

---

# Contents

---

Preface.....	xv
Authors .....	xix
<b>1. Introduction .....</b>	<b>1</b>
1.1 Different Types of Scientific Study .....	1
1.2 Relating Sample Results to More General Populations .....	3
1.3 Constructing Models to Represent Reality .....	4
1.4 Using Linear Models .....	7
1.5 Estimating the Parameters of Linear Models .....	8
1.6 Summarizing the Importance of Model Terms .....	9
1.7 The Scope of This Book .....	11
<b>2. A Review of Basic Statistics .....</b>	<b>13</b>
2.1 Summary Statistics and Notation for Sample Data .....	13
2.2 Statistical Distributions for Populations.....	16
2.2.1 Discrete Data .....	17
2.2.2 Continuous Data .....	22
2.2.3 The Normal Distribution .....	24
2.2.4 Distributions Derived from Functions of Normal Random Variables.....	26
2.3 From Sample Data to Conclusions about the Population.....	28
2.3.1 Estimating Population Parameters Using Summary Statistics .....	28
2.3.2 Asking Questions about the Data: Hypothesis Testing .....	29
2.4 Simple Tests for Population Means .....	30
2.4.1 Assessing the Mean Response: The One-Sample t-Test .....	30
2.4.2 Comparing Mean Responses: The Two-Sample t-Test.....	32
2.5 Assessing the Association between Variables .....	36
2.6 Presenting Numerical Results.....	39
Exercises .....	41
<b>3. Principles for Designing Experiments .....</b>	<b>43</b>
3.1 Key Principles .....	43
3.1.1 Replication .....	46
3.1.2 Randomization.....	48
3.1.3 Blocking.....	51
3.2 Forms of Experimental Structure .....	52
3.3 Common Forms of Design for Experiments .....	57
3.3.1 The Completely Randomized Design.....	57
3.3.2 The Randomized Complete Block Design .....	58
3.3.3 The Latin Square Design .....	59
3.3.4 The Split-Plot Design .....	60
3.3.5 The Balanced Incomplete Block Design .....	61
3.3.6 Generating a Randomized Design .....	62
Exercises .....	62

<b>4. Models for a Single Factor</b> .....	69
4.1 Defining the Model.....	69
4.2 Estimating the Model Parameters .....	73
4.3 Summarizing the Importance of Model Terms .....	74
4.3.1 Calculating Sums of Squares .....	76
4.3.2 Calculating Degrees of Freedom and Mean Squares .....	80
4.3.3 Calculating Variance Ratios as Test Statistics.....	81
4.3.4 The Summary ANOVA Table.....	82
4.4 Evaluating the Response to Treatments.....	84
4.4.1 Prediction of Treatment Means.....	84
4.4.2 Comparison of Treatment Means .....	85
4.5 Alternative Forms of the Model.....	88
Exercises .....	90
<b>5. Checking Model Assumptions</b> .....	93
5.1 Estimating Deviations.....	93
5.1.1 Simple Residuals .....	94
5.1.2 Standardized Residuals .....	95
5.2 Using Graphical Tools to Diagnose Problems .....	96
5.2.1 Assessing Homogeneity of Variances.....	96
5.2.2 Assessing Independence.....	98
5.2.3 Assessing Normality .....	101
5.2.4 Using Permutation Tests Where Assumptions Fail .....	102
5.2.5 The Impact of Sample Size.....	103
5.3 Using Formal Tests to Diagnose Problems.....	104
5.4 Identifying Inconsistent Observations .....	108
Exercises .....	110
<b>6. Transformations of the Response</b> .....	113
6.1 Why Do We Need to Transform the Response? .....	113
6.2 Some Useful Transformations.....	114
6.2.1 Logarithms.....	114
6.2.2 Square Roots .....	119
6.2.3 Logits .....	120
6.2.4 Other Transformations.....	121
6.3 Interpreting the Results after Transformation.....	122
6.4 Interpretation for Log-Transformed Responses .....	123
6.5 Other Approaches.....	126
Exercises .....	127
<b>7. Models with a Simple Blocking Structure</b> .....	129
7.1 Defining the Model.....	130
7.2 Estimating the Model Parameters .....	132
7.3 Summarizing the Importance of Model Terms .....	134
7.4 Evaluating the Response to Treatments.....	140
7.5 Incorporating Strata: The Multi-Stratum Analysis of Variance .....	141
Exercises .....	146

