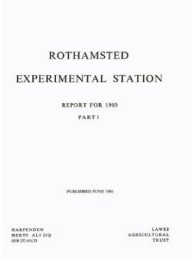
This work is licensed under a Creative Commons Attribution 4.0 International License.

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.

Report for 1980 - Part 1

Full Table of Content


ROTHAMSTED
EXPERIMENTAL STATION

REPORT FOR 1980
PART 1

Statistics Department

J. A. Nelder

J. A. Nelder (1981) *Statistics Department ;* Report For 1980 - Part 1, pp 265 - 277 **- DOI: https://doi.org/10.23637/ERADOC-1-137**

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

STATISTICS DEPARTMENT

J. A. NELDER

Staff

Head of Department J. A. Nelder, D.SC., F.R.S.

Senior Principal Scientific Officer
J. C. Gower, M.A.

Principal Scientific Officers
N. G. Alvey, M.I.S.
B. M. Church, B.SC.
J. H. A. Dunwoody, B.A.
D. A. Preece, PH.D.
G. J. S. Ross, B.A.
H. R. Simpson, M.A.

Senior Scientific Officers
Mrs Jill F. B. Altman, M.I.S.
R. J. Baker, M.SC.
R. W. Payne, B.A.
J. N. Perry, M.SC.
K. Ryder, B.TECH.
H. H. Spechter, PH.D.

Higher Scientific Officers
P. G. N. Digby, B.SC.
C. J. Dyer, B.SC.
P. W. Lane, B.A.
Miss Janet Riley, M.SC.
A. D. Todd, M.SC.
J. Wood, M.SC.

Scientific Officers
Mrs Janet I. Elsmere
Miss Catherine M. Garnett, B.SC.
Miss Penelope K. Leech
Miss Gina L. Smith, B.SC.
R. P. White, B.SC.

Visiting Scientists
J. E. Chiria, M.SC.
A. G. Constantine, PH.D.
J. D. Henstridge, PH. D.

Departmental Secretary
Mrs Claire M. McGlashan

Personal Secretaries
Mrs Doris A. C. Court
Mrs Susan Greenland

Senior Data Processors
Mrs Elsie Davies
Mrs Patricia K. Neill
Mrs Brenda J. Watler

Data Processors
Mrs Anne E. Arnold
Miss Elaine P. Bangs
Miss Sherryl L. Brunton
Miss Jane P. Dockeray
Mrs Geraldine F. Kirby
Mrs Janice R. Messer
Mrs L. Vera Wiltsher

Introduction

Work during the year has been hampered by staff shortages in all sections, but fortunately the customers for our service work have been understanding about the consequent inevitable delays. The introduction of new equipment for data entry, data validation, and word processing is already improving productivity in those areas, while significant progress has been made both in the content and distribution of our statistical computer programs, and in the development of new techniques of statistical analysis.

Practical applications

The applications of statistics are diverse, and this is reflected in the wide range of problems on which we are consulted. Any attempt at completeness would be tedious, and the following paragraphs do no more than illustrate some of the more substantial pieces of work.

Fertiliser experiments. Joint work on the relationship between nitrogen concentration in wheat grain and fertiliser nitrogen, using data from past Rothamsted experiments, has now been completed, and a paper accepted. (Lane, with Benzian, Soils and Plant Nutrition Department) Results from the Saxmundham Rotation II experiment from

https://doi.org/10.23637/ERADOC-1-137

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

1969 to 1976 have been analysed to determine the association between yield of the crops in the rotation (potatoes, sugar beet and barley) and the level of soil phosphorus. Due to the contrasting fertiliser treatments applied to the plots between 1899 and 1964, there is a wide range of bicarbonate-soluble P in the soil from about 5 to 65 ppm. The effects of fresh P fertiliser were also measured.

For sugar beet and potatoes, the association is best modelled by asymptotic regression, where yield increases strongly with soil P at low concentrations, but levels off to a plateau at about 20 ppm. However, there is evidence that potato yields continue to rise with soil P in some years. For barley, the strong increase of yield with soil P is only apparent for concentrations less than about 10 ppm when no fresh P is present, and there is not enough information to model this. Beyond 10 ppm, the yield seems to increase consistently in some years, but not in others, and is also affected by the previous crop in the rotation. (Lane, with Johnston and Mattingly, Soils and Plant Nutrition Department)

