

# Species Distribution Modelling Using R

(Version 2018.1)

---

Thomas C. Edwards

Research Ecologist and Professor

U.S. Geological Survey, Utah Cooperative Fish and  
Wildlife Research Unit, and

Department of Wildland Resources

Utah State University, Logan, UT 84322-5230

[t.edwards@usu.edu](mailto:t.edwards@usu.edu) (EMAIL)

A workshop on use of R for analyzing biological information

Hosted in collaboration with FWS National Conservation Training  
Center, the USGS Cooperative Research Units Program, and Utah  
State University



S.J. & JESSIE E. QUINNEY  
COLLEGE OF  
NATURAL RESOURCES  
UtahStateUniversity



# Species Distribution Modelling Using R

## Workshop Outline (Version 2018.1)

---

### Introduction to Species Distribution Modelling

---

- |     |     |   |
|-----|-----|---|
| Mod | 1.1 | Some History about Species Distribution Models (SDMs) |
|     | 1.2 | SDMs Compared to Other Modelling Forms                |
|     | 1.3 | SDMs: Question and Design Issues                      |
|     | 1.4 | Attributes of Data in SDMs                            |
|     | 1.5 | Why Do SDMs all Differ?                               |

### The Analytical Pathway for Species Distribution Models

---

- |     |     |   |
|-----|-----|---|
| Mod | 2.1 | Similarities Between Regression and Classification      |
|     | 2.2 | Selecting the Appropriate Classification Model to Apply |
|     | 2.3 | Purposive vs. Probabilistic Sample Frames               |
|     | 2.4 | Issues of Data Resolutions                              |
|     | 2.5 | Bounding Boxes for Presence-Only Data                   |
|     | 2.6 | Building the Training Data Set                          |
|     | 2.7 | Statistical Exploration of the Training Data Set        |

### Model Assessment and Uncertainty

---

- |     |     |   |
|-----|-----|---|
| Mod | 3.1 | The Validation Theaters of SDMs                           |
|     | 3.2 | A Validation Framework for Application to SDMs            |
|     | 3.3 | The Validation Process – Statistical Approaches           |
|     | 3.4 | The Validation Process – Less-Than-Statistical Approaches |

### Building Prediction Map Products from SDMs

---

- |     |     |  |
|-----|-----|--|
| Mod | 4.1 | Some historical (just 'cause it's fun) and current-day ecological uses of maps |
|-----|-----|--|



# Species Distribution Modelling Using R

## Workshop Outline (Version 2018.1)

---

### Logistic Regression GLM

---

- |     |     |  |
|-----|-----|--|
| Mod | 5.1 | When to Use the Logistic GLM Model   |
|     | 5.2 | The Statistical Model for the Logistic GLM   |
|     | 5.3 | Basic Assumptions of the Logistic GLM  |
|     | 5.4 | Related Techniques   |
|     | 5.5 | Example Analysis Using R:<br>Predicting the probability of bird nest occurrence as a function of habitat variables |

### Additive Logistic Regression GAM

---

- |     |     |   |
|-----|-----|---|
| Mod | 6.1 | When to Use the Additive Logistic GAM Model   |
|     | 6.2 | The Statistical Model for the Additive Logistic GAM   |
|     | 6.3 | Basic Assumptions of the Additive Logistic GAM  |
|     | 6.4 | Related Techniques  |
|     | 6.5 | Example Analysis Using R:<br>Discriminating lodgepole pine from a background of coniferous forest: the interplay of remote-sensing and environmental predictors |

### Maximum Entropy MAXENT

---

- |     |     |  |
|-----|-----|--|
| Mod | 7.1 | Characteristics of MAXENT analysis   |
|     | 7.2 | The Statistical Model for Maximum Entropy MAXENT   |
|     | 7.3 | Basic Assumptions of Maximum Entropy MAXENT  |
|     | 7.4 | Related Techniques   |
|     | 7.5 | Example Analysis Using R:<br>Modelling critical habitat for a plant species proposed for listing under the U.S. Endangered Species Act: Graham's Beardtongue <i>Penstemon grahamii</i> |



# Species Distribution Modelling Using R

## Workshop Outline (Version 2018.1)

---

### Random Forests RF

---

- Mod 8.1 Characteristics Random Forests RF (and Classification Trees)
- 8.2 The Statistical Model for Random Forests RF
- 8.3 Basic assumptions of the random forest RF
- 8.4 Related Techniques
- 8.5 Example Analysis Using R:  
Predicting epiphytic macrolichen presence across broad geographic ranges: how good are extrapolative models?

### Boosted Regression Trees BRT

---

- Mod 9.1 Characteristics of Boosted Regression Trees BRT
- 9.2 The Statistical Model for Boosted Regression Trees BRT
- 9.3 Basic assumptions of the boosted regression trees BRT
- 9.4 Related techniques
- 9.5 Example Analysis Using R:  
Setting the stage for climate-change projections: using SDMs to model the current distribution of the Utah juniper *Juniperus osteosperma*

### Ensemble Species Distribution Models

---

- Mod 10.1 What Constitutes an Ensemble Model
- 10.2 The Two Basic Classes of Ensemble Models
- 10.3 The Ensemble Process
- 10.4 Example Analysis Using R:  
An ensemble model predicting the spatial distribution of critical habitat for Graham's Beardtongue *Penstemon grahamii*



# Species Distribution Modelling Using R

## Workshop Outline (Version 2018.1)

---

### The Dilemma of Model Comparisons

---

- |     |      |  |
|-----|------|--|
| Mod | 11.1 | A comparison of results from the class exercises                   |
|     | 11.2 | Some (random and unstructured) thoughts on SDMs as forecast models |

### Appendices

---

- |            |                                    |
|------------|------------------------------------|
| Appendix 1 | Data Sources                       |
| Appendix 2 | Basic R Functions used in Workshop |

