ROTHAMSTED RESEARCH

Report for 1962

Full Table of Content

Statistics

Rothamsted Research

Rothamsted Research (1963) *Statistics ;* Report For 1962, pp 173 - 180 **- DOI: https://doi.org/10.23637/ERADOC-1-95**

Emily P. Simpson, G. W. Bonsall, P. Holgate and M. L. Hill left and R. T. Clarke and C. W. Fielding were appointed to the staff.

Yates acted as consultant to the Seminar on Sample Surveys relating to Passenger Transport by Road organised in Geneva by the Inland Transport Committee of the United Nations Economic Commission for Europe. Church visited Cairo with Dr. C. Potter at the request of the Egyptian Government to advise on the planning of a series of laboratory and field experiments to test the alleged build-up of resistance to insecticides.

Rees attended the 1962 Congress of the International Federation of Information Processing Societies, held in Munich, as a member of the British Programme Committee, and Gower attended the Symposium on Symbolic Languages in Data Processing held at the International Computing Centre in Rome and gave a paper (12.8).

Healy spent five weeks in the United States, during which he took part in a conference on time series analysis at Brown University, Providence, R.I.

Eight temporary workers, five from overseas, spent varying periods in the department during the year.

The 401 and 402 Computers

The 402 computer, presented to us by Elliott Brothers late in 1961, was put into working order early in 1962, but linking the 401 card-reader to the 402 proved difficult and took considerably longer than expected. Three new machine instructions had to be introduced, namely: (i) stop (necessary to synchronise the card-reader and the computer); (ii) switch to card-reader; (iii) switch to tape-reader. When this was done and the link-up was complete reading failures occurred with certain digit patterns. To overcome these failures part of the computer had to be redesigned, a task made more difficult because standard 402 units were unobtainable and 401 units had to be used. After Rees successfully resolved these difficulties card-reading was very reliable, indeed, considerably more so than with the corresponding system on the 401. This has enabled us to follow our original plan of using one operator for both computers. To make the operator's task easier a further instruction was incorporated and a bell rings when the computer needs attention, e.g., because of card-reading failure.

The 402 computer has proved reliable in operation, but the tape-readers and punches have given frequent trouble, and one punch has been largely rebuilt. The tape-readers are still troublesome, the more so because the 402 lacks the checking facilities we incorporated in the tape-reading system of the 401. The installation of the 402 necessitated some control of the temperature of the computer room, and three air-conditioning units were

https://doi.org/10.23637/ERADOC-1-95

installed. Because of the difficulty and delay in getting the 402 to read cards, no attempt was made to link magnetic tape to it.

Tests of the transcription for the 402 of Part 1 of the 401 General Survey Programme were completed at the end of July (H. R. Simpson). Part 1 calculates any required functions of the observed variates, and compiles the basic tables. The subsequent processes, i.e., such operations as inserting marginal totals, calculating tables of means by dividing tables of totals by tables of counts, and punching the results on paper tape for final printing are still done on the 401 by Part 2 of the programme. Communication between the 402 and the 401 is by paper tape.

The 401 continues to operate very reliably. The only modification made in 1962 was the redesign of a section of the track-6 switching control system; this has eliminated occasional switching failures.

TABLE 1

Record of machine operation for 1957, 1961 and 1962

	1957	1961	1962		
			401	402*	Total
Total hours worked	2,622	3,688	3,274	852	4,126
Productive work, hours	1,305	2,666	2,467	471	2,938
Percentage overtime	32·6	86·5	65·2†	—	—
Percentage of time for productive work	49·8	72·3	75·3	55·2	71·2
Programme development, etc.	22·6	11·8	10·5	12·4	10·9
Idle time	4·4	0·3	0·4	7·5	1·8
Maintenance, etc.	23·3	15·6	13·8	24·9	16·1

* August–December. Both computers were shut down for a week after Christmas for redecoration of the computer room.

