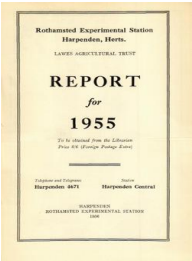
This work is licensed under a Creative Commons Attribution 4.0 International License.

Thank you for using eradoc, a platform to publish electronic copies of the Rothamsted Documents. Your requested document has been scanned from original documents. If you find this document is not readible, or you suspect there are some problems, please let us know and we will correct that.

ROTHAMSTED RESEARCH

Report for 1955

Full Table of Content



Statistics Department

F. Yates

F. Yates (1956) *Statistics Department ;* Report For 1955, pp 146 - 152 **- DOI: https://doi.org/10.23637/ERADOC-1-77**

This work is licensed under a Creative Commons Attribution 4.0 International License.

STATISTICS DEPARTMENT

F. YATES

Two members of the scientific staff left during 1955. G. E. Hodnett, the Colonial Office Statistician attached to the Department, has been appointed to the post of Statistician to the Regional Research Scheme based on the Imperial College of Tropical Agriculture, Trinidad, and Muriel E. Davis left to take up a post as Secretary to the Experimental Husbandry Farms of Reading University.

New appointments were: H. R. Simpson from Cambridge University, J. C. Gower from Cambridge and Manchester Universities, Mary G. Hills from the Central Research Establishment of the National Coal Board and Marjory G. Morris, who has in the past worked in the Department of Tsetse Control, Gold Coast.

Thirteen temporary workers were accommodated in the Department during the year, twelve of them from various countries of the Commonwealth.

THE ELECTRONIC COMPUTER

Arrangements have been concluded with the National Research Development Corporation for the Elliott 401 to remain at Rothamsted for a further period of 5 years, half the time of the machine being, as before, at the disposal of the National Research Development Corporation. Apart from a period in the early part of the year, when a faulty batch of valves gave rise to considerable trouble, and some trouble with the disc during the very hot weather, the machine itself has for the most part behaved well. The same cannot be said of the ancillary equipment, especially the tape reader and the typewriter.

During the year a good deal of work has been put into the reorganization of the input and output, and considerable progress has been made. The computer now uses a common input and output code in place of the two separate codes used previously. This has required a major electronic modification. The new code has the special feature that characters can be checked against errors in either reading or punching. When an error has been made during the input of information, the machine will stop. Errors in the output will lead either to a hold up on the typewriter or to the printing of a non-numerical character on the teleprinter. The programmer has the option to use either checked or unchecked input and output.

Plans to build a further tape reader have been abandoned in favour of replacement by a Ferranti high-speed reader, and steps are now being taken to secure one of these. A new Creed high-speed punch has been obtained to replace the present makeshift punch as the main output of the machine. This should be in operation very shortly and will nearly double the output speed of the machine. Tape from it will normally be fed directly to a teleprinter, thus obviating much of the present need for the typewriter.

https://doi.org/10.23637/ERADOC-1-77

This work is licensed under a Creative Commons Attribution 4.0 International License.

A new and much improved keyboard punch and teleprinter unit has also been obtained. This provides a printed record of all tapes as they are punched or reproduced, a feature which greatly facilitates the checking and editing of tapes prepared for the machine. Previously, input tapes could only be checked by reading the tape visually, by duplicate punching and visual or machine comparison, or by printing out via the machine. All these methods were crude and grossly inefficient, and led to much waste of operator and machine time. The old tape equipment is now being modified and reorganized so as to provide further punching and printing facilities to deal with the increasing load of productive work. The whole of our tape equipment can be utilized much more efficiently and rationally now we have a common input and output code.

The other main line of development with which we are most anxious to press ahead, namely input of data from punched cards and output of intermediate results back to cards for further processing by Hollerith equipment, has made slower progress than was hoped owing to numerous difficulties in obtaining the required equipment. The difficulties with regard to input from cards are now, however, well on the way to solution, though output to cards is not so well advanced.

