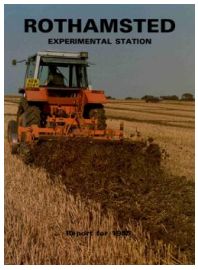
*This work is licensed under a Creative Commons Attribution 4.0 International License.*

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

# Rothamsted Experimental Station Report for 1985

*Full Table of Content*




## Biomathematics Division

### J. C. Gower

J. C. Gower (1986) *Biomathematics Division ;* Rothamsted Experimental Station Report For 1985, pp 75 - 89 **- DOI: https://doi.org/10.23637/ERADOC-1-26**

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

# BIOMATHEMATICS DIVISION

**STAFF**

**Head of Division J. C. Gower**

## Computing Unit

*Head of Unit*
P. J. Verrier, B.Sc.

*Senior Scientific Officers*
K. Bicknell, B.Sc.
W. J. Moore, B.Sc.

*Higher Scientific Officer*
Christine F. Thomas, B.Sc.

*Scientific Officer*
J. Summerfield, B.Sc.

*Assistant Scientific Officers*
J. W. Hipgrave
N. R. Shah

*Executive Officer*
R. J. Higgins, B.A.

*Honorary Scientists*
F. G. W. Jones, Sc.D.
F. Yates, Sc.D., C.B.E., F.R.S.

*Specialist Typist*
Dawn E. Dudley

*Senior Data Processor*
Denise V. Legg

*Data Processors*
Christine I. R. Alvis
Gwyneth Frier

## Statistics Department

*Head of Department*
J. C. Gower, B.A.

*Senior Principal Scientific Officer*
B. J. George, B.Sc.

*Principal Scientific Officers*
Rosemary A. Bailey, D.Phil.
J. N. Perry, M.Sc.
R. W. Payne, B.A.
Janet Riley, M.Sc.
G. J. S. Ross, B.A.
H. R. Simpson, M.A.

*Senior Scientific Officers*
P. G. N. Digby, B.Sc.
P. W. Lane, M.A.
K. Ryder, B.Sc.
H. H. Spechter, Ph.D.
A. D. Todd, M.Sc.

*Higher Scientific Officers*
Anne E. Ainsley, Ph.D.
C. J. Dyer, B.Sc.
Penelope K. Leech, B.Sc.
A. W. A. Murray, M.A.
R. P. White, B.Sc.

*Scientific Officers*
Janet I. Elsmere
S. A. Harding
S. P. J. Kirby, M.Sc.
Susan J. Lloyd, B.Sc.
Gina L. Smith, B.Sc.

*Assistant Scientific Officer*
Suzanne J. Clark

*Departmental Secretary*
Valerie Payne

*Personal Secretary*
Diane P. Dawkins

*Sandwich Students*
Anne-Marie Dillon
Rosemary F. Poultney

*Visiting Scientists*
S. Iwatsubo, D.Eng.
F. J. Rohlf, Ph.D.

*Specialist Typists*
Deborah Game
Angela C. Piears

*Senior Data Processors*
Anne E. Arnold
Brenda J. Watler

*Data Processors*
Wendy L. Barthorpe
Karen J. Hakesley
Belinda McVay
Belinda J. Smith
L. Vera Wiltsher

## INTRODUCTION

This year has been a busy one for both sections of the Division: the Computing Unit (RESCU) and the Statistics Department.

RESCU has continued to manage its section of AGRENET, consisting of two VAX 750's (now heavily overloaded) and an increasing number of microcomputers and terminals. The effort in supplying, installing and maintaining data-logging equipment and laboratory instruments has again increased, perhaps in part in response to reductions in scientific support staff. This year the Unit has also been concerned with the Station's word-processing policy and training, equipment for image analysis, the new telephone system and automation for the new glasshouses.



https://doi.org/10.23637/ERADOC-1-26 

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



The Statistics Department has continued its main functions of advising on statistical matters and analysing data from field-studies, surveys, laboratory and field experiments. This year the consulting services have been reorganized. The rota-statistician scheme for handling day-to-day queries has been abolished. Now every department (or group of departments) has been assigned a senior statistician supported by an assistant. It is hoped that this will provide a better service and provide a more stable environment for junior people to gain experience in statistical consulting and collaboration. It is planned to cycle junior people every two or three years to give them the opportunity of experiencing different kinds of scientific work and their varied statistical problems. Advice to and analyses for ADAS have continued as in previous years and the ODA-funded Unit successfully continues its work and on a trial basis is now also dealing with perennial crops. A system for discussion and approval of proposed experiments, similar to the Commodity Groups, has been set up to commence with the 1986 experimental programme.

Statistical research arising from these commitments has continued. This has been a particularly active year in the further development of statistical software. In Part 2 of this report, Gower (1985) describes the development of statistical computing at Rothamsted over the last 30 years.

## COMPUTING UNIT

The Computing Unit provides a design, development and advisory service for Rothamsted departments. A Systems Group supports data logging, small systems building, the running of two DEC VAX 11/750 supermini computers and word processing facilities throughout the campus. (Moore) An Applications Group provides software for a variety of computers to assist scientific research. (Bicknell) To continue existing work with fewer support staff many researchers have turned to computing methods. This has given the Unit a large workload in a variety of challenging and exciting fields. However, much remains to be done to catch up with outstanding requirements.

### Services

The Systems Group now provides service support for:

Twin DEC VAX 11/750 supermini computers, over 60 multi-access terminals, a local communications network, a Quantimet Image Analysis system, a Jacquard word processor system and terminals, several Scientex and DECmate stand alone word processors, the Statistics Department data-preparation systems, the Computing Unit data-preparation system, over 50 micro systems supporting research, media conversion tasks and data-preparation.

The Group was restructured during the year to cope with the ever-growing equipment base and the reduced need for a key-to-disc data-preparation service. The VAX 750s, (part of the AFRC computing network) are very much overloaded, each using more cpu time than any other in the AFRC. Although 1 Mbyte of memory and a floating point accelerator have been added to one VAX 750, either further on-site VAX resources are needed or arrangements must be made to shed the load to other AFRC sites.