† January–July, 76·5%; August–December, 48·0%. The 402 was only rarely worked when the 401 was not running.

Table 1 gives a record of machine operation for 1957 (the first year in which full records were kept), 1961 and 1962. The 401 and 402 are shown separately. Figures for the intermediate years are given in the 1961 Report. The total machine hours worked by the 401 fell by 11% in 1962 and the productive hours by 7½%, but the corresponding totals for both machines together rose by 12% and 10% respectively.

The installation of the 402 considerably relieved machine congestion, and we can now do more survey analysis with much less delay. Up to the beginning of August, when the 402 began to operate productively, we had done 189 hours of survey analysis, which occupied 12% of all productive time. Since then we have done 459 hours of survey analysis on the 402, and a further 167 hours (mainly Part 2) on the 401, equivalent to 45% of all productive time on the 401 and 402.

This great increase in survey analysis during the latter part of the year reflects the fact that several surveys (in particular the *Brucellosis* Survey, the Survey of Fertiliser Practice and an investigation into the classification of farms for the Ministry of Agriculture) were ready for analysis. It illustrates the importance of having adequate computer capacity to deal with exceptional demands. Over the whole year, survey analysis absorbed 28% of productive time and analysis of experiments 30%, proportions very similar to those of 1961.

https://doi.org/10.23637/ERADOC-1-95

The percentage overtime worked from January to July was 76%, whereas from August to December it was only 48%. Without the 402, we should have had to work about 120% overtime from August to December to handle this work. Moreover, delays in doing particular jobs would have been much longer.

It became apparent in 1961 that some system of controlling the acceptance and progress of computer jobs was required. Clearly such a system will be even more necessary for the Orion, and the matter was given attention in the spring of 1962. A system was put into operation in June that, in addition to improving the flow of work and shortening delays, enabled us to lessen demands for computer time, by eliminating computing of little or no value, to an extent that surprised us.

The Orion Computer

Further delays have occurred in the manufacture of this machine. Delivery is now promised for August 1963. There have also been regrettable delays in the development by the manufacturers of the essential programmes without which the computer cannot be operated. This has slowed our own programme development, but there are good prospects of our being able to test our programmes on some of the Orions which will be operating before ours.

A contract was placed for installing the air-conditioning plant for the Orion computer room; in drawing up the specification and in the consideration of the tenders we were greatly helped by Mr. R. J. Fiddes of H.M. Treasury Technical Support Unit. The off-line Flexowriter equipment was also ordered.

Courses were arranged, primarily for statisticians, to describe the operating and programming facilities we intend to provide on our Orion. We originally planned to hold one course at Rothamsted in January 1963 and another in Aberdeen in co-operation with the A.R.C. Statistics Unit there, but in all some 70–80 workers from other institutes have applied to attend these courses, and a second course is therefore being held at Rothamsted.

Programming Developments

The General Experiments Programme for the 401, which was described in the 1961 Report, was completed in March 1962, and a paper on it was presented at the Cardiff meeting of the British Computer Society (12.22). The programme has proved its value in providing means for analysing uncommon designs not covered by our existing set of programmes, e.g., 3^4 quasi-Latin squares with missing plots, and it enabled us to construct a general routine for randomised block designs, including covariance, possibly with a factorial set of treatments involving two or three factors, and possibly also with duplicated plots for certain treatments. Covariance analysis is not included in any of our other programmes, which provide only the sum of products of the residuals, leaving the remainder of the analysis to be completed on a desk calculator. The General Experiments Programme was also used effectively for the analysis of a uniformity trial

https://doi.org/10.23637/ERADOC-1-95

on maize (12.23), and for testing the correctness of a collection of cyclic solutions of balanced incomplete block designs, tasks which illustrate its versatility in dealing with problems for which it was not designed. The General Survey Programme was similarly used to test the orthogonality of a set of five 12×12 Latin squares.

