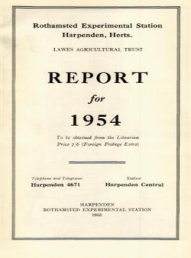
ROTHAMSTED RESEARCH

Report for 1954

Full Table of Content



Statistics Department

F. Yates

F. Yates (1955) *Statistics Department ;* Report For 1954, pp 137 - 142 **- DOI: https://doi.org/10.23637/ERADOC-1-76**

STATISTICS DEPARTMENT

F. YATES

The new temporary building adjoining Rivers Lodge was completed in March and has provided much-needed additional accommodation. The Hollerith equipment is now housed in this building, and this has considerably improved its efficiency of utilization. The lecture and conference room included in this building has also proved very useful, both to the department and to the Station generally.

Three members of the scientific staff left during 1954. P. M. Grundy has taken a post with the National Foundation for Educational Research, N. E. G. Gilbert has gone to work under Sir Ronald Fisher on an Agricultural Research Council grant and P. R. D. Avis has taken a post with the Air Ministry.

New appointments were : S. Lipton, from Liverpool University, A. J. Whitwell, from the R.A.F. Flying Training Command Research Branch, Sheila Cohen from Electrical and Musical Industries Ltd. and D. J. Knight, from the War Department, R.E.M.E., Leicester, as technician for the electronic computer. R. W. Sharp is spending 2 years in the Department working for a Ph.D. on a Ministry of Agriculture Post-Graduate Scholarship.

G. E. Hodnett spent 4 weeks in the West Indies visiting agricultural and statistical departments in Jamaica, Trinidad and British Guiana.

Twelve temporary workers were accommodated in the department during the year, three from Great Britain, one from Australia, seven from the Colonies and one from Belgium.

THE ELECTRONIC COMPUTER

The outstanding event of the year has been the installation of the N.R.D.C.—Elliott 401 electronic computer in the department. The machine, which is the property of the National Research Development Corporation, is being run on a joint basis, half of the time being at the disposal of Rothamsted and half at the disposal of the National Research Development Corporation. The machine was installed at the beginning of April, and was in running order by the middle of May. Since that time it has suffered no major breakdown, and has generally proved itself reasonably reliable. D. H. Rees, who with the initial assistance of Mr. H. G. Carpenter of the National Research Development Corporation has been responsible for keeping the machine in working order, deserves great credit for this.

It is often thought that the main feature of electronic machines is their speed. This is only part of the story. Equally important is their ability to perform long and complicated trains of operations without the intervention of the human operator; indeed, without this feature their speed would be of little value. It is thus possible to perform the *whole* of a complicated statistical analysis on the machine " untouched by human hand ". A good example is the analysis of randomized block experiments mentioned below. The

https://doi.org/10.23637/ERADOC-1-76

results are presented in the form of treatment number, yield and residual of each plot, arranged in the layout of the actual experiment (these can be suppressed if not required), the degrees of freedom and mean squares of the analysis of variance in the usual form, the standard error per cent per plot, the treatment means and their standard error, and the general mean. A conversion factor can be applied if required, and routines are being written to calculate quantities such as the plot yields of sugar from yields of dirty roots, sample weights of clean and dirty roots and sugar percentages. The whole operation consumes about 3 minutes of machine time for an experiment of about forty plots, of which 2 minutes is occupied by getting the data into the machine and punching out the results for printing on the teleprinter.

The programming of a series of operations of this type, i.e., the writing of the necessary instructions (orders) and testing their accuracy is, as might be expected, a time-consuming, if fascinating, task—the randomized block routine contains some 300 orders, most of them obeyed many times over in the course of the analysis of a single experiment—but once the programme has been constructed and tested it is available for use whenever required. All that is necessary is to feed the appropriate routine tape into the machine, an operation which, with the fast-input technique now being developed, will, for the randomized blocks routine, take between 1 and 2 minutes. The machine is then set up to analyse any required number of experiments. This may be contrasted with the time and skill required to plug a Hollerith plug board for even quite simple operations.

For those who are interested in technicalities it may be mentioned that the 401 is a serial computer with a word length of 32 binary digits and a word time of 100 μs. The main store is a magnetic disc of 23 tracks, each holding 128 words (orders or numbers); 7 of the tracks are always immediately available to the computer, but only 1 of the remaining 16 is accessible at any given time, switching between these tracks being by means of high-speed relays. There are 5 single-word immediate-access registers; one is the accumulator, which can be coupled with a second register for double-length working; the other 3 registers can each be used to modify orders (B modification) as well as for temporary storage of numbers. Input is by five-hole punched tape, and output by either electrical typewriter or teleprinter-punch. The machine has a two-address code to allow for optimum programming. Operations performed directly by the machine include addition, subtraction, multiplication with and without round off, collation, non-equivalence, discriminations on zero or negative values, and left and right shifting of single- or double-length numbers. Addition and subtraction take one word time (100 μs) to carry out, a shift of n places takes n word times and multiplication 31 word times (3·1 msec.).

