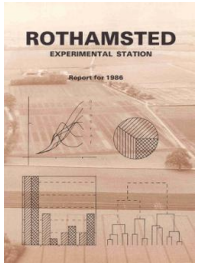
This work is licensed under a Creative Commons Attribution 4.0 International License.

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 1986

Full Table of Content



Biomathematics Division

J. C. Gower

J. C. Gower (1987) *Biomathematics Division ;* Rothamsted Experimental Station Report For 1986, pp 69 - 81 **- DOI: https://doi.org/10.23637/ERADOC-1-27**

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.

Assistant Scientific Officer
Caroline Mitchell
E. Wilson

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

Higher Professional and Technical Officer
R. N. Le Fevre, B.Sc.

Technical I
D. J. Peck, B.Sc.

Technical II
P. D. Compton

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

Specialist Typist
Nicole Ford

Sandwich Student
M. Gilbert

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 Officers (UG6)
Rosemary A. Bailey, D. Phil.
B. J. George, B.Sc.

Principal Scientific Officers (UG7)
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.
C. D. Kershaw, Ph.D.
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.
R. P. White, B.Sc.

Scientific Officers
Suzanne J. Clark
Janet I. Elsmere
S. A. Harding,
S. Langton, M.Sc.
Gina L. Smith, B.Sc.

Departmental Secretary
Valerie Payne

Personal Secretary
Diane P. Dawkins

Sandwich Student
J. Felgate

Visiting Scientists
B. M. Church, B.Sc.
G. Lampitt, M.Sc.
S. Iwatsubo, D.Eng.
T. Miwa, D.Eng.
P. R. Krajewski, M.Sc.

Specialist Typist
Angela C. Piears

Senior Data Processors
Anne E. Arnold
Brenda J. Watler

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

INTRODUCTION

The Computing Unit's services are in much demand. It expanded this year to provide electronic and computing skills needed to develop up-to-date laboratory equipment that is commercially unobtainable. Current computing resources now meet only one half that required by Rothamsted staff and the resulting poor response-times are leading to inefficiency and frustration. Steps are now being taken to acquire more computing power so that

https://doi.org/10.23637/ERADOC-1-27

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

the scientific research programme does not suffer. This computing power must be provided without placing too heavy a burden on the Unit's manpower resources, thus maximizing the staff-time available for development and innovation. With the computing hardware available today, the clear expansion path is to provide high-power, low-cost computing in the laboratory for the highly intensive users, leaving the main computing systems to provide support for the less intensive users. For Rothamsted, this means we must now provide some 10 to 15 distributed computers for graphics, modelling and statistical work plus approximately one and a half times the existing VAX systems for general purpose computing. It is the Unit's task to link such equipment while minimizing support requirements and maximizing flexibility.

The work of the Statistics Department has been as varied as usual and applications, theoretical developments and work on statistical computing are all described below. Closer contact with the other departments, consequent on the reorganization of our consulting services, reported last year, has benefitted both experimental design and sampling, and has indicated the importance of scientists consulting their liaison statisticians often and early. Both our policy of trying to make statistical software more accessible to non-statisticians and the spread of microcomputers, mean that more scientists are doing statistical analyses for themselves. This trend should not be discouraged but it must be recognized that it can sometimes lead to difficulties; ways must be found to monitor the quality of non-specialist work.

One of the purposes of the statistical liaison officer scheme is to provide continuity of advice, where the statistician concerned is well-known to his scientist clients and has built up a good knowledge of their research areas. The best statistical advice is not given in isolation and often involves close collaboration. A striking example of twenty years' collaboration on identification problems, especially as they relate to the identification of species of yeasts, of which there are about 500, is described in a recent paper. (Gower and Payne)

COMPUTING UNIT

The Computing Unit provides a computing and electronics design, development and advisory service for Rothamsted departments, as well as support for all the associated equipment. The Electronics Group (i) maintains and repairs existing laboratory and microcomputer equipment and (ii) designs, develops and makes equipment to support computing and instrumentation requirements. (Le Fevre) The Systems Group supports data-capture, development of small computer systems-integration, the running of two DEC VAX 11/750 supermini computers and word-processing facilities throughout the Station. (Moore) The Applications Group provides software for a variety of computing systems as required by the scientific research programme of the Station and also provides a programming and systems-analysis advisory service together with an image-analysis development advisory service. (Bicknell)

Services

The Systems Group now provides service support for the twin DEC VAX 11/750 supermini computer systems which connect to AGRENET, the AFRC computing network, which itself provides a gateway to JANET, the universities' academic computing network, plus PSS the British Telecom X25 UK network, and IPSS, the international X25 network. These networks allow Rothamsted users to access any computer system worldwide that is also connected to one or other of the networks; this they do to access specialized software packages and/or to collaborate with other research workers; it is also a means of shedding the local computer load to other systems, particularly within the AFRC.