<b>8. Extracting Information about Treatments</b> .....	149
8.1 From Scientific Questions to the Treatment Structure .....	150
8.2 A Crossed Treatment Structure with Two Factors.....	152
8.2.1 Models for a Crossed Treatment Structure with Two Factors.....	153
8.2.2 Estimating the Model Parameters .....	155
8.2.3 Assessing the Importance of Individual Model Terms.....	158
8.2.4 Evaluating the Response to Treatments: Predictions from the Fitted Model.....	160
8.2.5 The Advantages of Factorial Structure .....	162
8.2.6 Understanding Different Parameterizations .....	163
8.3 Crossed Treatment Structures with Three or More Factors .....	164
8.3.1 Assessing the Importance of Individual Model Terms.....	166
8.3.2 Evaluating the Response to Treatments: Predictions from the Fitted Model .....	171
8.4 Models for Nested Treatment Structures .....	173
8.5 Adding Controls or Standards to a Set of Treatments.....	179
8.6 Investigating Specific Treatment Comparisons.....	182
8.7 Modelling Patterns for Quantitative Treatments .....	190
8.8 Making Treatment Comparisons from Predicted Means .....	195
8.8.1 The Bonferroni Correction .....	196
8.8.2 The False Discovery Rate .....	197
8.8.3 All Pairwise Comparisons.....	198
8.8.3.1 The LSD and Fisher’s Protected LSD.....	198
8.8.3.2 Multiple Range Tests.....	199
8.8.3.3 Tukey’s Simultaneous Confidence Intervals .....	200
8.8.4 Comparison of Treatments against a Control.....	201
8.8.5 Evaluation of a Set of Pre-Planned Comparisons .....	201
8.8.6 Summary of Issues .....	205
Exercises .....	206
<b>9. Models with More Complex Blocking Structure</b> .....	209
9.1 The Latin Square Design.....	209
9.1.1 Defining the Model.....	211
9.1.2 Estimating the Model Parameters .....	211
9.1.3 Assessing the Importance of Individual Model Terms.....	212
9.1.4 Evaluating the Response to Treatments: Predictions from the Fitted Model.....	215
9.1.5 Constraints and Extensions of the Latin Square Design .....	217
9.2 The Split-Plot Design .....	220
9.2.1 Defining the Model.....	222
9.2.2 Assessing the Importance of Individual Model Terms.....	223
9.2.3 Evaluating the Response to Treatments: Predictions from the Fitted Model.....	225
9.2.4 Drawbacks and Variations of the Split-Plot Design.....	228
9.3 The Balanced Incomplete Block Design.....	232
9.3.1 Defining the Model.....	235
9.3.2 Assessing the Importance of Individual Model Terms.....	236