Any general programme is inevitably slower than one written to deal with one standard type of design. The General Experiments Programme is indeed very slow, the time factor for a randomised-block analysis, for example, being about 10. To a great extent, however, this is attributable to the lack of an immediate access store in the 401, to the additional delay in relay-switching to most of the tracks of the drum and to the absence of a divider (division has to be done by entering a sub-routine). Loss of time from these causes can be minimised in a programme for one standard type of design, with arbitrary limitation of plot numbers, etc., but not in a general programme which occupies a large part of the store, and which for convenience works with means instead of totals.

One of the main objects in writing the 401 General Experiments Programme was to provide a mock-up for a similar but more sophisticated programme for the Orion. It has served this purpose well. We have in fact decided to combine and extend the features included in our 401 General Survey Programme and General Experiments Programme, and have now worked out and written a complete specification for an Orion Survey and Experiments Programme. This is in essence a specialised autocode containing powerful instructions for manipulating and operating on tables, together with facilities for the easy input of data, and for printing the results in neat and well-annotated form. The work of programming the autocode has begun. (Gower, Simpson and Martin)

In addition to the Survey and Experiments Programme we are writing for the Orion a general programme for regression and multivariate analysis (Healy, 12.11) and one for probit analysis (Healy); also programmes to provide maximum likelihood solutions, and for classifying material from observations of several characteristics, on the lines of the 401 Classification Programme (Ross). An extended form of Mercury Autocode, which is being provided by Ferranti for Orion, will serve for programming special problems.

The organisation of work on a large computer such as the Orion, with magnetic tape on which both long-term records and standard programmes may be stored, and provision for time-sharing, i.e., working on two or more jobs simultaneously, is very much more complex than for a small computer such as the 401, where each job is self-contained and all that is necessary is to present the required programme and data tapes for the job to the computer. An operating system, to be satisfactory, must enable both programmes and previous records to be called up automatically by instructions at the head of each job tape, and must permit current data and results to be stored in retrievable form on magnetic tape if they are likely to be required later. Rees is working on these problems.

Apart from completing and testing the transcription for the 402 of Part 1 of the 401 General Survey Programme, no programmes have been written for the 402, as it was decided not to spend time programming this

https://doi.org/10.23637/ERADOC-1-95

computer but to use it only for survey analysis and for work for which programmes already existed. In addition to survey analysis, some crystallographic analysis was done on it for the Pedology Department, and Elliott Bros. have also used it occasionally.

Other programmes than the General Experiments Programme were written for the 401, including one for periodogram analysis (Martin), one for determining the minimum value of a function of two or three variables (Ross) and one to determine the distribution of the genetic correlation coefficient and heritability estimated by component analysis of full-sib families (Ross). We hope there will now be little further need for programming this computer.

Gower presented a paper to the Symposium on Symbolic Languages in Data Processing, held in Rome, on the concept of scanning and its introduction into autocodes to facilitate table manipulation (12.8). Healy gave a paper at the British Association meeting at Manchester on our contributions to the use of computers in research statistics.

Experiments

Slightly fewer experiments were analysed than in 1961 (Dunwoody, Emily P. Simpson, Frater). Table 2 shows the actual numbers. The fact that the increase of previous years was not maintained may reflect the diversion of the simpler types of analysis to computers recently installed in some universities.

TABLE 2

Numbers of replicated experiments analysed in the department

	Number of experiments			Number of variates on computer	Variates per experiment
	By hand	On computer	Total		
1934	115	—	115	—	—
1951	437	—	437	—	—
1955	384	419	803	834	2·0
1957	98	1,253	1,351	5,041	4·0
1959	67	2,649	2,716	11,102	4·2
1961	89	2,862	2,951	15,184	5·3
1962	107	2,613	2,720	13,423	5·1