Ecological survey of Phoridae. The analysis was completed. In comparing the phenology of several sites, or species, an analysis of variance is unhelpful since main effects and interactions have little meaning. This common problem in entomology was solved by fitting split-lines to transformed data, each parameter of which has meaningful biological interpretation. This allows timing and degree of peak abundances to be compared between sites, or species. Analyses of diversity showed all sites could be represented adequately by a single estimate of α, an index of diversity. Most species had values of the index of aggregation, b, between 1·5 and 1·7. (Perry, with Dr R. H. L. Disney, Malham Tarn Field Centre, Henderson, Entomology Department and Dr R. O. Clements, Grassland Research Institute)

Comparison of trapping methods. Analysis of data on lacewings (*Chrysopa carnea*) proceeds and data for 1979 and 1980 have recently become available. Similarities between six traps (two suction traps, two light traps in woodland, two light traps in open farmland) have been investigated for each of 7 years, using principal coordinates analysis. The agreement between the years has been examined by pairwise Procrustes rotations and by generalised Procrustes analysis. The results from the latter were compared with results from individual scaling, using macros now available in Genstat. Analysis is proceeding to relate the ordination of the traps on principal axes to features of the phenology of *C. carnea* as reflected by the traps. Differences between pairs of traps are clearly greater than within-pair differences. This effect was most noticeable in 1975 and 1976, when favourable weather resulted in large catches throughout the year. Other years, such as 1979 when abundances were low, produced configurations that differed radically from the concensus on the second principal axis. Individual scaling allows these year differences to be quantified. (Perry, with Bowden, Entomology Department)

Taxonomy of cyst nematodes. Assignation of cyst nematode populations to species may be aided by numerical taxonomy. Data concerning several known and nominated species were analysed by standard methods, including single-linkage cluster analysis, principal coordinate analysis and canonical variate analysis. Morphological information from both adults and juveniles indicates large separations between major solanaceous species, but it is not yet possible to assign new populations to known species with present data. (Perry, with Stone, Nematology Department)

Construction of diagnostic keys. Diagnostic keys were constructed, using our program Genkey, for: (i) species of the genus *Solanum* in Australia (Payne, with Dr D. E. Symon,

https://doi.org/10.23637/ERADOC-1-137

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

Waite Institute, S. Australia); (ii) species of the genus *Phytophthera* (Payne, with Dr C. J. Shepherd and Dr R. B. Cunningham, CSIRO); and (iii) for the identification of strains of the genus *Actinomadura*, bacteria that cause foot-rot in the tropics. (White, with Miss M. Athalye, University of Newcastle)

Work for ADAS. We continue to be concerned with the analysis of experiments originating with ADAS, and, more importantly, with the summarisation of groups of experiments from different sites and years. The work covers both arable crops and animals, and brief reports on some of the projects follow.

Crop experiments. A new series of trials has been designed to test nitrogen response in spring barley with more recent varieties. The use of seven levels of nitrogen will allow quite detailed examination of the form of the response curve, and pattern of application in time will also be tested. (Wood, with Mr P. A. Johnson, ADAS E. Midlands) A similar series has been designed for winter barley, but here timing includes the question of distribution of N between autumn and spring. (Wood, with Mr J. P. Grylls, ADAS Cambridge)

A report was prepared on the 1979 experiments on the timing of N for winter wheat, an adaptation of those carried out in 1977 and 1978. Twenty-two sites were involved. Response curves were fitted to individual sites to estimate the economic optimum value for N. As with the earlier series, the main interest lay in estimating whether it was better to apply N all at once or to split it into two or three separate dressings. Over the 22 sites, the average yield for the single dressing was about 0·1 t ha^{-1} less than when the dressings were split, almost the same as was found in the previous years' trials. The 1979 experiments suggested that with only a small total N (up to about 100 kg ha^{-1}) it was better applied all at once, whereas with 200 kg ha^{-1} three separate applications were better. (Dyer)

A paper has been prepared on the response of ryegrass to nitrogen on chalkland in the south of England. It describes trials carried out at Bridget's EHF during 1974–78; they were designed to study the yield response, and the effect of white clover, over the range 250–475 kg N ha^{-1}. In general the results showed that the optimum values of N were usually about 50 kg ha^{-1} higher in experiments without clover, with a value of about 375 kg ha^{-1}. (Dyer, with Mr A. L. Francis, Bridget's EHF)

Livestock experiments. The statistical service provided to ADAS staff involved in livestock experiments has been fully stretched with an increase of 25% on last year for both design enquiries and requests for analyses and interpretation. Examples of this work follow.

