

RESOURCE SELECTION BY ANIMALS

Resource Selection by Animals

Statistical Design and Analysis for Field Studies
Second Edition

by

Bryan F.J. Manly

*Western EcoSystems Technology Inc.,
Cheyenne, Wyoming, U.S.A.*

Lyman L. McDonald

*Western EcoSystems Technology Inc.,
Cheyenne, Wyoming, U.S.A.*

Dana L. Thomas

*Department of Mathematical Sciences,
University of Alaska, Fairbanks, U.S.A.*

Trent L. McDonald

*Western EcoSystems Technology Inc.,
Cheyenne, Wyoming, U.S.A.*

and

Wallace P. Erickson

*Western EcoSystems Technology Inc.,
Cheyenne, Wyoming, U.S.A.*

KLUWER ACADEMIC PUBLISHERS
NEW YORK, BOSTON, DORDRECHT, LONDON, MOSCOW

eBook ISBN: 0-306-48151-0
Print ISBN: 1-4020-0677-2

©2004 Kluwer Academic Publishers
New York, Boston, Dordrecht, London, Moscow

Print ©2002 Kluwer Academic Publishers
Dordrecht

All rights reserved

No part of this eBook may be reproduced or transmitted in any form or by any means, electronic, mechanical, recording, or otherwise, without written consent from the Publisher

Created in the United States of America

Visit Kluwer Online at: <http://kluweronline.com>
and Kluwer's eBookstore at: <http://ebooks.kluweronline.com>

TABLE OF CONTENTS

PREFACE	ix
LIST OF SYMBOLS	xi
1. INTRODUCTION TO RESOURCE SELECTION STUDIES	1
1.1 Motivation and Definitions	1
1.2 The Data for Resource Selection Studies	3
1.3 Sampling Protocols and Study Designs	4
1.4 Comparison of Designs	8
1.5 Indices of Selection	9
1.6 Hypothesis Tests and Confidence Intervals	11
1.7 Discussion	14
Chapter Summary	15
2. STATISTICAL MODELLING PROCEDURES	16
2.1 Simple Sample Comparisons	16
2.2 Linear Regression	17
2.3 Logistic Regression	17
2.4 Log-Linear Models	18
2.5 Proportional Hazards Models	18
2.6 Generalized Linear Models	19
2.7 Tests Used in Modelling	20
2.8 Model Selection Using AIC	21
2.9 Analysis of Residuals	22
2.10 Multiple tests and confidence intervals	22
2.11 Bootstrap Methods	23
Chapter Summary	24
3. EXAMPLES OF THE USE OF RESOURCE SELECTION FUNCTIONS	27
3.1 Introduction	27
3.2 Examples	28
3.3 Sample Designs	42
3.4 Assumptions	43
Chapter Summary	44
4. STUDIES WITH RESOURCES DEFINED BY SEVERAL CATEGORIES	46
4.1 Introduction	46
4.2 Sampling Designs and Protocols	46
4.3 Choice of Estimators for Selection Ratios	47
4.4 Variance of a Ratio and the Difference of Ratios	48
4.5 Chi-Squared Tests	49
4.6 Design I with Known Proportions of Available Resource Units	50
4.7 Tests on Proportions of Used Units	52
4.8 Inferences Concerning Selection Ratios	55
4.9 Comparison of Selection Ratios	57

4.10 Design I with Estimated Proportions of Available Resource Units	59
4.11 Design II with Sampling Protocol A	64
4.12 Census of Available Resource Units	65
4.13 Sample of Available Resource Units	72
4.14 Design III with Sampling Protocol A	76
4.15 Census of Available Resource Units	76
4.16 Sample of Available Resource Units	78
4.17 Discussion	78
Chapter Summary	80
Exercise	81
5. RESOURCE SELECTION FUNCTIONS FROM LOGISTIC REGRESSION	83
5.1 Census Data	83
5.2 Use With a Random Sample of Resource Units	84
5.3 Separate Sampling	99
5.4 Separate Samples of Available and Used Units	99
5.5 Separate Samples of Available and Unused Units	102
5.6 Separate Samples of Used and Unused Units	103
5.7 Variances for Estimators and Their Differences	106
5.8 Discussion	109
Chapter Summary	110
Exercises	110
6. RESOURCE SELECTION OVER SEVERAL TIME PERIODS	118
6.1 Census Data	118
6.2 Sample Data	123
6.3 Small Proportions of Used Resource Units Sampled	127
6.4 Samples of Unused Resource Units Only	132
Chapter Summary	135
Exercises	135
7. LOG-LINEAR MODELLING	139
7.1 General Log-linear Modelling of Count Data	139
7.2 Examples	140
Chapter Summary	148
Exercise	149
8. DISCRETE CHOICE MODELS WITH CHANGING AVAILABILITY	150
8.1 The Theory of Discrete Choice Models	150
8.2 Two Units Per Choice Set	152
8.3 Examples	153
8.4 Availability Sampled	161
Chapter Summary	162
Exercises	162
9. APPLICATIONS USING GEOGRAPHIC INFORMATION SYSTEMS	164
9.1 Defining Use and Availability	164
9.2 Variable and Model Selection	165
9.3 Mapping of Resource Selection Functions	165
9.4 Examples	165
Chapter Summary	170