### Development

Much design and development work was undertaken this year concerning VAX and microcomputer applications, data-loggers and equipment control applications.

Significant developments in graphics applications have brought improvements in SURFACE II and GHOST for Entomology and Soils and Plant Nutrition, AREAS for Soil



https://doi.org/10.23637/ERADOC-1-26 

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



Survey of England and Wales (SSEW), Genstat Mk. 4 and Mk. 5 graphics utilities, R$TRANPLOT (an overhead transparency production package), R$DIGITIZE, R$SIG-LIB (a library of routines for the Sigma colour graphics display) and an interactive colour mapping display system. (Bicknell, Summerfield, Hipgrave and Thomas) Investigative work has been undertaken with GKS, now the graphics display international standard, and with image analysis techniques—in preparation for a proposed replacement image analyser for Rothamsted. (Bicknell and Hipgrave)

The availability of good database packages on the VAX systems and on many microcomputers has led to developments of an insect data system, equipment inventory, work scheduling, laboratory data collection and an administration data-base. (Thomas, Higgins, Summerfield and Verrier) The Columbia University file transfer program Kermit, widely used for transferring data between many types of computing equipment, has been implemented on the HX20. (Moore) Changes in the financial fortunes of Rothamsted require rapid reappraisal and forecasting of the Station's budget. The Administration's TI 990/20 computer software proved inadequate for the purpose and a program suite to provide accurate current and forecast staffing costs was written for a TI PC microcomputer. (Verrier)

Varied data-logging activities have included the following. A system based on a CP/M microcomputer which prompts a remote logger placed in a field at the National Vegetable Research Station to send data over a telephone link to the microcomputer at Rothamsted. Data transmission errors are noted and re-transmitted automatically until correct. Soils and Plant Nutrition collect some 10000 characters of data daily using this system. (Verrier) The fertilizer nitrogen and nitrogen fixation experiments of Soils and Plant Nutrition have been aided by automating the weighing and validation of samples. The researcher prompts for missing items and produces an appropriate analysis on completion of sample input. (Hipgrave) Two scintillation counter logging systems based on BBC microcomputers have been developed for Biochemistry and Insecticides. (Moore) A data capture system based on HX20 microcomputers and an Apple microcomputer, has been developed for Physiology and Environmental Physics to measure leaf area and weigh samples. The Apple automatically loads the HX20 systems with the appropriate control software for its current task and collects the data into a matrix; a prompt is given to supply any values found to be missing. (Moore)

Automation of the SSEW laboratory is nearly completed. Based on a 68000 microprocessor running the Unix-like operating system Idris, this system monitors soil samples through a complete laboratory analysis, controlling turntables and Epson HX20 systems and reads balances, a pH meter and a colorimeter. Data captured is held in a database which after validation is merged with the main SSEW database held on the VAX systems. (Summerfield)

A system (based on an Apple with colour graphics) has been designed and developed for the constant environment cabinets, which enables cabinet temperatures to be set up and monitored. Out-of-range values initiate temperature logging until temperature falls into range; trials indicate further refinements. (Verrier) A Victor Automation environment control system was installed in the new glasshouses to control and monitor house conditions and to provide an interface for data-logging; this has now been proved. (Verrier)

## Word processing

Almost all the Station's word processing has been performed in a multi-user environment on two large Jacquard systems one of which also provided a data-preparation system for the Statistics Department. The inadequacies and age of the Jacquard systems led to consideration of their replacement. After considerable effort the Scientex software running on Sperry



https://doi.org/10.23637/ERADOC-1-26 

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



PC stand alone microcomputers was selected for scientific word processing (and has now become an AFRC standard) and the DECmate system was selected for non-scientific word processing.

Five DEC Rainbow systems (for Statistics Department data preparation), four Scientex and two DECmate word-processors have now been deployed, enabling one Jacquard to be removed. The remaining Scientex systems should be installed before April 1986 enabling the second Jacquard to be removed. A DEC laser printer provides a high quality printing facility for all word processor and VAX users. (Dudley, Higgins and Verrier)

### Staff

Marjorie Bellingham retired on 28 February 1985 after 19 years at Rothamsted. D. Jessop (six-month sandwich studentship) joined the Unit in July 1985 to work on graphical software. J. Summerfield gained a B.Sc. in Computing Science from Hatfield Polytechnic and is now proceeding to Honours level. Dr F. Yates was made an Honorary Member of the International Statistical Institute. K. Bicknell is a member of the GHOST Technical Committee and a member of the Eurographics (UK) Executive Board.

## STATISTICS DEPARTMENT

### Practical applications

In this section a sample of the work done for Rothamsted departments and for some outside bodies is described.

A screening procedure for assessing the results of tests of new compounds to control slugs has been designed. A large number of compounds is assayed for their potential on controlling slug damage to germinating cereals. Assessments are made in test-tube cultures and, later, in trays of soil. The presence or absence of damage is noted. Binomial distributions have been used to determine threshold values for acceptance of compounds for further investigation. (Murray, with Woodcock and Pickett, Insecticides and Fungicides)

Differences in resistance of eight strains of fly to a set of nine pyrethroids have been characterized. The strains include susceptible, resistant and highly resistant genotypes. Previously unknown patterns of cross-resistance to pyrethroids were suspected, and multivariate analysis of a matrix of LD50 values has provided additional evidence in support of these new and interesting phenomena. (Murray and Digby with Denholm, Farnham, and Sawicki, Insecticides)

Many problems of perspective effects were solved in the video records of tracks of pea moth (see last year's annual report). A system has now been programmed to translate the records into two-dimensional tracks, and three tracks have been processed. The ground track and the track relative to wind direction were derived. Further video records will be examined. (Perry and Clark with Wall, Entomology)

