

Shiny Application to Iowa DNR MSIM-SGCN Modeling

Xiaodan (Annie) Lyu, Tyler M. Harms

September 23, 2016

Introduction

- ▶ Period: Oct 2015 - May 2016
- ▶ State Wildlife Grants - Iowa Department of Natural Resources
- ▶ Run predictive models using R package RMark
- ▶ Produce predictive maps using ArcGIS
- ▶ Develop an interactive web application using Shiny

Outline

- ▶ Project overview: data source and objectives
- ▶ Methods: model fit and covariates
- ▶ Model Validation: AUC
- ▶ Results: example of estimates table and predictive maps
- ▶ Demonstrate Shiny application

Data Collection

- ▶ Multiple Species Inventory and Monitoring (MSIM) Program
 - ▶ Iowa DNR and Iowa State University
- ▶ Sampling Interval
 - ▶ Primary: 2006-2014
 - ▶ Secondary: survey occasions (days)
- ▶ Survey Objects
 - ▶ 414 SGCN, only 69 SGCN where sufficient data were available
 - ▶ Birds, mammals, reptiles, amphibians, odonates and butterflies

Objectives

- ▶ Predict the distribution of species of conservation need
 - ▶ robust design occupancy model
- ▶ Create predictive species maps for priority SGCN
 - ▶ ArcGIS raster files
 - ▶ Shiny interactive web application
- ▶ Prioritize areas of conservation action for SGCN
 - ▶ habitat restoration and management

Methods

- ▶ Robust design occupancy model (MacKenzie et al. 2003)
 - ▶ package RMark
- ▶ Parameters of Interest
 - ▶ probability of occupancy (ψ)
 - ▶ probability of colonization (γ)
 - ▶ probability of extinction (ϵ)
 - ▶ probability of detection (p)
- ▶ Model types
 - ▶ RDOccupEG, RDOccupPE, RDOccupPG, ...

Statistical Model

$$Pr(\mathbf{X}_i) = \phi_0 \left\{ \prod_{t=1}^{T-1} D(\mathbf{p}_{X,t}) \phi_t \right\} \mathbf{p}_{X,T}$$

- ▶ $\mathbf{p}_{X,t}$: vector denoting probability of observing the detection history $X_{i,t}$ in primary period t
- ▶ ϕ_t : matrix of transition probabilities between states of occupancy from primary period t to $t + 1$

$$\phi_t = \begin{bmatrix} 1 - \epsilon_t & \epsilon_t \\ \gamma_t & 1 - \gamma_t \end{bmatrix}, t = 1, \dots, T-1$$

$$\phi_0 = \begin{bmatrix} \psi_1 & 1 - \psi_1 \end{bmatrix}$$

Example

$$Pr(\mathbf{X}_{i,1} = 010) = \psi_1(1 - p_{1,1})p_{1,2}(1 - p_{1,3})$$

$$Pr(\mathbf{X}_{i,2} = 000 | \mathbf{X}_{i,1}) = (1 - \epsilon_1) \prod_{j=1}^3 (1 - p_{2,j}) + \epsilon_1$$

$$\begin{aligned} Pr(\mathbf{X}_i = 010 \ 000) &= \phi_0 D(\mathbf{p}_{010,1}) \phi_1 \mathbf{p}_{000,2} \\ &= \begin{bmatrix} \psi_1 & 1 - \psi_1 \end{bmatrix} \begin{bmatrix} (1 - p_{1,1})p_{1,2}(1 - p_{1,3}) & 0 \\ 0 & 0 \end{bmatrix} \\ &\quad \times \begin{bmatrix} 1 - \epsilon_1 & \epsilon_1 \\ \gamma_1 & 1 - \gamma_1 \end{bmatrix} \begin{bmatrix} \prod_{j=1}^3 (1 - p_{2,j}) \\ 1 \end{bmatrix} \\ &= \psi_1(1 - p_{1,1})p_{1,2}(1 - p_{1,3}) \left[(1 - \epsilon_1) \prod_{j=1}^3 (1 - p_{2,j}) + \epsilon_1 \right] \end{aligned}$$