10. DISCRIMINANT FUNCTION ANALYSIS	171
10.1 Linear Discriminant Function Analysis	171
10.2 Samples of Available and Used Resource Units	173
10.3 Samples of Available and Unused Resource Units	173
10.4 Samples of Unused and Used Resource Units	174
10.5 Quadratic Discriminant Function Analysis	177
10.6 Discussion	177
Chapter Summary	177
Exercise	178
11. ANALYSIS OF THE AMOUNT OF USE	179
11.1 Introduction	179
11.2 Analysis of Counts of the Amount of Use	179
11.3 Analysis of Continuous Measures of the Amount of Use	187
11.4 Mixtures of Zeros and Positive Measures of the Amount of Use	189
Chapter Summary	190
Appendix: Estimation of a Modified Log-Linear Model	190
12. SOME OTHER TYPES OF ANALYSIS	192
12.1 Compositional Analysis	192
12.2 Mahalanobis Distance	195
12.3 Estimation Using Machine Learning Methods	197
12.4 Chapter Summary	199
13. RISK ASSESSMENT AND POPULATION SIZE ESTIMATION	200
13.1 Risk Assessment	200
13.2 Estimating Population Size	203
Chapter Summary	204
14. COMPUTING	205
14.1 General Considerations	205
14.2 Logistic Regression Example Code	205
14.3 Log-linear Model Example Code	207
14.4 Discrete Choice Model Example Code	208
Chapter Summary	209
REFERENCES	210
NAME INDEX	217
SUBJECT INDEX	220

PREFACE

We have written this book as a guide to the design and analysis of field studies of resource selection, concentrating primarily on statistical aspects of the comparison of the use and availability of resources of different types. Our intended audience is field ecologists in general and, in particular, wildlife and fisheries biologists who are attempting to measure the extent to which real animal populations are selective in their choice of food and habitat. As such, we have made no attempt to address those aspects of theoretical ecology that are concerned with how animals might choose their resources if they acted in an optimal manner.

The book is based on the concept of a resource selection function (RSF), where this is a function of characteristics measured on resource units such that its value for a unit is proportional to the probability of that unit being used. We argue that this concept leads to a unified theory for the analysis and interpretation of data on resource selection and can replace many ad hoc statistical methods that have been used in the past.

This second edition differs from the first edition in several ways. There is much new material including the uses of the Akaike information criterion (AIC) for model selection, a considerable rearrangement of the material on logistic regression, and completely new chapters on discrete choice models, the analysis of data from geographical information systems (GIS), approaches to studying resource selection other than using RSFs, the uses of RSFs for assessing the risks involved to animals if habitats are changed or estimating population sizes, and computer software that can be used to carry out the calculations for RSF analyses.

The book has the following structure: Chapters 1 and 2 provide a review of statistical methods that have been used in the past to study resource selection and some examples of how different types of study can be analysed in terms of RSFs; Chapter 3 gives a brief introduction to a range of statistical techniques that are used for modelling data in the later chapters; Chapter 4 covers the particularly important special case where the resources being studied are in categories of habitat or food types; Chapter 5 covers the use of logistic regression for estimating a RSF; Chapter 6 covers situations where samples of used or unused resource units are taken over a period of time; Chapter 7 covers the uses of log-linear modelling; Chapter 8 (completely new) covers the uses of discrete choice models; Chapter 9 (completely new) covers the use of GIS data; Chapter 10 covers the uses of discriminant function analysis as an alternative to logistic regression for estimating RSFs; Chapter 11 gives some suggestions on how to analyse data when the amount of use of resource units is recorded rather than simply whether or not they are used; Chapter 12 covers methods such as compositional analysis that do not use RSFs for studying resource selection; Chapter 13 covers the use of RSFs for assessing the risks involved in habitat changes, and for estimating population sizes; and finally Chapter 14 discusses some software that is available to carry out RSF calculations. The book requires much special notation, and readers may find the list that follows this preface to be useful in keeping track of this notation.

We acknowledge the help of many of our colleagues in writing this book. Particularly, we acknowledge Dan Reed who assisted with early versions of Chapter 4, Ed Arnett, C.R. Bantock, S. Farley (chair of the Interagency Brown Bear Study Team), W.F. Harris, Gregory Golet, and Tom Ryder who provided us with raw data from their studies, John Payne, David Yokel, Robert Skinner, Ed Merritt, Fritz Reid, Robb

Macleod and Dick Kempka for support and funding for the development of Chapters 9 and 12, Steve Arthur for figures in Chapter 8, Diana Craig who read over the final manuscript for the first edition of the book, Ryan Nielson and Kimberly Bay who did the same for the second edition of the book, Shay Howlin for help with analyses, and Nadine Wilson who made many of the final changes to the text. Many users of the first edition of the book also made comments that have helped us to improve the second edition. However, any errors or omissions in this edition are entirely our responsibility.