Models of host-parasitoid interaction based on the empty-cell probability assuming a negative binomial distribution with constant parameter, k, needed modification when examination of four sets of field data indicated that k depended on population density. An alternative model has been proposed, which differs from the classical Nicholson-Bailey model in that movement of both parasitoid and host is allowed. (Perry, with Dr L. R. Taylor, retired member of Entomology)

The diffusion model developed to explain the effect of cultivation on the movement of soil has been further improved and extended to two dimensions. Predictions from the model of metal content in the Woburn Market Garden experiment now agree tolerably well with soil analyses over the past 40 years. (Lane with McGrath, Soils and Plant Nutrition)



https://doi.org/10.23637/ERADOC-1-26 

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



As part of the recently started oilseeds research programme, the multidisciplinary rape trial has been redesigned and enlarged, to permit better control of field variation; this applies not only to the main part of the trial but also to its subsidiary studies. (Digby, with Rawlinson, Plant Pathology and McEwen, Field Experiments)

There has been a revival of interest in spatial models for the analysis of experiments. Rothamsted has an unrivalled collection of data on field experiments done on known plots; these data can be used to check assumptions used in a variety of spatial models. Series of Rothamsted experiments done on the same plots with the same crops, but in different years, were examined with the aim of discovering patterns of covariance on the plots. Preliminary results indicated that spatial models could not be distinguished from the classical model of independent additive block and plot effects, but much more work is needed before definitive statements can be made. (Bailey, Dillon and Smith)

Most tests used in the identification of yeasts take 7–14 days to give a result and therefore it is usual to perform simultaneously all the tests that may be needed. However, this requires more tests than a sequential method, such as an identification key, would need. As a compromise between quick identification and redundant testing, an initial set of four tests was taken, all of which can be observed within a few hours; sets of tests were selected for each combination of possible results from these initial tests, as well as for various other characteristics that may be observed without specifically being tested. Use of this strategy decreases the number of tests required from 61 to between four and 53 depending on the various outcomes.

A microcomputer program was written to allow yeasts to be identified from observed test results. If there are several yeasts with the given results, the program selects what further tests are needed to complete the identification, optionally allowing for possible errors in testing. Alternatively, the program can list all yeasts with a given set of characteristics. (Payne with Dr J. A. Barnett, retired from AFRC funding at the University of East Anglia, and Dr D. Yarrow, University of Delft)

***Routine analysis.*** In July the Jacquard J300 was transferred to take over word-processing for the Station following the installation and commissioning of new data entry equipment centred on five DEC Rainbow microcomputers; a twin Midas microcomputer will also be installed. Various pieces of equipment, including an Epson HX-100 printer, and Epson HX-20 and Comart microcomputers, can be interconnected via a plug board, which also has access to the VAX system. These connections permit interchange of data transferred from ADAS sources via floppy discs on the Comart, transferred from Rothamsted sources by micro-cassette on the Epson, or entered directly on the Rainbow machines.

New software had to be developed for this system. Using CBASIC running under the CP/M operating system, this software was modelled on that previously written by Todd for verification of data on the Jacquard, but it uses the DEC editor for the initial data entry. One advantage of using a widely available operating system was that the software could be developed and tested on the Comart and transferred to the Rainbows as soon as they arrived. The changeover period was thereby reduced to a few weeks, and retention of the keyboard operations used on the Jacquard minimized the operator retraining required. In the new system each machine has access only to the directories of its own discs rather than, as previously, to a common file. Initially, this caused some confusion, but a workable method has now been devised for recording the progress of jobs and creating local archives on floppy discs. Problems with the occasional loss of disc files should disappear now that higher grade discs are in use. (George and Todd)

At one stage jobs submitted for batch processing were blocking interactive use during the day, even though they were not currently being processed. They are now submitted to a queue for overnight processing and promoted one by one to the day queue if previous jobs



https://doi.org/10.23637/ERADOC-1-26 

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



have been run. Delays in printing output in the mornings have also proved troublesome but have been partly overcome by spacing jobs during the day, and should be further relieved by the installation of a medium-speed printer, now on order. (Todd)

Routine work, from both Rothamsted and ADAS, has been later than usual this year due to the delayed harvest. The total volume is expected to be normal with an average turn-round time of five days; postal delays often extend the total time by as much as five days in each direction.

## Work for ADAS

A training course in Statistics for Experimental Husbandry Farms/Experimental Horticulture Stations (EHF/EHS) staff was run at the St Albans College of Agriculture, and a three day Genstat course was run for ADAS staff at the Cambridge office. (George, Dyer, Murray)

There has been a drop of about 10% in the number of jobs processed this year. Most of this can be explained by the increasing number of experimenters who are writing their own Genstat programs for their simpler experiments. This is borne out by the increase in the number of Genstat queries received this year and also by the larger proportion of experiments received for analysis that are not straightforward. (Dyer)

***Crop experiments.*** Further analysis was done on the results from trials of response of winter wheat to nitrogen in the ADAS Eastern region from 1981–84. Several variates have been analysed; regression analysis has been used to explore the relationships among these and with various site characteristics. Retrospective comparisons were made of proposed schemes for predicting nitrogen requirement. Current ADAS recommendation practices appear to give very reasonable results. Results for 1985 are to be incorporated into the dataset. (Murray)

Further refinements have been made to the curve-fitting program developed for ADAS clients. The program can now be used to fit 'Linear over Quadratic' models as well as the 'Linear plus Exponential', $y=a+br^N+cN$ (with or without fixed $r$). Programs to fit grain nitrogen offtake responses have been incorporated into the ADAS macro library. (Murray)

The SB06 (Nitrogen on Spring Barley) and SB10 (Autumn and Early Spring Nitrogen for Wheat) trials were summarized using the curve fitting program developed by Murray. Further work needs to be done on the SB10 data. (Dyer)

