEWIC (2002). *Canadian species at risk*. Ottawa, ON, Canada: Canadian Wildlife Service, Committee on the Status of Endangered Wildlife in Canada.
- Czech, H. A., & Parsons, K. C. (2002). Agricultural wetlands and waterbirds: A review. *Waterbirds*, 25, 56–65.
- Darrah, A. J., & Kremetz, D. G. (2008). Distribution and habitat use of king rails in the Illinois and Upper Mississippi River valleys. *Journal of Wildlife Management*, 73, 1380–1386. <https://doi.org/10.2193/2008-561>
- Darrah, A. J., & Kremetz, D. G. (2010). Occupancy and habitat use of the least bittern and pied-billed grebe in the Illinois and Upper Mississippi River valleys. *Waterbirds*, 33, 367–375. <https://doi.org/10.1675/063.033.0314>
- Darrah, A. J., & Kremetz, D. G. (2011). Habitat use of nesting and brood-rearing king rails in the Illinois and Upper Mississippi River valleys. *Waterbirds*, 34, 160–167. <https://doi.org/10.1675/063.034.0204>
- Diario Oficial de la Federacion. (2002). *Norma Oficial Mexicana NOM-059-ECOL-2001, Proteccion ambiental-Especies nativas de Mexico de flora y fauna silvestres-Categoriosde riesgo y especificaciones para su inclusion, exclusion o cambio-Lista de especies en riesgo*. Secretaria de Medio Ambiente y Recursos Naturales.
- Eddleman, W. R., Knopf, F. L., Meanley, B., Reid, F. A., & Zembal, R. (1988). Conservation of North American rallids. *Wilson Bulletin*, 100, 458–475.
- Elith, J., & Franklin, J. (2013). Species distribution models. In S. A. Levin (Ed.), *Encyclopedia of biodiversity* (pp. 692–705). Waltham, MA: Academic Press.
- Elith, J., & Leathwick, J. R. (2009). Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, 40, 677–697.

