

Mapping species distributions

Spatial inference and prediction

JANET FRANKLIN

*School of Geographical Sciences and School of Life Sciences, Arizona
State University*

*Formerly at the Departments of Biology and Geography, San Diego
State University*

With contributions by

JENNIFER A. MILLER

*Department of Geography and the Environment, University of Texas,
Austin, Texas, USA*



CAMBRIDGE
UNIVERSITY PRESS

Contents

<i>Preface</i>	page xv
<i>Acknowledgments</i>	xvii

Part I: History and ecological basis of species distribution modeling

1	Species distribution modeling	3
1.1	Introduction	3
1.2	What is in a name?	5
1.2.1	Niche models	5
1.2.2	Habitat suitability models	6
1.3	Heightened interest in species distribution modeling	7
1.4	What is species distribution modeling and how is this book organized?	9
1.5	Why model species distributions?	11
1.5.1	Reserve design and conservation planning	12
1.5.2	Impact assessment and resource management	14
1.5.3	Ecological restoration and ecological modeling	14
1.5.4	Risk and impacts of invasive species including pathogens	15
1.5.5	Effects of global warming on biodiversity and ecosystems	17
2	Why do we need species distribution models?	21
2.1	Introduction	21
2.2	Mapping species – atlas projects and natural history collections	22

2.2.1	Grid-based atlases of species distributions	22
2.2.2	Species locations from natural history collections	30
2.3	Direct interpolation of species data	30
2.4	Summary – what do we really want?	32
3	Ecological understanding of species distributions	34
3.1	Introduction	34
3.2	The species niche concept	34
3.2.1	The species niche in environmental and geographical space	35
3.2.2	The species niche in evolutionary time	39
3.2.3	Niche or resource selection function?	40
3.3	Factors controlling species distributions	41
3.4	Environmental gradients and species response functions	41
3.5	Conceptual models of environmental factors controlling species distributions	44
3.5.1	Heat, moisture, light, nutrients, and the distribution of plants	44
3.5.2	Hierarchical and nested scales of factors affecting species distributions	46
3.5.3	Environmental factors affecting species diversity and life form	49
3.6	Summary	51

Part II: The data needed for modeling species distributions

4	Data for species distribution models: the biological data	55
4.1	Introduction – the species data model	55
4.2	Spatial prediction of species distributions: what is being predicted?	55
4.3	Scale concepts related to species data	57
4.4	Spatial sampling design issues related to species data	59
4.4.1	Probability sample designs	60
4.4.2	How many observations?	62
4.4.3	Species prevalence	63

4.4.4	Sample resolution	64
4.4.5	Study area extent and sampling environmental gradients	66
4.4.6	Using existing data for modeling	66
4.4.7	Species presence-only data	71
4.5	Temporal sampling issues and species data	73
4.5.1	Species detectability	73
4.5.2	Historical species data	73
4.6	Summary	74
5	Data for species distribution models: the environmental data	76
5.1	Introduction	76
5.2	Spatial data representing primary environmental regimes	77
5.2.1	Climate maps	77
5.2.2	Digital terrain maps	84
5.2.3	Soil factors and geology maps	89
5.3	Other environmental data for SDM	91
5.3.1	Vegetation maps	91
5.3.2	Disturbance and disturbance history	93
5.3.3	Remote sensing	94
5.3.4	Landscape pattern	99
5.3.5	The distributions of other species	100
5.4	Environmental data for aquatic and marine species	101
5.5	Summary	103

Part III: An overview of the modeling methods

6	Statistical models – modern regression (Janet Franklin and Jennifer A. Miller)	113
6.1	Introduction	113
6.2	The linear model	114
6.3	Generalized linear models	115
6.3.1	Transformations of the predictors	117
6.3.2	Model estimation	119
6.3.3	Model selection and predictor collinearity	122
6.3.4	Use of GLMs in species distribution modeling	124
6.3.5	Summary	125

6.4	Generalized additive models	126
6.4.1	Use of GAMs in species distribution modeling	127
6.4.2	Summary	130
6.5	Multivariate adaptive regression splines	130
6.5.1	Use of MARS in species distribution modeling	134
6.6	Multivariate statistical approaches to SDM	134
6.7	Bayesian approaches to SDM	136
6.8	Spatial autocorrelation and statistical models of species distributions	138
6.8.1	Consequences of SAC data	139
6.8.2	Solutions to SAC data	142
	Autoregression	143
	Applications of autoregression methods in SDM	147
	Generalized estimating equations and generalized linear mixed models	148
	Geographically weighted regression	149
	Spatial filtering methods	151
6.8.3	Summary	152
7	Machine learning methods	154
7.1	Introduction	154
7.2	Decision tree-based methods	155
7.2.1	How decision trees work	155
7.2.2	When are decision trees useful?	161
7.2.3	A note about multivariate decision trees	165
7.2.4	Application of decision trees in species distribution modeling	165
7.3	Ensemble methods applied to decision trees – bagging, boosting, and random forests	165
7.4	Artificial neural networks	170
7.5	Genetic algorithms	173
7.6	Maximum entropy	174
7.7	Support vector machines	174
7.8	Ensemble forecasting and consensus methods	176
7.9	Summary	178

8	Classification, similarity and other methods for presence-only data	180
8.1	Introduction	180
8.2	Envelope models and similarity measures	181
8.2.1	Environmental envelope methods	182
8.2.2	Environmental distance methods	183
8.3	Species presence versus habitat availability	187
8.3.1	Resource selection functions using discriminative models	188
8.3.2	Ecological niche factor analysis	191
8.3.3	Genetic algorithms for rule production (GARP)	194
8.3.4	Maximum entropy	196
8.4	Habitat suitability indices and other expert models	200
8.5	Summary	203

Part IV: Model evaluation and implementation

9	Model evaluation	209
9.1	Introduction	209
9.2	Data for model evaluation	211
9.3	Measures of prediction errors	213
9.3.1	Threshold-dependent measures of accuracy	214
9.3.2	Choosing a threshold for classification	216
9.3.3	Threshold-independent measures of accuracy	222
	AUC	222
	Correlation	224
	Calibration	224
9.3.4	Evaluating presence-only models	226
9.3.5	Spatial distribution of model uncertainty and error	228
9.4	Summary	233
10	Implementation of species distribution models	235
10.1	Introduction	235
10.2	Species attributes	237
10.3	Species data	246
10.4	Environmental data and scale	248

10.5 Modeling methods	253
10.6 Model evaluation	259
10.7 Summary – beyond species distribution modeling	260
<i>References</i>	262
<i>Index</i>	318