

2736-2886

DATA ANALYSIS IN  
COMMUNITY AND  
LANDSCAPE ECOLOGY

*Edited by*

R. H. G. JONGMAN, C. J. F. TER BRAAK &  
O. F. R. VAN TONGEREN



**CAMBRIDGE**  
UNIVERSITY PRESS

# Contents

<i>List of Contributors</i>	page	xii
<i>Preface to first edition</i>		xiii
<i>Acknowledgement</i>		xv
<i>List of symbols</i>		xvi
<i>Dune Meadow Data</i>		xix
<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Types of ecological research	1
1.2	Steps in research in community and landscape ecology	2
1.3	Implementing ecological research	4
1.4	Terminology	7
1.5	Historical notes	8
<b>2</b>	<b>Data collection</b>	<b>10</b>
2.1	Data collection, not an isolated or undirected activity	10
2.2	Objectives and data collection, start by defining the finish	11
2.3	Planning of data collection	12
2.3.1	Study design	12
2.3.2	Study type	14
2.3.3	Sampling strategy	15
2.4	Analysis of data and data collection	16
2.4.1	Introduction	16
2.4.2	Measurement scales	16
2.4.3	Frequency distributions	18
2.4.4	Transformations	20
2.5	Interpretation of results	22
2.6	Sources of misinterpretation	23
2.6.1	Incorrect use of statistical analysis	23
2.6.2	Taking detection to be proof	24
2.6.3	Misinterpretation caused by biased sampling	24
2.6.4	Other sources of misinterpretation	25
2.7	Complexity of ecological research	25
2.8	Bibliographical notes	26
2.9	Exercise	27
2.10	Solution to the exercise	27

<b>3</b>	<b>Regression</b>	29
3.1	Introduction	29
3.1.1	Aim and use	29
3.1.2	Response model and types of response variables	30
3.1.3	Types of explanatory variables and types of response curves	31
3.1.4	Outline of the chapter	33
3.2	Regression for quantitative abundance data: least-squares regression	33
3.2.1	Nominal explanatory variables: analysis of variance	33
3.2.2	Straight lines	37
3.2.3	Parabolas and Gaussian curves	40
3.3	Regression for presence-absence data: logit regression	43
3.3.1	Nominal explanatory variables: chi-square test	43
3.3.2	Sigmoid curves	44
3.3.3	Gaussian logit curves	47
3.4	Regression for abundance data with many zero values	49
3.5	Multiple regression	51
3.5.1	Introduction	51
3.5.2	Multiple least-squares regression: planes and other surfaces	51
3.5.3	Multiple logit regression: logit planes and Gaussian logit surfaces	53
3.5.4	Interaction between explanatory variables	56
3.5.5	Nominal explanatory variables	58
3.6	Model choice and regression diagnostics	60
3.7	The method of weighted averaging	61
3.8	Bibliographic notes	63
3.9	Standard errors of estimated optimum and tolerance; confidence interval for the optimum	65
3.10	Exercises	66
3.11	Solutions to exercises	72
<b>4</b>	<b>Calibration</b>	78
4.1	Introduction	78
4.2	Maximum likelihood calibration using response functions	79
4.2.1	Introduction	79
4.2.2	Predicting a nominal environmental variable	79
4.2.3	Predicting a quantitative environmental variable	81
4.3	Weighted averaging using indicator values	83
4.4	Inverse regression	86
4.5	Bibliographic notes	87
4.6	Exercises	87
4.7	Solutions to exercises	89
<b>5</b>	<b>Ordination</b>	91
5.1	Introduction	91
5.1.1	Aim and usage	91