- Flores, R. E., & Eddleman, W. R. (1995). California black rail use of habitat in southwestern Arizona. *Journal of Wildlife Management*, 59, 357–363. <https://doi.org/10.2307/3808949>
- Gelman, A., Hwang, J., & Vehtari, A. (2014). Understanding predictive information criteria for Bayesian models. *Statistics and Computing*, 14, 997–1016. <https://doi.org/10.1007/s11222-013-9416-2>
- Gerber, B. D., Kendall, W. L., Hooten, M. B., Dubovsky, J. A., & Drewien, R. C. (2015). Optimal population prediction of sandhill crane recruitment based on climate-mediated habitat limitations. *Journal of Animal Ecology*, 84, 1299–1310. <https://doi.org/10.1111/1365-2656.12370>
- Gergely, K. J., & McKerrow, A. (2013). *Species data - National inventory of range maps and distribution models: U.S. Geological Survey fact sheet 2013-3087*. Boise, ID: U.S. Geological Survey.
- Glisson, W. J., Brady, R. S., Paulios, A. T., Jacobi, S. K., & Larkin, D. J. (2015). Sensitivity of secretive marsh birds to vegetation condition in natural and restored wetlands in Wisconsin. *Journal of Wildlife Management*, 79, 1101–1116. <https://doi.org/10.1002/jwmg.937>
- Glisson, W. J., Conway, C. J., Nadeau, C. P., Borgmann, K. L., & Laxson, T. A. (2015). Range-wide wetland associations of the king rail: A multi-scale approach. *Wetlands*, 35, 577–587. <https://doi.org/10.1007/s13157-015-0648-0>
- Gu, W., & Swihart, R. H. (2004). Absent or undetected? Effects of non-detection of species occurrence on wildlife-habitat models. *Biological Conservation*, 116, 195–203. [https://doi.org/10.1016/S0006-3207\(03\)00190-3](https://doi.org/10.1016/S0006-3207(03)00190-3)
- Guillera-Aroita, G., Lahoz-Monfort, J. J., Elith, J., Gordon, A., Kujala, H., Lentini, P. E., ... Wintle, B. A. (2015). Is my species distribution model fit for purpose? Matching data and models to applications. *Global Ecology and Biogeography*, 24, 276–292. <https://doi.org/10.1111/geb.12268>
- Guillera-Aroita, G., Lahoz-Monfort, J. J., MacKenzie, D. I., Wintle, B. A., & McCarthy, M. A. (2014). Ignoring imperfect detection in biological surveys is dangerous: A response to 'fitting and interpreting occupancy models'. *PLoS ONE*, 9, e99571. <https://doi.org/10.1371/journal.pone.0099571>
- Guisan, A., & Thuiller, W. (2005). Predicting species distribution: Offering more than simple habitat models. *Ecology Letters*, 8, 993–1009. <https://doi.org/10.1111/j.1461-0248.2005.00792.x>
- Guisan, A., Tingley, R., Baumgartner, J. B., Naujokaitis-Lewis, I., Sutcliffe, P. R., Tulloch, A. I. T., ... Buckley, Y. M. (2013). Predicting species distributions for conservation decisions. *Ecology Letters*, 16, 1424–1435. <https://doi.org/10.1111/ele.12189>
- Harms, T. M., & Dinsmore, S. J. (2013). Habitat associations of secretive marsh birds in Iowa. *Wetlands*, 33, 561–571. <https://doi.org/10.1007/s13157-013-0414-0>
- Heikkinen, R. K., Marmion, M., & Luoto, M. (2012). Does the interpolation accuracy of species distribution models come at the expense of transferability? *Ecography*, 35, 276–288. <https://doi.org/10.1111/j.1600-0587.2011.06999.x>
- Hobbs, N. T. (2003). Challenges and opportunities in integrating ecological knowledge across scales. *Forest Ecology and Management*, 181, 223–238. [https://doi.org/10.1016/S0378-1127\(03\)00135-X](https://doi.org/10.1016/S0378-1127(03)00135-X)
- Hobbs, N. T., & Hooten, M. B. (2015). *Bayesian models: A statistical primer for ecologists*. Princeton, NJ: Princeton University Press.
- Hooten, M. B., & Hobbs, N. T. (2015). A guide to Bayesian model selection for ecologists. *Ecological Monographs*, 85, 3–28.
- Houlahan, J. E., McKinney, S. T., Anderson, T. M., & McGill, B. J. (2017). The priority of prediction in ecological understanding. *Oikos*, 126, 1–7. <https://doi.org/10.1111/oik.03726>
- Jackson, H. B., & Fahrig, L. (2015). Are ecologists conducting research at the optimal scale? *Global Ecology and Biogeography*, 24, 52–63. <https://doi.org/10.1111/geb.12233>
- Kéry, M. (2002). Inferring the absence of a species: A case study of snakes. *Journal of Wildlife Management*, 66, 330–338. <https://doi.org/10.2307/3803165>
- Kéry, M. (2010). *Introduction to WinBUGS for ecologists*. Burlington, MA: Academic Press.
- Kéry, M. (2011). Towards the modeling of true species distributions. *Journal of Biogeography*, 38, 617–618.
- Kremetz, D. G., Willard, K. L., Carroll, J. M., & Dugger, K. M. (2016). King rail (*Rallus elegans*) nesting and brood-rearing ecology in a managed wetland in Oklahoma, USA. *Waterbirds*, 39, 241–249.
- Lahoz-Monfort, J. J., Guillera-Aroita, G., & Wintle, B. A. (2014). Imperfect detection impacts the performance of species distribution models. *Global Ecology and Biogeography*, 23, 504–515. <https://doi.org/10.1111/geb.12138>
- Legare, M. L., & Eddleman, W. R. (2001). Home range size, nest-site selection and nesting success of black rails in Florida. *Journal of Field Ornithology*, 72, 170–177. <https://doi.org/10.1648/0273-8570-72.1.170>
- Levin, S. A. (1992). The problem of pattern and scale in ecology. *Ecology*, 73, 1943–1967.
- Lipsey, M. K., Naugle, D. E., Nowak, J., & Lukacs, P. M. (2017). Extending utility of hierarchical models to multi-scale habitat selection. *Diversity and Distributions*, 23, 783–793. <https://doi.org/10.1111/ddi.12567>
- Lor, S., & Malecki, R. A. (2006). Breeding ecology and nesting habitat associations of five marsh bird species in western New York. *Waterbirds*, 29, 427–436. [https://doi.org/10.1675/1524-4695\(2006\)29\[427:BEANHA\]2.0.CO;2](https://doi.org/10.1675/1524-4695(2006)29[427:BEANHA]2.0.CO;2)
- Ma, Z., Cai, Y., Li, B., & Chen, J. (2010). Managing wetland habitats for waterbirds: An international perspective. *Wetlands*, 30, 15–27. <https://doi.org/10.1007/s13157-009-0001-6>
- 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. [https://doi.org/10.1890/0012-9658\(2002\)083\[2248:ESORW\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2002)083[2248:ESORW]2.0.CO;2)
- MacKenzie, D. I., Nichols, J. D., Royle, J. A., Pollock, K. H., Bailey, L. L., & Hines, J. E. (2006). *Occupancy estimation and modeling: Inferring patterns and dynamics of species occurrence*. Burlington, MA: Academic Press.
- Mayor, S. J., Schneider, D. C., Schaefer, J. A., & Mahoney, S. P. (2009). Habitat selection at multiple scales. *Ecoscience*, 16, 238–247. <https://doi.org/10.2980/16-2-3238>
- McGarigal, K., Wan, H. Y., Zeller, K. A., Timm, B. C., & Cushman, S. A. (2016). Multi-scale habitat selection modeling: A review and outlook. *Landscape Ecology*, 31, 1161–1175. <https://doi.org/10.1007/s10980-016-0374-x>
- Miguet, P., Fahrig, L., & Lavigne, C. (2017). How to quantify a distance-dependent landscape effect on a biological response. *Methods in Ecology and Evolution*, 8, 1717–1724. <https://doi.org/10.1111/2041-210X.12830>
- Miller, J. R., Turner, M. G., Smithwick, E. A. H., Dent, C. L., & Stanley, E. H. (2004). Spatial extrapolation: The science of predicting ecological patterns and processes. *BioScience*, 54, 310–320. [https://doi.org/10.1641/0006-3568\(2004\)054\[0310:SETSOP\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2004)054[0310:SETSOP]2.0.CO;2)
- Nadeau, C. P., & Conway, C. J. (2015). Optimizing water depth for wetland-dependent wildlife could increase wetland restoration success, water efficiency, and water security. *Restoration Ecology*, 23, 292–300. <https://doi.org/10.1111/rec.12180>
- Nadeau, C. P., Conway, C. J., Smith, B. S., & Lewis, T. E. (2008). Maximizing detection probability of wetland-dependent birds during point-count surveys in northwestern Florida. *The Wilson Journal of Ornithology*, 120, 513–518. <https://doi.org/10.1676/07-041.1>
- Naugle, D. E., Johnson, R. R., Estey, M. E., & Higgins, K. F. (2001). A landscape approach to conserving wetland bird habitat in the prairie pothole region of eastern South Dakota. *Wetlands*, 21, 1–17. [https://doi.org/10.1672/0277-5212\(2001\)021\[0001:ALATCW\]2.0.CO;2](https://doi.org/10.1672/0277-5212(2001)021[0001:ALATCW]2.0.CO;2)
- Nice, C. C., Forister, M. L., Harrison, J. G., Gompert, Z., Fordyce, J. A., Thorne, J. H., ... Shapiro, A. M. (2019). Extreme heterogeneity of