We were again consulted on many problems of design and interpretation, and did much miscellaneous work on the computer on data arising from experiments. Patterson designed modifications for the Rothamsted and Woburn ley–arable experiments and the Woburn Market-garden Experiment to provide information on questions arising from previous results. He also completed the report started by the late Dr. H. H. Mann on the 1944–60 results of the Woburn Market-garden Experiment, and T. W. Barnes and Clarke summarised the 1955–62 results of the Woburn Green Manuring Experiment (see this Report, pp. 186 and 193). Boyd analysed the results of investigations by the National Agricultural Advisory Service soil chemists on the value of various methods of soil analysis, and the effect of block size on the errors of cereal experiments was reported (12.18). Vessey reported an investigation for treatment of acetonaemia in

https://doi.org/10.23637/ERADOC-1-95 pp 6

cows (12.10), and Leech, in collaboration with the Veterinary Laboratory, Weybridge, analysed some experimental records on Johne's disease.

In collaboration with Mr. J. A. Nelder of the National Vegetable Research Station, and members of other institutes, we are investigating methods of recording data in a form that can be fed directly into a computer, either by using some form of marking on cards that are subsequently punched automatically or by direct use of paper tape punches. An automatic digitising recorder is already in operation at the Glasshouse Crops Research Institute. For many purposes, fully automatic recording should be the aim, but suitable apparatus is not as yet commercially available.

Surveys

Field work for the large survey of fertiliser practice in 30 districts planned for 1962 is complete; results from 21 districts were received, of which 12 were analysed. A paper on the use of fertilisers on grassland based on the surveys done in England and Wales in 1957–60 (12.4) and two papers on the results of herbicide surveys (12.5, 12.6) were completed.

The National *Brucellosis* survey was analysed and a report prepared (Leech, Vessey). The main results of the cow mastitis survey were presented at a meeting arranged by the Agricultural Research Council, and a final report is being agreed for publication. Reports are also being prepared on the results of other animal-disease surveys. Field work for the national survey of calf husbandry and losses of calves started in October.

A report on the 1960 Interim Census of the National Dairy Herd analysed in 1961 for the Milk Marketing Board was published by the Board (12.17). Leech and Vessey contributed substantially to this report. We are also assisting the Board in an analysis of a survey of farm costs.

At the request of the Ministry of Agriculture we did a series of trial analyses on a sample of 6,000 holdings, using information provided by 4 June Returns. This work was undertaken to compare the relative merits of different methods of classifying farms into farming types. It is hoped that such a classification will provide a better basis for sampling farms for various purposes. This work was done at Rothamsted instead of on the Ministry's Deuce computer at Guildford, because our General Survey Programme is particularly suited to an exploratory investigation of this nature. (Church, Simpson and Yates)

Commonwealth and Overseas

The results of various experiments from overseas were analysed and we were consulted on various problems of experimental design. A Southern Rhodesian uniformity trial was analysed to find the optimal plot size for maize variety trials (12.23). A similar analysis of a much larger trial in Nigeria is in progress (Vernon). The analysis of 15 years' records from a cocoa uniformity trial on a 6,000-tree farm was completed. (Vernon and Morris)

Further data relevant to the forecasting of cocoa yields during the

https://doi.org/10.23637/ERADOC-1-95

next decade were analysed for Miss N. Tanburn, Department of Agriculture, Ghana, using the programme prepared in 1961.

Bonsall completed his analysis of the data dealing with dairy cattle in Israel; the amount of analysis required was much larger than expected. A report is being prepared by Dr. M. Soller. A report on the long-term experiment on coconuts for the Coconut Research Institute, Ceylon, was finished (12.7).

In co-operation with Mr. R. H. D. Sandford, and at the request of the E.C.A./F.A.O. Joint Agriculture Division, Church prepared a report for the Ethiopian Government on the livestock survey planned by him in 1961.

Leech advised on a serological survey of sheep in Tierra del Fuego designed to confirm the absence of foot-and-mouth disease.

