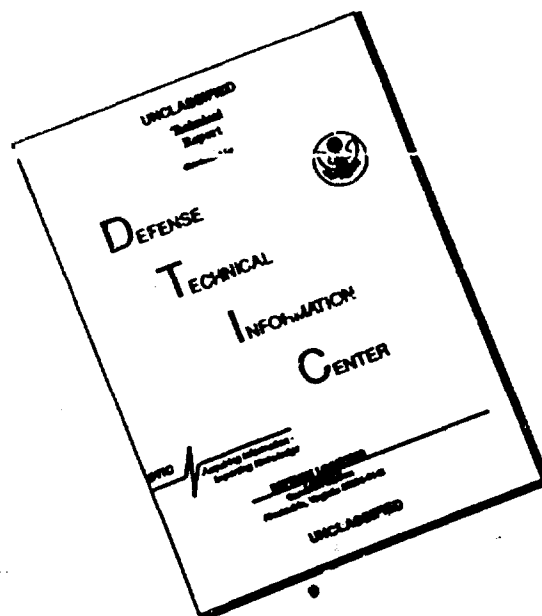


DISCLAIMER NOTICE



**THIS DOCUMENT IS BEST
QUALITY AVAILABLE. THE COPY
FURNISHED TO DTIC CONTAINED
A SIGNIFICANT NUMBER OF
PAGES WHICH DO NOT
REPRODUCE LEGIBLY.**

THE USE OF SPREADSHEETS IN CHEMICAL LABORATORIES

R.A. Coad, DSTO, Materials Research Laboratory,
PO Box 147, Scottsdale, Tasmania, 7260, Australia.

ELECTE
AUG 23 1993

Abstract

The use of standard spreadsheets in chemical laboratories is examined. The main benefits offered by such programs are ease and accuracy of data manipulation and accuracy of record keeping. Data may be entered manually or data acquisition software may be utilised. Spreadsheets can be introduced to the laboratory without the need for intensive training courses. A spreadsheet is defined and a brief history of the subject is presented. The application of these programs to chemical laboratories is discussed with respect to the benefits to the laboratory, quality assurance, data and formulae entry, data acquisition, calculations, graphics and purchasing guidelines.

Introduction

Prior to the advent of commercially available electronic spreadsheet programs in 1980, calculation of results, presentation of raw data and results, checking of calculations and information storage were usually separate manual tasks. This was time consuming, laborious and prone to human error. Spreadsheet programs provide a quicker, more accurate means of data manipulation.

MRL-Tasmania forms the Food Science Branch of the Protective Chemistry Division of the Materials Research Laboratory, a part of the Defence Science and Technology Organisation within the Commonwealth Department of Defence. The laboratory makes extensive use of electronic spreadsheets for the calculation of results, collation and presentation of raw and final data, storage of data and formulae used for calculations, preparation of presentation graphics and as a convenient aid to quality assurance. Spreadsheets have been created for moisture, salt, protein, fat, energy, carbohydrate, metals and niacin determinations.

Definition

An electronic spreadsheet may be

defined as an array of individually accessible cells which may contain data, alphanumeric labels, "macro" commands or formulae. Formulae are constructed from mathematical operators, constants, references to other cells and built in functions, including those for scientific and statistical calculations (1). Data is entered via a keyboard or read from a file created using data acquisition software.

History

In 1980 an electronic spreadsheet program called *VisiCalc* (Software Arts) became available for microcomputers. The program recreated the accountant's spreadsheet on the computer screen. It had many applications in the sciences, including chemical calculations. Imitations followed and one, *Supercalc*, developed by Sorcim, met with market success. This program was cheaper than *VisiCalc* and improved on some of its features.

In 1983 Lotus Development Corporation released *Lotus 1-2-3*. For the first time the three features, electronic spreadsheet, business graphics, and database management, were combined into one package (2). It was developed to run on the IBM PC and other MS-DOS hardware. *Lotus 1-2-3* and its revisions have continued to be very popular, particularly in large corporations.

