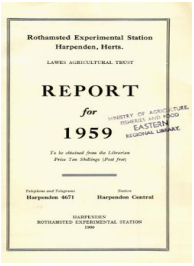
ROTHAMSTED RESEARCH

Report for 1959

Full Table of Content



Statistics Department

F. Yates

F. Yates (1960) *Statistics Department ;* Report For 1959, pp 151 - 157 **- DOI: https://doi.org/10.23637/ERADOC-1-92**

STATISTICS DEPARTMENT

F. YATES

The following members resigned during the year; S. Lipton, on appointment as Senior Lecturer in the School of Mathematics, University of New South Wales, Sydney, for the development of computer work there; O. B. Chedzoy, to Electrical and Musical Industries Ltd.; Edith Avis (Spetch) and Joan Parle (Anderson), on marriage.

New appointments were: G. W. Bonsall, from Armstrong Siddeley Ltd., and C. W. Fearne, from De Havillands Ltd. Seven temporary workers were accommodated during the year, four from overseas.

Yates visited India again at the end of the year on a United Nations Food and Agriculture Organisation assignment to work with the Indian Institute of Agricultural Research Statistics. He also contributed to a conference on mathematics in industry and its relation to mathematical education held at Liverpool University (12.5). Healy is spending a year in the United States at the invitation of Dr. J. W. Tukey of Princeton University. Rees attended the U.N.E.S.C.O. First International Conference on Information Processing in Paris. Church visited Ghana for a joint symposium of the Commission de Cooperation Technique en Afrique au Sud de Sahara and the U.N. Food and Agriculture Organisation, and contributed a paper describing the pilot survey of livestock in Eastern Ethiopia.

THE ELECTRONIC COMPUTER

The electronic computer was very fully used during the year, and continued to give reliable service. In addition to work for Rothamsted, much was done for other research organisations in this country

TABLE 1

Record of machine operation for 1957–59

Percentage of time for:	1957	1958	1959
System modification	0·6	2·4	0·0
Scheduled maintenance	12·4	13·9	9·4
Unscheduled maintenance:			
Computer	6·2	2·3	3·1
Peripheral	1·8	1·5	1·1
Programme checking	17·8	13·3	8·9
Programme tape preparation	4·8	3·5	1·4
Production runs	49·8	58·4	73·9
Abortive efforts:			
Computer	1·3	1·4	0·7
Peripheral	1·0	1·7	0·8
Idle time	4·4	1·7	0·7
Total	100·1	100·1	100·0
Total hours worked	2,622	2,869	2,993
Percentage overtime	32·6%	45·1%	50·8%

https://doi.org/10.23637/ERADOC-1-92

and overseas. The machine was operated for a total of 2,993 hours, i.e., 51% in excess of normal laboratory hours, compared with 45% excess in 1958. The longest continuous run was one of 81 hours. Although the total operating time increased by only 4% over 1958, there was a 32% increase in the time devoted to production runs. This reflects the shorter time devoted to maintenance and to programme checking and tape preparation. Details are shown in Table 1.

The output of work per hour of production time also improved through better organisation of machine operations. We now have a very effective team of machine operators.

TABLE 2

Numbers of replicated experiments analysed in the Department

	Number of experiments: By hand	On computer	Total	Number of variates on computer
1934	115	—	115	—
1951	437	—	437	—
1955	384	419	803	834
1956	181	683	864	1,701
1957	98	1,253	1,351	5,041
1958	182	1,664	1,846	6,260
1959	67	2,649	2,716	11,102

The number of replicated experiments analysed on the computer again increased substantially. The number of experiments handled increased by 60% and the number of variate analyses by nearly 80%. Details are shown in Table 2. (Dunwoody, E. P. Simpson, Turner, Gower.) A paper summarising our methods of experimental analysis was presented at a meeting of the German Region of the Biometric Society held in Leipzig (12.13). (Healy.)

The computer was also very fully used for many other types of computation, including the analysis of a number of surveys.

No major modifications were made to the machine, but further tape-handling equipment was acquired and the tape readers were made more reliable. The Hollerith card reader gave good service, and the tape-to-card equipment was also extensively used.

It was decided early in the year to abandon work on the Pye magnetic tape unit, and sanction was obtained to buy a Decca twin magnetic tape unit. This was delivered in December, and work is now proceeding on linking it up with the machine. (Rees and Knight.)

