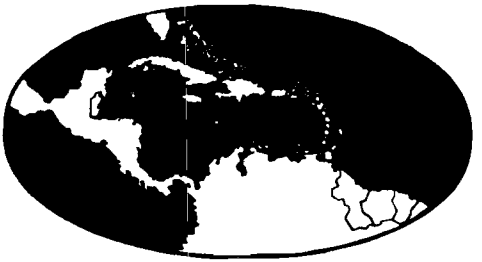
CARIBBEAN FOOD CROPS SOCIETY



THIRTEENTH ANNUAL MEETING
ST. AUGUSTINE TRINIDAD, W. I.
JULY 6-12, 1975

PUBLISHED WITH THE COOPERATION
OF THE
UNIVERSITY OF PUERTO RICO
MAYAGUEZ CAMPUS
1980

VOLUME XIII

AN INFORMATION RETRIEVAL SYSTEM FOR RESULTS OF FIELD EXPERIMENTS IN THE COMMONWEALTH CARIBBEAN

A.C. BREWER*
Caribbean Agricultural Research and Development Institute,
University Campus, Trinidad.

SUMMARY

Experience with a project to make a survey of coefficients of variation from field experiments in the region has highlighted many deficiencies in the present organization of recording results. In many cases results from completed experiments have been lost altogether and in other cases the statistical analysis of results have been incorrect in part or in whole. This paper discusses the uses and advantages of setting up a computer-oriented information retrieval system and outlines the administrative and data-processing procedures required to form an efficient and useful system.

INTRODUCTION

Most, if not all, large research institutes have an administrative system where the previous work of that institute is readily obtainable for any purpose for which it might be used, providing many obvious advantages. Agricultural research work in the Caribbean is fragmented into many separate bodies and a centralized system of administration is obviously not possible. However, experience with a project to survey coefficients of variation and other features of field experiments in the region has made evident the advantages of having an information retrieval system for results of field experiments, not only for the above project but also for other research and administrative reasons.

PRESENT SITUATION

The Ministries of Agriculture of the Regional Governments, the Regional Research Centre of UWI (now CARDI), the old Regional Field Experimental Programme of RRC, the UWI Faculty of Agriculture and various other organizations of more specialized research into various food crops each hold their own data individually, sometimes in systems where information is reasonably easy to retrieve and in other cases where it is not. In fact, in may places the results are held solely by the indivi-

---

* Seconded on technical assistance by the Ministry of Overseas development (UK)

dual researcher and although this establishes security of the work it makes the job of data collection very difficult. Also in this situation previous work is often lost when a research worker leaves the organization and useful data is sometimes mislaid for other reasons. Frequently statistical analysis of results from field experiments with a statistical design have not been performed, or may have been calculated incorrectly in whole or in part, mostly because of a previous lack of routine statistical data processing facilities. Results are presented sometimes in imperial units, sometimes in mixed imperial and metric units and sometimes in metric units. It would be preferable to have all information in standardized metric measurements. The flow of information within the region is by personal communication and by means of reports and scientific papers. This can and does in many instances work quite well, but it could usefully be supplemented with an organized system of information retrieval.

USES OF A RETRIEVAL SYSTEM

The Biometrics Research Project to assess features of field experiments which affect the precision of the results involves collection of a large body of data on a variety of features of completed experiments, including the results of the statistical analysis. A retrieval system is essential for this purpose and is the prime reason for its instigation although there are other auxiliary purposes for which it can be used. Many of the deficiencies of the present situation will be appreciably alleviated by it. An experimenter would have the results of his experiment routinely analysed with the results presented in standard metric units, irrespective of the units in which the observations were measured. This, however, does not alleviate the need for adequate statistical consultancy services to advise on and interpret results of statistical problems. A researcher wanting to know the type of work that has been previously done in his field could obtain the relevant information with a request for a listing of the previous work classified in that field, specifying the type of information that will be useful for his pruposes. This necessitates an adequate classification both of the details of the experiment which was performed, and the type of results that have been recorded. The system could also conceivably be used for the administration of research, deciding upon useful projects and defining them in such a way as to avoid repetition of work previously performed.

ORGANIZATION OF THE SYSTEM

The product of the system would be a magnetic tape file on the computer containing classified information on all the experimental results received together with an associated physical file containing information which is not easily handled on the computer. An example of

such information would be the field plan which would be needed for statistical studies. The system, when in operation, would be a continuous process where incoming results would be processed and added to the file as they are received. A body of past data has already been collected and is in the process of organization in order to be incorporated in the system.