An investigation has been made of factors affecting the response of cut grass to the nitrogen content of dairy cow-slurry. The efficiency of the slurry N was calculated by comparing herbage dry matter yields from slurry applications with those from fitted response curves for $NH_4$-N. Although location, rainfall, slurry DM and soluble N content all appeared to influence the efficiency of slurry N on occasions, rate and time of application were the most important factors at each site. The efficiency of slurry N averaged over all experiments ranged from 38% for a low rate of application in spring down to 17% for a high rate in summer. In general the combined effects of slurry N and fertilizer N were additive. The residual effects of slurry N on herbage yields later in the same season were consistently low. (Dyer with Mr B. F. Pain, Animal and Grassland Research Institute, Shinfield and Mr K. A. Smith, ADAS Wolverhampton)

***Data collection.*** The ADAS software for recording and transferring data was little used; no data was submitted by the EHFs using this program. Since the number of data processors engaged on ADAS work had already been reduced, in anticipation of less need for direct data entry, it became imperative to produce an acceptable alternative. A trial program was produced in time for the EHF/EHS Statistics Course at Easter and left running for anyone to



https://doi.org/10.23637/ERADOC-1-26 

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



try it. It became clear that an extended version would be welcomed, and a first working version was issued to most EHFs and some EHSs in May. Further versions have since been issued, and most of the field experimental data from EHFs is now sent in on discs written by this program. Although not designed for livestock experiments, the program is usable for them with a little ingenuity, and some livestock data is now submitted in this way. The overall effect is that some direct data-entry has been transferred to the EHFs, but there will be a more positive saving when advantage can be taken of data-loggers. (George)

***Nutritional chemistry.*** Contacts with the nutrition chemists have increased considerably and have included the estimation of inter-laboratory variation in the chemical analysis of animal feeds, fitting linear functions to chemical data to predict the metabolizable energy (ME) value and other characteristics of feeds, and fitting non-linear functions to experimental data to estimate the protein digestibility. The work is continuing and involves liaison with MAFF, the NFU (representing farmers) and UKASTA (representing suppliers). Several equations used by MAFF and the industry for prediction of ME have now been amended on the basis of our recommendations, and better error limits have been established that allow for variations between laboratories. The digestibility work is of considerable interest to the ADAS Feed Evaluation Unit, who are now being pressed to evaluate commercial products. With the introduction of new EEC regulations on the labelling of animal feeds the accurate assessment of variation in ME is of immediate importance. Currently, each nation is left to determine its own tolerances on declared values, with a review in two years' time to determine common values. No analytical data exist at present to govern the selection of suitable tolerances, and our data on inter-laboratory variation will be fundamental to this study; the final values established will be enforceable by law. (George and Kirby)

***Livestock experiments.*** A partially-confounded split-plot design was used to study interactions between several factors known to affect laying-hen performance. Light source and light intensity were also studied. The information was sought in order to improve the predictive equations used in 'HEN', the ADAS empirical model of egg production. In this trial fluorescent lighting encouraged higher feed intake than did tungsten lights. Several two- and three-factor interactions were indicated which involved the factors room temperature, light source, feeding system, light intensity and protein level. Light intensity had some effect on egg yield during early lay. As a result of this new knowledge, some changes to the 'HEN' model are needed. (Bailey and Spechter with Dr D. Charles, ADAS)

Piglet mortality is one major cause of economic loss in pig production. Research therefore continued into the causes of mortality and ways of minimizing this loss. The Medata Blowaway system was investigated. The device triggers a fan when the sow gets up, blowing air across the crate. This encourages piglets to leave the crate area, thereby reducing the risk of overlaying. Analysis of the data showed some reduction of mortality due to overlaying. A trial was run to compare the effect of providing extra heat for younger piglets in forward creep farrowing pens with unsupplemented controls. There was a small advantage for luxury creeps as regards mortality during the 0–48 h period. In a study of pre-weaning mortality, piglet mortality was found to be strongly related to litter size and sow parity. Timing and reasons for deaths were investigated. Both stillbirths and subsequent mortality showed a progressive decrease with increasing birth weight. (Spechter with Dr S. Edwards and Mr P. Smith, ADAS)

## Surveys

***Fertilizer practice.*** Surveys of fertilizer practice done in England and Wales and in Scotland during 1984 were reported. This series of surveys continues as a collaborative



https://doi.org/10.23637/ERADOC-1-26 

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



project between ADAS soil scientists, representatives of the Fertiliser Manufacturers' Association and Rothamsted, but, because of reduced resources, ADAS withdrew from the field work in 1985. The size of the survey was, however, almost maintained at previous levels, all field work (1200 farms in England and Wales and 250 in Scotland) being done in June–July by Farm Research Ltd on behalf of the Fertiliser Manufacturers' Association. This new arrangement has slightly reduced the organizational workload at Rothamsted and allows more flexibility in data preparation, validation and processing. Preliminary results for England and Wales (all even-numbered farms) were available by 1 October and those for the full samples by the year end; further improvement in timeliness is looked for in subsequent years. (Church, Leech, Elsmere)

***Sugar beet surveys.*** Data from the 1984 Specific Field Survey were received from British Sugar and translated for use on the VAX. As in recent years, tables showing relationships between yields, pest and disease incidence, and growers' practices were prepared to supplement those routinely produced by British Sugar. This work, encompassing several years' data, was done to continue critical examination of survey evidence on factors influencing yield, and in response to specific requests from British Sugar and Broom's Barn, who were also provided with subsets of the data for further analysis. (Church, Leech)

***Representative soil sampling.*** A paper was completed summarizing information from Representative Soil Sample Surveys done by ADAS Soil Science Staff between 1969 and 1983. These surveys provide estimates of the proportions of agricultural soils in England and Wales in different pH categories and by nutrient index. The P, K and MG status of most soils is satisfactory, but an appreciable proportion of both arable and grassland may be regarded as too acid for maximum production. The data, based on resampling the same fields at intervals of five years, show that there was some increase between 1969–73 and 1979–83 in the proportion of grassland at low pH and a reduction in the proportion of arable land at low P and K indices. However, despite large estimated changes on individual fields, these general trends have been quite small.

ADAS soil scientists have decided to resume this small-scale monitoring exercise of agricultural soils from 1985, and amendments to sampling procedures, record forms and instructions for the next five-year cycle have been implemented. (Church, Elsmere).