Computers have now been installed in a number of Universities, and are consequently available to university research workers for handling their statistical problems. Before a computer can be effectively used for statistical research, however, an appropriate set of programmes has to be written. To make the experience gained on the Rothamsted 401 computer generally available, a start was made on the construction of a " communication language " suitable for statistical programmes. It is hoped that it will be practicable, without too much labour, to re-write the more important 401 programmes in this language. A trial on the programme for the

https://doi.org/10.23637/ERADOC-1-92

analysis of 2^n experiments showed that such re-writing, if done by the author of the original programme, can be effected without much trouble. It also appears that new programmes for the 401 can be conveniently written initially in the communication language and then translated into the 401 order code. (Yates.)

Further experience is required to assess the utility of a communication language of this kind, but should it prove of real value to programmers of other types of machine the re-writing of existing programmes and the writing of new programmes in this language will not only make them available to users of other machines but will also be of great assistance to us when the time comes to replace the 401 with a new machine.

This work is linked with similar work that is now proceeding in many quarters on the evolution of a universal programme language. The philosophy behind such a language is that, for any particular computer, a translation routine or programme can be constructed which will enable the computer to accept programmes written in the universal language, and translate them into the language (i.e. order code) of that computer.

A further piece of similar work was the improvement of the automatic programming routine of the 401, mentioned in the 1956 Rothamsted Report (p. 169), to enable the quantities entering into the computation to be referred to by letter symbols, or a combination of letters and numbers. This simplifies the writing of a programme, and makes the programme, as written, more intelligible, because the algebraic symbols used to state the problem can be written directly in the programme wherever required. The routine was also improved in other ways to make it more convenient for the programmer. (Yates and Bonsall.)

The opportunity was taken, in the course of writing a chapter on the use of electronic computers in survey work for the third edition of *Sampling Methods for Censuses and Surveys*, of working out a general system for the specification of survey analyses (12.1). All the types of analysis that are commonly required can be simply specified in this system, and this in itself will greatly simplify their programming. The ultimate aim is to write an interpretative routine which will interpret these specifications directly, so that no special programming will be required. A full routine of this type will not be practicable for the 401, because of its limited speed and storage capacity, but an attempt will be made to construct a routine which will at least deal with the simpler types of analysis without much special programming. (Yates.)

The problem of handling data from continuous recording apparatus, which concerns many agricultural research institutes, received some preliminary study. The first step is to arrange for the data to be recorded at suitable intervals in digital form so that the record may be fed to the computer without any manual transcription. A specification of suitable equipment for the Glasshouse Crops Research Institute was agreed with a manufacturer. The problem arises in many branches of research and in production control, so it is likely that a number of types of suitable recording instrument will be developed in the near future. A possible alternative solution, which would not involve the installation of special

https://doi.org/10.23637/ERADOC-1-92

recording instruments, would be optical scanning of records. (Rees and Gower.)

Experience was also gained in programming the analysis of continuous or quasi-continuous data by the analysis of a series of meteorological observations taken at the Vegetable Research Station, in which the temperature and radiation determinations were obtained from continuous daily records. (Rees.)

Experimental Design and Analysis

The analysis of the 1956–57 results of a set of co-operative fertiliser trials on rice in the State of Bihar, India, referred to in the 1958 Report, was completed successfully and the methods followed were published (12.21). The 1957–58 and 1958–59 results have now been analysed in a similar manner. It is hoped that we shall soon be provided with the results of the 1959–60 experiments, when a combined analysis of the 4 years' results will be undertaken, so that cultivators can be provided with an improved set of recommendations. (Yates, Lipton and Gower.) A similar analysis of a simpler set of fertiliser trials from Pakistan was also made. (Gower.)

The examination of the results of the Rothamsted 1933–58 Three-Course Rotation experiment, on the effects of straw ploughed in or composted, was completed (16). (Patterson.) A brief summary of these results was given in the 1958 Report (p. 167). A paper on the manuring of potatoes, summarising the results of recent factorial fertiliser trials on this crop, was presented to the Society of Chemical Industry. (Boyd.)

