The SAS® System for personal computers (PCs) allows researchers to use SAS software at locations where mainframe access is difficult or impossible. In combination with electronic data recorders, PC SAS software provides an efficient system for collecting, editing, and analyzing study data. Data can be collected on electronic recorders and transferred by asynchronous communications software to a PC for editing and analysis. By configuring the communications port on the PC appropriately, data can be read directly into the Display Manager of PC SAS software.

Attempts to read observations into the DATA Step from the recorders failed, but if available, this capability would make the SAS System for personal computers even more useful in research data management and analysis.

Introduction

Collecting research data is costly and time consuming. In the past data sheets with carefully labeled columns and rows were prepared manually. Or, data were recorded on makeshift forms. In some studies, recording instruments such as rain gauges, hygrothermalographs, and gas chromatographs recorded observations on coordinate paper. Currently, many of these instruments are being replaced with those capable of making observations machine readable in the form of analog or digital output (Cooney 1986). These signals can be received by other types of recording instruments, including microcomputers and data recorders.

Microcomputers and electronic data recorders created new and efficient methods of collecting research data, a welcomed development because scientists are continually seeking ways of increasing efficiency. Scientists who conduct forestry research are no exception.

This paper briefly describes some of our past methods of data recording and analysis, and discusses current methods of combining electronic data recorders and PC SAS software into a more efficient system for collecting, editing, and analyzing study data.

Manual Data Collection

Our data are usually collected in a field situation, where we are measuring the response of a given treatment or collecting observations on biological processes in the outdoors. We often take precise measurements of trees, such as height, diameter, branch length, root weight, and number of root tips. It is not uncommon for us to collect 40,000 to 60,000 80-column lines of data in one field season.

Before data recorders became available, the normal practice was to prepare field sheets for each of our studies. These sheets were reproduced on rainproof paper, because our field work often must be done in adverse weather. Hundreds of data sheets would be used in a normal season. The filled data sheets would be checked for completeness and errors. Depending on skill of field personnel, the need for error checking would vary considerably. After collection, the data would be entered (keypunched) into a mainframe computer. Depending on the amount of data collected during a given season, the time between collection and keypunching ranged from 6 months to a year.

In addition to the large amount of time consumed in gathering observational data, the costs of inputting data into a mainframe computer were high. Also, errors, which could go undetected, were generated during transcription. When errors were detected after keypunching, they often could not be corrected, because the experiment was completed or the experimental site was under 6 feet of snow. Thus, large amounts of hard-earned data could be lost. Occasionally key parts of studies or entire experiments were lost because errors were not detected in a timely manner.

Data Recorders

During the past few years designers of electronic data recorders have made large strides in improving performance and storage capabilities (Rogerson 1982; MacCracken and others 1984). At first the food industry began using recorders to track product inventory, but soon other businesses used the devices for timekeeping, quality control, remote sales ordering, and similar applications (Cooney 1985). Small, hand-held computers and pocket calculators showed some promise for use in a field environment, but were not weatherproof and lacked adequate memory.

Several manufacturers make data recorders each with its own merits (Cooney 1985). Some are small, waterproof computers with large memories that are fully programmable; others are more limited in their application.

We use hand-held recorders that are waterproof, programmable, and capable of receiving digital, analog, and keyboard input.
They have built-in communications, are easily formatted to page images, and contain 32K of memory. Newer recorders have more memory, and are smaller and lighter than models we are presently using.

Some adjustments were made in our data collection when recorders came into use. As with most equipment, the user must adapt somewhat to its peculiarities. In the past we prepared data sheets; the recorder is prepared with prompts to allow the user to enter data. With data sheets, a user works across the page entering data in each column one line at a time. The same method is used with the data recorder, but the entire page is not visible; thus, making random access to the recorder less efficient than a systematic approach. When each prompt is satisfied, a new prompt appears. With a line completed, a new line begins, with a repeat of the prompts of each column heading. Most recorders allow moving about the file so observations can be checked and changed if necessary. Depending on available memory, several files can be stored at the same time.

Programs can be written for the recorders to check entries or calculate new variables from entered variables. Also, recorders can be loaded with previous measurements, which can be compared to current entries to check for validity. For example, in our work it is nice to be sure that data are not representing trees growing backwards or growing after they are dead! These programs of course require memory. The more complex the program, the more memory utilized. Therefore, a balance is required between data checking and data storage when using electronic recorders. If the recorders are used to supply data readily into a microcomputer (PC), minimal programming or loading with previous measurements is desirable.

