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 Report for 1966



Full Table of Content

Statistics Department

F. Yates

F. Yates (1967) *Statistics Department*; Rothamsted Report For 1966, pp 216 - 227 - DOI: https://doi.org/10.23637/ERADOC-1-14

F. YATES

N. G. Alvey, A. J. B. Anderson and F. B. Lauckner were appointed. Professor S. Lipton of the University of New South Wales, Australia, a member of the department from 1954 to 1959, spent six months of his sabbatical year with us. Mr. M. Weiss of the Canada Department of Agriculture visited us for two months. Four other workers spent various periods in the department.

F. Yates visited the Canada Department of Agriculture to advise on the development of computer use in research and on research statistical organisation generally. D. H. Rees was a member of the technical assessment committee set up to advise on the formation of the Edinburgh Regional Computer Centre, which the Agricultural Research Council is supporting. In connection with this work, he visited the I.B.M. research and development laboratories in the United States as part of a study of multi-access computer systems. Rees was appointed technical secretary of the recently formed Agricultural Research Council Computer Advisory Committee (Chairman: Professor M. V. Wilkes, F.R.S.), of which Yates is also a member.

General

The computer installation was considerably extended, first by the addition of a line printer, and later by components made available by the replacement of the Orion belonging to the Rutherford High Energy Laboratory. These extensions increased throughput per hour by more than 50%. Were it not for this, the computer would already be seriously overloaded. It is now run regularly till 10.30 p.m. and proper double-shift working will have to be organised if we are to continue our present standard of service and meet the rapidly increasing demands, exemplified by the fact that the throughput per working day during the last three months of 1966 was two and a half times the average for 1965. This increase in demand arises partly because our growing repertoire of programmes enables us to do more varied work and partly because of spreading awareness of the value of computers in agricultural and biological research, for both statistical and other work.

Some valuable contributions were made to multivariate theory and to the theory of experimental design, and good progress was made with the analysis of long-term experiments on the computer. Survey analyses and analogous work also increased substantially. There was a large survey of fertiliser practice in 1966, the preliminary results of which are discussed and related to the results of earlier surveys in a special article (p. 339).

The Ministry of Overseas Development provided a special post for a statistician to work on overseas problems, and we hope that this will enable us to extend such work, which is greatly needed.

The Orion computer

Additions to the installation. The Analex 1,000-line/min line-printer, fitted in March, proved most reliable and the quality of the printing is well maintained. Later we had an unexpected enquiry from the Rutherford High Energy Laboratory as to whether we could take over their large Orion installation, which was being replaced. Transfer to Rothamsted was impracticable because of lack of space and services. The possibility of installing it in the Edinburgh area for the Scottish research institutes was considered but turned down. As no other institute or university could house the complete installation, we were ultimately offered such parts as we could use. For the modest sum of less than £10,000, which included reinstallation and recommissioning, we added the following items to our installation:

1 module of core store (4,096 words)

2 drums (16,384 words each)

4 TM2 Ampex tape decks

1 magnetic tape switching unit

1 TR7 1,000-character per second paper tape reader

1 Control flexowriter

Additionally we were given extended loan of 50 magnetic tape reels. The reinstallation and reconditioning were completed on schedule by Orion maintenance specialists within three week-ends and five normal working days. This was a splendid effort.

Unfortunately lack of space prevented us from accepting the extra lineprinter, which would have further helped us to increase throughput and shorten turnaround time, as two time-shared programmes could then use line printers simultaneously. As it is, the results of many jobs have to be punched on paper tape and then printed on the line-printer by a separate run through the computer. We would also have liked to have installed an extra character control unit, as this would have enabled us to make experiments with the use of flexowriters on line for direct communication with the computer.

As the theoretical total heat load with the additional equipment is above the maximum specified loading of the air-conditioning plant, the adequacy of the air conditioning for this additional equipment was in doubt. Before accepting the offer we made some preliminary tests with added heat load. These indicated that the plant would be adequate under almost all circumstances; if in exceptional conditions trouble occurs, "load-shedding" can be adopted by shutting down one or more tape decks.

The air-conditioning plant was also recalibrated, and this greatly improved its performance. Despite reference in the original detailed specification of the plant to the troubles likely to be encountered because of the mineral content of the local water supply, the suppliers under-estimated their effect on operational efficiency. These troubles have now been cleared up by fitting a resin-water-exchange purifier.