Statistical Model (cont'd)

- ▶ Including Covariates

$$\theta = \frac{\exp(\mathbf{Z}'\beta)}{1 + \exp(\mathbf{Z}'\beta)}$$

- ▶ θ is the probability of interest
- ▶ \mathbf{Z} is the matrix of covariate information
- ▶ β is the vector of logistic model coefficients to be estimated

Covariates

- ▶ Landscape habitat variables ¹
 - ▶ Radius of sampled site: 200m, 500m and 1km
 - ▶ Land use classification:
Water, Wetland, Grassland, Woodland and Agriculture
 - ▶ Landscape configuration:
percentage of landscape (PLAND), large patch index (LPI),
edge density (ED), patch density (PD) and
interspersed-juxtaposition (IJI)
- ▶ Climate variables ²
 - ▶ Wind speed (km/h), Cloud cover (%), Temperature (°C)

¹site-specific, modeled on ψ, γ and ϵ

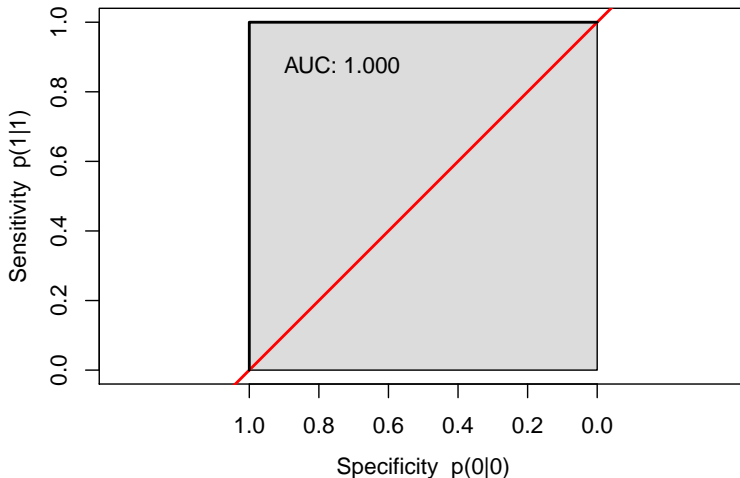
²time-varying, modeled on p

Model Selection and Validation

- ▶ Akaike's Information Criterion adjusted for small sizes (AIC_c)
 - ▶ Models with $\Delta AIC_c \leq 2 \Rightarrow$ strong support (Burnham and Anderson 2002)
- ▶ Area under the receiver operating characteristic curve (AUC)
 - ▶ evaluate performance of predicting occupancy
 - ▶ training data set: survey years 2006-2012, 2014
 - ▶ test data set: survey year 2013 (better representative)
 - ▶ package pROC
 - ▶ $AUC = 0.5 \Rightarrow$ random guess
 - ▶ $AUC = 1.0 \Rightarrow$ perfect prediction (Jimenez-Valverde 2012)

ROC Curve

ROC Curve: Random Guess VS Perfect Prediction



Results

Table 1: Results Summary

Statistics	Value/Range
Number of SGCN modeled	64 out of 69
Occupancy Prob $\hat{\psi}$	0.001(0.0003)~0.995(0.004)
Colonization Prob $\hat{\gamma}$	0.0003(0.0001)~0.999(0.00007)
Detection Prob \hat{p}	0.030(0.028)~0.998(0.006)
AUC values	0.426~0.921
Number of AUC > 0.5	61 out of 64