9.3.3	Drawbacks and Variations of the Balanced Incomplete Block Design.....	237
	Exercises .....	238
<b>10.</b>	<b>Replication and Power .....</b>	<b>241</b>
10.1	Simple Methods for Determining Replication.....	242
10.1.1	Calculations Based on the LSD.....	242
10.1.2	Calculations Based on the Coefficient of Variation .....	243
10.1.3	Unequal Replication and Models with Blocking .....	244
10.2	Estimating the Background Variation .....	245
10.3	Assessing the Power of a Design .....	245
10.4	Constructing a Design for a Particular Experiment .....	249
10.5	A Different Hypothesis: Testing for Equivalence.....	253
	Exercise.....	256
<b>11.</b>	<b>Dealing with Non-Orthogonality .....</b>	<b>257</b>
11.1	The Benefits of Orthogonality .....	257
11.2	Fitting Models with Non-Orthogonal Terms.....	259
11.2.1	Parameterizing Models for Two Non-Orthogonal Factors .....	259
11.2.2	Assessing the Importance of Non-Orthogonal Terms: The Sequential ANOVA Table .....	265
11.2.3	Calculating the Impact of Model Terms.....	269
11.2.4	Selecting the Best Model.....	270
11.2.5	Evaluating the Response to Treatments: Predictions from the Fitted Model.....	270
11.3	Designs with Planned Non-Orthogonality.....	272
11.3.1	Fractional Factorial Designs.....	273
11.3.2	Factorial Designs with Confounding.....	274
11.4	The Consequences of Missing Data .....	274
11.5	Incorporating the Effects of Unplanned Factors .....	277
11.6	Analysis Approaches for Non-Orthogonal Designs.....	280
11.6.1	A Simple Approach: The Intra-Block Analysis.....	281
	Exercises .....	284
<b>12.</b>	<b>Models for a Single Variate: Simple Linear Regression .....</b>	<b>287</b>
12.1	Defining the Model.....	288
12.2	Estimating the Model Parameters .....	292
12.3	Assessing the Importance of the Model .....	296
12.4	Properties of the Model Parameters.....	299
12.5	Using the Fitted Model to Predict Responses .....	301
12.6	Summarizing the Fit of the Model .....	305
12.7	Consequences of Uncertainty in the Explanatory Variate .....	306
12.8	Using Replication to Test Goodness of Fit.....	308
12.9	Variations on the Model.....	313
12.9.1	Centering and Scaling the Explanatory Variate.....	313
12.9.2	Regression through the Origin.....	314
12.9.3	Calibration .....	320
	Exercises .....	321

<b>13. Checking Model Fit</b> .....	325
13.1 Checking the Form of the Model.....	325
13.2 More Ways of Estimating Deviations .....	328
13.3 Using Graphical Tools to Check Assumptions.....	330
13.4 Looking for Influential Observations .....	332
13.4.1 Measuring Potential Influence: Leverage.....	333
13.4.2 The Relationship between Residuals and Leverages .....	335
13.4.3 Measuring the Actual Influence of Individual Observations .....	336
13.5 Assessing the Predictive Ability of a Model: Cross-Validation.....	338
Exercises .....	342
<b>14. Models for Several Variates: Multiple Linear Regression</b> .....	345
14.1 Visualizing Relationships between Variates.....	345
14.2 Defining the Model.....	347
14.3 Estimating the Model Parameters .....	350
14.4 Assessing the Importance of Individual Explanatory Variates .....	352
14.4.1 Adding Terms into the Model: Sequential ANOVA and Incremental Sums of Squares.....	353
14.4.2 The Impact of Removing Model Terms: Marginal Sums of Squares ...	356
14.5 Properties of the Model Parameters and Predicting Responses.....	358
14.6 Investigating Model Misspecification.....	359
14.7 Dealing with Correlation among Explanatory Variates.....	361
14.8 Summarizing the Fit of the Model .....	365
14.9 Selecting the Best Model.....	366
14.9.1 Strategies for Sequential Variable Selection.....	369
14.9.2 Problems with Procedures for the Selection of Subsets of Variables .....	376
14.9.3 Using Cross-Validation as a Tool for Model Selection.....	377
14.9.4 Some Final Remarks on Procedures for Selecting Models .....	378
Exercises .....	378
<b>15. Models for Variates and Factors</b> .....	381
15.1 Incorporating Groups into the Simple Linear Regression Model.....	382
15.1.1 An Overview of Possible Models .....	383
15.1.2 Defining and Choosing between the Models.....	388
15.1.2.1 Single Line Model .....	388
15.1.2.2 Parallel Lines Model .....	388
15.1.2.3 Separate Lines Model .....	390
15.1.2.4 Choosing between the Models: The Sequential ANOVA Table.....	391
15.1.3 An Alternative Sequence of Models.....	396
15.1.4 Constraining the Intercepts.....	398
15.2 Incorporating Groups into the Multiple Linear Regression Model.....	399
15.3 Regression in Designed Experiments .....	406
15.4 Analysis of Covariance: A Special Case of Regression with Groups .....	409
15.5 Complex Models with Factors and Variates.....	414
15.5.1 Selecting the Predictive Model .....	414
15.5.2 Evaluating the Response: Predictions from the Fitted Model.....	417