The incremental plotter was fitted in October, but still requires further

operational programmes to make it really useful. Martin is working on this.

The establishment of a data link with the National Vegetable Research Station was delayed in the hope that equipment capable of making use of the higher transmission rates now possible on telephone lines might become available. However, this has not occurred, and it has now been decided to instal an ordinary Telex link. A similar link will also be arranged with East Malling Research Station.

Performance. Demands on the machine are increasing rapidly, and much more overtime was worked than last year. Although the maintained time was slightly less in 1966 than 1965, because of time required for modifications to the system, the unmaintained time was more than doubled, resulting in an increase of 16% in useful time (Table 1). The machine is

TABLE 1
Orion performance, 1965-6 (hours worked)

	Maintained		Unmaintained		Total		%
	1965	1966	1965	1966	1965	1966	(1966)
Useful time	1,543	1,426	374	804	1,917	2,230	90
Computer faults	159	146	9	23	168	169	7
Restarts after faults	35	52	4	15	39	67	3
Total	1,737	1,624	387	842	2,124	2,466	100
Maintenance					*	752	
Total running time					*	3,218	

^{*} Not available.

now run every day until 10.30 p.m. to the great credit of the staff concerned.

The time lost by computer faults and subsequent restarts was 10% of nominally available time, very similar to 1965. The maintenance time required by the engineers amounted to 23% of the total time during which the machine was switched on. As previously, most of the faults were transient, so that use of the machine could continue after restarting procedure. This makes unmaintained late running practicable.

Troubles with magnetic tapes increased somewhat, partly from increased use of magnetic tapes and partly from wear of tapes. There is no practical way of forecasting the life of a magnetic tape, but detailed records of the performance of all tapes are maintained so that defective tapes can be copied and examined for faulty sections.

Volume of work. The increase in the size of the installation has resulted in a corresponding increase in the amount of work that can be handled in a given time. The best measure of work done is provided by the accounting costs, which are calculated for each job by the special programme used by I.C.T. for their commercial costing. The actual cost at the present rate of working, allowing for depreciation, maintenance and operating staff, is about half the commercial accounting cost. Table 2 shows the nominal value of work done in 1965 and 1966. (The total figure for 1965 was incorrectly given in the 1965 Report as £180,000.) There are several 218

TABLE 2
Work done, 1965-6 (Nominal value, £,000)

	1965		1966	
		JanSept.	OctDec.	Total
Development	85	88	33	121
Production	69	86	38	124
Systems	9	8	3	11
Total	163	182	74	256
Work per hour of useful time Production per working day	£85 £270	£104 £460	£155 £670	£115 £510

interesting points about these figures. The amount of productive work in 1966 increased by 80% over 1965. During the last three months of 1966 the productive work per working day increased by 148% over that for 1965. For the first time also in this period, productive work exceeded programme development. With our wider range of programmes and the increasing use of the machine on non-statistical work, the demand can be expected to continue to grow rapidly. Organisation of regular maintained two-shift working is therefore necessary to meet the demands. We are now negotiating for an additional maintained shift, arranged to make the computer available from 9.0 a.m. until 11.0 p.m. daily. Extra staff will be required to provide a continuous operating service.

The work done per hour of useful time also increased greatly in the last three months of 1966 with the larger installation (82% more than 1965). Were it not for this we could not have met demands. The smaller increase per hour of useful time in the first nine months of 1966 is mainly attributable to the installation of the line-printer. To some extent this reflects a transfer of costs from off-line flexowriter printing to the computer, but it also reflects some real increase in work done resulting from faster output with the line-printer.