Table 2: Best Models for Each Species

Species	Model
Red-shouldered Hawk	$\psi(\text{Wod1KPLND})\gamma(\text{Ag1KPD})p(\text{Cld})$
Yellow-billed Cuckoo	$\psi(\text{Wod500PLND})\gamma(\text{Wod1kLPI})p(\text{Wind})$
Red-headed Woodpecker	$\psi(\text{Wod500PLND})\gamma(\text{Ag500PD})p(\text{Cld})$
Eastern Wood-pewee	$\psi(\text{Wod500PLND})\gamma(\text{Wod1KPLND})p(\text{Wind})$
Acadian Flycatcher	$\psi(\text{Wod500PLND})\gamma(\text{Wod500PLND})p(\text{Wind})$
Veery	$\psi(\text{Wod1kED})\gamma(\text{Ag500LPI})p(\text{Wind})$

Table 3: AUC and Coefficients Estimates under Best Model

Species	AUC	Psi.Cov	Gam.Cov	p.Cov
Red-shouldered Hawk	0.798	0.034	0.002	-0.008
Yellow-billed Cuckoo	0.706	0.060	0.042	-0.076
Red-headed Woodpecker	0.610	0.017	-0.002	-0.004
Eastern Wood-pewee	0.906	0.128	0.071	-0.110
Acadian Flycatcher	0.892	0.092	0.070	0.086
Veery	0.551	0.005	-0.578	0.179

Yellow-billed Cuckoo Colonization

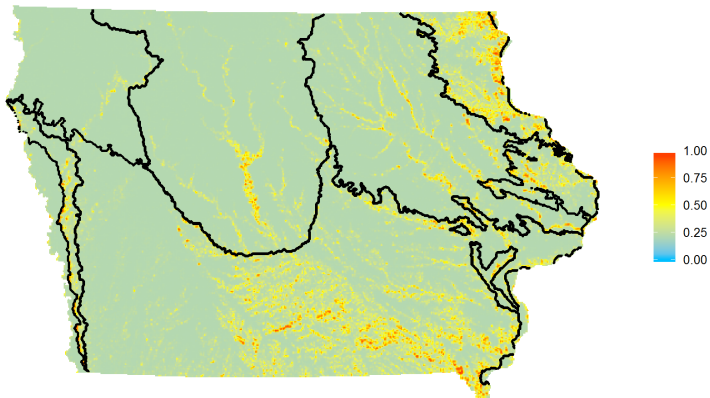


Figure 1: Predictive Map

Red-shouldered Hawk Occupancy

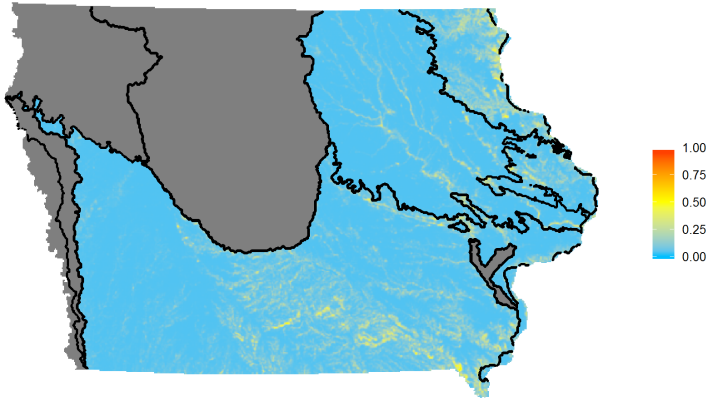


Figure 2: Predictive Map of Range-Restricted Species

Shiny

- ▶ **Shiny** by RStudio is a web application framework for R.
- ▶ No HTML, CSS or JavaScript knowledge required to turn your analyses into interactive web applications.

```
install.packages("shiny")  
library("shiny")  
runExample("01_hello")
```

- ▶ ui.R defines the page layout and user interface
- ▶ server.R contains the R code to create any output
- ▶ More information available at [Shiny Webpage](#)

About Our Shiny Application

- ▶ Interactive web application to display predictive maps and parameter estimates for each SGCN
- ▶ Easy to download personalized data and maps
- ▶ Available to researchers and managers across Iowa (credentials needed)
- ▶ Hosted by CSSM for 1-2 years
- ▶ The URL for accessing the application is <https://dnrswg.cssm.iastate.edu/>