Microcomputer Use in Experimental Aquaculture: What Software is Needed?

J. MICHAEL VAKILY
Institut für Meereskunde
an der Universität Kiel
Düsternbrooker Weg 20
D-2300 Kiel 1
Federal Republic of Germany

The first two articles of this series dealt with hardware requirements to be considered by potential users of microcomputers (Vakily 1988, 1989). This contribution will discuss some general aspects of selecting appropriate software.

Compared to other technical aspects of experimental aquaculture, decisions on appropriate software are difficult. If information was needed to set up e.g. a larval rearing unit, one would go to a library or a book store, look through a couple of relevant publications, check the contents and the form of presentation, and then decide which one to buy.

Obtaining information on software is much more tedious. The best solution, of course, would be the possibility to "test-drive" a range of programs before making a final decision. However, test driving takes a lot of time, particularly if several packages are to be compared; e.g. Lotus 1-2-3, Multiplan, Supercalc, etc. This opportunity might best be given on a general computer training course. For more information, see Vakily (1987).

Browsing through computer magazines or software manuals is either frustrating (especially if you are not familiar with "Computerese") or of little use because the examples given in these publications usually apply to problem solving tailored to the business world. Newcomers to the world of bits and bytes should not hesitate to contact computer-experienced colleagues or friends for more information. By explaining as precisely as possible the tasks where assistance from a computer is expected, the chances are good to find somebody who went through very similar problems and has solved it with the appropriate software.

This article is a short overview of types of software and presents some general thoughts related to strategies in data management, with emphasis on software integration. A quite extensive list of software products with a short description of their features details of aquaculture-related software was published in the April 1987 issue of NAGA (obtainable from ICLARM): an issue entirely devoted to microcomputer use in fisheries and aquaculture. It is hoped that the present article will encourage readers to use Aquabyte as their forum to present personal experiences with software applications in aquaculture management and research.

Software can be roughly divided into the following three categories: system software; programming tools; and application software.

System software is of interest to the researcher only insofar that it is needed to start the computer and to do the "house keeping". New and promising system software already exists on the market (OS/2, UNIX), but it can be safely assumed, that DOS will stay with us for some time, as it has become the standard for PC systems.

Programming tools are software packages that allow development, testing and implementation of custom-made programs. Various high-level programming languages exist (BASIC, C, Pascal, Cobol, etc.). They differ in degree of complexity and in algorithms and functions already built in, giving each of them their own advantages in specific areas of application.

Application software comprises custom-tailored programs written in a high-level programming language as well as the vast amount of general purpose software packages that can be bought off-the-shelf.

Custom-tailored programs, as stand-alone software, are written for a specific purpose and might handle within predefined limits all necessary steps from recording data to their analysis and the presentation of results. Usually, they are completely menu-driven, allowing efficient use of their features without any specific computer knowledge. A disadvantage is their limited flexibility. Adaptation of the program structure to changes in the general working set-up might prove either impossible or difficult enough to require the assistance of an experienced programmer.

Typically, custom-tailored programs are of great interest in the management of highly developed aquaculture production systems, where they are used for the implementation of optimization techniques in stocking, feeding and harvesting. The anticipated reduction of production costs and the capability of producing an aquaculture product on a timely basis according to the needs of the target group can, indeed, pay off the initial costs of developing such a program. In reality, though, custom-tailored programs with

their rigidity in data requirements seem of only little use in most aquaculture projects (especially if these have substantial research components).

The second category of application software, comprises the well-known general purpose software packages developed by commercial software houses. From a researcher's point of view, many of these programs can be considered sophisticated "tool boxes" that offer great flexibility in data management in experienced hands. The principal areas of application for these tools are: collection, storage, and analysis of data, and presentation of results. In contrast to the custom-tailored programs mentioned above, however, they require the initial investment of a substantial amount of time to study manuals and tutorials. As with any sophisticated tool, it is only experience that gives the maximum benefits of use.

Figure 1 gives a schematic overview of how application software can be used for different tasks in data management. The arrows indicate the routes that data might take within the process of managing the flow of information from the collection of data to the presentation of final results. It should be noted that, especially on the level of data storage, this data flow is not necessarily one-way, but can lead, by means of software integration, to a free exchange of information between programs.

In many research projects, **data collection** is done by field observations which are recorded on paper and later entered manually into the computer. However, experiments might require the frequent reading of a number of variables over an extended period of time, e.g. environmental parameters in fishponds (temperature, DO, pH, unionized ammonia, etc.). In such circumstances DA&C systems (Data Acquisition and Control systems) come in very handy.

DA&C systems consist of the appropriate hardware for measuring and recording the variables and of software that communicates the data to a computer. Two different set-ups exist:

1. The monitoring unit is housed in its own case together with some device to store the incoming data. These data are then later transferred to a host computer through its serial port by means of a communication program.