Table 3 shows the use of the computer by the various institutes. Direct

TABLE 3
Use of computer by institutes, 1966 (% total accounting costs)

Institute	Development	Production	Total
Rothamsted Experimental Station	31.9	21.2	53.1
Scottish Group	4.4	0.3	4.7
National Vegetable Research Station	1.5	2.9	4.4
East Malling Research Station	4.6	1.1	5.7
National Institute of Agricultural Engineering	1.5	2.0	3.5
Glasshouse Crops Research Institute	0.3	1.2	1.5
Grassland Research Institute	0.6	1.3	1.9
National Institute for Agricultural Botany	0.0	0.7	0.7
Others	2.5	3.4	5.9
Total	47.3	34.1	81.4
Experiments (Production, Rothamsted program Systems	mes)		14·2 4·4
			100.0

use by other institutes was slightly less than in 1966 because of much less use by the National Vegetable Research Station, attributable to Mr. J. A. Nelder being in Australia and loss of other staff. These figures, of course,

do not reflect the amount of work done for other institutes. Any work controlled by Rothamsted appears under the Rothamsted heading. Further, for operational reasons all analyses of replicated experiments run on Rothamsted programmes are grouped together.

We now keep separate records of work requiring the use of survey analysis programmes, including development of the steering programmes for specific surveys; this amounted in 1966 to 17% of the total throughput, i.e. very similar in amount to that used for the routine analysis of experiments. This item can be expected to increase faster than experimental analyses, at least until we succeed in making our survey programme usable on other installations.

Systems development. As a result of continued pressure on I.C.T., and with support from other Orion users, the E.M.A. Compiler has at last been modified to provide extended facilities for integer working. This is an important addition. Previously there were only 24 integer registers, and these could not be treated as a series with modifiable addresses. Now an indefinite number, with modifiable addresses, can be made available.

We are organising the reading of E.M.A. programmes from magnetic tape. This is now working for routines, and a library of routines is being prepared; previously the use of paper tape as the only means of storage made the organisation of any central library of routines impracticable. We hope shortly to store on magnetic tape programmes that are being developed, with facilities for correction and for assembly of the component parts. All this has at present to be done on paper tape, a time-consuming and tiresome task.

The extension of the installation by the components from the Rutherford Laboratory enables us to use the Fortran compiler written by the staff of that laboratory with assistance from I.C.T. We are now gaining experience with it. The ubiquity of Fortran as a programming language makes this compiler a valuable addition to our installation, though one that will undoubtedly increase the load on those responsible for systems organisation.

Programming developments

General. The problem of making effective use of computers in statistical research work was discussed in the First Fisher Memorial Lecture (12.14), which stressed the need for better general programmes for standard statistical analyses of the many varied types required in practical statistical work. Fortunately this need is becoming increasingly recognised. The design of such programmes, however, requires thorough knowledge both of the statistical problems involved and of computer techniques; their writing and development is a major task, requiring hard work and skill in programming. Clearly, as was emphasised by the Flowers Committee, every effort should be made to co-ordinate this work to avoid different groups writing programmes for much the same tasks for computer installations with similar facilities. Such co-ordination implies that programmes must be made available to other installations, and this frequently necessitates translation into another language.

Rothamsted can make a useful contribution to this co-operative effort in the branches of statistics in which we specialise. We are therefore encouraging others to translate some of our more important programmes into other languages, particularly Fortran, and are also attempting to get them made available on university installations with Extended Mercury Autocode Compilers. This will have the additional advantage that much routine work from university departments that at present is sent here can be dealt with by the university concerned.

Programmes for the analysis of experiments. The programme for the analysis of general factorial designs (12.15) abundantly proved its worth, and we have had enquiries about it from various other computer departments, both in this country and abroad. The Canada Department of Agriculture Statistical Research Service, in particular, is hoping to get a Fortran version running on an I.B.M. 360, and Mr. Milton Weiss of that Department spent two months here working on this task. An extensive revision was made that incorporates several improvements and additions, including storage of results on magnetic tape so that sets of experiments can be subsequently summarised. A supplementary programme will be written for making such summaries. Apart from the final organisation of the magnetic-tape facilities and correcting a few minor errors, the revised version is now finished and is doing productive work. (Yates, Preece and Alvey)

A new version of the General Input for Experimental Designs (G.I.E.D.) is nearly complete and will be incorporated in the general factorial programme and other experimental programmes. This provides for reading in data from cards as well as paper tape and for calculating linear functions and orthogonal polynomials. (Martin)