We continued to do a great deal of work for the National Agricultural Advisory Service. The electronic computer is now used for the routine analysis of almost all the one-year experiments, and we also handle the results of the long-term experiments and co-operate in the planning of experimental programmes and the design of experiments. (Boyde, Patterson, Lessells and E. P. Simpson.) The results so far obtained from two important series of experiments on the N.A.A.S. Experimental Husbandry Farms, namely the experiments on straw disposal and on residual effects of phosphorus fertilisers, were summarised and a report presented to the Crop Experiments Sub-Committee of the Agricultural Improvement Council. (Patterson.)

Several investigations were made for the Soil Fertility Committee of the Closed Conference of Soil Chemists and some reports prepared. (Lessells.) Work for the National Institute of Agricultural Botany included the examination of the results of the N.A.A.S./N.I.A.B. cereal variety trials and the N.I.A.B. fodder-beet trials. (Boyd and Lessells.)

G. C. Chisci, of the Stazione Sperimentale di Praticoltura, Lodi, Italy, while working in the Department, studied the relative productivity of varieties of lucerne in Northern Italy (12.9), and the value of leaf measurements for discriminating between varieties.

Work for the Advisory Entomologists Conference and the Pest Assessment Committee continued, and a paper was prepared on soil sampling for potato-root eelworm (12.11). (Church.)

https://doi.org/10.23637/ERADOC-1-92 *pp 5*

Surveys

The analysis of the results of the 1958 Survey of Maincrop Potatoes, made in co-operation with the National Institute of Agricultural Engineering and the Potato Marketing Board, was completed; a report was prepared and will be published shortly. (Church, H. R. Simpson and Hills.)

During 1959 the N.A.A.S. undertook surveys of fertiliser practice in four districts, using a modified sampling and recording technique. The results from these districts and those outstanding from 1958 were analysed. We also analysed the results of similar surveys of fertiliser practice for six districts in Scotland on behalf of the A.R.C. Statistics Unit, Aberdeen. A paper summarising information on recent trends in cereal manuring and describing the methods of fertiliser practice surveys in England and Wales was published (12.10), and also one on survey techniques of interest to the N.A.A.S. (12.8). (Boyd, Church, H. R. Simpson and Hills.)

A pilot survey of herbicide practice was undertaken in conjunction with the A.R.C. Unit of Experimental Agronomy, Oxford, to discover how much information could be obtained on herbicide practice and its effects by interviewing farmers; a paper was prepared describing the results (12.12). Information on the proportions of crop acreages treated with different herbicides was readily obtained, but farmers were usually unable to give details of techniques and times of treatment. The limited information obtainable may, however, be of sufficient value to justify extending this work now that herbicides are so widely used. (Church and Hills.)

The Grassland Survey of England is being analysed on the computer for the Grassland Research Institute. (H. R. Simpson and Mr. C. D. Kemp.)

The results of the First National Survey of Disease in Dairy Cattle covering the year 1957–58 are being analysed on behalf of the Ministry of Agriculture, Fisheries and Food, Animal Health Division, and provisional reports were prepared. This survey follows some pilot surveys of which the results have been published (12.20). Data were obtained from 1,158 farms, and combined techniques involving both the electronic computer and the Hollerith tabulating equipment were used for analysis. This survey provides the first unbiased estimates of the national incidence and relative economic importance of disease in farm animals. The results will be published by H.M.S.O. in the form of a bulletin. A second survey was made during 1958–59. (Leech, H. R. Simpson and Davis.)

The losses of breeding ewes during the 1958–59 breeding season were surveyed by the N.A.A.S. for the Veterinary Investigation Service and will be analysed as soon as complete records have been received. The same flocks will be surveyed in more detail during the current breeding season. (Leech and Davis.)

Plans were worked out in conjunction with the Ministry of Agriculture, Fisheries and Food for a survey of the incidence of *Brucella abortus* infection in dairy cows, which will start in the autumn of 1960. (Leech.)

A study was made by the N.A.A.S. of the effects of various factors on the content in milk of solids other than fat. The data came from

https://doi.org/10.23637/ERADOC-1-92

a co-operative investigation on ten farms, seven of which had a history of producing milk low in such solids, and the others were normal farms (12.7). As a result of this work, requests were received to analyse similar sets of data on milk composition from the National Institute for Research in Dairying and the Hannah Dairy Research Institute. These analyses are in progress. Plans were also prepared for the Joint Milk Quality Control Committee for a fuller investigation, and the field work is proceeding. (Leech.)

A report was presented to the Agricultural Research Council's 1959 Conference on Hypomagnesaemia on the incidence of grass tetany estimated from random sample field surveys. (Leech.)