***Straw incorporation.*** In 1984 the NFU, in collaboration with ADAS and AFRC, set up a register of farmers who had already been incorporating cereal straw or who would be shortly experimenting with this practice. Forms to record details of their experience were designed and sent to interested farmers; the data were analysed at Rothamsted and reported to the first NFU conference for members of the register. Information from this small non-random sample was not representative of cereal growers as a whole but aroused sufficient interest for this exercise to be repeated at the end of 1985. (Church, Elsmere)

## Theory

**Analysis and design.** Analysis of covariance combining information on treatment effects estimated in several strata has previously been described for a variety of specific multi-stratum designs but not in general. Equations were developed to allow combination of information to be done from analysis of covariance of any generally-balanced design, and these will be incorporated into Genstat. (Payne with Mr R. D. Tobias, University of North Carolina)

In the theory of factorial design, it has been shown that many commonly-used asymmetrical designs are entirely specified by their symmetric components; this greatly simplifies the



https://doi.org/10.23637/ERADOC-1-26 

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



search for good asymmetrical designs. A published algorithm for constructing loosely balanced fractional factorial designs has been shown to be very sensitive both to the size of the fraction (for example, 2/5 instead of 3/8) and to the way the size is expressed (for example, 8/10 instead of 4/5): appropriate amendments to the algorithm have been devised. The structure of rectangular lattice designs has been elucidated, so they may be analysed by Genstat. Restricted randomization procedures have been developed for some common two-dimensional layouts. (Bailey)

**Discrimination between models.** The sample size needed to discriminate between the Poisson distribution and other two-parameter distributions has been studied. Similar studies of the likelihood ratio criterion for a sample of size $N$ from a distribution with frequency function $f(x, \theta, \phi)$ when fitted by the simpler model $f(x, \theta, 0)$show that the function

$$N(\theta, \phi)=L_c \Big/ \int (x, \theta, \phi) \log \{f(x, \theta, \phi)/f(x, \theta, 0)\},$$

where $L_c$ is a critical value of the log-likelihood, is a measure of the minimum sample size required; this function can be computed independently of data. Contour charts can be prepared showing when discrimination is feasible. The method was applied to mixtures of exponential distributions and showed that discrimination often could not be expected with fewer than several thousand samples. (Ross)

The use of different response curves for estimating optimum levels of nitrogen fertilizer application for cereals have been compared. It is of concern that different models (e.g. 'Linear over Quadratic' and 'Linear plus Exponential') may fit equally well but give quite different estimates of optima, and an understanding is needed of why, and under what circumstances, this occurs. Simulation techniques are being used to study this problem. (Murray)

**Inference.** The information obtained by fitting a model to data compared with the information from a single observation is termed 'effective replication'. Effective replications calculated at interpolated or extrapolated design points show clearly the range over which the fitted model may be trusted. The concept was generalized to handle non-normal data, correlated errors and to apply to general functions. (Ross)

Recently there have been a number of articles by statisticians on abuses of multiple-range and multiple-comparison tests. A recent article in one of the journals of the Entomological Society of America (Jones, *Environmental Entomology* (1984) **13,** 635–649) exemplified the undue emphasis put on significance-testing at the expense of estimation. A riposte was prepared demonstrating that the use of multiple-comparison procedures is inappropriate for data with structured treatments and is often unhelpful even for totally unstructured data. (Perry)

**Estimation.** Estimates of variance from random samples can be very different from their theoretical values, especially in long-tailed distributions. The incomplete variance function,

$$V(x)=\int_{-\infty}^{x} \{t-m(x)\}^2 f(t)\, \mathrm{d}t \qquad \text{where } m(x)=\int_{-\infty}^{x} tf(t)\, \mathrm{d}t$$

in the form $V(x)/V(\infty)$ may be plotted against $x$, the proportion of the ordered sample included. For some log-normal distributions the largest 1% of the observations contained over 90% of the variance. This throws doubt on the credibility of mean-variance relationships computed from small samples. (Ross)

A method to estimate, $\beta$, the exponent of the power relationship between variance and mean, outlined by Perry (*Applied Statistics* (1981) **30,** 254–263) involves applying a transformation to make the variance of the transformed value independent, to first order, of its



https://doi.org/10.23637/ERADOC-1-26 

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



mean. By regressing this variance on this mean and re-applying a transformation based on the value of the slope, an iterative scheme is defined which leads to an estimate of $\beta$. A substantial number of simulations have been run to compare this with the usual estimate, the coefficient from regression of log variance on log mean. (Perry)

The algorithm developed last year for fitting families of models with some common and some specific parameters has been improved. The likelihood of the common parameters is now optimized not with fixed values of the specific parameters, but with values adjusted using an empirical relationship between the two sets. This avoids the slow convergence typical of the EM style of algorithm and has many possible applications. (Ross)

**Multivariate analysis.** The theory of robust ordination between and within groups has been developed. Units that are far from their group mean, using either Euclidean or Mahalanobis distance, are given reduced weight. Weighted sums-of-squares-and-products and/or means are then calculated iteratively. In this way outlying values have little effect on estimates. (Digby)

## Statistical computing

**Genstat 4.** External converters completed conversions to eight more machine ranges: Burroughs B6700(MCP), CDC 6000(Scope, NOS), DGC Eclipse, Harris, Honeywell DPS6, IBM(VMS), IBM(CMS) and Perkin-Elmer. An IBM Personal Computer version was produced by J. Coursol, Université de Paris-Sud. (Simpson) Genstat now runs in 388 installations, 27 conversions and 35 countries.

Arrangements have been made for NAG to take over the more formal aspects of future conversions.