5.1.2	Data approximation and response models in ordination	93
5.1.3	Outline of Chapter 5	95
5.2	Correspondence analysis (CA) and detrended correspondence analysis (DCA)	95
5.2.1	From weighted averaging to correspondence analysis	95
5.2.2	Two-way weighted averaging algorithm	97
5.2.3	Diagonal structures: properties and faults of correspondence analysis	103
5.2.4	Detrended correspondence analysis (DCA)	105
5.2.5	Joint plot of species and sites	108
5.2.6	Block structures and sensitivity to rare species	109
5.2.7	Gaussian ordination and its relation with CA and DCA	110
5.3	Principal components analysis (PCA)	116
5.3.1	From least-squares regression to principal components analysis	116
5.3.2	Two-way weighted summation algorithm	119
5.3.3	Best lines and planes in m-dimensional space	125
5.3.4	Biplot of species and site scores	127
5.3.5	Data transformation	130
5.3.6	R-mode and Q-mode algorithms	131
5.4	Interpretation of ordination with external data	132
5.5	Canonical ordination	136
5.5.1	Introduction	136
5.5.2	Canonical correspondence analysis (CCA)	137
5.5.3	Redundancy analysis (RDA)	144
5.5.4	Canonical correlation analysis (COR)	147
5.5.5	Canonical variate analysis (CVA)	148
5.5.6	Interpreting canonical axes	150
5.5.7	Data transformation	151
5.6	Multidimensional scaling	151
5.7	Evaluation of direct gradient and indirect gradient analysis techniques	153
5.8	Bibliographic notes	156
5.9	Ordination methods in terms of matrix algebra	158
5.9.1	Principal components analysis (PCA)	158
5.9.2	Correspondence analysis (CA)	160
5.9.3	Canonical correlation analysis (COR)	161
5.9.4	Redundancy analysis (RDA)	163
5.9.5	Canonical correspondence analysis (CCA)	163
5.10	Exercises	165
5.11	Solutions to exercises	169
<b>6</b>	<b>Cluster analysis</b>	<b>174</b>
6.1	Introduction	174
6.1.1	Aims and use	174
6.1.2	Types of cluster analysis	175
6.2	Agglomerative methods	175

6.2.1	Introduction	175
6.2.2	Similarity and dissimilarity	176
6.2.3	Properties of the indices	180
6.2.4	Transformation, standardization and weighting	184
6.2.5	Agglomerative cluster algorithms	186
6.3	Divisive methods	191
6.3.1	Introduction	191
6.3.2	Association analysis and related methods	193
6.3.3	Two Way INdicator SPecies ANalysis	193
6.4	Non-hierarchical clustering	197
6.5	Optimality of a clustering	198
6.6	Presentation of the results	199
6.7	Relation between community types and environment	199
6.7.1	The use of simple descriptive statistics	201
6.7.2	The use of tests of significance	202
6.8	Bibliography	203
6.9	Exercises	204
6.10	Solutions to exercises	207
<b>7</b>	<b>Spatial aspects of ecological data</b>	<b>213</b>
7.1	Introduction	213
7.2	Trend-surface analysis	215
7.2.1	Introduction	215
7.2.2	One-dimensional trends	215
7.2.3	Extensions to non-normally distributed variables	216
7.2.4	Non-linear dependence of $Z$ and $X$	216
7.2.5	Two-dimensional surfaces	217
7.2.6	Trend-surfaces as interpolators	218
7.2.7	Use of trend surfaces	219
7.2.8	Local trends	222
7.3	Spatial autocovariation	222
7.3.1	Introduction	222
7.3.2	Stationarity	224
7.3.3	Statistical moments of spatial series	224
7.3.4	Autocovariograms and semivariograms	226
7.3.5	Example of the use of autocorrelation analysis	227
7.3.6	Considerations of non-stationarity	228
7.3.7	The semivariogram, its properties and uses	230
7.3.8	Isotropic and anisotropic variation	234
7.3.9	The semivariogram as a structural tool	235
7.3.10	An example of using semivariograms to analyse spatial variation in soil	235
7.4	Spatial interpolation	237
7.4.1	Introduction	237
7.4.2	Weighted moving averages	237
7.4.3	Optimum interpolation methods using spatial autocovariance	239

7.5	Optimum sampling	242
7.6	Discussion	245
7.7	Bibliography and historical overview of interpolation	245
7.8	Exercises	246
7.9	Solutions to exercises	248
<b>8</b>	<b>Numerical methods in practice: case-studies</b>	<b>252</b>
8.1	Spatial variation in vegetation, soil and precipitation across the Hawkesbury Sandstone Plateau in New South Wales, Australia: an analysis using ordination and trend surfaces	252
8.1.1	Data collection and analysis	252
8.1.2	Deviations from the trends	256
8.1.3	Conclusions	257
8.2	Processing of vegetation relevés to obtain a classification of south-west European heathland vegetation	258
8.2.1	Introduction	258
8.2.2	Preparation of the set of data	258
8.2.3	Operations applied to the set of data	259
8.2.4	Preliminary results	263
8.2.5	Concluding remarks	265
8.3	Restoration and maintenance of artificial brooks in the Veluwe, the Netherlands	266
8.3.1	Introduction	266
8.3.2	Ecological inventory	268
8.3.3	Data analysis	268
8.3.4	Results	269
8.3.5	Application of the results	272
	<i>References</i>	275
	<i>Index</i>	292