reasonably well fitting model will do. The author gives us no evidence by which to assess this seemingly philistine statement.

There are other shortcomings. I wish more had been made of the connections between Spjøtvoll's test, the mean-squared error of prediction, AIC, C_p and their generalizations. For example, some of the published work on the consistency of generalized Akaike information criteria provides guidance on the choice of the parameter in model selection procedures using generalizations of C_p.

There is a very dramatic figure on p. 148 of the difference between least squares and conditional maximum likelihood estimators for an example in which the first of two variables is selected. But the author admits that the changes are largest in the region where both variables would be selected, so this reader (at least) is left wondering whether the appreciable improvements achieved depend on the use of an unsatisfactory model selection procedure.

I hope that I have said enough to indicate that this is not one of the more distinguished additions to the Chapman and Hall monograph series. However, the subject is interesting and of practical relevance. I hope that the book may stimulate other statisticians to solve some of the problems suggested by Dr Miller's work.

Anthony Atkinson
London School of Economics and Political Science

9. Optimal Scaling of Time Series. By S. Van Buuren. ISBN 90 6695 040 3. DSWO, Leiden, 1990. 200 pp.

This attractive and deceptively slim volume presents a mathematically detailed account of the techniques necessary to analyse categorical time series data by using conventional time series and multivariate methods. Many social science data sets consist of (often multivariate) observations on categorical variables, where the actual size and relative ordering of the values may not be important. Some modifications are needed before we can apply conventional statistical techniques which rely on scale and ordering. This book outlines many relevant statistical techniques in a lucid and relaxed style, using some non-standard notation and peculiar English, but nevertheless in sufficient detail to be useful to a social scientist who is not already familiar with them (though preferably with a strong mathematical background). A few nice pictorial ideas appear, such as the three-dimensional 'correlation box' and scatter (trajectory) plots where sequential dots are joined. These are presumably standard exploratory tools, but the graphical presentation makes them appealing, even if their usefulness is not immediately clear.

The main contents of this work are the extensions of some time series and multivariate methods to incorporate the so-called optimal scaling of the categorical variables involved. Optimal here means with respect to a certain loss function, which in this case is one familiar and local to the author (Gifi, 1981) and is just the sum of squared differences of the values assigned to the categories from target values (which are defined by prior beliefs, constraints to ensure identifiability etc.). It is shown that the solution to this optimization can be incorporated into that already used for certain techniques and the necessary algorithms are given in the appendix. For instance, in a first-order autoregression $x_t=\phi x_{t-1}+\epsilon_t$, we seek a coefficient ϕ to minimize $\Sigma\epsilon_t^2$. With optimal scaling of the x_t (which come from, say, k categories), we seek ϕ and $\mathbf{y}$ (a $(k\times 1)$-vector of category loadings—subject to certain normalization constraints) to minimize the same sum of squared residuals, where now $x_t=\mathbf{i}_t^{\mathrm{T}}\mathbf{y}$ ($\mathbf{i}_t$ is an indicator vector with a 1 in the position corresponding to the category in which x_t lies and 0s elsewhere). This is a much larger optimization problem and has not yet been solved for general autoregressive moving average models. In addition, the scaling alters some useful model specification tools (such as the autocorrelation function), so that certain standard methods of model identification cannot be applied. A few interesting and persuasive examples are presented including one of raw psychological data from which some sensible conclusions are drawn. Another uses Box and Jenkins's series D—310 chemical process readings measured on a numerical scale. It may seem perverse to force this into a categorical scale, but the observations on this variable happen to fall into a small number (26) of discrete values—a histogram shows that most of the readings are from the middle or right of the range. The resulting optimal scaling applied is a monotonic increasing transformation which maximizes the lag 1 autocorrelation and brings the extreme values closer to the centre of the range (while preserving the ordering). This appears to be a useful tool to linearize data and to control outliers.

To conclude, this is a very attractive and interesting book, full of a wealth of mathematics necessary for the statistical analysis of some social science data. It is by no means the last word on the matter—the author recognizes and points out many of the failings and directions for further research and I am sure that Dr Van Buuren will have more to say in the future. The book should appeal to anyone with an interest in applying time series methods to more than just real number data sets or even just in novel ways of looking at such 'standard' data.

Reference

Gifi, A. (1981) *Nonlinear Multivariate Analysis*. Department of Data Theory, University of Leiden.

Howard Grubb
University of Bath