Genstat 4 courses were given at Rothamsted, the Meat and Livestock Commission, the AFRC Computer Centre and the Institute of Industrial Research and Standards in Dublin. (Digby, Lane and Payne)

**Genstat 5.** A major effort this year has been put into the development of Genstat 5. Programming is now well advanced and many features were demonstrated live during the Fourth Genstat Conference, held in York in September (see below). Genstat 5 is a major revision of previous versions. It differs in having a consistent, easily learned syntax, a fully interactive mode of operation, procedure calls (rather than macros) that look identical to standard Genstat commands, and high-quality graphics. New statistical facilities include directives for fitting a variety of standard curves. Genstat 5 will be under test at Rothamsted during early 1986, ready for release for conversion to various supported machine ranges towards the end of the year.

***Documentation.*** The Implementer's Manual has been revised for Genstat 5. (Simpson and Harding) The whole question of future documentation has been considered. It is hoped to publish the next Genstat Manual professionally and every effort will be made to make it readily accessible to a wide range of readers.

The *Introduction to Genstat* by Alvey, Galwey and Lane (1982) originally written for Genstat 4 has already been revised for the new syntax and awaits one new chapter. A new text, *Genstat—a second course,* has been prepared, again apart from a single chapter. (Digby, Lane and Dr N. W. Galwey, University of Cambridge)

**MLP.** New facilities include: (i) comprehensive treatment of missing values in curves, general models and distributions, (ii) the calculation of the determinant of the dispersion matrix of parameters, useful in studying experimental designs; (iii) estimates derived from



https://doi.org/10.23637/ERADOC-1-26 

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



individual frequency data rather than grouped observations; (iv) improved convergence of algorithms for parallel model analysis. (Ross)

The Italian Science Research Council has approved funding of a three-year project for studying expert systems in Statistics. (Gower and Professor Carlo Berzuini, University of Pavia) As a first step a menu facility is being constructed for MLP, that leads to a sequence of questions and answers that elicit the data and option-settings for curve-fitting. Special features include: (i) a 'no idea' response, which causes data to be plotted and various forms of curves to be suggested; (ii) short-cut routes through the menu on the second use after a fit; (iii) facilities for editing and correcting data. The basic specification (Ross) has been adapted using the system EXPERT. (Professor C. Berzuini)

**Genkey.** An irredundant set of tests is one that is sufficient to distinguish all taxa but would be insufficient if any one test were omitted. To allow for errors in testing, *n*th order irredundant sets may be defined in which taxa cannot be distinguished when up to *n* tests are omitted. An extension of the usual algorithm was devised to construct these, and has been incorporated into Genkey. (Payne)

**Glim.** Nearly all work on Glim has ceased because of staff losses. The facilities for fitting generalized linear models in Genstat and the availability of Glim itself allow the continued use of the methodology for practical applications.

**Data-entry.** Much time has been spent on developing a field data-entry program, primarily for the use of ADAS scientists at EHFs and possibly EHSs. Information on unit labels, block and treatment factor values, previously put onto paper forms, may now be entered directly at the keyboard, followed by the data which may be entered at the keyboard or transferred from data-loggers. Various checks are made, and values may be corrected, inserted or deleted. Simple graphs, derived variates and other calculations are available.

The program offers enough options to be useful but not so many that the user may become frustrated. It is robustly written so that inconsistent or unwanted choices are trapped and may be easily corrected. As a safety measure the original files can be deleted only when the new ones are complete and closed. Files produced may be used directly for input into Genstat. (George)

There is clearly much to be gained from standardizing data-entry procedures and the files they produce for all field experiments; a working party has been set up to coordinate data-entry programs from ADAS and the AFRS institutes. (George)

## Overseas

The Overseas Development Administration continued to fund the Biometric Unit. (Riley, Preece, Ryder, Lloyd and Poultney) Despite staff changes in the Unit, the amount of work from annual crop experiments and from animal trials has remained steady, coming largely from the following countries: Belize, Botswana, the Dominican Republic, the Falkland Islands, Kenya, Malawi, Pakistan, Paraguay, Sierra Leone, Swaziland, Sudan, Tanzania, Thailand, Turkey, Tuvalu, Zambia and Zimbabwe.

Work from experiments on perennial crops has been directed to the Unit (from East Malling Research Station) and the amount is being monitored to see whether the present number of staff is sufficient to handle the extra work. Requests for statistical advice and anlayses of perennial crop data were received from Honduras, Kenya, Kiribati, Malawi, Malaysia, Nepal, the Solomon Islands, St Lucia and Tanzania.

A small increase in the number of personal visits to the Unit was noted; efforts are being made to encourage a further increase so that more efficient collaboration can be achieved



https://doi.org/10.23637/ERADOC-1-26 

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



despite the great distances involved. Particularly encouraging is the interest of overseas researchers who wish to visit the Unit to obtain statistical advice when writing end-of-contract reports.

Statistical packages available for microcomputers continued to be surveyed; nearly 250 have so far been examined. Many of these packages prove undesirable from the points of view both of statistical content and of accuracy. Each one must be carefully assessed before any recommendation for its use can be given to scientists working in remote parts of the world. Despite the very many packages available, none has yet been found that combines the desired range of statistical techniques with the necessary degree of accuracy. (Ryder)

Work on intercropping continued, and collaboration began with the Commonwealth Forestry Institute in Oxford and the International Council for Research in Agroforestry in Nairobi on statistical methods for agroforestry. (Riley)

Genstat programs and macros were prepared for the analysis of data from intercropping experiments. These programs permit more efficient analysis of complex experiments involving more than one species. (Poultney and Riley)

### Staff and visiting workers

B. M. Church retired after 38 years with the Department, many of them spent in ably running the Survey Section. D. A. Preece left to become Head of Statistics at East Malling Research Station and was succeeded as Head of the ODA Section by Janet Riley. Lilian B. Robertson retired as Divisional Secretary and R. J. Baker, T. J. Dixon and A. Gnanasakthy also left. Anne Ainsley, Susan Lloyd and Belinda McVay were appointed. Rosemary Poultney and Anne-Marie Dillon each spent six months in the Department as sandwich students.