A programme for the analysis of long-term experiments is now working. This deals with experiments containing factorial treatments, possibly with split plots, provided there is sufficient orthogonality between plots and years to permit the plots and years × plots components of error variance to be estimated from the unadjusted plot totals with reasonable accuracy. and provided also that the design gives complete years x treatments tables of yield values. The programme is in two parts which can operate independently. The first reads the data from paper-tape documents (including outputs from other programmes), rearranges them to give years X treatments (or other factors), tables of yields and external variates and stores the results on magnetic tape, together with external variates such as meteorological factors. Tables of means can be updated as required. The second section deals with the analysis and presentation of results. Basically the analysis consists of the estimation over specified years of mean effects, linear, quadratic, etc., trends and orthogonal regressions on external variates together with the appropriate analysis of years × treatments variance. Standard errors of means are calculated from estimates of the plots and years × plots error variances; standard errors of regressions also include appropriate components of the residual years x treatment variance. The main results of the analysis are stored on magnetic tape so that printing and further analysis can be done when required. (Patterson, Lowe)

Preece revised a programme written by Mr. D. H. Rees, East Malling Research Station, for the analysis of balanced incomplete block designs, Youden squares and balanced lattices, to provide for recovery of interblock information. Ross wrote a lattice and lattice-square programme with hierarchical analysis of variance and covariance.

Survey and experiments programme. The Survey and Experiments Programme (SEP) is now sufficiently complete to be used effectively on many types of job, both survey analysis and more general statistical problems. It is proving very useful for special jobs, and enabled us to write a general hierarchical analysis of variance programme very quickly. The programme was described at a meeting at the Atlas Laboratory, Chilton (12.9). It gives us useful experience on what is required in a general statistical language for future machines (12.8).

General survey programme. As mentioned in the 1965 Report, the Survey and Experiments Programme is too slow for large survey jobs. For these our General Survey Programme (G.S.P.) written in Extended Mercury Autocode is used. Although this programme is fast, the analysis of each survey requires much detailed specification, and the provision for headings and layout of the tables of results is poor. An extensive revision is therefore being made, and is well advanced. The revised programme will greatly simplify the specification of the analysis of particular surveys, will provide for better titling and layout of tables with a minimum of trouble to the user and will simplify successive analyses on the same survey. Some additional facilities are also being added. (Yates, Anderson and Prof. S. Lipton)

The revised programme is being translated by the A.R.C. Statistics Group, Edinburgh, into Atlas autocode for use on the Edinburgh KDF9 computer, and the Atlas Laboratory have expressed interest in getting it running on their computer, which has an E.M.A. compiler. It is hoped also that a Fortran version will be made available.

Fitting constants to multiway tables. This programme was successfully extended to provide for the fitting of constants to quantal data, and both the original programme and the extension proved of great value. They successfully handle tables produced by G.S.P. or S.E.P. stored on magnetic tape (12.11).

Multivariate and discriminant analysis. A programme was written to evaluate canonical variates on crop data by the Q-technique, using the method evolved by Gower. The latent roots and vectors programme was extended to include discriminant analysis on grouped multivariate data. (Lewis)

Construction of experimental designs. The construction of designs for large multifactorial experiments with confounding is a tedious task and liable to error. Patterson wrote a programme that produces designs in which the number of levels of all factors is given by powers of a single 222

prime number, including split plot designs. The randomisation required for the full plan can be constructed by the computer with optional restricted randomisation for blocks of 8, 9 or 16 plots. Ross wrote a programme to generate large permutations of up to 5,000 items and their inverses.

Miscellaneous. Various programmes written for general use and for special jobs included the following: orthogonal polynomials, data plotting on the line printer, display of the likelihood function, special maximum likelihood routines for the general maximum likelihood programme. (Ross)

Non-statistical programmes. A set of crystallographic programmes was written by J. H. Rayner of Pedology Department and is proving of value both to the Pedology Department and to the new A.R.C. Unit of Structural Chemistry.

Various programmes were written for members of the Physics Department, and for workers at other institutes, particularly the Institute of Animal Physiology, Babraham. These last are concerned with the transmission of impulses along nerves, amino-acid determinations, gaschromatographic techniques and cortisol metabolism. (Rees and Bicknell)

Experiments

There was further considerable increase in the demand for routine analyses of experiments (Table 4). The number of experiments analysed on the

TABLE 4
Numbers of replicated experiments analysed in the department

	Number of experiments			Number of variates	Variates per
	By hand	On computer	Total	on computer	experiment
1951	437	-	437	_	B
1957	98	1,253	1,351	5,041	4.0
1963	72	2,770	2,842	14,357	5.2
1964	88	3,383	3,471	18,054	5.3
1965	69	4,751	4,820	28,663	6.0
1966	109	6,162	6,271	40,826	6.6

computer increased by 30%, and the number of variate analyses by 42%. Rothamsted programmes were used for 95% of the experiments.