State of the art features include multiple spreadsheets, high resolution graphics, interfacing with word processing, import of automatic data acquisition files, data base managers and dynamic data exchange (DDE). Multiple spreadsheets and spreadsheet linking is a feature being marketed by many software companies. *Lotus 1-2-3 Release 3* (Lotus Development Corporation, 1989) allows three dimensional spreadsheeting by permitting multiple spreadsheets in memory at the same time, multiple spreadsheets per file, and links between spreadsheets, even if they reside in different files (3). Worksheet linking is one of *Quattro Professional's* (Borland, 1989) most important features. Up to 32 worksheets can be consolidated and

Moisture Determination Date	
Filename	
Weighted by: Ross Coad	14/1/91
moist14.jan	
Calculated by: Ross Coad	16/1/91
Checked by:	

Figure 1 The spreadsheet header should include information to ensure traceability.

even put on the screen (4).

Most spreadsheeting software will export files in ASCII format acceptable to most word processors. *PlanPerfect 5.0* (WordPerfect Corporation, 1989) offers word processor file compatibility and will import and export WordPerfect 4.2 and 5.0 files. The ability to read data files from other software is a feature expected of spreadsheet packages. *20/20* (Access Technology, 1989) reads, writes and consolidates Excel and Lotus 1-2-3 WKS and WK1 files into *20/20* and vice versa (5).

Within the Windows environment, Microsoft's *Dynamic Data Exchange* makes it possible to share information between spreadsheet and database applications (6). Once the link has been made between a spreadsheet and a copy of that spreadsheet in a word processing document any changes made in the source spreadsheet will be updated in the word processing copy. Another Windows capability, *Object Linking and Embedding* (OLE), allows the spreadsheet to be run from within the word processing document (7).

Benefits

The benefits gained from using spreadsheets will depend largely on the suitability of the package to the application. Benefits obtained should at least include speed of data manipulation, ease of accurate record keeping and traceability of results.

Prior to the development of electronic spreadsheets data was obtained from instrumentation via a printout or by visually reading the display. This was then transcribed to a result book, calculations performed using a calculator, results plotted if necessary and a report

Approved for public release
Distribution Unlimited

REPORT DOCUMENTATION PAGE			
1 AGENCY USE ONLY		2 REPORT DATE	
4 TITLE AND SUBTITLE THE USE OF SPREADSHEETS IN CHEMICAL LABORATORIES.		3 TYPE/DATES COVERED	
6 AUTHOR R A COAD		5 FUNDING NUMBERS	
7 PERFORMING ORG NAMES/ADDRESSES DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION, MATERIALS RESEARCH LABORATORY, P.O. BOX 147, SCOTTSDALE, TASMANIA, 7260 AUSTRALIA		8 PERFORMING ORG. REPORT NO	
9 SPONSORING/MONITORING AGENCY NAMES AND ADDRESSES			
11 SUPPLEMENTARY NOTES			
12 DISTRIBUTION/AVAILABILITY STATEMENT UNCLASSIFIED <i>Unlimited A</i>		12B DISTRIBUTION CODE	
13 ABSTRACT (MAX 200 WORDS) THE USE OF STANDARD SPREADSHEETS IN CHEMICAL LABORATORIES IS EXAMINED. THE MAIN BENEFITS OFFERED BY SUCH PROGRAMS ARE EASE AND ACCURACY OF DATA MANIPULATION AND ACCURACY OF RECORD KEEPING. DATA MAY BE ENTERED MANUALLY OR DATA ACQUISITION SOFTWARE MAY BE UTILISED. SPREADSHEETS CAN BE INTRODUCED TO THE LABORATORY WITHOUT ANY NEED FOR INTENSIVE TRAINING COURSES. A SPREADSHEET IS DEFINED AND A BRIEF HISTORY OF THE SUBJECT IS PRESENTED. THE APPLICATION OF THESE PROGRAMS TO CHEMICAL LABORATORIES IS DISCUSSED WITH RESPECT TO THE BENEFITS TO THE LABORATORY, QUALITY ASSURANCE, DATA AND FORMULAE ENTRY, DATA ACQUISITION, CALCULATIONS, GRAPHICS AND PURCHASING GUIDELINES.			
14 SUBJECT TERMS		15 NUMBER OF PAGES 3	
		16 PRICE CODE	
17 SECURITY CLASS. REPORT UNCLASSIFIED	18 SEC CLASS PAGE	19 SEC CLASS ABST.	20 LIMITATION OF ABSTRACT UL

DTIC QUALITY INSPECTED 3

Accession For	
NTIS CRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	
By _____	
Distribution /	
Availability Codes	
Dist	Avail and/or Special
A-1	

generated. The results would be approved by the laboratory manager, or his delegate, after checking the figures.