Bailey, Gower and Ryder attended the Centenary Session of the International Statistical Institute held in Amsterdam, where Bailey was an invited speaker. Lane demonstrated Genstat at the '17th Symposium on the Interface' in Lexington, Kentucky and visited AT & T Bell Laboratories, where he gave a seminar. Gower and Payne attended the Fourth International Data Analysis Conference in Versailles, where Genstat was demonstrated. Gower visited INRA sites in Montpellier and in the Paris region and Bailey was invited to participate in the INRA workshop on 'Modèls Linéaires et construction de plans d'experiences', held in Pech Rouge. Bailey was also an invited speaker at a workshop on 'Design and related inference' held in Los Angeles. At the invitation of the University of Copenhagen, RECKU (the Danish Computer Centre for Research and Education) and the Danish Society for Theoretical Statistics, Gower visited Copenhagen and Aalborg where he gave two talks; Bailey gave three talks in Denmark at the Universities of Aarhus and Aalborg.

In the British Isles the Fourth European Conference of the Psychometric Society and Classification Societies was held in Cambridge (organized by Gower), preceded by a workshop on 'Multidimensional data analysis' at which Gower gave an invited paper. Gower was also an invited participant at the Royal Statistical Society/SERC workshop on 'Statistics and pattern recognition' held in Edinburgh. Bailey was an invited speaker at the Cambridge meeting of the British Mathematical Colloquium. The fourth Genstat Conference (organized by Payne with J. Byrne and A. J. Weekes as local organizers) was held at York and attended by 92 participants from 11 countries. Bicknell, Digby, Gower, Lane, Murray, Payne, Riley and Ross gave talks, while Bicknell, Digby, Harding, Lane, Payne, Simpson and White demonstrated a prototype of Genstat 5. The Genstat meeting in York was followed by the 6th AFRC Statisticians' Meeting, the first since 1980; 61 persons attended, 16 from Rothamsted. Gower, jointly with Professor J. Copas of Birmingham University, organized an Anglo-French workshop on 'Data-analysis' held at Gregynog, near Newtown,



https://doi.org/10.23637/ERADOC-1-26 

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



Powys. Lane organized a SERC course for postgraduate students on 'Applying statistics in agriculture' at Oaklands, the Hertfordshire College of Agriculture and Horticulture; several members of the Department contributed talks. The Biometric Society had its Summer Meeting at Rothamsted, organized by Perry and Clark; 53 people attended. Members of the Department contributed to many other seminars, workshops, society meetings etc.

R. D. Tobias of the University of North Carolina visited (January–July), T. Hancock of the Waite Institute (April) and J. Abington was with the ODA Unit in October. Dr K. MacRae continued with us until July and Professor L. Underhill until June. Dr S. Iwatsubo from the National Centre of University Entrance Examination, Tokyo, arrived in October and Professor F. J. Rohlf of the State University of New York at Stony Brook arrived in November.

## PUBLICATIONS

### Computing Unit

THESIS

SUMMERFIELD, J. (1985) *The design and implementation of a core for a computer based soil analysis system.* B.Sc. Project, Hatfield Polytechnic.

GENERAL PAPER

JONES, F. G. W. (1985) Modelling multigenic resistance to potato cyst nematodes. *EPPO Bulletin* **15,** 155–166.

PROGRAM AND SYSTEMS GUIDES

RESCU/077/85 Accessing local weather data via DATATRIEVE
RESCU/078/85 The Hi-Pad digitiser
RESCU/079/85 R$SIGSURF
RESCU/080/85 VAX fault logging and monitoring system
RESCU/081/85 Statistics Department Rainbow systems for data preparation
RESCU/082/85 Scientex word processor introductory user guide
RESCU/083/85 Using the Epson HX20 to record root data
RESCU/084/85 User Guide for R$DIGITISE
RESCU/085/85 User Guide for how to use the Epson HX20 program for aphid data capture
RESCU/086/85 User Guide for Report Program for Acornsoft Database File
RESCU/087/85 User Guide for Append Program for Acornsoft Database File
RESCU/088/85 Remote logging from CR7 unit
RESCU/089/85 Glasshouse automation system
RESCU/090/85 Rothamsted Departmental inventory system
RESCU/092/85 How to use the aphid monitoring program
RESCU/093/85 Storage and retrieval of insect data
RESCU/094/85 Scientex word processor advanced user guide
RESCU/095/85 SSEW Laboratory automation system user's guide
RESCU/096/85 SSEW Laboratory automation system manager's guide
RESCU/097/85 SSEW Laboratory automation system software guide
RESCU/098/85 User Guide for R$SIGLIB

### Statistics Department

BOOK

DIGBY, P. G. N. & GOWER, J. C. (1984) *Ordination and classification.* Séminaire de Mathématiques Supérieure, Vol. 98, Montréal: Les Presses de l'Universté de Montréal.

PROGRAM GUIDE

(BARNETT, J. A.), PAYNE, R. W. & (YARROW, D.). *Yeast identification program.* (Disc with booklet.) Cambridge: Cambridge University Press.



https://doi.org/10.23637/ERADOC-1-26 

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



## Paper in Rothamsted Report, Part 2

Gower, J. C. (1985) A history of statistical computing at Rothamsted. *Rothamsted Experimental Station. Report for 1985,* pp. 221–235.

## Research Papers

Bailey, R. A. (1985) Factorial design and Abelian groups. *Linear Algebra and its Applications* **70,** 349–368.

Bailey, R. A. (1985) A note on loosely balanced incomplete block designs. *Computational Statistics and Data Analysis* **3,** 115–117 and 121–122.

Bailey, R. A. (1985) Restricted randomization versus blocking. *International Statistical Review* **53,** 171–182.

Baker, R. J. (Clarke, M. R. B.) & Lane, P. W. (1985) Zero entries in contingency tables. *Computational Statistics and Data Analysis* **3,** 33–45.

Baker, R. J. (Clarke, M. R. B.), Lane, P. W. (1985) Further comments on M. B. Brown and C. Fuchs 'On maximum likelihood estimation in sparse contingency tables'. *Computational Statistics and Data Analysis* **3,** 125–127.

