Computer Software Reviews

J. Am. Chem. Soc., Vol. 110, No. 13, 1988 4477

Computer Software Reviews

MicroModeller. By I. Salmon (University of Kent). IBM-PC/XT/AT (128K or 256K). US $150.00. January 1987. ISBN 0-947946-94-2.

Micromodeller is a program for graphically demonstrating kinetic models of biological and chemical reactions. It allows for the entry of the constants, differential equations, and mathematical relationships or measurable parameters that describe the model and then calculates concentrations of the time-dependent variables at user-defined intervals.

It is available for the IBM PC and compatables and requires a color graphics adapter. It contains no direct provision for printing graphical results; however, these results can be screen printed if an appropriate driver and printer are available.

The program is written in Turbo Pascal and although the version used for this review came on a copy-protected disk, the most recent release is unprotected and is sold under license. The list price covers one installation, but problems associated with copy protection have been eliminated.

This program is designed as a teaching tool with limited applications in research. It provides for the graphical presentation of kinetic data based on a user-defined model and is capable of generating data for demonstration purposes and class analysis. It does a good job of performing this function; however, the entry format is relatively clumsy and not explicitly clear at times. The experienced modeller should have no trouble but a student, unfamiliar with the terminology and techniques of setting up differential equations, would have some difficulty in benefitting from use of the program.

As a research tool it is not very useful. Although data for a model could be generated and compared to experimental data this is relatively useless since the constants used to generate the data must be derived from experimental results based on a specific kinetic model.

The version of the program used for this review does an adequate job of graphically representing the effect of modifying kinetic parameters in a defined model. A more recent release has a much improved graphic output. I found this program more useful from the teacher's standpoint, providing a means for display and manipulation of predefined models, but feel it requires to much knowledge of modelling techniques and procedures to be used effectively by students.

Lawrence Poulsen, *University of Texas at Austin*

Statistix. Version 1.1. NH Analytical Software: 1958 Eldridge West, Roseville, MN 55113. List Price $75.00.

Statistix is a statistical analysis and data manipulation system for microcomputers. Five different versions of the program are available: for the IBM PC/XT/AT series and compatibles (either with or without a math coprocessor), for other MS-DOS machines, and for the Apple II+/e/c series (either 64K or 128K of memory). Each program sells for $75.00. The version tested and reported here is the IBM PC version without the coprocessor option. The product literature states that the major differences between the various versions are (1) speed and (2) the number of data items permitted. For the MS-DOS version, machines with 256K will hold about 30000 data items. For the Apple machines, the 128K version will hold about 8000 data items and the 64K version will allow 3000 data items. The Apple versions run under either Apple DOS or Apple Pascal, but you do not need Apple Pascal to run the program. The program requires neither graphics monitors nor printers, although a printer is very helpful during operation. The PC-DOS version requires either two 5.25 in. floppy disk drives or a hard disk; it will also run from a RAM disk. None of the versions are copy-protected.

This program is a comprehensive statistical analysis package emphasizing sound statistical principles. It delivers nearly everything most chemists would want in a statistics software. It is modular in nature and has sections devoted to Data Management, File Management, Summary Statistics (including Scatter Plots, Frequency Distributions, and Crosstabulations), Probability Functions (Binomial, Poisson, Chisquare, and others), Goodness-of-Fit Association Tests (Chi-square test, Spearman Rank Correlation, and more), One-, Two-, and Multi-sample Tests, Randomness/Normality tests, and Linear Models (Multiple and All-subset regressions, Analysis of Variance, Principal Components, plus others). The Linear Models Section has full residuals analysis functions and Analysis of Variance options.

Statistix is claimed to be remarkably easy to use, and indeed, it is. The functions are called from a series of menus, with each choice requiring a single-letter keystroke. All of the expected data manipulation functions are present, and the program can import data from ASCII files generated by other programs (spreadsheets, etc.), as well as read its own internal data format. The users manual is very complete, without giving a course in statistics. It treats each topic in enough depth that even non-statistician users (such as this reviewer) should have little trouble understanding when a particular test can (or cannot) be used. Extensive reference is made to the statistical literature for the algorithms used, and to current books for those who desire to read further on any topic.

The literature accompanying the package succinctly states the niche the authors have intended for the package. "[It] is a very easy-to-use system for general analyses of moderate size data sets. It is not: 1. a number-cruncher for large multivariate problems, 2. a programming language, 3. a graphics program, 4. a batch system, 5. the answer to all statistical computing needs, 6. a hacker's delight, or 7. a solution to poorly designed experiments." These claims are generally fair ones, and on the basis of its power and ease of use, I would enthusiastically recommend it for use within the above constraints.

One particularly useful feature of the package is its ability to easily perform transformations. Using this feature one can generate other variables from the raw data that was originally entered. For example, kinetic data may be entered as time and percent remaining then be transformed to molar concentrations and finally to the logarithms of molar concentration. This transformation is performed in "spreadsheet" fashion, with a new variable generated which corresponds to each old variable. This feature is used to perform the extensive residuals analysis functions, with any of several quantities being assigned as new variables. These can then be subjected to further analysis.

As with any software package, one can suggest improvements. With this particular package I found myself wishing it could label cases (records) with an alphanumeric label. This would be particularly useful for QSAR analysis of biological data and associated physicochemical parameters, where one would like to assign a functional group identifier to each record in the data set. According to the company, there are no plans to do this, but they took the suggestion under advisement. Also, the program does not have the ability to continuously update transformation variables (those defined by other variables). This becomes inconvenient when performing multiple regressions. If a data point is omitted on the basis of residuals analysis, the residuals variables are not updated when a new regression expression is computed. The interactive nature of the package forces one to continuously retype all the keystrokes necessary to redefine the variables derived from an analysis whenever a data point is omitted from the analysis. When I inquired about this problem, the authors explained that implementing this change would complicate other aspects of the operation of the program. They did, however, suggest a solution. By using a keystroke macro generator program a series of keystrokes can be assigned to a single key and this could effectively accomplish the same objective. This suggestion proved to be invaluable, and I strongly recommend that the package be used in conjunction with such a program.

In summary, the package is fast, useful, and performs well within the claims made for it. User support was prompt and helpful. When price is considered, it seems to be an excellent value.

NOTE: Version 2.0 is now being provided by the vendor at $169 (Apple II version $99). The vendor has indicated that the package will do time series identification and estimation, stem leaf and box plots, and full screen data entry and editing.

James G. Henkel, *University of Connecticut*