- population responses to climatic variation and the limits of prediction. *Global Change Biology*, 25, 2127–2136.
- O'Hara, R. B., & Sillanpää, M. J. (2009). A review of Bayesian variable selection methods: What, how, and which. *Bayesian Analysis*, 4, 85–118.
- Pickens, B. A., & King, S. L. (2012). Predicting the spatial distribution of king rails in an agricultural landscape. *The Condor*, 114, 113–122.
- Pierluissi, S., & King, S. L. (2008). Relative nest density, nest success, and site occupancy of king rails in southwestern Louisiana rice fields. *Waterbirds*, 31, 530–540.
- Quesnelle, P. E., Fahrig, L., & Lindsay, K. E. (2013). Effects of habitat loss, habitat configuration and matrix composition on declining wetland species. *Biological Conservation*, 160, 200–208. <https://doi.org/10.1016/j.biocon.2013.01.020>
- Roach, N. S., & Barrett, K. (2015). Managed habitats increase occupancy of black rails (*Laterallus jamaicensis*) and may buffer impacts from sea level rise. *Wetlands*, 35, 1065–1076. <https://doi.org/10.1007/s13157-015-0695-6>
- Roach, N. S., Hunter, E. A., Nibblelink, N. P., & Barrett, K. (2017). Poor transferability of a distribution model for a widespread coastal marsh bird in the southeastern United States. *Ecosphere*, 8, e017175. <https://doi.org/10.1002/ecs2.1715>
- Royle, J. A., & Dorazio, R. M. (2008). *Hierarchical modeling and inference in ecology: The analysis of data from populations, metapopulations, and communities*. London, UK: Academic Press.
- Rush, S. A., Woodrey, M. S., & Cooper, R. J. (2010). Variation in the nesting habits of clapper rails in tidal marshes of the northern Gulf of Mexico. *The Condor*, 112, 356–362. <https://doi.org/10.1525/cond.2010.090078>
- Rushton, S. P., Ormerod, S. J., & Kerby, G. (2004). New paradigms for modelling species distributions? *Journal of Applied Ecology*, 41, 193–200.
- Schneider, D. C. (2001). The rise of the concept of scale in ecology. *BioScience*, 51, 545–553.
- Scott, J. M., Heglund, P. J., Morrison, M. L., Haufler, J. B., Raphael, M. G., Wall, W. A., & Sampson, F. B. (2002). *Predicting species occurrences: Issues of accuracy and scale*. Washington, DC: Island Press.
- Sequeira, A. M. M., Mellin, C., Lozano-Montes, H. M., Vanderkluft, M. A., Babcock, R. C., Haywood, M. D. E., ... Caley, M. J. (2016). Transferability of predictive models of coral reef fish species richness. *Journal of Applied Ecology*, 53, 64–72. <https://doi.org/10.1111/1365-2664.12578>
- Stevens, B. S., & Conway, C. J. (2019). Predicting species distributions: Unifying model selection and scale optimization for multi-scale occupancy models. *Ecosphere*, 10, e02748. <https://doi.org/10.1002/ecs2.2748>
- Tockner, K., & Stanford, J. A. (2002). Riverine flood plains: Present state and future trends. *Environmental Conservation*, 29, 308–330.
- U.S. Fish and Wildlife Service (2005). *The U.S. fish and wildlife service's focal species strategy for migratory birds*. Arlington, VA: USDI Fish and Wildlife Service, Division of Migratory Bird Management.
- U.S. Fish and Wildlife Service (2008). *Birds of conservation concern 2008*. Arlington, VA: USDI Fish and Wildlife Service, Division of Migratory Bird Management.
- Valente, J. J., King, S. L., & Wilson, R. R. (2011). Distribution and habitat associations of breeding secretive marsh birds in Louisiana's Mississippi alluvial valley. *Wetlands*, 31, 1–10. <https://doi.org/10.1007/s13157-010-0138-3>
- Wenger, S., & Olden, J. D. (2012). Assessing transferability of ecological models: An underappreciated aspect of statistical validation. *Methods in Ecology and Evolution*, 3, 260–267. <https://doi.org/10.1111/j.2041-210X.2011.00170.x>
- Wheatley, M., & Johnson, C. (2009). Factors limiting our understanding of ecological scale. *Ecological Complexity*, 6, 150–159. <https://doi.org/10.1016/j.ecocom.2008.10.011>
- Wiens, J. (1989). Spatial scaling in ecology. *Functional Ecology*, 3, 385–397. <https://doi.org/10.2307/2389612>
- Winstead, N. A., & King, S. L. (2006). Least bittern distribution among structurally different vegetation types in managed wetlands of northwest Tennessee, USA. *Wetlands*, 26, 619–623.
- Yates, K. L., Bouchet, P. J., Caley, M. J., Mengersen, K., Randin, C. F., Parnell, S., ... Sequeira, A. M. M. (2018). Outstanding challenges in the transferability of ecological models. *Trends in Ecology & Evolution*, 33, 790–802.
- Zelder, J. B., & Kercher, S. (2005). Wetland resources: Status, trends, ecosystem services, and restorability. *Annual Review of Environment and Resources*, 30, 39–74.
- Zuckerberg, B., Desrochers, A., Hochachka, W. M., Fink, D., Koenig, W. D., & Dickinson, J. L. (2012). Overlapping landscapes: A persistent, but misguided concern when collecting and analyzing ecological data. *Journal of Wildlife Management*, 76, 1072–1080.