Additionally, this group provides support for 16 Scientex stand-alone word processor

https://doi.org/10.23637/ERADOC-1-27

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

systems, four DECmate stand-alone word processor systems, eight data preparation key-to-disc work stations, over 60 stand-alone user microcomputer systems, over 80 multi-access terminals and associated printers, five graph plotters, line printers and a media conversion service. All these systems are distributed throughout the campus linked to an associated local network which also must be provided and supported.

This demanding service requires the Unit's staff to respond rapidly to problems and to have considerable knowledge and experience in order to deal with any aspect of hardware or system software failure that may arise.

Development

This year has seen the completion of many projects and the start of many new ones. The following gives a brief summary of the activities covered.

An image analysis system was purchased to replace the existing 12 year-old equipment. Current developments are mainly with software requirements but problems concerned with leaf damage, soil sample and gel analysis are already being tackled on this equipment. (Bicknell and Hipgrave)

Database systems still provide considerable software and development work. Their power to marshall both small and large, simple and complex data collections, enabling rapid comparisons and searches to be performed, is becoming more appreciated. The 1032 database package has been used to develop the large (60 Mbyte) insect database of the Entomology Department. (Summerfield and Wong, Entomology) The ESQL package, mounted on a UNIX system is being used for the Broom's Barn sugar beet laboratory analysis database. (Summerfield and Verrier) The Datatrieve package has been used to develop the Insecticides Department database system. (Summerfield) The microcomputer package dBASE III has been used for additional development of cost-forecasting for administrative purposes and for handling the daily meteorological data collection for PEP. (Verrier)

Genstat spread its influence to the Unit this year for the development of high level graphics, allowing Genstat users to obtain graphical representations of their data and analyses by interfacing through the NAG JO6 graphics supplement and the Culham GHOST package. (Bicknell and Harding, Statistics) An editor has been provided within Genstat to enhance the interactive mode of working. (Verrier)

Further graphics activities have included additional development of the SURFACE II package which allows for contouring; and a substantial development of the R$GRAPH package for simple-to-use graphics. (Bicknell, Hipgrave, Thomas)

Field and laboratory data collection is an important activity at Rothamsted and further labour saving software has been provided for Soils and Plant Nutrition and Plant Pathology, utilizing Epson HX20 and Huskey Hunter systems. (Hipgrave and Thomas) Further systems have been developed for equipment-control and data-capture, including that for Physiology and Environmental Physics (Moore), Soils and Plant Nutrition microcalorimetry (Higgins) and Plant Pathology (Moore).

Development work must be undertaken in support of the local network and to enable inter-machine communication. This year has seen local development of Kermit, the internationally accepted protocol for machine to machine file transfer, for the Epson HX20 (Moore); this implementation is now in use throughout AFRC and has spread as far as New Zealand. The problem of machine connectivity has been addressed by the hardware and software development of a contention switch (Moore) to provide a further 32 terminal or microcomputer access points to the twin VAX systems. This also required specialized software for disc-less computer system development, which is much needed for instrumentation projects and monitoring tasks.

https://doi.org/10.23637/ERADOC-1-27

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

Considerable activity has taken place to develop control systems. In particular, a system was developed in a UNIX-based computing engine to control additional glasshouses via specialized interface devices. (Gilbert and Verrier) A mobile shelter control system for the Farringdon site (Le Fevre) and equipment-failure alarm-units have been developed. (Le Fevre and Compton)

A high precision, linear voltage-to-current transducer has been designed and developed for the measurement of high currents via clamp-on probes. (Le Fevre and Peck) Instrumentation has been designed for electroporation research, requiring high voltage discharges through plant cells. (Le Fevre)

Development work in support of word processing systems has included laser printer software, character-set development and specialized printer interfacing in order to meet the demand for quality reproduction and disparate needs of the different departments. (Higgins)

Development of the local network to connect the many microcomputers and peripherals located around the campus has led to the decision to install a campus Ethernet for data transmission at up to 10 million bits per second. This decision requires the Unit to install low loss coaxial cable coupled to fibre optic cable segments with appropriate transceivers and repeaters to IEEE 802·3 standards. (Le Fevre and Verrier)

Staff

N. R. Shah, Dawn Dudley, J. W. Hipgrave and J. Summerfield left. Nicole Ford, R. N. Le Fevre, P. D. Compton, D. J. Peck and E. Wilson joined. M. Gilbert joined us on a studentship for six months; G. Wong worked in the Unit for six months while involved with the Entomology database project. K. Bicknell is a member of the GHOST Technical Committee and a member of the Eurographics (UK) Executive Board. P. J. Verrier was invited to Mauritius Sugar Research Institute to examine their computing requirements.

STATISTICS DEPARTMENT

Practical applications

The following practical applications are examples of the many statistical analyses done for Rothamsted departments throughout the year.