Upon completing a task with the recorder, the data are transferred to a PC using asynchronous software. We use a simple program that copies the recorder contents to a floppy disk. In addition, we make a hard copy of the data while transferring. When the data are safely on the disk, and a hard copy is made, the recorder can be cleared.

In addition to keyboard input, we also use recorders with electronic balances. The balance outputs a digital signal, which is received by the recorder. A program in the recorder places the readings from the balance in the proper format. The recorder has the capability to store formats and programs for different applications. But again, any memory used for programs or formats is not available for storing data. Therefore, it is desirable to maintain unused application formats and programs on a PC. They can be downloaded to a recorder when needed.

Spreadsheet software is used to sort, graph, and summarize field data. We can check for errors in data that has been sorted and summarized more readily and efficiently than by checking data sheets. If an observation is incomplete, in error, or missing we have the ability to make the correction immediately and not wait for the data to be keypunched. We have made many corrections in our data not possible with manual methods.

SAS System For Personal Computers

PC SAS software opened new possibilities for managing our data and for performing analysis, and we foresee additional improvements. We can use the SAS System on site, and most of the scientists and foresters in our organization are familiar with the System so it is not difficult for them to use.

Three years of data collection have been completed using electronic recorders and PCs at remote field locations. Analysis of these data has been accomplished with the SAS System on a mainframe computer. Because of poor and expensive telephone communications, we transfer field data to the mainframe through the asynchronous link at our main laboratory rather than from our field location. This procedure is often time-consuming and data can be lost in the transfer.

PC SAS software can replace the spreadsheet we are now using for data management. The ability to sort, merge, calculate, and check observations makes SAS software very applicable in meeting our data management needs. The SAS system is able to sort on more than two fields and divide fields—a task that cannot be accomplished with the spreadsheet. In addition, the PC SAS System is better adapted to management of large data sets. The easiest way to apply SAS software to our needs was to replace the spreadsheet with the PC SAS System, and continuing to use asynchronous software to transfer the data from the recorders to the PC.

Data Into The Display Manager

The Display Manager in the PC SAS System was designed to read files from PC disk drives. In addition, it can read observations directly from the communications port on the PC.

Most electronic recorders have communication parameters (protocol) as part of their operating system. These usually include such things as baud rate, parity, databits, stopbits, turn around characters, end of line sequence, and end of file sequence. The communications port on the PC has similar protocol which can be set by the mode command in DOS. The mode command can be invoked before a SAS session is started or during a SAS session. From our experience, it is advisable to set the communications port before starting the SAS session. The data recorder and PC should have the same protocol.
In addition, the data recorder should place a carriage return at the end of each input line. Column delimiters such as commas are not required, but they can be used if desired.

After the PC and data recorder are set to communicate, the SAS System is started on the PC. A file in the recorder is "included" into the Display Manager. The command INCLUDE 'COM1' is typed on the command line. After issuing this command the transfer is started on the recorder. Using this method the asynchronous communications program is eliminated from the process.

Once the data have been loaded into the Display Manager the information can be saved on the disk and printed. Our custom is to never change raw data files. The raw data can be checked by SAS programs. After checking, the data can be stored as SAS data sets ready for analysis.

Often the success of our research projects depends on quick processing of data. The combination of the PC SAS System and electronic data recorders provides this ability.

DATA Step

The PC SAS System has the ability to read files from different sources, such as the A or C disk into the DATA Step. Thus, we hoped data could be read directly into the DATA Step similarly to the method used for reading a file into the Display Manager. The code was written to read variables into the DATA Step directly from the electronic recorder. This procedure failed. No matter how it was tried, with different protocols and start-up sequences, a device error was received. The SAS System would crash and have to be restarted. Therefore, our attempts to shorten the process by one additional step failed, but we believe this shortcut would be a useful option to have.

Many of our applications in the future will be with equipment capable of producing digital output, such as electronic balances, gas chromatographs, or pH meters. This equipment can easily communicate with a PC. Reading directly into a SAS program from the communications port would increase the flexibility and utility of the PC SAS System.

Conclusions

To continue improving research efficiency, we need methods to reduce the amount of resources spent on managing data, and to enhance timeliness. The more efficiently we are able to design studies, collect data, and report results, the more our productivity will be increased. Electronic data recorders and the PC SAS System appear to be two tools that can help us achieve these goals.