2. Sensors feed an interface board directly installed inside the computer's case, which has the advantage of almost instant access to the incoming data.

Whether a remote monitoring unit is used or not depends largely on the circumstances. Whereas a computer can easily control instruments in a laboratory, few people want to see their machines being set up on a table somewhere in a field between ponds.

Before data can be analyzed by a computer, they have to be stored either on floppy diskettes or (if available) on the computer's hard disk, in order to be accessible. With respect to **data storage**, one should adhere to the rule that primary data are to be entered **only once**. Repetitious entering of data increases the chances of typing errors. Therefore software is needed that creates data files that can be read (or translated) by all other programs that later may possibly be used for specific analysis routines.

The idea of using the same data files for different purposes has been especially promoted with the advent of the so-called "integrated packages": software that offers under one "roof" capabilities for data management, communication, graphics and word processing capabilities. These programs, such as Lotus Symphony or Framework, are popular because of the relative ease with which one can access the different parts of the program. One should, however, keep in mind that some parts of these programs necessarily appear limited compared to full-fledged programs developed for just one specific purpose. As a consequence, most of the recent software offers now built-in utilities to use or transform common external data files, allowing software integration without compromising on available features. It should also be emphasized that most software allows data to be stored in a common format (e.g. ASCII files for text files, DIF files for spreadsheets, and PIC or META files for graphics) accessible by the others without the need for data transformations. Moreover, there exists memory resident software, e.g., "Sidekick" (for text) or "HiJaak" (for graphics) which can capture part or all of a screen for including in other programs, be it directly (text) or through file conversion (graphics).

Among general purpose programs there are two types of software frequently used to store data: spreadsheet and database programs. Spreadsheet programs store their data in cells arranged in rows and columns. Cells can be filled either with text, numbers or formulas that compute results in dependence of the contents of other cells. Database programs store data in records. Records consist of a series of defined fields containing information related to one key field.

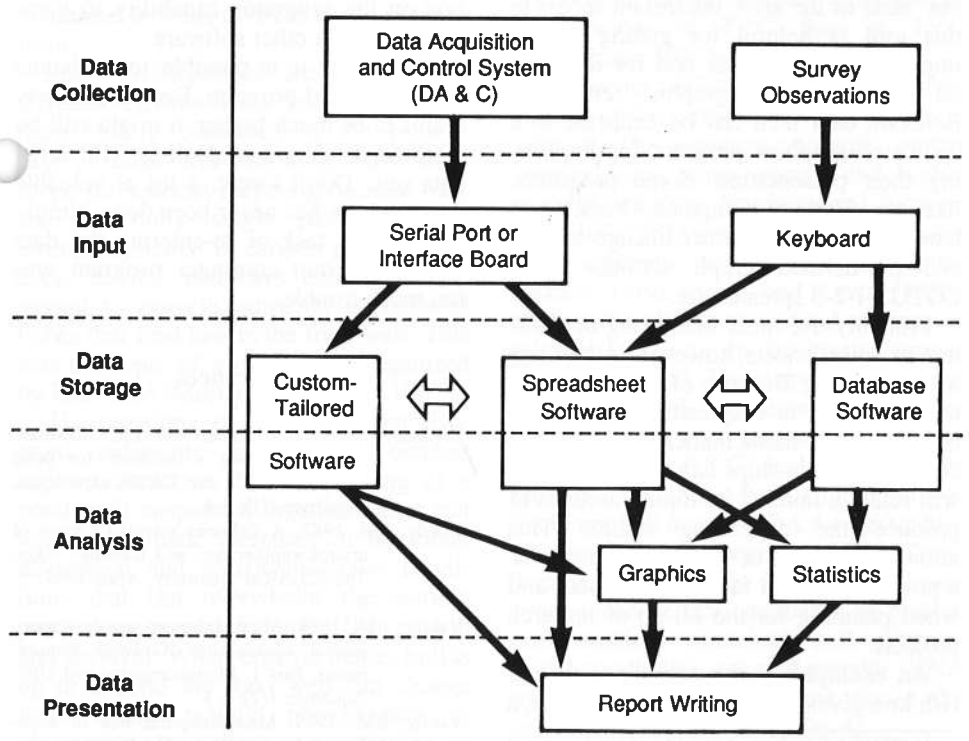

Figure 1. Data flow and software application.

One major difference between these two data storage systems concerns the file handling. Spreadsheet programs have to load the whole file into the computer's memory, before one can work with the data. Database programs access data in their files sequentially. Only that part of the data file that is actually needed is loaded into memory. In spreadsheet programs, therefore, the size of data files is limited by the amount of **internal** memory available, while in database programs the **external** mass storage device (i.e. floppy or hard disk) is the limiting factor.