Analyses of several sets of data, corresponding to periods of heavy rainfall, provided figures for flow rates and nitrogen concentration in the drainage water from plots of the Brimstone Experiment. From these figures, it was decided that the number of observations made could not be reduced if a good estimate of the total nitrate loss were to be retained. Work continues on automating the analysis of the observations and estimating the lost nitrate. (Lane with Goss, Soils and Plant Nutrition)

The project to investigate similarities between the catches of aphid in Insect Survey suction-traps in different regions of the UK continued. Weekly catches of *Metapolophium dirhodum,* at 18 sites over ten years, were matched by using generalized and orthogonal Procrustes rotations. Based on these analyses, the traps were divided into six distinct geographical areas. Investigations into trap relationships have begun for another species of aphid, *Sitobion avenae.* (Clark and Perry with Woiwod and Tatchell, Entomology)

A pot experiment for granular boron fertilizer was designed to assess the uptake of boron. The original factors were six different fertilizer formulations, two particle sizes (fine/medium), two soil types, two rates of application (single and double), two methods of incorporation (as a top dressing or mixed with the soil), two levels of pH (acidic and neutral) and two watering regimes, plus control pots. One of the two watering regimes was abandoned. Some treatment combinations were unavailable, e.g. three of the fertilizer formula-

https://doi.org/10.23637/ERADOC-1-27

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

tions were produced only as fine particles, and the medium particle size permitted only the double rate of application, thus confounding particle size and rate of application with fertilizer formulation. The final design was reduced to two replicates, each of 51 pots. The test crop was clover which could be cut at regular intervals. Despite considerable imbalances in the design, early results are encouraging. Small effects on harvested yield (fresh weight and dry weight) were found at the first two cuts; soil pH seemed to affect the yield at cut one, probably due to differences in establishment and differences between soil-types appeared at cut two. Rate of application, method of incorporation, particle size and different fertilizer formulations all seem to have a large effect on boron uptake. (White with McGrath and Ewens, Soils and Plant Nutrition)

A program was written to calculate a modification of Cornish-Bowden's measure of homology between protein sequences. Unlike Cornish-Bowden's original formula, the modified form is specific to the number of residues whose quantities are precisely determined by experiment. Cornish-Bowden's measure does not have the statistical properties commonly claimed for it and so work is being done to construct a more useful measure of homology. (Bailey with Burgess, Biochemistry)

The fit of families of Adès distributions (*Rothamsted Report for 1983,* 77 and *1984,* 80) to 16 sets of mainly insect data was generally better than that of constrained families of negative binomial distributions with the same number of parameters. Practical applications include the derivation of a sequential sampling plan and use in sampling schemes involving incidence counts; a limited role in the analysis of spatial pattern was demonstrated. (Perry)

Analysis has begun on a large set of data on survival times of strains of house fly with differing resistance to pyrethroids exposed to a range of doses. (Murray with Denholm, Insecticides)

Routine analysis. The Rainbow data-entry system has functioned satisfactorily. Some disc reading problems were eventually traced to sub-standard discs whose replacement by better quality discs has removed most of the difficulties. Data submission on Epson tapes has increased but the majority of data is still hand-written. Changes in the VAX queues had little effect on turnaround. The main difficulty is still the inability to run trial jobs interactively; a whole day may be lost when jobs need to be revised. The volume of Rothamsted work has increased slightly but turnaround has remained steady at about five days. (Smith and Todd)

Work for ADAS

Staff illness has resulted in a slight increase in the turnaround times for ADAS jobs since October. The work processed through the year for ADAS has dropped by roughly 10%. Virtually all the experimental results received from the Experimental Husbandry Farms (EHFs) are now sent on floppy discs, and the users on the farms appear to be getting used to the FIELD program. (Dyer)

Staff losses and the delays in getting approval for posts to be refilled coupled with the paucity of good statisticians, has inhibited our work for ADAS. Nevertheless the bulk of work has been completed successfully and we look forward to improvements in the coming months.

Crop experiments. The SB15 trails on the effect of two different sources of nitrogen fertilizer on the yield of winter wheat and winter barley have been analysed. At 26 out of 32 trials there was no evidence of different responses for urea and ammonium nitrate. However, ammonium nitrate performed better than urea at the remaining six trials, where it appeared that one unit of nitrogen from ammonium nitrate is equivalent to 1·2 units of nitrogen from urea. This ratio is robust to the choice of curve relating yield response to applied nitrogen. (Bailey, with Mr A. Lloyd, ADAS, Bristol)

https://doi.org/10.23637/ERADOC-1-27

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

The five most recent years of ADAS Eastern Region data on response of winter wheat to applied nitrogen fertilizer has been used to compare current and potential systems of nitrogen recommendation for farmers. Results suggest that for the majority of sites where soil nitrogen reserves are low current recommendations for these soils appear to be adequate and there is little scope for farmers to benefit from recommendation systems based on soil analysis. There may be scope for improving recommendations where soil nitrogen reserves are expected to be high. (Murray with Dr R. Sylvester-Bradley, ADAS)