Apart from the improvement of input and output facilities, the main task of the year has been the writing of programmes for various types of statistical analysis. Yates, Healy, Lipton, Rees, and since their arrival, Gower and H. R. Simpson have been mainly concerned with this. Experience has shown that to be of real use in routine statistical analyses, the machine should carry out the whole of the analysis, being presented with the raw data and producing typed copies of the results in the required form. Thus in the analysis of replicated experiments the preliminary conversion of yields and calculation of such quantities as yields of sugar from yields of dirty roots, dirt tares and sugar percentages is performed on the machine, which then carries out the analysis of variance and prints out the analysis of variance table, together with the necessary tables of mean yields and their standard errors. An added feature, which is not practicable with desk computers, is the calculation of the residuals, i.e., what is left after allowing for effects of treatments and blocks (or rows and columns of a Latin square). These residuals are very useful in spotlighting the sources of excessive error. Residual sums of products (for covariance analyses) can also be obtained when required, and missing data (" missing plots ") can be dealt with by means of small supplementary programmes (201).

Programmes of this type are necessarily fairly elaborate, and are practicable only for designs which are in common use. Those at present available are for randomized blocks (up to 64 blocks or 64 treatments, with a maximum of 128 plots), Latin squares (up to 10×10), the $3 \times 3 \times 3$ factorial design in blocks of 9 with a single replicate, and one form of randomized blocks with split plots (this last being for an extensive series of variety trials containing N as an additional factor, carried out by the National Institute of Agricultural Botany in 1955).

The following analyses of experiments were carried out on the computer during 1955 (each variate analysed being reckoned as one unit) :

https://doi.org/10.23637/ERADOC-1-77

This work is licensed under a Creative Commons Attribution 4.0 International License.

	Randomized blocks	Latin squares	3 × 3 × 3
Rothamsted, etc.	187	121	—
National Agricultural Advisory Service...	248	12	16
National Institute of Agricultural Botany	250	—	—
	685	133	16

Another standard type of statistical analysis for which a programme giving a complete solution has been written is the fitting of probit lines. This provides the equation to a fitted probit line together with the LD50 and its fiducial limits. A χ^2 value measuring the goodness of fit of the line is printed, and allowance can be made for heterogeneity and non-zero response in the controls. A similar programme for fitting a probit plane has also been prepared. These programmes have proved of considerable value to the Insecticides Department, particularly on account of the speed with which the work can be completed.

For the less standardized types of statistical analysis the parts which are of standard form and which require heavy computation can be carried out on the machine. Much multivariate work is of this type, and the following general programmes for such work have been written :

Sums of squares and products
Multiple regression
Matrix inversion

These programmes have already been used on multivariate work of various kinds.

Other miscellaneous programmes written for special purposes during the year are :

Calculation of a discriminant function (for tuberculin testing of cattle).
Asymptotic regression (used for fitting growth curves of children and the study of response curves for fertilizers).
Solution of the differential equation $\nabla^2\phi = 2e^{\phi}$ (required by the Physics Department in the study of the properties of colloids).

The programming of various types of Fourier analysis required by the Pedology Department for their crystallographic studies is now in progress.

A programme for the solution of coupled non-linear differential equations was written for an important piece of work on non-linear regenerative extraction of synchrocyclotron beams, reported by Dr. K. J. Le Couteur and Lipton (205). This work was begun while Lipton was at Liverpool University. All the necessary computational work has been carried out on the Elliott computer.

The study of sampling errors, which is required in the evolution of satisfactory sampling techniques, will be greatly aided by electronic computation. A programme has already been written for the determination of sampling errors in samples of various types from serial data. This was required for the study of sampling errors in nutritional surveys (193).

While the writing of programmes does not occupy machine time,

https://doi.org/10.23637/ERADOC-1-77

This work is licensed under a Creative Commons Attribution 4.0 International License.

the associated testing necessary to locate and eliminate all errors can consume very substantial amounts of both machine and programmers' time. We have given a good deal of thought during the past year to devising ways for lightening the programmers' task, both by reducing the actual work of coding, improving the accuracy of the initial programme and speeding up the location of errors. One obvious way is to use the machine to carry out the simpler parts of its own programming. In the order code for the Elliott, as in most computing machine codes, there are a number of minor arithmetical conditions between the different parts of the order which have to be satisfied. The code is also of the type in which each order specifies the address of the next order, so that the orders can be dotted about the store in any manner that appeals to the programmer. This enables speed of operation to be increased (so-called "optimum" programming), but obviously leads to more complex programming than do codes in which all the orders are located serially in the store in the order they are to be obeyed. These details can now be relegated to the machine by means of an automatic programming routine. This enables the programme to be written in a simplified code, without specification of store positions. The routine then assigns timing numbers and store positions so that all the required arithmetical conditions are satisfied and the timing is reasonably optimal, draws attention to any blunders on the part of the programmer that can be detected from internal evidence, prints a diagram of occupied store positions and punches out a final order tape. This automatic programming routine has already proved its value, both in saving programmers' labour and in producing programmes which contain far fewer errors. In response to developing needs it has recently been modified in order to deal with larger programmes, and programmes written on more than one "Track 7". (See Lipton (208) for a description of this feature.)