Before it came to Rothamsted the machine was in the Mathematical Laboratory at Cambridge for 6 months, but the first magnetic disc proved unsatisfactory, and a new one was fitted only a short time before it left Cambridge. Consequently the machine came to us with practically no sub-routines, and a considerable amount of work has consequently been required to build up an adequate library for our purposes.

https://doi.org/10.23637/ERADOC-1-76

The machine has, however, already done some useful work. A routine for analysing randomized block experiments (M. J. R. Healy) is in operation, and has been used to obtain results for several large series of experiments. It is expected that by the end of the present winter about 300 separate analyses will have been carried out for the National Institute of Agricultural Botany and about 100 for the National Agricultural Advisory Service, as well as miscellaneous analyses for Rothamsted and other research centres. Work on the routines for the analysis of commonly occurring designs of more complicated types—factorial experiments, Latin squares, split-plot designs, etc.—is under way; it is expected that a number of these routines will shortly be in working order.

A number of routines for multivariate work have already been completed, and others are being prepared. A problem of discriminating between the teeth of men and of great apes by means of measurements is being investigated, using a large body of data collected by Dr. E. H. Ashton of the Department of Anatomy, Birmingham University (M. J. R. Healy). It is hoped that this investigation will throw some light on the nature of the australopithecine fossils.

A thorough investigation of the efficiency of the " combination of probabilities " test of significance has also been completed (F. Yates and Averil M. Munns). This required a large volume of computing on the machine, and has given illuminating and conclusive results. The problem is of interest, since it has become customary, particularly amongst geneticists, to use this test as a method of combining the results of a number of experiments of the quantal (" all-or-nothing ") type involving two treatments only, e.g., differences in mutation rates under two different treatments. Other methods of combining the data which are more efficient and also more satisfactory in other ways are available, but it is difficult to get such methods adopted unless convincing proof of the inefficiency of the combination of probabilities method can be advanced. The results of these computations will enable us to do this.

The general routines already developed include a " programming " routine (F. Yates and S. Lipton) and a set of routines for floating binary operations (S. Lipton). The former inserts the appropriate address in a skeleton programme and punches out the final routine for immediate use, thereby considerably lessening the work required to write a programme and reducing the potentialities of error; this is of considerable importance with the 401, as with the two-address system used in the machine the ordinary system of programming is a tiresome and tedious operation and one which is peculiarly liable to give rise to errors. The routines for floating binary operations allow the full 9-significant-decimal accuracy of the machine to be utilized and accommodate numbers in the range $10^{\pm 8000}$.

On the mechanical and electronic side a good deal of development work, much of it associated with the ancillary equipment, has been undertaken (D. H. Rees and H. G. Carpenter). The facility to type or punch out the results according to instructions contained in the programme is now available. Reorganization and supplementation of the existing input and output equipment is under way. We are also investigating methods of reading information into the

https://doi.org/10.23637/ERADOC-1-76

machine directly from Hollerith cards and of punching out information directly from the machine on to Hollerith cards. These two features are of great importance in statistical work, first, because many simple arithmetical operations require to be performed on data which can then be efficiently analysed by punched-card machines, and secondly, because a great deal of data which can best be analysed on electronic machines at present exists on punched cards. When these developments have been completed the computer should play an important part in the better analysis of survey data.

As a result of the first 9 months' experience it may be said that :

(*a*) Having an electronic machine on the spot has made all the difference to developing its applications to research statistical problems. In this respect our experience is exactly parallel with our experience with Hollerith equipment, where we found that it was only by having equipment on the spot, so that research workers could themselves use it, test out different methods and examine the results as they were obtained, that we could exploit its full potentialities.

(*b*) The introduction of electronic methods of computation will make available for regular use statistical methods which at present are scarcely used because of the heavy numerical work involved. This in turn is likely to lead to major developments in method. It will also facilitate and speed up the routine analyses which are at present done on desk machines, but which are of a sufficiently standard type to be programmed electronically, and enable a much more thorough preliminary examination of the data to be made (to check for gross errors, inconsistencies, etc.) than is at present customary or possible.

Design and Analysis of Experiments

The work on the amount of experimentation that is economically justified was brought to a conclusion by P. M. Grundy before he left, but the final details of the full paper on the subject have yet to be settled. M. J. R. Healy and D. H. Rees have co-operated in this work. The joint paper by Grundy, Healy and Rees which was presented to the Biometric Conference in 1953 has now been published (180).

A good deal of work has been done on the reorganization of the long-term and classical Rothamsted and Woburn experiments. H. D. Patterson has been particularly concerned with this work, and with the design and analysis of long-term experiments at Rothamsted and elsewhere. E. M. Crowther's sudden death has placed an added burden on the Statistics Department in connection with this and other work.