Consulting work for Soil Science was handled in the earlier part of the year largely by Murray. Since he left, little has been possible but immediate advice on problems and queries; some long-term analyses are delayed until a replacement is appointed. A curve-fitting routine for cereal trials, written in Genstat 4 and issued for ADAS users, now needs revision and extension but not until Genstat 5 is in use by ADAS. (Murray)

Randomized plans, sometimes incorporating restricted randomization, have been supplied as follows. Spring barley trial (Mr P. Harris, Bridgets EHF); multifactorial hops trial (Mr F. J. Dickens, Rosemaund EHF); restricted randomizations for 20 sites for the SB38 milling wheat trials (Mr P. M. R. Dampney, Soil Science, Newcastle); six plans for factorial trials where inclusion of extra varieties at smaller replication and only some fungicide levels complicated the design (Mr D. Inns, Starcross); factorial systems trial and restricted randomization for SB38 milling wheat trial (Mr I. Ap Dewi, Rosemaund). (Bailey)

Data collection. Further versions of the FIELD data entry program for use on Comart microcomputers have been issued to both EHFs and Experimental Horticultural Stations (EHSs). The current version provides a flexible and user-friendly means of entering and modifying experimental details and data, and allows a range of summaries and analyses. Most of the data from field experiments at EHFs are now entered and submitted via this program. An increasing number of livestock experiments is now submitted in this way, although some are still submitted on paper and others on Supercalc files. The program accepts data from CP/M files produced by Husky or Epson machines and similar files produced by data-loggers but cannot yet accept formatted files from Supercalc which need manual editing to remove extraneous characters.

Three one-day courses on the use of FIELD have promoted usage and generated useful suggestions for extension and improvement. A primitive analysis of variance module has been issued on a trial basis. This allows general treatment models but does not yet cater for missing values, or adequate error checking. Further generalization will be deliberately curtailed to reject any data-sets with non-orthogonal terms or with more than one error stratum. (George)

Nutritional chemistry. Regrettably, contacts with Nutrition Chemistry have had to be reduced following the departure of S. Kirby. Advice has been given on the analysis and interpretation of animal feedstuffs data and on industry tolerance levels for declared metabolic energy (ME) and other values required by EEC legislation. A program written for fitting non-linear functions to experimental data to estimate protein digestibility has been transferred to the Feed Evaluation Unit and implemented on their Prime; further work is needed. (George)

Livestock experiments. The analysis of ruminant experiments was taken over by Smith in addition to her other duties. Delays in turnaround have been inevitable and liaison with clients inadequate despite considerable effort on her part. We urgently need to, and hope to, restore a better service when Kirby's replacement is in post.

A trial was conducted to assess the egg-yolk pigmenting abilities of substitutes for the Roche colour standard. The yolk colouring (xanthophyll) potential of the test products was

https://doi.org/10.23637/ERADOC-1-27

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

established by fitting a curve to the responses to various levels of the test substances 'Oroglo' and 'Lucantin' *vis-à-vis* the Roche standard. The parallel curve analysis given by MLP showed that there were differences in location, but not in shape, of fitted logistic response curves. A difficulty with this experimental approach was that a value for the colouring potential of the test product had to be assumed initially in order to fit the response curve; an alternative approach has been suggested for future investigations. (Spechter with Mr S. Jackson, ADAS Gleadthorpe)

Since recommendations vary, a trial investigated the effect of different space allowances on the performance of growing pigs on a standard diet and in a controlled environment. Increasing space gave increased liveweight gains and better food conversions. The results indicated that a space allowance of at least 0·27 m^2 kg^{-1} metabolic bodyweight should be used in fully slatted housing. (Spechter with Mr S. A. Edwards and Mr A. W. Armsby, ADAS)

A trial investigated the effect of temperature and volume of pre-rinse on the efficiency of circulation-cleaning in milking parlours. The investigation took place on five farms with herringbone parlours, all of which had acceptable standards of hygiene. Five rinse treatments were tested on five occasions in a 5×5 Latin square design. An eight litre hot (>75°C) pre-rinse gave significantly lower total bacterial counts than rinses at lower temperatures but it was the most costly. There were substantial unattributable farm differences. (Spechter with J. Runnalls, ADAS, Cambridge)

Surveys

Fertilizer practice. As usual, Part 2 of this report contains a summary of the findings of the year's survey of fertilizer practice. (Elsmere) This year, Part 2 also contains an account of the Department's work on agricultural surveys over the past 50 years. (Church and Kershaw)