The Orion began to make a useful contribution to this work towards the end of 1964. Over the last three years the number of experiments has more than doubled, and the number of variate analyses has nearly trebled. Much of the increase in 1966 was in experiments for which the data were punched at other institutes. The number of experiments punched at Rothamsted increased by 12%, those punched at other institutes by 80%. The tendency for requests to be made for more variate analyses per experiment in experiments punched elsewhere continues; those punched at Rothamsted averaged 5·2 variates per experiment in 1966, those punched elsewhere 9·2 variates per experiment. This is in part accounted for by horticultural experiments, but it may well be that some unnecessary analyses are being demanded by other institutes.

The development of a general programme for the analysis of long-term experiments has enabled better progress to be made with such analyses. Particular experiments dealt with this year were the forestry long-term experiments for the Chemistry Department, a long-term nematological experiment for the Nematology Department and the Woburn Green Manuring experiment. (Patterson and Lowe)

A complicated experiment on the reactions of sheep to excessive cold was analysed for the Animal Breeding Research Organisation by Ross, using a special programme.

National Agricultural Advisory Service

In addition to routine analysis of experiments on Experimental Husbandry Farms, Experimental Horticulture Farms and commercial farms (Dunwoody, Hill and Lessells) we co-operated in various special investigations. Of these, several were in connection with recommendations for manuring agricultural and horticultural crops being prepared by National Agricultural Advisory Service Soil Scientists. Arising from this work, a paper was prepared on the effect on yield of maincrop potatoes of different methods of fertiliser application (12.2). Different methods of determining the N content of soils were also compared.

A final report on the 44 trials of Charolais cattle was prepared. A paper on the results of National Agricultural Advisory Service experiments on the progeny testing of beef bulls (12.10), and various notes on the design, conduct, analysis and interpretation of animal experiments were written; also reports assessing the effects of environmental factors on relative yields in the winter wheat and spring barley variety trials done jointly by the National Agricultural Advisory Service and National Institute of Agricultural Botany, with suggestions for modifying the experimental programmes. (Lessells and Macfarlane)

Surveys

The results of a large survey of fertiliser practice covering 31 districts, done in collaboration with the Regional Soil Chemists of the National Agricultural Advisory Service and representatives of the Fertiliser Manufacturers Association, are being analysed. Fifteen of the districts were also surveyed in 1962, and consequently the present survey should give good information on recent changes in practice. Specially detailed information on grassland was obtained, to relate fertiliser practice to the way the grass was used (see pp. 346–348).

The analysis of the Calf Wastage Survey is now nearly finished, and most of the report written. The fitting constants programme for quantal data proved useful in disentangling the effects of different factors on calf mortality rates. (Leech)

Surveys of the disposal of sheep-dip residues were made in 1965 and 1966 by the Ministry of Agriculture's Animal Health Division. The analysis of the first survey is almost complete, and most of the data from the second survey have been received. The main purpose of these surveys 224

is to find out whether or not sheep-dip residues are likely to be a hazard to wild life. They will also provide some useful incidental information on current practices in sheep dipping. The data for the long-term trial on Johne's disease (on about 50,000 cards) have been received from the Ministry of Agriculture and are being analysed. (Leech)

The analysis of the National Economic Survey of the Wheat Crop organised by the Farm Economics Branch of the Cambridge School of Agriculture was completed; the Farm Economics Branch are now writing the report. We are analysing a survey on the use of insecticides and fungicides on soft fruit for the Ministry of Agriculture's Plant Pathology Laboratory, and a quality assessment survey for the Potato Marketing Board. The investigations for the Economics Division of the Ministry of Agriculture are progressing well. They are being done on the Orion in co-operation with Mr. I. Sadler because the Ministry's computers have no programmes suitable for exploratory investigations. (Church and Hills) A survey of farm buildings by the National Institute of Agricultural Engineering is being analysed. (Simpson)

The analysis of surveys still takes longer than it should, and compares poorly with the service we can offer for the analysis of replicated experiments. The revision of the General Survey Programme now being undertaken should do much to simplify the specification of the analysis and so speed the machine processing of simple surveys, which is frequently delayed by trivial programming errors that can be detected and put right only by repeated test runs. However, we still have much to learn about the best techniques for analysing complex investigational surveys. There is much more to good survey analysis than producing tables of counts and means, which, taken at their face value, are often very misleading.