Some laboratories had been able to improve their data manipulation and presentation using custom written software or programmable calculators, but the former was expensive, and therefore available to only the larger laboratories, and the latter did little more than speed up the calculations.

Spreadsheet programs provide a quick and accurate means of data manipulation. A single keystroke may be used to evaluate all formulae; each cell

required.

Spreadsheets also offer benefits in terms of checking of results and archiving; these are discussed in the section on quality assurance.

Automatic data acquisition software linked to a spreadsheet eliminates operator error in reading and entering data, drawing standard curves and plotting results. It saves time normally spent on these tasks and producing a final report.

The Food Science Section, MRL-Tasmania, as a National Association of Testing Authorities (NATA), registered

a title describing the calculations performed by the spreadsheet.

- Filename; used to record the spreadsheet data.
- Weighed by; name of the person who weighed the sample for the analysis or for a particular parameter of the analysis. Separate entries must be made for different parameters if they involve different people.
- Weighed date; which is the date the sample was weighed out for the analysis. Separate entries must be made if more than one person, or more than one date, is involved.

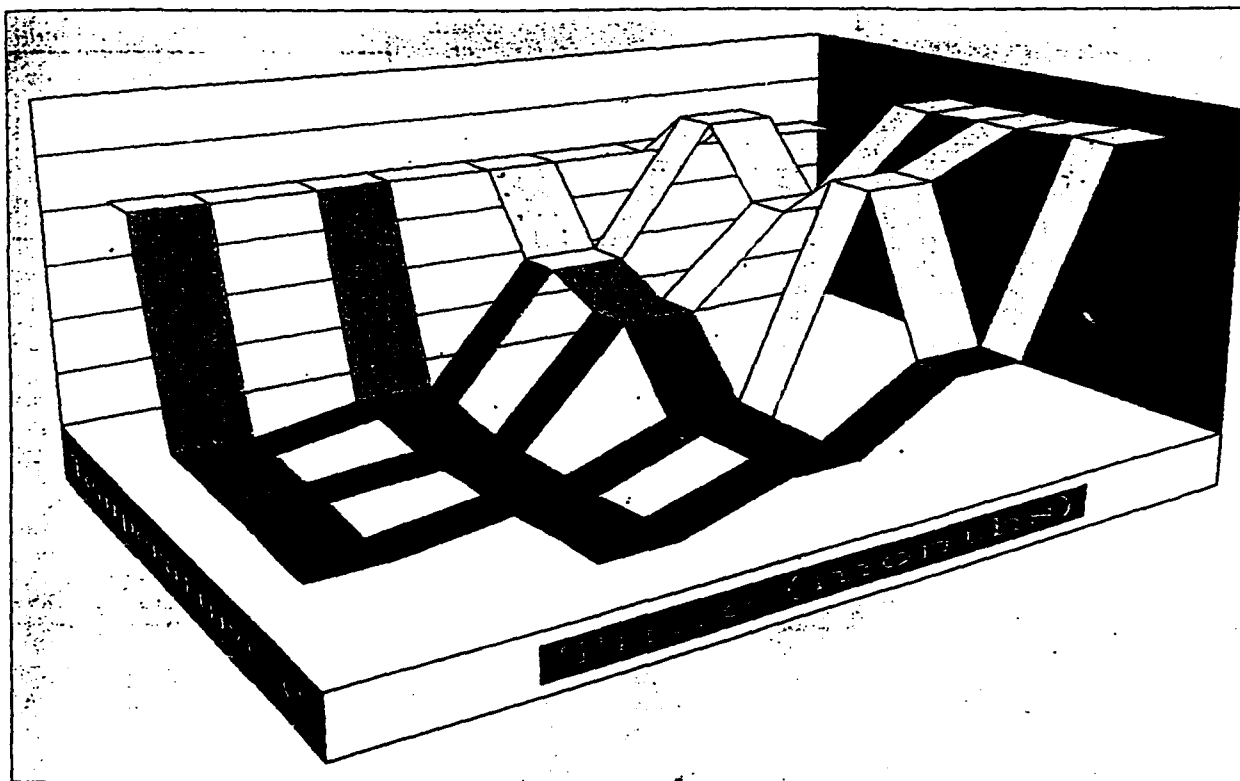


Figure 2 Three-dimensional graphics are possible with some spreadsheeting packages.

does not need to be calculated individually.