Methodology

In addition to our work on computer methodology and programme development, the following contributions to general statistical methodology may be mentioned.

Healy collaborated in two important papers on time-series analysis (12.3, 12.14), resulting from research he did at the Bell Telephone Laboratories, New Jersey, in 1959/60. He also prepared two sets of tables: (*a*) a table of the fiducial limits of a variance component, the distribution of which was established by R. A. Fisher in 1936; this was computed on the IBM 7090 at the Atomic Weapons Research Establishment, Aldermaston (12.12); (*b*) tables for power law transformations, i.e., transformation of the form $z = (x + c)^p$, of which the square-root and logarithmic transformations are special cases (12.16).

Gower evolved a systematic method for computing the coefficients of variance components in a hierarchical analysis of variance (12.9). Yates collaborated with the late Sir Ronald Fisher in the preparation of a 6th Edition of *Statistical Tables*. This was delayed by other work and by Sir Ronald's death, but is now almost complete.

Healy prepared a subject index to the Kendall–Doig bibliography of statistical literature (12.13). A subject index for statistics has long been required, and publication of the bibliography has made this possible without excessive labour. It is intended to maintain the index on magnetic tape on the Orion.

Other Work

The many miscellaneous statistical investigations in which members of the Department collaborated include:

(*a*) A summary of records of wheat yields obtained by sampling in forty districts of England from 1821 to 1859. (Healy and Dr. E. L. Jones, Oxford University (12.15))

(*b*) A comparison of sampling methods for weevils in grain; two machine methods sampled adequately, but hand methods were unsatisfactory. (Ross and Mr. R. H. Howe, Pest Infestation Laboratory)

(*c*) An analysis of factors affecting conception rate in cows. This

https://doi.org/10.23637/ERADOC-1-95

involved fitting of constants to quantal data. (Leech and the Reading Cattle Breeding Centre)

(*d*) Derivation of rules for interpreting the tuberculin test. Discriminant functions were used. (Leech, Martin and M.A.A.F. Tolworth)

(*e*) Relation of lamb birthcoat and subsequent fleece type, etc., and the heritability of birthcoat. (Ross and Mr. J. Wilcox, Trawscoed Experimental Husbandry Farm)

(*f*) Estimates of heritabilities and genetic correlations for Large White bacon pigs. (Gilbert and Animal Breeding Research Organisation (12.21))

(*g*) Classification problems. Our Classification Programme was used to investigate the possibility of classifying soils by physical and chemical characteristics (see Pedology Department Section, p. 76) and by vegetation (Department of Agriculture, Nigeria). It was also used for the classification of *Aerononas* and *Achromobacter* (Medical Research Council Low Temperature Research Unit, Cambridge), *Oplonia* (British Museum), *Aerononas* and *B. Salmonicida* (Marine Laboratory, Torry), *Brucella* (Veterinary Laboratory, Weybridge), and psychiatric patients (M.R.C. Unit, Maudsley Hospital).

(*h*) Examination of various alternative hypotheses regarding mate-killer (*Mu*) particles in *Paramecium aurelia* (Ross and Dr. E. C. R. Reeve, Edinburgh University (12.19, 12.20)).

(*i*) Discriminant analysis of measurements of Primate scapulae. This investigation, on similar lines to the earlier one on measurements of teeth of human beings and the great apes, also gave illuminating results (Healy and Dr. E. H. Ashton, Birmingham University). A similar investigation on barnacles was started for Dr. H. Barnes, of the Marine Station, Millport.

(*j*) Pilot analysis of origin and destination data of road traffic, for planning of motorways, was done at Rothamsted because of the difficulty of getting it programmed on any other computer. It was handled by our General Survey Programme without any special programming. (Yates, Rees, Healy, and the Ministry of Transport)

(*k*) Miscellaneous work for the National Survey of Health and Development of Children, M.R.C. Units, and the National Foundation for Educational Research.

https://doi.org/10.23637/ERADOC-1-95