15.6	The Connection between Factors and Variates .....	417
15.6.1	Rewriting the Model in Matrix Notation .....	421
	Exercises .....	423
<b>16.</b>	<b>Incorporating Structure: Linear Mixed Models .....</b>	<b>427</b>
16.1	Incorporating Structure .....	427
16.2	An Introduction to Linear Mixed Models .....	428
16.3	Selecting the Best Fixed Model .....	430
16.4	Interpreting the Random Model.....	432
16.4.1	The Connection between the Linear Mixed Model and Multi-Stratum ANOVA .....	434
16.5	What about Random Effects? .....	435
16.6	Predicting Responses .....	436
16.7	Checking Model Fit.....	437
16.8	An Example.....	438
16.9	Some Pitfalls and Dangers.....	444
16.10	Extending the Model .....	445
	Exercises .....	447
<b>17.</b>	<b>Models for Curved Relationships .....</b>	<b>451</b>
17.1	Fitting Curved Functions by Transformation.....	451
17.1.1	Simple Transformations of an Explanatory Variate.....	451
17.1.2	Polynomial Models.....	456
17.1.3	Trigonometric Models for Periodic Patterns.....	460
17.2	Curved Surfaces as Functions of Two or More Variates.....	463
17.3	Fitting Models Including Non-Linear Parameters.....	472
	Exercises .....	476
<b>18.</b>	<b>Models for Non-Normal Responses: Generalized Linear Models.....</b>	<b>479</b>
18.1	Introduction to Generalized Linear Models .....	480
18.2	Analysis of Proportions Based on Counts: Binomial Responses.....	481
18.2.1	Understanding and Defining the Model.....	483
18.2.2	Assessing the Importance of the Model and Individual Terms: The Analysis of Deviance.....	487
18.2.2.1	Interpreting the ANODEV with No Over-Dispersion.....	489
18.2.2.2	Interpreting the ANODEV with Over-Dispersion .....	490
18.2.2.3	The Sequential ANODEV Table .....	493
18.2.3	Checking the Model Fit and Assumptions .....	494
18.2.4	Properties of the Model Parameters.....	496
18.2.5	Evaluating the Response to Explanatory Variables: Prediction.....	498
18.2.6	Aggregating Binomial Responses .....	500
18.2.7	The Special Case of Binary Data.....	501
18.2.8	Other Issues with Binomial Responses .....	501
18.3	Analysis of Count Data: Poisson Responses .....	502
18.3.1	Understanding and Defining the Model.....	503
18.3.2	Analysis of the Model .....	506
18.3.3	Analysing Poisson Responses with Several Explanatory Variables.....	509
18.3.4	Other Issues with Poisson Responses.....	512

18.4	Other Types of GLM and Extensions .....	512
	Exercises .....	513
<b>19.</b>	<b>Practical Design and Data Analysis for Real Studies .....</b>	<b>517</b>
19.1	Designing Real Studies .....	518
19.1.1	Aims, Objectives and Choice of Explanatory Structure .....	518
19.1.2	Resources, Experimental Units and Constraints .....	519
19.1.3	Matching the Treatments to the Resources.....	520
19.1.4	Designs for Series of Studies and for Studies with Multiple Phases....	521
19.1.5	Design Case Studies .....	523
19.2	Choosing the Best Analysis Approach .....	535
19.2.1	Analysis of Designed Experiments.....	536
19.2.2	Analysis of Observational Studies .....	537
19.2.3	Different Types of Data .....	538
19.3	Presentation of Statistics in Reports, Theses and Papers .....	538
19.3.1	Statistical Information in the Materials and Methods .....	539
19.3.2	Presentation of Results.....	540
19.4	And Finally. . . ..	543
	<b>References .....</b>	<b>545</b>
	<b>Appendix A: Data Tables .....</b>	<b>551</b>
	<b>Appendix B: Quantiles of Statistical Distributions .....</b>	<b>559</b>
	<b>Appendix C: Statistical and Mathematical Results .....</b>	<b>563</b>
	<b>Index .....</b>	<b>569</b>