The first requirement for participation in the system would be to use a standard field recording sheet and experiment description sheet, both with carbon copies. Many organizations already have their own standard sheets which would be almost adequate for the system's purpose. The sheets should be in a format which could be easily coded for input to the computer. One set of copies would be kept by the experimenter whilst the other set would be sent for the analysis and incorporation on the file. The computer programmes used would be designed to input the classified information onto the file, write the data to the file, perform the statistical analysis, and record the results of this on the file as well as printing them out. The printout would be sent back to the experimenter whilst portions of the field recording sheets and experiment description sheets would be placed in the physical file together with any comments.

Programmes would also be available to retrieve information from the magnetic tape file according to classifications that are given to it, such as crop, territory, which results are required, etc.

INFORMATION COLLECTION

An attempt has been made to consider what information should be handled by the system, where it would come from and where it would go to. We should not try to collect too much information as this would make the requirements arduous for the experimenter and unmanageable for the operators of the system. Thus, we would collect information which is of obvious use for the purposes of the system and this conforms to a large degree with the information recorded on experiment description sheets already existing in the region. The items of information can broadly be broken up into four categories under the headings of:

(a) A general description of the experiment;

(b) Details of the experimental plan;

(c) Data and statistical results;

(d) Extra information which would be useful in and information retrieval system.

All quantitative information to be output to the file would be converted by the programmes to standard metric units beforehand. The following table lists the units of information which would be handled, their source and their destination in the system. Codes are used for source and destination which are interpreted below the table:

	Use of information	Source	Destination
A.	**Description of Experiment**		
1.	Title of the experiment	EDS	CF, PF, PO
2.	Experiment code for reference in the system	SC	CF, PF, PO
3.	Crop(s) involved	EDS	CF
4.	Variety used	EDS	CF
5.	Island and location of experiment	EDS	CF
6.	Time of planting	EDS	CF
7.	Time of harvest	EDS	CF
8.	Growing season (wet or dry)	EDS or 5, 6, 7	CF
9.	Climatic factors:- total rainfall, temperature. Only brief summaries that have been measured	EDS	CF
10.	Soil characteristics: Soil type and any other known information	EDS	CF
11.	Type of treatments used	EDS	CF
B.	Details of Field Plan		
12.	Experimental design used (coded)	EDS	CF, PF, PO
13.	No. of replicates, No. of treatments etc.	EDS	CF, PF, PO
14.	Field plan	EDS	PF
15.	Dimensions of plot (net)	EDS	CF
16.	Area of net plot	EDS or 15	CF

	Use of information	Source	Destination
17.	Depth of guard rows	EDS	CF
18.	Area of guard rows	EDS or 17	CF
19.	Plant spacing	EDS	CF
20.	Plants/plot	EDS or 16, 19	CF
21.	Treatments used	EDS or DRS	CF, PF, PO
22.	Percent plant stand	DRS	CF, PF
C.	Data and Statistical Results (Repeated for each observation taken)		
23.	Observation taken	DRS	CF, PF, PO
24.	Unit of measure	DRS	CF, PF
25.	Raw data	DRS	CF, PF
26.	Transformation used in statistical analysis	SC	CF, PO
27.	Treatment means	SA	CF, PO
28.	Standard errors of means	SA	CF, PO
29.	Overall mean yield	SA	CF, PO
30.	Standard deviation	SA	CF, PO
31.	Coefficient of variation	SA	CF, PO
32.	Results of special techniques of analysis (e.g. Covariance analysis)	SC, SA	CF, PO
D.	Extra Information		
33.	Experimenter's name	EDS	CF, PF
34.	Reference to written up report on paper	C	CF
35.	Land preparation (general comments)	EDS	PF
36.	Cultural methods (general comments)	EDS	PF
37.	Topography (general comments)	EDS	PF
38.	Other comments	EDS & C	PF
39.	Statistician's comments	SC	PF
40.	Whether permission is granted by the experimenter to use the material in the system	EDS	

Key to Codes Used:

Source

EDS = Experiment description sheet sent by the experimenter.

DRS = Data Recording Sheet sent by the experimenter.

SA = Programme to perform statistical analysis of the results.

SC = The Statistical consultant.

C = Communications with the experimenter.

Numbers refer to previous units of information from which the present unit can be derived if it is not on the experiment description sheet.

Destination

CF = Magnetic tape file on the comupter

PF = Physical file

PO = Computer print-out to be sent back to the experimenter.

CONCLUSION

The considerable advantages of having this system in operation have already been described. The problems involved in setting up such a scheme have yet to be experienced but are likely to be more than trivial and will increase in magnitude according to the scale of the system that is envisaged. It is also highly dependent on the degree of cooperation that exists between the statisticians and the experimenters. After evolving the system with data which has already been collected, we plan to operate the system initially on a small scale and hope that it will eventually prove its usefulness and will receive the further cooperation of experimenters in the region.