Surveys of fertilizer practice done in England and Wales and in Scotland during 1985 were reported. This series of surveys is a collaborative project between ADAS soil scientists, representatives of the Fertiliser Manufacturers' Association (FMA) and Rothamsted. In 1986 the size of these surveys was maintained at previous levels with all field work (1200 farms in England and Wales and 250 in Scotland) being done in the period June to mid-August by Farm Research Ltd on behalf of the Fertiliser Manufacturers' Association. It was agreed at the annual meeting with the FMA and ADAS that preliminary figures on a subsample of the data were no longer required because tabulations of the full data are now made available in October rather than in the following year. (Elsmere, Kershaw and Leech)

Sugar beet surveys. The 1985 data-tape was received from British Sugar and translated for use on the VAX. Standard analyses were run on this data, and lists of grid references of fields with their virus yellows and powdery mildew incidence were produced for the years 1981–85 to enable Dewar and Asher (Broom's Barn) to map the disease levels. (Leech)

Consultation with British Sugar has been concerned with an examination of relationships between sugar beet yields, growers' practices and other recorded factors in survey data, and with the analysis and interpretation of series of experiments. (Church)

Representative soil sampling. A paper appeared summarizing results of the Representative Soil Surveys done by ADAS Soil Science staff between 1969 and 1983. (Church and Mr R. J. Skinner, ADAS) After a break in 1984, ADAS soil scientists decided to resume this small-scale monitoring exercise of agricultural soils. The 1985 data is now being processed and results will be available early in the new year. Soil samples have been collected for the 1986 survey. (Kershaw and Elsmere)

https://doi.org/10.23637/ERADOC-1-27

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

Straw incorporation. In 1984 the National Farmers' Union (NFU), in collaboration with ADAS and the AFRC, set up a register of farmers who had already incorporated cereal straw or would shortly do so. Forms have been sent to farmers on the register over the last three years. The 1984 and 1985 data were analysed and the 1986 data is now being received. Annual NFU conferences for members of the register have been organized at which results have been reported. (Kershaw and Elsmere)

BBC Domesday Project. Summary tables of N, P and K usage in Scotland and in England and Wales during 1984 were provided for inclusion in the BBC Domesday project. These were based on appendix tables from the appropriate Rothamsted reports. (Leech)

Theory

Analysis and design. Many designs, whether ordinary block designs or multi-stratum designs, have a structure which is effectively unchanged by a permutation of the treatments. In fact, many common designs, such as generalized cyclic designs, are constructed in such a way that a given group of treatment permutations preserves the design. Those properties of the group that ensure general balance of the design have been characterized and given formulae for the efficiency factors in terms of the group characters and initial blocks of the design. (Bailey with Dr C. A. Rowley, Open University)

Previous work on valid randomizations for orthogonal designs has recently been extended to cover certain classes of non-orthogonal designs. (Bailey with Dr D. A. Preece, Institute of Horticultural Research)

Non-linear models. Correlation coefficients obtained by standardizing the dispersion matrix of parameter-estimates of non-linear models are useful indicators of dependencies. However, when there are three or more parameters dependencies may be large even when correlations appear small. A simple measure of the dependency of a parameter θ_i on the remaining parameters in a model is the *variance multiplier,*

$$\lambda_i = V_{ii} V_{ii}^{-1}$$

where V is proportional to the dispersion matrix. When a parameter is independent of the others then $\lambda_i = 1$, otherwise $\lambda_i > 1$ and measures the relative increase in variance due to uncertainties in the values of the other parameters. (Ross)

Racine *et al.* (*Applied Statistics* (1986) **35,** 93–150) proposed the use of Bayesian estimation to obtain a more precise estimate of LD50 than that given by maximum likelihood. The assumption of a uniform prior has been shown to give increased weight (compared to likelihood estimates) to large estimates of slope, and therfore may be quite unacceptable in a non-linear model. Graphical methods for illustrating the phenomenon were devised, and methods of performing the necessary integrations using MLP were derived. (Ross)

When comparing samples from different members of the same family of distributions some parameters may be common to all samples while the others are specific to each sample (e.g. normal samples may have different means but common variance). For discrete two-parameter models, such as the negative binomial or Neyman's Type A, it is by no means obvious that the conventional defining parameters provide the best way of expressing constraints among samples. Instead it is proposed that the *stable* parameter, (Variance—mean)/mean^2 should be constrained to be equal between samples, and the defining parameters fitted accordingly. This allows distributions of similar shape but different genesis to be compared. Another possibility is to constrain variance/mean. (Ross)

Recent experiments from Rothamsted were selected to provide a range of types of data suitable for analysis based on generalized linear models—counts, proportions and percent-

https://doi.org/10.23637/ERADOC-1-27

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

ages. These were analysed both by using linear methods, after suitable transformation, and by corresponding generalized linear models. Parameter estimates are generally similar for both methods but it is clear that for some data, particularly when counts or proportions are small, acceptable estimates of variability require the full generalized linear model approach. (Lane and Felgate)

Modelling. As is clear from the above accounts, statisticians have more than a passing interest in modelling. Apart from models developed by statisticians themselves, we have a general interest in models developed by all scientists especially in relating models to data. Despite an overstretched staff we are able to make some contribution to this area, and would like to do more.