At the request of Dr. K. Rasmussen (Head of the Department of Agricultural Economics, Nottingham University) the relation of gross output of farms to various inputs (fertilisers, labour, etc.) was studied, using 4 years' data on 11 variates for 1,654 English farms and 3 years' data on 13 variates for 1,139 Irish farms. There were over 10,000 cards containing the original data. As a preliminary every variate had to be converted to logarithms, involving 117,197 conversions. Output and the various input factors were found to be very closely related. (Gower.)

Commonwealth Work

We have continued our advisory work for Commonwealth countries and have analysed many experiments on the computer, including a large set of 2^5 manurial trials on a number of crops in Fiji, extensive analyses for the Gambia, maize variety trials for the Maize Rust Research Unit, Nigeria, and experiments of the Tobacco Board, Tanganyika. Advice was given to the Kumasi Technical College, Ghana, on the design of some related long-term experiments on cocoa and to the Nigerian Federal Department of Agricultural Research on the planning of a rotation experiment (with H. Greene). (Vernon.)

Three Commonwealth workers worked in the Department for varying periods.

Other Work

Of the many miscellaneous problems dealt with for various research workers and institutes by members of the Department, the following may be mentioned; most involved the use of the computer:

1. A comparison of various methods advocated for fitting an exponential curve (12.15, 12.17). (Patterson and Lipton.)
2. The development for Dr. C. B. Williams of an asymptotic formula for the remainder terms

$$S(x, R) = \sum_{r=R}^{\infty} \frac{x^r}{r}$$

where $x \sim 1 - 10^{-13}$ and $R \geqslant 10^6$. This formula gives the total number of species in a population with more than R individuals per species. The value of x is related to the possible frequency distribution of number per species for all the insects in the world (10^{18} individuals in 2×10^6 species). (Gower.)

https://doi.org/10.23637/ERADOC-1-92

3. A study of the problem of determining the degree of imperfection in mixing batches of seed, for the International Seed Testing Association (12.18). (Westmacott.)

4. Biological studies of the wheat-bulb fly (*Leptohylemyia coarctata* (Fall.)) in conjunction with the Entomology Department. (Dobson, Long and Morris.)

5. The development of a multiple transfer method for plant virus (and other disease) transmission studies, in conjunction with the Plant Pathology Department. (Gibbs and Gower.)

6. Further predator–prey investigations, for Dr. P. H. Leslie of the Bureau of Animal Population, Oxford (12.14). (Gower.)

7. Programming of the Jackson method for the analysis of capture–recapture data, for Mr. J. A. Nelder of the National Vegetable Research Station. This will also be of value for the analysis of similar data being collected by the Rothamsted Entomology Department. (H. R. Simpson.)

8. A study for the Animal Virus Diseases Research Institute, Pirbright, of the efficiency of an approximate method of estimating genetic variance in susceptibility of unweaned mice to foot-and-mouth disease, compared with the maximum likelihood method of estimation. (Leech.)

9. Analysis of comparative ventilation data, for the Glasshouse Crops Research Institute. (Rees and Bonsall.)

10. Further analyses of data from the National Survey of the Health and Development of Children, for the Joint Committee of the Institute of Child Health, Society of Medical Officers of Health and the Population Investigation Committee. (H. R. Simpson.)

11. Studies of growth measurements of children from the Harpenden Survey of the Institute of Child Health. (H. R. Simpson.)

12. A discriminant analysis of data on test reactions for the diagnosis of tuberculosis and leprosy, for the South Metropolitan Cancer Registry. (H. R. Simpson.)

13. Principal component analysis of recidivism data, for Mr. L. T. Wilkins, and prison data for Mr. A. Straker, both of the Home Office. This work was brought to us because of our experience in this type of analysis, after requests were made to several other computer installations. (H. R. Simpson.)

Mr. N. E. G. Gilbert of the John Innes Horticultural Institution continued his work for the Animal Breeding Research Organisation. He also analysed data from the progeny-testing stations of the Pig Industry Development Authority.

H. R. Simpson wrote up his work on the effect of insecticidal applications on tsetse populations (12.22), and Yates contributed a section on methods of sampling human and animal populations to the Report of the World Health Organisation Study Group on Immunological and Haematological Surveys (12.20).

https://doi.org/10.23637/ERADOC-1-92