There has been the usual volume of routine work on the design and analysis of field and laboratory experiments for Rothamsted and other research stations and for the National Agricultural Advisory Service.

https://doi.org/10.23637/ERADOC-1-76

Surveys

During the autumn of 1953 and the first half of 1954, at the request of the Agricultural Research Council and the Ministry of Agriculture, a survey was undertaken by officers of the Ministry's Advisory Land Service and the National Agricultural Advisory Service of a sample of farms assisted under the Hill Farming and Livestock Rearing Acts, and of other farms not assisted under these Acts. The main aims of the survey are to estimate the returns being obtained from public and private expenditure already incurred under the Acts, and to discover what scope there is for further expenditure on upland farms. The survey has been designed by D. A. Boyd and B. M. Church in collaboration with Professor Ellison, University College of Wales, Aberystwyth, and the results have been analysed by the Department (B. M. Church, W. J. Lessells and Muriel E. Davis). The findings will be published in due course.

The survey of fertilizer practice continued in 1954 in twelve districts, the field work being undertaken by the staff of the Advisory Chemists of the National Agricultural Advisory Service. Most of these districts had already been surveyed in recent years, and the combined results will therefore give useful information on trends in fertilizer use. The completion of field work has been delayed; a report will be produced early in 1955. B. M. Church is in charge of this work.

In 1952 estimates were obtained by survey methods of the extent of the loss in yield of winter wheat due to the grazing of rabbits. This was shown to be most severe when the rabbits grazed in January. The estimates were, however, somewhat imprecise because of the restricted size of the survey, and in addition they could not be regarded as national averages because several cereal-growing areas had not been surveyed. A second survey, planned by this Department (M. H. Westmacott) in collaboration with the Ministry of Agriculture and Fisheries Plant Pathology Laboratory, will show whether similar results are obtained in the two seasons, and will improve the precision of estimates by providing a greater body of data. The 1954 survey covered the whole of England and Wales except for the areas where very little winter cereal is grown. Results will be available for analysis during 1955. In view of the very variable incidence of myxomatosis and its incidental effect on rabbit trapping, it may well be advisable to repeat the survey in 1956.

The survey of diseases of dairy cattle (F. B. Leech in collaboration with Dr. F. W. Withers) has been extended to cover Salop and the principal dairying areas of Ayrshire and Lanarkshire, besides continuing the work in Wiltshire, Devon, Surrey and Berkshire. The results of work in Wiltshire, Surrey and Berkshire in 1952–53 have been analysed, and Withers is writing a report.

The results of the 1953 survey of losses in pregnant ewes in Yorkshire have been analysed and a report submitted for publication by Dr. H. C. Sellers and F. B. Leech (186). They showed that although losses were not on the average high, a number of flocks were seriously affected. Flocks in the Holderness region were seriously affected by deaths during pregnancy and at parturition; flocks in the Wold by abortion; flocks in the North-Eastern Moors by barrenness. The survey was repeated during the 1954

https://doi.org/10.23637/ERADOC-1-76

lambing season, and the data have arrived for analysis. Weather during the 1953 lambing season was exceptionally favourable; losses during the severe 1954 season may therefore be much greater if weather is a contributory factor. Similarly, if the incidence of barrenness is affected by weather conditions during tupping time there should be relatively little barrenness in 1953–54, for the autumn of 1953 was very mild and the autumn of 1952 very severe.

A paper on some results of the survey of methods of milk production has been completed by F. B. Leech, with Messrs. J. W. Egdell, P. Heskin and S. B. Thomas (185).

On the theoretical side P. M. Grundy prepared a paper on a method of sampling with probability exactly proportional to size (181), and B. M. Church completed an extensive investigation into problems of sample allocation and estimation in an agricultural survey using the data of the survey of fertilizer practice (178).

Colonial Work

A training course in sampling theory and practice, mainly for officers in the Colonial Service, was held at Rothamsted from 23 August to 8 October. This course was organized by the Rothamsted Statistics Department and the Division of Research Techniques of the London School of Economics. The course appears to have been highly successful.

G. E. Hodnett has continued his work in advising and assisting Commonwealth workers, particularly in the Colonial territories, in statistical problems. Seven colonial workers have stayed in the department for periods up to 3 months, and nine others have paid brief visits.

Other Work

M. J. R. Healy has assisted in a number of miscellaneous statistical investigations (183, 184, 121). He contributed a paper on statistical techniques for inspection sampling to *Tropical Agriculture* (182). He has also co-operated with the Meteorological Office at Dunstable on the analysis of long-term weather data on our Hollerith equipment. B. M. Church assisted Mr. A. H. Strickland in a problem of sampling cabbage aphid populations on Brussels sprouts (179).

Various members of the department have co-operated in refereeing scientific papers. In many cases direct advice and assistance have been given to the authors concerned on the better statistical treatment of their data.

https://doi.org/10.23637/ERADOC-1-76