The Department has been awarded a New Initiative contract to investigate spatial models for field trails. Semivariograms are used in spatial interpolation models, e.g. kriging, and in other spatial models. The use and estimation of semivariograms, relating variance of differences to the distance between plots, is under investigation. Six programs, written either in Gentat 5 or in Fortran, which do various one-dimensional neighbour analyses have been assembled as the beginning of a comprehensive collection. Work to make the programs comparable in their treatment of end-plots and the input parameters required is in progress. (Ainsley)

The spatial pattern of individuals within a honeycomb tesselation of hexagons may be made to conform locally to any desired variance-mean relationship by allowing an individual in a given hexagon to move, according to a simple rule, to one adjacent. This involves computing $3[f(m_i)-s_i^2]+x_i-1$ where x_i is the count in the individual's own hexagon, m_i and s_i^2 are local sample mean and variance in the seven hexagons centred on it and f( ) is the expected variance-mean relationship. A stochastic model has been built to study such a process. Also, two population dynamic models have been developed, each with a small Poisson immigration rate and a 'carrying-capacity' for each 'patch' which varies as a lognormal distribution. The first is driven purely by environmental heterogeneity whilst the second is based on Pielou's demographic discrete-time difference equation analogue of the logistic. All three models simulate realistic spatial variance-mean data with a power-law relationship. Furthermore, the exponent of this relationship can range from under 1·5 to over 2·5, matching observed values. This shows that several mechanisms can account for such observed relationships. (Perry)

Models are often simulated to explore the consequences of different parameterizations and starting conditions etc. Sometimes a direct mathematical anlaysis is possible and when it is this can be more informative and more efficient. One such investigation of this kind was concerned with a model describing annual fluctuations in populations of wild oats. The qualitative mathematical approach supported most of the conclusions reached from simulations, showing that these conclusions are independent of choice of initial state, but disagreed with some of the others. Other mathematical aspects of the model such as chaos and instability require further analysis. (Bailey)

Work was continued on stochastic modelling of DNA sequences and protein sequences, addressing such questions as: (i) are there more repeated amino acid pairs in a given sequence than would occur by chance? (ii) is the match between two given sequences attributable to chance? (iii) what is the most likely evolutionary tree for a given set of DNA sequences? Some partial answers have been obtained, but much remains to be done. (Bailey)

Statistical computing

Genstat 4. Most of the work relating to the earlier version of Genstat is now handled, at least initially, by the Numerical Algorithms Group (NAG). Two new conversions made

https://doi.org/10.23637/ERADOC-1-27

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

available during the current year were for Norsk Data Nord 500 (Mr R. Paynter, Thames Polytechnic) and Whitechapel Workstation (Mr B. Francis, Lancaster University). Genstat now has 419 installations at 326 sites in 41 countries and has been converted to run on 40 ranges of computers. The 100th VAX Genstat was sent out in March; during the year 15 copies were sent to new sites—the total number is now 112. (Simpson)

Genstat 5. Genstat 5 differs from Genstat 4 in having a revised syntax that is more consistent and easier to use, particularly by non-statisticians and those without training in computing. The new procedure structure enables the use of frequently required analyses to be made much easier for those without detailed technical knowledge. Output has been redesigned to make its examination on VDUs more convenient and the syntax facilitates interactive use, such as is expected by users of PCs and workstations (and even the VAXs!). New facilities include the fitting of standard growth and exponential curves, the randomization of generally balanced experimental designs, facilites for editing and manipulating text, and high quality graphs, histograms and contour plots.

Work has continued throughout the year on this new and much extended version of Genstat and we now have a version available for general release. Much is owed to the enthusiasm and devotion of the Genstat team, many of whom have worked long hours, at night and over weekends, without financial reward. This has been necessary not only to get through the work but also to compensate for the appalling response-times of the Station's VAX computers (see above). (Ainsley, Bicknell, Digby, Harding, Lane, Leech, Payne, Simpson, Todd, Verrier, White)