## BIOSKETCHES

**Bryan S. Stevens** is a research scientist with the Idaho Cooperative Fish and Wildlife Research Unit and Department of Fish and Wildlife Sciences at University of Idaho. His research focuses on the intersection of quantitative ecology and fish and wildlife conservation, and on developing decision-support tools for applied conservation and management.

**Courtney J. Conway** is Leader of the Idaho Cooperative Fish and Wildlife Research Unit and Professor in the Department of Fish and Wildlife Sciences at University of Idaho. Courtney's laboratory group pursues questions related to behavioural ecology and management of bird and mammal populations.

Author contributions: C.J.C. conceived the study, and B.S.S. and C.J.C. planned and designed the analyses used to make inferences. C.J.C. coordinated the organization of field data. B.S.S. conducted the statistical analyses and wrote the initial manuscript draft, where C.J.C. provided additional contributions and feedback that resulted in the finished manuscript.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**How to cite this article:** Stevens BS, Conway CJ. Predictive multi-scale occupancy models at range-wide extents: Effects of habitat and human disturbance on distributions of wetland birds. *Divers Distrib*. 2020;26:34–48. <https://doi.org/10.1111/ddi.12995>

© 2020. This work is published under <http://creativecommons.org/licenses/by/4.0/>(the “License”). Notwithstanding the ProQuest Terms and Conditions, you may use this content in accordance with the terms of the License.