Operators without programming experience can manipulate chemical data and perform repetitive calculations. Most modern packages include on-line help and a tutorial, usually based on an accounting application. Therefore spreadsheets can be introduced to the laboratory without the necessity for intensive training courses.

The ability to tailor each spreadsheet, firstly to the specific requirements of a particular analysis, and secondly to the more particular requirements of an individual run, is a useful feature. It offers the user a combination of fixed and flexible formats. A master spreadsheet may be set up and used for protein determinations for example. In individual cases the spreadsheet may be altered if a different number of standards or blanks are used or if a different protein factor is

laboratory, is required to maintain accurate analysis records, including people and dates involved, which can be traced back to the raw data. This can be effectively achieved by the implementation of spreadsheet quality assurance procedures.

Quality assurance

Quality assurance of spreadsheets is particularly important if maximum benefit is to be obtained from this technology. This may be achieved by ensuring that a standard format is adopted for setting out spreadsheet information, master sheets are protected from overwriting and spreadsheet checking procedures are followed. MRL-Tasmania's Laboratory Quality Manual requires that the following information be included in the spreadsheet header to provide a clear audit trail (Figure 1):

- Spreadsheet application name; which is

- Calculated by; name of the person who entered the data into the spreadsheet and undertook the calculations.
- Calculated date; which is the date on which calculations were done.
- Checked by; name of the person who checked the accuracy of the entries and formulae used for the calculations.
- Checked date; which is the date the check was made.
- Filename for other data; a cross reference must be included to any other spreadsheets used to provide data for the current spreadsheet.

All spreadsheets must include details of calibration calculations, sample number and name for each calculation. Constants which are to be used in calculating results should be listed in the header. Log books must reference any spreadsheet file used for calculations.

A master spreadsheet can be set up with the necessary formulae which are

verified against a manual calculation before release to operators. Data can then be entered for statistical and arithmetical calculations to be performed. A procedure should be established whereby alterations to formulae may only be made with authorisation. Individual cells can be locked so that formulae cannot be altered accidentally. In addition, the integrity of the master spreadsheet may be safeguarded by marking it as a "read only" file in MS-DOS which cannot be accidentally overwritten by saving another spreadsheet in the same name.

The spreadsheet checking procedure should involve a person other than the analyst. Checking spreadsheets consists of ensuring that the raw data was accurately entered, the format has not been changed, all name and date fields have been filled in and formulae are correct. There is no need to laboriously recalculate individual figures. A procedure which may have taken an hour can now be completed in five or ten minutes.

A spreadsheet file may be archived once all checks have been completed and the results released. Archiving has the usual advantages inherent in diskette storage, namely space saving and ease of retrieval. NATA requires all data to be retained by the laboratory for five years. The ability to retrieve files at a later date and manipulate the data in those files is of particular value if a critical error or long term trend is observed.

Data and formula entry

The entry of labels, formulae and data varies amongst the various programs. Older spreadsheet programs require labels or text to be enclosed in inverted commas. Formulae need to be identified to distinguish them from text and data. This is usually achieved by an identifying keystroke, usually a mathematical operator, immediately prior to entering the formula. Data can usually be entered by typing in the figure followed by either the "enter" key or an arrow key; the latter will simultaneously move the cursor. Movement about the spreadsheet is achieved by use of the four arrow keys and commands which specify particular locations on the spreadsheet. Methods of viewing spreadsheets vary according to the quality of the monitor and the software.

A macro may be used to automate tasks which are performed repetitively. It is a set of instructions in the form of a sequence of keystrokes and commands. A macro is usually identified by a one word name which, when invoked, causes a set of instructions to be carried out. The advantage of macros is that time and effort is saved by using one command to replace a series of commands. They may be used to set up worksheet defaults in a certain way,

prompt for input, handle errors and other applications.

Calculations and graphics

Before a calculation can be performed the formula for that calculation must be entered into the cell in which the result is to be displayed. The formula uses labels to identify cells containing data upon which the calculation is to be performed. The calculation process is initiated by a single keystroke. Formulae may be copied to a range of cells so that a range of data points may be calculated without having to type the formula into each cell. When using a constant its value should be entered into a cell and that cell identified by its location or reference. Absolute cell referencing is usually used to identify cells containing constants, because the formula can then be copied without the constant's cell reference changing.