Ball, B. V. (Overton, H. A., Buck, K. W., Bailey, L.) & Perry, J. N. (1985) Relationships between the multiplication of chronic bee-paralysis virus and its associate particle. *Journal of General Virology* **66,** 1423–1429.

(Bartlett, P. W.) & Murray, A. W. A. (1986) Modelling adult survival in the laboratory of diapause and non-diapause Colorado Beetle *Leptinotarsa decemlineata* (Coleoptera: Chrysomelidae) from Normandy, France. *Annals of Applied Biology* **108,** 487–501.

Bowden, J., Sherlock, P. L. & Digby, P. G. N. (1985) Studies of elemental composition as a biological marker in insects. III. Comparison of the elemental composition of apterous and alate cereal aphids, especially Rhopalosiphum padi (L.) from oats and wheat, and from oats infected with or free from barley yellow dwarf virus. *Bulletin of Entomological Research* **75,** 477–488.

Bowden, J., Sherlock, P. L. & Digby, P. G. N. (1985) Studies of elemental composition as a biological marker in insects. IV. The influence of soil type and host plant on elemental composition of *Agrotis segetum* (Denis & Schiff) (Lepidoptera: Noctuidae). *Bulletin of Entomological Research* **75,** 675–687.

Brobyn, Patricia J., Wilding, N. & Clark, Suzanne, J. (1985) The persistence of the infectivity of conidia of the aphid pathogen *Erynia neoaphidis* on leaves in the field. *Annals of Applied Biology* **107,** 365–376.

Bunce, N. A., White, R. P. & Shewry, R. P. (1985) Errors in determining the $M_r$s of cereal prolamins by SDS-PAGE. *Journal of Cereal Science* **3,** 131–142.

Gower, J. C. (1985) Properties of Euclidean and non-Euclidean distance matrices. *Linear Algebra and its Applications* **67,** 81–97.

Perry, J. N. & (Taylor, L. R.). (1985) Adès: New ecological families of species-specific frequency distributions that describe repeated spatial samples with an intrinsic power-law variance-mean property. *Journal of Animal Ecology* **54,** 931–953.

Perry, J. N. & Wall, C. (1985) Orientation of male pea moth, *Cydia nigricana,* to pheromone traps in a wheat crop. *Entomologia Experimentalis et Applicata* **37,** 161–167.

Prew, R. D., Church, B. M., Dewar, A. M., Lacey, J., Magan, N., Penny, A., Plumb, R. T., Thorne, Gillian N., Todd, A. D. & Williams, T. D. (1985) Some factors limiting the growth and yield of winter wheat and their variation in two seasons. *Journal of Agricultural Science, Cambridge* **104,** 135–162.

Riley, J. (1985) Examination of the staple and effective land equivalent ratios. *Experimental Agriculture* **21,** 369–376.

Ross, G. J. S. & Preece, D. A. (1985) The negative binomial distribution. *The Statistician* **34,** 323–336.

Ryder, K. (1986) Numerical tables: Guidelines for presentation. *Journal of Applied Statistics* **13,** 77–88.



https://doi.org/10.23637/ERADOC-1-26 

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



(WEBSTER, A. J. F., SAVILLE, Claire), CHURCH, B. M., GNANASAKTHY, A. & (MOSS, R.) (1985) The effects of different rearing systems on the development of calf behaviour. *British Veterinary Journal* **141,** 249–264.

(WEBSTER, A. J. F., SAVILLE, Claire), CHURCH, B. M., GNANASAKTHY, A. & (MOSS, R.) (1985) Some effects of different rearing systems on health, cleanliness and injury in calves. *British Veterinary Journal* **141,** 472–483.

## GENERAL

BAILEY, R. A. (1985) Nesting and crossing in design. In: *Encyclopedia of Statistical Sciences*. Vol. 6. Eds N. L. Johnson, S. Kotz & C. B. Read, New York: J. Wiley, pp. 181–185.

BAILEY, R. A. (1985) Partially balanced designs. In: *Encyclopedia of Statistical Sciences*. Vol. 6. Eds N. L. Johnson, S. Kotz & C. B. Read, New York: J. Wiley, pp. 593–610.

BAILEY, R. A. (1985) Restricted randomization for neighbour-balanced designs, statistics and decisions. *Proceedings 16th European Meeting of Statisticians,* pp. 237–248.

CHURCH, B. M. & LEECH, P. K. (1985) Fertilizer use on farm crops in Scotland, 1984. Harpenden: Rothamsted Experimental Station, 24 pp.

(DREW, S. B., RANDALL, E. M., CHALMERS, D. I., PELLATT, D. G.) & KIRBY, S. P. J. (1985) Dystococia in Friesian/Holstein dairy heifers. *Animal Production* **40,** 521.

GOWER, J. C. (1984) Distance matrices and their Euclidean approximation. *Data Analysis and Informatics, III*. Eds E. Diday, M. Jambu, L. Lebart, J. Pagès and R. Tomassone, Amsterdam: Elsevier Science, pp. 3–21.

LANE, P. W. & PAYNE, R. W. (1985) Genstat 5: a general-purpose interactive statistical package, with facilities for generalized linear models. In: *Proceedings GLIM 85 Conference*. Eds R. Gilchrist, B. Francis & J. Whittaker, New York: Springer Verlag, pp. 105–111.

PAYNE, R. W. (1985) Genkey, a general program for constructing aids to identification. *Informatique et Sciences Humaines* **64,** 45–62.

PREECE, D. A. (1985) Illustrative examples. Illustrative of what? In: *Proceedings of the Meeting held at the University of Kent at Canterbury to mark the 70th birthday of Professor Clifford Pearce,* (unpaginated), Royal Statistical Society East Kent Local Group.

WALL, C. & PERRY, J. N. (1985) Effect of habitat on orientation behaviour to pheromone sources. In: *Proceedings 2nd ISCE Conference, Madison,* Abstract Volume, p. 41.



https://doi.org/10.23637/ERADOC-1-26 