Six trial releases were formed on RESB during the year. Release 5.00 (February) was given to a number of interested AFRS sites. Releases 5.02 (June) and 5.04 (September) were distributed by the Computing Centre to most of the AFRS sites. The two intermediate releases (April and August) were only made available for local testing. Release 5.02 was also sent to Dr McLaren at Cambridge, Dr Tunnicliffe-Wilson at Lancaster and other trial sites outside the AFRS. We owe a great deal to McLaren who not only succeeded in putting the program on to the IBM 3081 but also read at least half of the Reference Summary with penetrating discernment and sent forth a stream of highly critical but extremely useful comments. His concern over the Fortran Input/Output facilities and the handling of character data led to dramatic changes in this area which will make the program much more transportable. Release 5.05 is being sent out, via NAG, to four 'beta' sites for conversion. They are Honeywell (Arizona), Prime (AFRUS, Edinburgh), IBM (Cambridge) and IBM PC (Digby with Professor J. Rohlf, State University of New York at Stony Brook); the latter is the first version to require overlaying and to take advantage of windowing facilities. The conversion program (*Rothamsted Report for 1978,* Part 1, 311) has been rewritten (in Fortran 77) to meet new requirements and several 'service' programs have been prepared as aids to conversion. (Simpson)

Courses. The new syntax of Genstat 5 requires changes to existing Genstat programs. A one-day Conversion Course for Genstat 4 users has been devised and presented at Rothamsted (twice) and at the Plant Breeding Institute. Beginners' courses have been given at the AFRC Computing Centre, East Malling Research Station, Long Ashton Research Station and the National Vegetable Research Station. (Digby, Lane, Payne and Simpson)

A summary document, showing the connections between Genstat 4 and Genstat 5 directives and listing new directives, was prepared for use in conversion courses. (Ainsley)

Documentation. A completely new manual is required for Genstat 5 and much of this has been written. (Ainsley, Bicknell, Digby, Gower, Harding, Lane, Leech, Payne, Simpson,

https://doi.org/10.23637/ERADOC-1-27

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

Todd, Verrier, White with Dr Tunicliffe-Wilson, Lancaster University and Dr L. Paterson, Heriot-Watt University)

A preliminary version of a Conversion Manual has been written and sent to the 'beta' sites. This will undoubtedly need considerable revision in the light of comments from convertors. (Simpson)

A revision of the book 'Genstat: an introduction' has been completed and titled 'Genstat 5: an introduction' (Lane with Mr N. Alvey and Dr. N. Galwey, University of Cambridge) and all chapters of a completely new and more advanced text 'Genstat 5: a second course' have been written, but this needs revision before publication. (Lane, Digby, and Dr N. Galwey, University of Cambridge)

Macros and procedures. The existing Genstat Macro Library will become a Genstat Procedure Library and the necessary conversions and new documentation have been begun. An editorial committee has been established to control the library. (Payne and Digby) A new feature that has been developed is the automatic attachment and searching of procedure libraries; this allows use of both system libraries and personal libraries of procedures. (Harding and Todd)

MLP. Sixty-eight licences were current in November 1986, and conversions of version 3.08 were complete or in progress for over 20 machine ranges; a version for the IBM Personal Computer was produced. (Ross and Digby) MLP was demonstrated at Compstat 1986 in Rome and at the London School of Economics; it formed the basis for a course on non-linear modelling given at Preston Polytechnic.

New facilities developed, but not yet released, include (i) extension of the interactive menu to control general user-models, frequency distributions and assay analysis, (ii) diagnostics for outliers, influential observations and inefficient model formulation, (iii) residual plots and difference curves the latter produced by subtraction of two different models fitted to the same data, (iv) specification of elements in arrays, greatly simplifying the definition of models for frequency distributions and differential equations, (v) modifications to the basic optimization algorithm to enable more parameters to be fitted. The manual for release 3.08 was greatly extended and is awaiting publication by NAG. (Ross)

Work on the project to study the possibility of producing an Expert System for non-linear modelling continues and was described at Compstat 86. The first stage has been to identify the information required to define models and to describe data so that the necessary instructions for running a job in MLP could be created interactively. A prototype version to run MLP jobs has been written in Pavia using an expert system 'shell' to construct an operational program from basic logical elements given as lists of compatible responses to questions. The second stage involves the program commenting on the suitability of the model or the parameterization, and allows the user to explore modifications that might be suggested; diagnostics for this purpose have been studied. (Ross, with Professor Carlo Berzuini, University of Pavia)

Rothamsted General Survey Program (RGSP). This program has been used for all our survey analyses since 1965 (see Gower, *Rothamsted Report for 1985,* Part 2, and Church and Kershaw, *Rothamsted Report for 1986,* Part 2). Since 1965 many programs for survey analysis have been produced and these have been assessed for their suitability for our work. They have all been found unsuitable for scientific survey work, though it is recognized that many are easier to use and are better documented than is RGSP. It has consequently been decided to investigate the feasibility of modernizing RGSP by providing an up-to-date syntax and a unified program. Work has started on producing a Personal Computer (PC) version. Certain aspects of survey work may be handled conveniently by database manage-

https://doi.org/10.23637/ERADOC-1-27

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

ment systems and the table manipulation by an extended version of the facilities in Genstat 5. These possibilities too are being investigated with the hope that appropriate interfaces can be provided. (Kershaw)