Two further routines of considerable use in checking and correcting programmes have been written. The first is a fault-finding routine which enables a computation to be interrupted at any point in the programme and the contents of the registers and any chosen store positions printed out with subsequent return to the programme at the correct point. The second is a routine for printing out a programme in logical order. This enables a programme which has been considerably corrected or modified to be reprinted in logical order, and also reveals any errors in addresses which lead to errors in the logical order. (These errors are in fact usually the most troublesome to locate, since the machine "runs out of programme" and may indeed mutilate its programme in the process.)

Design and Analysis of Experiments

The paper on the amount of experimentation (Grundy, Healy and Rees (197)) has been completed and presented at a meeting of the Research Section of the Royal Statistical Society. This deals with the case in which it is necessary to decide, after a preliminary set of experiments, whether to recommend one of two alternative processes or to carry out a further set of experiments, and if so what the size of this further set shall be, assuming that on its completion a decision

https://doi.org/10.23637/ERADOC-1-77

This work is licensed under a Creative Commons Attribution 4.0 International License.

must be taken. The work has aroused considerable interest amongst statisticians concerned with empirical experiments to test out new processes.

Work has continued on the reorganization of the long-term and classical Rothamsted and Woburn experiments. Patterson has been particularly concerned with this work and with the design and analysis of long-term experiments at Rothamsted and elsewhere. Boyd has co-operated with W. B. Haines in summarizing the factory series of sugar-beet fertilizer trials, and a report on this work is now almost complete.

There has been a growing 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. Reorganization of the computing side has been effective in considerably speeding up the production of the annual report of the results of the Rothamsted and Woburn field experiments. The electronic computer has played an increasing part in the analysis of standard types of design.

Miscellaneous work includes a paper by Hodnett on the analysis of an unusual experimental design (204) and one by Healy on the analysis of a factorial experiment with additional treatments (199). Hodnett has also discussed the number of levels that are best included in fertilizer trials designed to establish the response curve (203). Patterson has investigated the fitting of response curves of the exponential type (209). After some delay the report by Yates and Finney to the Government of India on statistics in agricultural research has now been published by the Food & Agriculture Organization of the United Nations (215).

The Department is represented not only on the Crop Experiments Sub-Committee of the Experimental Husbandry Committee of the Agricultural Improvement Council but also on the Animal and Poultry Sub-Committees, and has recently established closer links with the Agricultural Research Council Statistics Group at Cambridge, which is mainly concerned with animal problems. Work on the collation of the results of experiments on antibiotics for pigs is substantially complete, and a paper on the results of 3 years of a poultry experiment at Gleadthorpe Experimental Husbandry Farm is in course of preparation in co-operation with members of the National Agricultural Advisory Service. A report on the results of five experiments at three different laboratories on the assessment of bacterial populations in soil is being prepared with the Provincial Advisory Bacteriologist, Bristol, and a paper has been completed on the feeding of fodder beet to pigs (212). Westmacott has continued in charge of this work.

We co-operated with the National Institute for Research in Dairying, Shinfield, and the Advisory Bacteriologists of the National Agricultural Advisory Service in a series of trials to compare the standard hypochlorite disinfectant with some new detergent sterilizers that had been submitted to the Ministry of Agriculture, Fisheries and Food for approval (Leech). Leech also completed the analysis of a large experimental assay of new tuberculins on cattle and tests of their power to discriminate between tuberculous and tubercle-free cattle.

A fortnight's course for members of the National Agricultural

https://doi.org/10.23637/ERADOC-1-77

This work is licensed under a Creative Commons Attribution 4.0 International License.

Advisory Service, mainly on experimental design and analysis, was given in the spring. Sixteen N.A.A.S. officers attended.