Graphics functions may be used to produce histograms, pie charts or graphs from spreadsheet information. Three dimensional graphics (Figure 2) are available with some packages. Resolution, labelling and the ability to display error bars are areas where programs vary in their capabilities and each new version offers improvements.

Data acquisition software

There is a large range of these programs designed to be run on various instruments and interfacing with a range of spreadsheets. **Labtech Notebook** (Laboratory Technologies Corporation, 1984) is a data acquisition software package which has been designed to interface with programs such as **Lotus 1-2-3** (8). **Lotus Measure** (Lotus Development Corporation, 1987) collects data from instruments and stores the information directly in a **Lotus 1-2-3** or **Lotus Symphony** spreadsheet for analysis, storage, and display (9).

Software can be developed for specific applications. For example, an Apple II based data acquisition system has been developed for use with an autoanalyser used for the chemical analysis of soil, plant and water samples. The system provided fully automated data capture, computer data manipulation, preparation of printed report sheets and automatic data storage on magnetic disk (10).

At MRL-Tasmania chromatography data is acquired using the **Nelson 2600 Series Chromatography Data System** (Nelson Analytical, 1986). This data can then be converted to a form which can be read and spreadsheeted by the **Enable System** (The Software Group, 1986).

Purchasing guidelines

There is tremendous variation in the range of spreadsheets currently available. They vary in their ability to carry out calculations, macro commands, ease of use, compatibility with hardware and, of course, price. The most expensive

program is not necessarily the best for a particular application and there is little point in purchasing a program which has capabilities far in excess of those required.

The decision to purchase must be based on a realistic confidence in the ability of the software to perform as required. When considering the purchase of a spreadsheet program it is important to establish selection criteria; firstly to identify the desirable features and secondly to ensure that the spreadsheet eventually purchased will perform as expected.

If there is a need to display three dimensional data, for example, the selection criteria should include the ability to produce 3D graphics. Another selection criterion should relate to the compatibility of software and hardware. This can be checked with the company marketing the software. Some offer a demonstration version for a prospective buyer to evaluate.

Once the required capabilities and compatibility of the package have been identified a decision must be made on how much money one is willing to spend. It is advisable to obtain some pricing prior to making this decision unless one is already well informed in these matters.

Conclusions

Spreadsheets offer significant benefits to the laboratory and may be introduced without difficulty. Good quality assurance of laboratory data and results can be maintained if the laboratory manager develops clear guidelines for the creation of a spreadsheet, maintenance of master spreadsheets and provision of an audit trail through recording of source documentation or files. Careful selection of a spreadsheet package will ensure satisfaction with its performance.

References

- (1) Filby, W.G., Bauer, B., Kahus, W.H. and Weinstein, K. (1988). "Kinetics Data Acquisition Using a Low-Cost Personal Computer and a Commercial Spreadsheet 'Add-In' Program", *Intelligent Instruments & Computers*, September/October, 281-284.
- (2) Bimes, W.J. (ed.) (1985). *Personal Computer Programming Encyclopedia*. McGraw-Hill Book Company, New York, 443.
- (3) Davies, I. (1989). "Lotus 1-2-3 Release 3: the wait is over", *Australian Personal Computer*, July, 55-72.
- (4) Dubash, M. (1989). "Spreadsheet line-up", *Australian Personal Computer*, November, 129-145.
- (5) Anon. (1990). *20/20 by Access Technology*. Access Technology, Natick, MA.
- (6) Schatzman, B. (1991). "Linking spreadsheets with databases; how the Windows DDE links spreadsheets to databases, and a look at how SQLVision makes the link. (Dynamic Data Exchange)", *DBMS*, June, 60-64.
- (7) Radding, A. (1991). "PC Spreadsheets", *Computerworld*, September, 69-70.
- (8) Anon. (1984). *Labtech Notebook*, Laboratory Technologies Corporation.
- (9) Anon. (1987). *Lotus Measure*, Lotus Development Corporation.
- (10) Winkelman, G.E. and Leyshon, A.J. (1990). "An Apple II-Based Data Acquisition System for the Technicon AutoAnalyzer", *Intelligent Instruments & Computers*, September/October 211-219.