BFJM, LLMc, TLM, DLT and WPE
Laramie, March, 2002

LIST OF SYMBOLS

I	Number of categories of resource units
m_+	Size of a sample of available resource units
m_i	Number of available units in category i in a sample of available resource units
π_i	Proportion of the population of available units that are in category i
$\hat{\pi}_i$	m_i/m_+ , sample proportion of available units in category i
u_+	Size of a sample of used resource units
u_i	Number of units in category i in a sample of used units
o_i	u_i / u_+ , sample proportion of used units in category i
f_i	Proportion of the available category i items that are used.
$\hat{w}_i$	$o_i / \hat{\pi}_i$, the forage ratio (also called the selection ratio and the preference index)
E_i	$(o_i - \hat{\pi}_i) / (o_i + \hat{\pi}_i)$, Ivlev's electivity index
L_i	$o_i - \hat{\pi}_i$, Strauss' linear selection index
Q_i	$[o_i (1 - \hat{\pi}_i)] / [\hat{\pi}_i (1 - o_i)]$, Jacobs' first selection index
D_i	$(o_i - \hat{\pi}_i) / (o_i + \hat{\pi}_i - 2o_i \hat{\pi}_i)$, Jacobs' second selection index
α_i	$(o_i / \hat{\pi}_i) / \Sigma(o_i / \pi_i)$, Chesson's selection index
B_{i1}	$(u_i/m_i) / \Sigma(u_i/m_i)$, Manly's standardized selection index with used resource units replenished
B_{2i}	$\log(1 - f_i) / \Sigma \log(1 - f_i)$, Manly's standardized selection index with used resource units not replenished
W_i	$f_i / \Sigma f_i$, Vanderploeg and Scavia's first selection index
E^*_i	$(W_i - 1/I) / (W_i + 1/I)$, Vanderploeg and Scavia's second selection index
SI	$\text{Max}[(\Sigma m_i/m_+) - (\Sigma o_i/o_+)]$, Rondorff <i>et al.</i> 's selection intensity for continuous data

$w^*(\mathbf{x})$	Resource selection probability function for a single period of selection where resource units are characterized by their values $\mathbf{x} = (x_1, x_2, \dots, x_p)$ for variables X_1 to X_p
$w(\mathbf{x})$	Resource selection function, which is $w^*(\mathbf{x})$ multiplied by an unknown positive constant
$w^*(\mathbf{x}, t)$	Resource selection probability function for selection from time 0 to time t
$w(\mathbf{x}, t)$	Resource selection function for selection from time 0 to time t
$\phi^*(\mathbf{x}, t)$	Probability that a resource unit is not used ('survives') in the time interval 0 to t
$\phi(\mathbf{x}, t)$	$\phi^*(\mathbf{x}, t)$ multiplied by an unknown positive constant
β_i	Coefficient of the variable X_i in a resource selection or survival function
$\hat{\beta}_i$	Estimate of β_i
$se(\hat{\beta}_i)$	Standard error of the estimator of β_i
X_L^2	Log-likelihood chi-squared statistic for measuring the goodness of fit of a model for data
X_p^2	Pearson chi-squared statistic for measuring the goodness of fit of a model for count data
R_i	Standardized residual for the i th observation in a set of data
μ_x	Population mean of the random variable X
s_x	Sample standard deviation of the variable X
r_{XY}	Sample Pearson correlation between variables X and Y
$\text{var}(X)$	Variance of the random variable X
$\text{cov}(X, Y)$	Covariance of random variables X and Y
A_+	Size of a finite population of available resource units
A_i	Number of the A_+ units that are in category i
w^*_i	Proportion of the A_i available resource units that are used
$\hat{w}^*_i$	An estimate of w^*_i
z_α	Value that is exceeded with probability α by a standard normal random variable
$E(X)$	Expected (mean) value of a random variable X

u_{ij}	Number of category i resource units used by animal j
u_{i+}	Number of category i resource units used by all animals
u_{+j}	Number of resource units used by animal j
u_{++}	Total number of units used by all animals
$\hat{w}_{ij}$	Selection ratio for the j th animal and the i th resource category
π_{ij}	Proportion of the resources available to animal j that are in category i
$\hat{\pi}_{ij}$	Sample estimate of π_{ij}
S	Number of selection episodes
Θ_{ij}	Probability that a type i resource unit will be used in the time interval from t_{j-1} to t_j
P_a	Probability of sampling an available resource unit
$P_u(t)$	Probability of sampling a used resource unit at time t
$P_{\bar{u}}(t)$	Probability of sampling an unused resource unit at time t
$u_i(t)$	Number of resource units of type i found in a sample of used units taken at time t
$\bar{u}_i(t)$	Number of resource units of type i found in a sample of unused units taken at time t
P_u	Probability of sampling a used resource unit when only one sample of these units is taken
$P_{\bar{u}}$	Probability of sampling an unused resource unit when only one sample of these units is taken
μ	Population mean vector
Σ	Population covariance matrix
$\Omega(\mathbf{x})$	A function such that selecting individuals with measurements $\mathbf{x} = (x_1, x_2, \dots, x_p)$ with probabilities $\Omega(\mathbf{x})$ from one multivariate population will produce a second population
ρ	Proportion of available units that are used