Surveys

Surveys of fertilizer practice were carried out by the Provincial Soil Chemists of the National Agricultural Advisory Service in four districts of South-West England and in South-West Lancashire, and analysis is in progress. A report on the 1954 surveys has been prepared and will appear shortly. Eight of the districts surveyed during 1954 were previously surveyed in 1951; the results show average increases of 0·04 cwt. N and 0·05 cwt. K_2O/acre crops and grass in these districts. Phosphate consumption had changed very little except in two districts. The proportionate increases in use of fertilizers were greatest on grassland.

Church has summarized information from fertilizer surveys (1950–54) on the manuring of cereal crops (192). More than a third of the cereal acreage in England and Wales receives no nitrogen fertilizers, and the average dressings given are 0·25 and 0·35 cwt. N/acre, compared with a calculated optimum dressing of 0·6 cwt. N/acre. Profitable increases in the use of nitrogen on cereals might increase total cereal yields in England and Wales by about 6 per cent. Average potash dressings are about optimal, on the available experimental evidence, but phosphate fertilizers are used on much of the cereal acreage at rates which cannot be profitable. At variance with current advice, as much as a third of the nitrogen used on winter cereals is applied in the seedbed in autumn, and use of complete fertilizers results in the wasteful application of phosphate and potash to winter cereals in spring top dressings.

On the animal side, collaboration between F. B. Leech and Dr. F. W. Withers has continued in the survey of diseases of dairy cattle. Samples of farms from Devon, Salop, Yorkshire (W. Riding), and a region in Ayrshire and Lanarkshire have been supplied to the field workers. Results obtained in the 1953–54 season have been analysed. A combined report on the two years' work is planned. The results of a second survey of losses in pregnant ewes in Yorkshire, carried out on the same farms as the first survey, are being analysed by Leech and will be written up with Dr. K. C. Sellers, who carried out the survey. Leech has also advised on the analysis by punched-card methods of the results of a survey of animal husbandry and diseases now being carried out in Northern Ireland. The analysis will be done in Belfast.

Emily P. Simpson and Lessells have examined data on the errors associated with soil sampling for chemical analysis at the request of a committee of the Conference of Soil Chemists of the National Agricultural Advisory Service.

Other publications include a report on the results of a survey (described last year) of progress under the Hill Farm and Livestock Rearing Acts, undertaken by officers of the Agricultural Land Service and the National Agricultural Advisory Service (195), a paper by Church on the use of weed-killers and insecticides (191), and a report by Church and Westmacott in collaboration with Dr. F. H. Jacob on the results of the 1954 survey of rabbit damage to winter cereals (194).

https://doi.org/10.23637/ERADOC-1-77

This work is licensed under a Creative Commons Attribution 4.0 International License.

Colonial Work

Twelve Commonwealth workers have stayed in the Department for periods up to three months, and others have paid brief visits. Advice and assistance on the design and analysis of experiments and sample surveys has been continued. As already mentioned, Hodnett has now been appointed to the post of Statistician on a Regional Research Scheme based at the Imperial College of Tropical Agriculture, Trinidad and the future of this section is now under discussion with the Colonial Office. Hodnett completed his study of the responses of sugar-cane to fertilizers, and a general report is in the press (202).

Other Work

An investigation by Yates into the methods of analysis appropriate to a set of quantal experiments involving two treatments, which was substantially completed in 1954, was written up (213). A further note on a secondary point arising in the course of this investigation has also been prepared (214). The work on the electronic computer on the efficiency of the combination of probabilities test mentioned in the 1954 report originally arose out of this investigation. A report on this latter work is substantially complete, but pressure of other electronic work has delayed publication.

Healy has prepared notes on the test of significance for the difference in efficiency between two predictors (198), on weighted probits allowing for a non-zero response in the controls (200), on the prediction of adult body measurements from measurements taken in each year from birth to 5 years (in collaboration with others concerned in the Aberdeen Growth Study (211)) and on the chemical control of wireworms, with Potter and Raw (137).

Gower has prepared a note on a periodogram of the Beveridge Wheat Price Index arising from work carried out on the Manchester electronic computer (196).

Boyd and Lessells have prepared a note on the value of setting potato seed arising from their earlier work in this field (190).

Leech has contributed two papers on the subject of vital statistics in the study of cattle diseases (206, 207). He has also assisted in the elucidation of the sampling errors in observations on the behaviour of grazing cattle (210).

https://doi.org/10.23637/ERADOC-1-77