Data-entry. Development of the FIELD and the Feed Evaluation Unit digestibility programs have taken up considerable time. Discussions have been held with Miss B. Chadburn of the ADAS Computer Development Unit on interfacing FIELD with her programs written in dBASEII. Users now wish an extension of the program to include the amalgamation of data from similar experiments or the splitting of the data into cross referenced subfiles. Both requirements would be better implemented via a database system and this is now a priority. Transfer to other machine ranges such as Apricot and PCs has also been discussed as a means of making the program available to other ADAS clients who would then be able to transfer the data to the ADAS Prime system and thence to AGRENET. The EHFs will also be on-line to the Prime within the next two years and may wish to move over to direct entry to this instead of using the risky and time-consuming postage of discs produced on the Comart micros. This cannot be implemented fully without a better link between AGRENET and the MAFF Prime system. (George)

Overseas

The Overseas Development Administration (ODA) continued to fund the Biometric Unit (Riley, Ryder, Dutton and Langton). Riley visited Sri Lanka and the West Indies to assess proposals for ODA-funded projects; she visited India to provide statistical advice to an ODA tree-crops project in Karnataka.

The amount of annual crop and animal work has remained steady, coming from the following countries: Barbados, Belize, Bolivia, Botswana, Dominican Republic, Ecuador, the Falkland Islands, India, Jamaica, Nepal, Sierra Leone, South Yemen, Sudan, Thailand, Zambia and Zimbabwe. (Riley, Ryder and Dutton)

Now in its second year, requests for statistical advice and analyses of perennial crop data have increased markedly; these requests came from Bangladesh, Dominica, India, Indonesia, Kenya, Kiribati, Malawi, Malaysia, Nigeria, Papua New Guinea, the Solomon Islands, Sri Lanka, Tanzania (Zanzibar) and Thailand. (Riley, Ryder and Dutton)

The number of personal visits to the Unit remained steady. Particularly encouraging are the numbers of visitors interested in seeking the Unit's advice about microcomputers and associated statistical packages. An extensive survey of statistical packages for microcomputers has been prepared and summarized. (Ryder)

A major simplification of the very complicated current methods for the analysis of experiments on mixtures of varieties of agronomic interest has been made. Improvements in design have been worked out and graphical plots devised for assessing the effect of differences in sole yields on apparent aggressivity. (Dutton and Dr F. Yates, RESCU)

An ODA Research Grant for the development and application of statistical methods for intercropping research in agroforestry began with a survey of available statistical methods; 100 overseas research workers were approached for data to use in the testing of new methods. (Langton and Riley)

Staff and visiting workers

We congratulate Rosemary Bailey on her well-deserved promotion to Senior Principal Scientific Officer (Unified Grade 6) on individual merit.

Susan Dutton (neé Lloyd), S. Kirby, A. W. A. Murray, Belinda Smith and Deborah Game left during the year; C. D. Kershaw and S. Langton were appointed.

https://doi.org/10.23637/ERADOC-1-27

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

Gower was an invited speaker at both the NATO Advanced Research Workshop on Numerical Ecology at Roscoff, France and at the Classification Society of North America's 17th Annual Meeting at Columbus, Ohio; he also visited the State University of New York at Stony Brook. Bailey, Perry and Riley attended the 13th International Biometrics Conference at Seattle. Bailey also attended the International Congress of Mathematicians at Berkeley. Digby, Gower, Harding, Lane, Payne and Ross attended Compstat 86 in Rome where most contributed papers and where both Genstat 5 and MLP were demonstrated; Ross and Gower also visited the University of Pavia to discuss progress on the MLP expert system project with Professor C. Berzuini. Lane went to INRA, Jouy-en-Josas, France, to discuss ways of interfacing Genstat with other packages. In December Bailey started a three months' visit to various organizations in Australia, beginning with the University of Sydney.

An Anglo-Nordic meeting of the Biometrics Society was held in Cambridge at which Murray and Perry read papers; Digby, Dutton, and Gower also attended. Gower presented an invited paper and Lane and Ross attended the 4th Sheffield Conference whose topic was Influential Data Analysis. Members of the Department contributed to many other seminars, workshops, colloquia, society meetings etc.

Professor J. Rohlf returned to the USA in January but came back for a few weeks in July; Dr S. Iwatsubo returned to Tokyo in July. Mr G. Lampitt of the University of Zimbabwe was with us from March to September and Mr M. Tavallai of the Soil and Water Institute, Teheran, Iran from July to September. Mr D. Baird of DSIR, Auckland, New Zealand came during September and October and Mr C. Brien of Roseworthy Agricultural College, South Australia, September to November. Dr T. Miwa of the National Institute of Agro-Environmental Science, in Ibaraki, Japan arrived in September and Mr P. Krajewski of the Institute of Plant Genetics, Poznan, Poland, arrived in October.

https://doi.org/10.23637/ERADOC-1-27