



## Brevia

### BOOK REVIEW

#### Vital Statistics

Swan, A. R. H. and Sandilands, M. 1995. *Introduction to Geological Data Analysis*. Blackwell Science, Oxford, UK. 446 pp. ISBN 0-632-03224-3. Price: paperback US\$39.95.

Via the computers on our desks we now enjoy access to a wealth of software for the presentation and evaluation of geological data. The spreadsheets, presentation and statistical packages installed on our machines make it possible to generate tables, statistical parameters, graphs and block diagrams, all to publication standard literally at the press of a button. In principle, the availability of these new tools should mean that more secure scientific deductions than ever before can be extracted from our precious raw data. These new capabilities do, however, bring dangers with them. It is easy to become mesmerised by the flashy 4-dimensional high-resolution displays in 256 colours, and easy to be over-impressed by how much better our data looks now than it did in hand-written form in our field notebooks. Never before has it been so easy to massage the figures; to select the interpolation method which enhances the pattern on the map you had hoped for, or to smooth away those awkward bumps. There is a general danger that these tools will be misused if the user is unaware of the theoretical basis of the techniques employed. In this context, this book by Swan and Sandilands provides a clear and very readable description of the principles behind a wide range of techniques for statistical analysis and data presentation.

The substantial text is built up of twelve chapters, each one well illustrated with figures and worked examples. Chapter 1, entitled Data Collection and Preparation, deals with different classes of data, the quality of data, and issues concerned with sampling. This is followed by a meaty chapter (Chapter 2) on univariate statistics. This important section explains how invariate data can be represented graphically, not only by the familiar histogram but also by the dot-plot, stem-and-leaf diagram and the box-and-whisker plot. The usual numerical descriptors of central tendency, variance and skewness are explained. An introduction to probability theory is followed by a discussion of various probability distributions, statistical inference and analysis of variance. In this chapter and elsewhere concepts are explained in a manner which makes them accessible to the non-mathematician.

Chapter 3 discusses the statistics of two variables. Correlation coefficients are explained, together with the problems of induced correlations in closed systems. Regression in its various forms follows with numerical examples taken from the fields of palaeontology, petroleum geology and geochemistry. Non-parametric statistics is the subject of the next chapter (Chapter 4). Worked examples abound, and these are placed in text boxes where all working stages are to be found. Even in these days of the computer, students need to have performed the calculations by hand at least once.

The chapter on Directional Data and Circular Statistics will be of particular interest to structural geologists. Many will be disappointed that three-dimensional orientation data receives only brief mention. In a book that is otherwise thorough in its coverage of ways to represent data, the stereogram, and the statistical questions it raises, is conspicuous by its absence.

The chapter on Data Through Time will appeal to stratigraphers, since it deals with the statistical analysis of data consisting of two variables, one being time. Here, and elsewhere in the book, occasional mention is made of specific computer programs which can do the necessary number crunching, but no particular packages are promoted. This is a wise policy on the part of the authors; such programs will come and go whereas most of the content of the book will remain relevant to geologists for many years to come.

The following chapter, entitled Geographically Distributed Data, gives an excellent guide to the representation of three or more variables simultaneously, by means of block diagrams and contour maps. Valuable insight is gained into how contour lines are drawn and how spurious contour patterns can arise. Kriging is briefly mentioned, as is the topic of fractals. The final chapter follows naturally on from this, and is concerned with Multivariate Analysis. The necessary mathematics (vectors, matrices, eigenvalues) are provided in an appendix. Another appendix provides handy statistical tables, making this textbook self-contained.

As you have probably already gathered, I like this book. I like it for the way it handles tricky statistical concepts in clear and simple language, for its comprehensive coverage, for its provision of plentiful worked examples and for the practical index. I would recommend this book to students and researchers of all ages. The book is the ideal companion to many geological software manuals, and is much better written than most.

Richard J. Lisle

Cardiff, Wales