---

## *Preface*

---

This book provides an introductory, practical and illustrative guide to the design of experiments and data analysis in the biological and agricultural plant sciences. It is aimed both at research scientists and at students (from final year undergraduate level through taught masters to PhD students) who either need to design their own experiments and perform their own analyses or can consult with a professional applied statistician and want to have a clear understanding of the methods that they are using. The material is based on courses developed at two British research institutes (Rothamsted Research and Horticulture Research International [HRI – then Warwick HRI, and now the School of Life Sciences, University of Warwick]) to train research scientists and post-graduate students in these key areas of statistics. Our overall approach is intended to be practical and intuitive rather than overly theoretical, with mathematical formulae presented only to formalize the methods where appropriate and necessary. Our intention is to present statistical ideas in the context of the biological and agricultural sciences to which they are being applied, drawing on relevant examples from our own experiences as consultant applied statisticians at research institutes, to encourage best practice in design and data analysis.

The first two chapters of this book provide introductory, review and background material. In Chapter 1, we introduce types of data and statistical models, together with an overview of the basic statistical concepts and the terminology used throughout. The training courses on which this book is based are intended to follow preliminary courses that introduce the basic ideas of summary statistics, simple statistical distributions (Normal, Poisson, Binomial), confidence intervals, and simple statistical tests (including the t-test and F-test). Whilst a brief review of such material is covered in Chapter 2, the reader will need to be comfortable with these ideas to reap the greatest benefit from reading the rest of the book. Some readers may feel that their knowledge of basic statistics is sufficiently comprehensive that they can skip this review chapter. However, we recommend you browse through it to familiarize yourself with the statistical terminology that we use.

The main body of the book follows. Chapters 3 to 11 introduce statistical approaches to the design of experiments and the analysis of data from such designed experiments. We start from basic design principles, introduce some simple designs, and then extend to more complex ones including factorial treatment structures, treatment contrasts and blocking structures. We describe the use of analysis of variance (ANOVA) to summarize the data, including the use of the multi-stratum ANOVA to account for the physical structure of the experimental material or blocking imposed by the experimenter, introduce simple diagnostic methods, and discuss potential transformations of the response. We explain the analysis of standard designs, including the randomized complete block, Latin square, split-plot and balanced incomplete block designs in some detail. We also explore the issues of sample size estimation and the power of a design. Finally, we look at the analysis of unbalanced or non-orthogonal designs. Chapters 12 to 18 first introduce the idea of simple linear regression to relate a response variable to a single explanatory variable, and then consider extensions and modifications of this approach to cope with more complex data sets and relationships. These include multiple linear regression, simple linear regression with groups, linear mixed models and models for curved relationships. We also extend related themes from the earlier chapters, including diagnostic methods specific to regression. We emphasize throughout that the same type of models and principles are used for

both designed experiments and regression modelling. We complete the main body of the book with a discussion of generalized linear models, which are appropriate for certain types of non-Normal data.

We conclude with a guide to practical design and data analysis (Chapter 19), which focuses on the selection of the most appropriate design or analysis approach for individual scientific problems and on the interpretation and presentation of the results of the analysis.