Though a large number of different programs exist in each category of software, some programs have gained such popularity that they have become a sort of standard in their respective category. In spreadsheet software, LOTUS 1-2-3 is widely used. There are, however, other products such as Framework or Borland's Quattro. LOTUS 1-2-3 facilitates simple data exchange with many other software packages. In database software, Ashton Tate's dBase III+ became very popular but seems to have lost ground to other products such as Data Ease, Reflex or PC File that excel in "user friendliness". A new version, dBase IV, is now available. For organizing data in a large number of different but interrelated databases, Data Ease is very convenient. No doubt newer and better versions of these and other programs will continue to emerge.

One of the clear advantages of spreadsheet programs is that data filing can be done in a fashion very similar to that which most researchers are used to, using pencil and paper. The screen can easily be organized like an ordinary table with headings and explanatory descriptions preceding the areas receiving data input. This is termed "full screen editing". If well organized, a hardcopy of the (empty) table can be used as tabular data sheet in the collection of field data, while the table with the filled-in data can also be printed out and used as primary data tables in the appendix of reports.

As there is no problem in exchanging data between LOTUS 1-2-3 and most database programs, one of the many ways to organize data management could be to use spreadsheet software like LOTUS 1-2-3 as the primary tool for entering data into files, especially if this is to be done by personnel not too familiar with computer applications. Such data could be e.g., regularly sampled farm data. At a later date, these data then can be combined with other data of similar structure in an appropriate database program. This approach was e.g., applied in the CRSP database program for the monitoring of aquaculture pond dynamics (Hopkins et al. 1988).

The abovementioned software also plays an important role in **data analysis**. Whereas database software is very suitable for filtering out subsets of data using specific search criteria, spreadsheet software like LOTUS 1-2-3 has proven extremely helpful when mathematical manipulation of the data is required. The built-in mathematical, logical and statistical functions (including regression analysis) will meet almost all needs of standard data analysis with the option of easily exporting, for example, those data that need further analysis by more refined statistical methods to advanced statistical packages. A relatively simple programming language ("Macros") allow the development of menu-driven spreadsheets for automated data input and analysis. In this respect, spreadsheet programs (as well as database programs) can be utilized like any other programming language to develop custom-tailored applications.

Another handy feature of most spreadsheet programs is their integrated graphics capability, which allows the visualizing of recorded data. Though LOTUS 1-2-3 graphics capabilities are not 'state of the art'[1], the instant access to this tool is helpful for getting a first impression of the data and for deciding on a preliminary graphic setting-up. Relevant data then can be exported to a dedicated graphics program for improving their presentation. Some programs, like e.g. Harvard Graphics (Vers. 2 or later) even allow a direct linkage to data and pre-defined graph settings of a LOTUS 1-2-3 spreadsheet.

Probably the most interesting application of spreadsheets, however, is their use in simulations. Because of the possibility to link a cell in a spreadsheet with any other cell by mathematical expressions, complex relationships can be built up that will relate a number of input variables to produce the depending results. This allows the testing of "what - if" questions, a powerful tool in farm management and when planning for the set-up of research projects.

An example: if the average yield per fish in a given pond system was known, a spreadsheet could be set up to calculate total yield in dependence of stocked number of fish. By varying the input (stocking density), one would obtain resulting total yields. This simple model could then, of course, be extended to include other variables such as feeding rate or rearing time. In an additional step, a variable such as optimum stocking density could be included to allow the simulation of optimization techniques. Principally, modelling by means of a spreadsheet is only limited by the number of variables that can be measured and the ability to formulate their influence on the model in a mathematical expression.

[1] LOTUS 1-2-3, Vers. 3 is said to have improved substantially in this respect, offering now even logarithmic scaling.

Conclusions

If the purchase of a microcomputer system is planned, spreadsheet and database programs should be considered an essential part of the whole arrangement. In addition, the work station should be equipped with appropriate statistics, graphics and word processing software. If applicable, DA&C hardware and software should be added.

When selecting software, it is more advantageous to choose well known "market leaders" to facilitate data exchange between programs. Examination of the chapter on "supported external file formats" in the manuals supplied with the software will give valuable information on the programs capability to communicate with other software.

Sometimes it is possible to exchange data via a third program. Even if this way seems to be much bother, it might still be worthwhile if one is dealing with large data sets. Don't forget: a lot of valuable data analysis has never been done, simply because the task of re-entering the data into a different computer program was too much trouble.

References

Hopkins, K.D., J.E. Lannan and J.R. Bowman. 1988. Managing a database for pond research data - the CRSP experience. Aquabyte 1(1): 3-4.

Vakily, J.M. 1987. A fisheries scientist's view of microcomputer use and training. Naga, The ICLARM Quarterly, April 1987: 5-6.

Vakily, J.M. 1988. Microcomputer use in experimental aquaculture: Hardware requirements, Part 1: Microcomputer and UPS. Aquabyte 1(2): 2-3.

Vakily, J.M. 1989. Microcomputer use in experimental aquaculture: Hardware requirements, Part 2: Printers. Aquabyte 2(1): 2-3.