Commonwealth and overseas

P. Walker was appointed to a special post provided by the Ministry of Overseas Development for a statistician to work on overseas problems and make short-term visits overseas. This appointment will enable us to give better attention to the needs of developing countries, and in particular to develop further the use of the computer in the analysis of agricultural experimental data from overseas.

Much work was done for Commonwealth countries. In addition to the routine analysis of straightforward experiments (Dunwoody), Ross handled various more complicated problems, in particular the analysis of the 1966 results of the lattice-square cotton-variety trials for the Empire Cotton Growing Corporation in East Africa, crop—weather regressions (Barbados), insecticide trials (Nigeria), crop—soil regressions (Uganda), cocoa regressions (Uganda), bagworm observations (Natal). We analysed and reported on a set of fertiliser trials from Ceylon organised by the F.A.O. Freedom from Hunger Campaign. (Boyd and Hill) Similar data came from Ghana for analysis. There was also an urgent job on the relation between solar radiation and transpiration from the leaves of a cotton crop for Mr. D. Rijks of the Cotton Research Corporation of Uganda. (Martin)

P 225

Theory

Gower continued his investigations of multivariate analysis, and evolved a Q-technique, which has some advantages over the usual R-technique for evaluating canonical variates (12.6). Papers on some distance properties of latent root and vector methods and multivariate analysis and multi-

variate geometry were prepared (12.5, 12.7).

Patterson considerably extended the theory of serial factorial designs, that is designs, cyclic over the years, for long-term experiments in which the same treatment factor or factors are included as separate factors in successive years, e.g. with a single factor N at three levels and 27 plots, all combinations of the three levels of N in each of any three successive years will be included. Methods were found for generating suitable designs, with confounding. This advance will considerably simplify the design and analysis of long-term experiments of various kinds.

Preece continued his investigations on balanced incomplete block designs, etc. (12.12, 12.13). Ross developed the theory of large permutations for the Department of Genetics, Birmingham University, and made some further investigations into maximum likelihood estimation. He also solved the problem of fitting new objects into a basic classification for which the distances or similarities are defined. Occupancy theory was

used to simulate nematode behaviour. (Macfarlane)

Various theoretical points in connection with fitting constants to quantal data were cleared up. (Yates and Lewis)

Other work

Many statistical problems arising in other Rothamsted departments, and some from other institutes, were dealt with, including an investigation on population dynamics of the potato cyst nematode Heterodera rostochiensis. (Ross) Classification problems, using Gower's classification programme, were investigated for the British Museum (Phthiracaroidea), National Agricultural Advisory Service (Bacterial phages), Wright-Fleming Institute of Microbiology (Streptococci), Portsmouth College of Technology (Lepidoptera), Torry Research Station (Corynebacteria, Achromobacter), Institute of Larynology and Otology (Auditory and Odour Maladies), University of Birmingham (Psychiatric Patients), Medical Research Council Microbial Systematics Research Unit (several small jobs), Hartley Laboratories, Liverpool (National Fruit Trials), and the Royal Free Hospital, Hampstead (Epileptic symptoms). (Gower and Ross) We continued the routine processing of results of herbicide screening trials for the Weed Research Organisation, and gave them general statistical assistance. (Patterson, Preece and Lowe)

Simpson assisted in the analysis of further data from the National Survey of Health and Development of Children and is co-operating with Dr. J. W. B. Douglas in a book on the survey results relating to secondary education. Church is assisting Mr. T. H. Hollingsworth of Glasgow

University in an analysis of Scottish population movements based on National Registration records and Ministry of Health records. Mr. Hollingsworth turned to us because the survey programme on the Chilton Atlas was not versatile enough for his needs. The job is of interest as a test of the adequacy of our own General Survey Programme.