Most chapters include exercises which we hope will help to consolidate the ideas introduced in the chapter. In running the training courses from which this book has been developed, we often find that it is only when students perform the analyses themselves that they fully appreciate the statistical concepts and, most importantly, understand how to interpret the results of the analyses. We therefore encourage you to work through at least some of the exercises for each chapter before moving to the next one. There are fewer exercises in the earlier chapters and the required analyses build in complexity, so we expect you to apply knowledge gained throughout the book when doing exercises from the later chapters. All of the data sets and solutions to selected exercises are available online. Some of the solutions include further discussion of the relevant statistical issues.

We have set up a website to accompany this book ([www.stats4biol.info](http://www.stats4biol.info)) where we show how to do the analyses described in the book using GenStat®, R and SAS®, three commonly used statistical packages. Whilst users familiar with any of these packages might not refer to this material, others are encouraged to review it and work through the examples and exercises for at least one of the packages. Any errors found after publication will also be recorded on this website.

By the time you reach the end of the book (and online material) we intend that you will have gained

- A clear appreciation of the importance of a statistical approach to the design of your experiments,
- A sound understanding of the statistical methods used to analyse data obtained from designed experiments and of the regression approaches used to construct simple models to describe the observed response as a function of explanatory variables,
- Sufficient knowledge of how to use one or more statistical packages to analyse data using the approaches that we describe, and most importantly,
- An appreciation of how to interpret the results of these statistical analyses in the context of the biological or agricultural science within which you are working.

By doing so, you will be better able both to interact with a consultant statistician, should you have access to one, and to identify suitable statistical approaches to add value to your scientific research.

This book relies heavily on the use of real data sets and material from the original courses and we are hence indebted to many people for their input. Particular thanks go to Stephen Powers and Rodger White (Rothamsted Research) and John Fenlon, Gail Kingswell and Julie Jones (HRI) for their contributions to the original courses; also to Alan Todd (Rothamsted Research) for providing many valuable suggestions for suitable data sets. The majority of real data sets used arose from projects (including PhDs) at Rothamsted Research, many in collaboration with other institutes and funded from many sources; we thank Rothamsted Research for giving us general permission to use these data. We also thank, in alphabetical order, R. Alarcon-Reverte, S. Amoah, J. Baverstock, P. Brookes,

J. Chapman, R. Curtis, I. Denholm, N. Evans, A. Ferguson, S. Foster, M. Glendining, K. Hammond-Kosack, R. Harrington, Y. Huang, R. Hull, J. Jenkyn, H.-C. Jing, A.E. Johnston, A. Karp, J. Logan, J. Lucas, P. Lutman, A. Macdonald, S. McGrath, T. Miller, S. Moss, J. Pell, R. Plumb, P. Poulton, A. Salisbury, T. Scott, I. Shield, C. Shortall, L. Smart, M. Torrance, P. Wells, M. Wilkinson and E. Wright, for specific permission to use data from their own projects or from those undertaken within their group or department at Rothamsted. Rothamsted Research receives grant-aided support from the Biotechnology and Biological Sciences Research Council of the United Kingdom. We thank various colleagues, past and present, at Horticulture Research International, Warwick HRI and the School of Life Sciences, University of Warwick, for permission to use data from their research projects, particularly Rosemary Collier and John Clarkson. We thank M. Heard (Centre for Ecology and Hydrology), A. Ortega Z. (Universidad Austral de Chile) and R. Webster for permission to use data. Examples and exercises marked ‘\*’ use simulated data inspired by experiments carried out at Rothamsted Research or HRI. The small remainder of original examples and exercises (also marked ‘\*’) were invented by the authors but are typical of the type of experiments we are regularly asked to design and the data we analyse as part of our consultancy work. In the few cases where we have not been able to find examples from our own work we have drawn on data from published sources. We would like to thank Simon Harding for technical help in setting up a repository for our work and our website and Richard Webster, Alice Milne, Nick Galwey, James Bell and Kathy Ruggeiro and an anonymous referee for reading draft chapters and providing many helpful comments and suggestions.

Finally, we would like to make some individual acknowledgements. SJW, SJC and SAG thank Rothamsted Research, and in particular Chris Rawlings, for support and encouragement to pursue this project. AM thanks his colleagues at Horticulture Research International and the University of Warwick, particularly John Fenlon, for support in the development of the original training courses, and hence the development of this project, and his co-authors for the invitation to join this project. SJW thanks Simon Harding for his support, help and long-term forbearance. SAG thanks Emma Weeks for her encouragement, and the other co-authors for their patience and the fruitful discussions we had on this project. SJC thanks Tony Scott for his patience and support, Elisa Allen for her contribution to the presentation of our courses and useful comments on some chapters, and past students for their enthusiasm and constructive feedback which led to improvements in our courses and ultimately this book. AM also thanks his family, Sara and Tom, for their continuing support and understanding.

**S J Welham**

*Welwyn Garden City, UK*

**S A Gezan**

*Harpندن, UK and Gainesville, Florida, USA*

**S J Clark**

*Harpندن, UK*

**A Mead**

*Leamington Spa, UK*



---

## *Authors*

---

**Suzanne Jane Welham** obtained an MSc in statistical sciences from University College London in 1987 and worked as an applied statistician at Rothamsted Research from 1987 to 2000, collaborating with scientists and developing statistical software. She pursued a PhD from 2000 to 2003 at the London School of Hygiene and Tropical Medicine and then returned to Rothamsted, during which time she coauthored the in-house statistics courses that motivated the writing of this book. She is a coauthor of about 60 published papers and currently works for VSN International Ltd on the development of statistical software for analysis of linear mixed models and presents training courses on their use in R and GenStat.

**Salvador Alejandro Gezan, PhD**, is an assistant professor at the School of Forest Resources and Conservation at the University of Florida since 2011. Salvador obtained his bachelor's from the Universidad of Chile in forestry and his PhD from the University of Florida in statistics-genetics. He then worked as an applied statistician at Rothamsted Research, collaborating on the production and development of the in-house courses that formed the basis for this book. Currently, he teaches courses in linear and mixed model effects, quantitative genetics and forest mensuration. He carries out research and consulting in statistical application to biological sciences with emphasis on genetic improvement of plants and animals. Salvador is a long-time user of SAS, which he combines with GenStat, R and MATLAB® as required.

**Suzanne Jane Clark** has worked at Rothamsted Research as an applied statistician since 1981. She primarily collaborates with ecologists and entomologists at Rothamsted, providing and implementing advice on statistical issues ranging from planning and design of experiments through to data analysis and presentation of results, and has coauthored over 130 scientific papers. Suzanne coauthored and presents several of the in-house statistics courses for scientists and research students, which inspired the writing of this book. An experienced and long-term GenStat user, Suzanne has also written several procedures for the GenStat Procedure Library and uses GenStat daily for the analyses of biological data using a wide range of statistical techniques, including those covered in this book.

**Andrew Mead** obtained a BSc in statistics at the University of Bath and an MSc in biometry at the University of Reading, where he spent over 16 years working as a consultant and research biometrician at the Institute of Horticultural Research and Horticulture Research International at Wellesbourne, Warwickshire, UK. During this time, he developed and taught a series of statistics training courses for staff and students at the institute, producing some of the material on which this book is based. For 10 years from 2004 he worked as a research biometrician and teaching fellow at the University of Warwick, developing and leading the teaching of statistics for both postgraduate and undergraduate students across a range of life sciences. In 2014 he was appointed as Head of Applied Statistics at Rothamsted Research. Throughout his career he has had a strong association with the International Biometric Society, serving as International President and Vice

President from 2007 to 2010 inclusive, having been the first recipient of the 'Award for Outstanding Contribution to the Development of the International Biometric Society' in 2006, serving as a Regional Secretary of the British and Irish Region from 2000 to 2007 and on the International Council from 2002 to 2010. He is a (co)author of over 80 papers, and coauthor of *Statistical Principles for the Design of Experiments: Applications to Real Experiments* published in 2012.