Four Experimental Husbandry Farms have been involved in a coordinated trial to evaluate the suitability of pure-bred Canadian Holstein steers for beef production under two systems: intensive 12-month barley beef at Boxworth, Drayton and High Mowthorpe, and 18-month grass/cereal beef at Drayton and Liscombe. Separate farm analyses have been completed for data on both live and carcass weight, whilst summaries over farms for the two systems have been completed only for the carcass data. Results indicate that the main problem with the Holstein is poor conformation, especially under the intensive system, and this leads to few acceptable carcasses.

The amount of saleable meat from each of 18 joints was measured for ten Holstein and ten Friesian carcasses from Drayton, and results showed that there was less saleable meat in a side from a Holstein carcass so that although the price per kilogram of saleable meat was similar the retail value of the Holstein carcass was 3p kg^{-1} less than that of the Friesian. (Altman and Garnett)

https://doi.org/10.23637/ERADOC-1-137

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

The proposed new system for deciding the protein requirement for ruminants will be based on the degradability of crude protein in the rumen instead of its digestibility. One method of determining this degradability is to insert polyester bags, containing the feed of interest, into the rumen of fistulated sheep or cattle and then to measure the remaining dry matter and nitrogen over periods of time. Data from 55 trials on cattle have been received so far, and many of the feeds involved were forages. Before attempting to fit curves to this data, the components of variance between and within animals were estimated to see whether they were similar to those found by other workers using sheep and concentrated feeds. These components have proved to be unacceptably high for the forages and at present a trial is being designed to try to pinpoint the sources of error. (Altman)

Work has continued with Gleadthorpe EHF on the analysis of poultry and pig experiments. These included temperature and nutrition trials with turkeys, where increasing room temperatures over the range 12–26°C progressively depressed food intake and weight gain, while improving food efficiency. High-energy diet gave better food conversions than medium-energy diet. (Spechter)

The effects of light intensity and light wavelength were studied on the growth and efficiency of food conversion of male and female broilers. The wavelength treatments, which included coloured light ranging in peak wavelength from 450 to 650 nm and a broad-spectrum white light of median wavelength 560 nm, did not affect production. Light intensities in the range 0·7–46·5 lux were tested, and food intake, of female birds in particular, was depressed at the highest intensity. A behavioural study checked to what extent the effect of treatment on production could be explained by observed differences in bird activity. A joint paper is being prepared. (Spechter, with Dr C. Wathes, ADAS, Shardlow)

An investigation has been started with the aim of predicting egg grades from bulked egg weights, and vice versa. In order to derive a suitable method, we have studied the distribution and standard deviation of egg weight for two breeds at various ages. Egg weights for birds of the same breed and age-group are approximately Normally distributed. Egg grades can, therefore, be predicted by the Normal probability method for various assumed values of the standard deviation, and the goodness of fit for the predicted and actual grades compared. The prediction of grades by Normal probability from weighted mean egg weights averaged over the whole laying cycle was inadequate. An alternative distribution-free method for bulked egg samples is currently being investigated. (Spechter)

Surveys

Fertiliser practice. In this continuing series of surveys done in collaboration with ADAS Soil Scientists and representatives of the Fertiliser Manufacturers Association, a representative sample of about 1300 farms in England and Wales was used during 1980. Preliminary estimates based on a representative subsample were produced by the end of September and basic tabulations by December, despite delays in fieldwork in arable regions accentuated by difficulties with harvesting spring cereals. (Church, Elsmere and Leech)

A general report was prepared on the 1979 survey, and tabulations for individual farm-type regions were produced for circulation to ADAS Regions and for reference. (Church and Leech) More detailed reports on methods and timing of individual fertiliser applications to cereal crops, and the associations between previous cropping and practice were also provided. (Church)

Representative soil sampling. Data for a complete 10-year cycle of this survey, in

https://doi.org/10.23637/ERADOC-1-137

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

which representative samples of fields in England and Wales are re-sampled for soil analysis at intervals of 5 years, were available by the summer, and a report was prepared on the pH and lime status of soils in England and Wales between 1969 and 1978. There is little evidence of any general change in soil acidity between 1969–73 and 1974–78. However, about a tenth of spring barley crops were grown at pH less than 6 (current advice gives 5·9 as a critical pH for barley), and use of lime has declined since the period covered by these data. (Carter and Church)

Other surveys. The main computer analysis of a survey on the use of anthelmintics in cattle was completed (Leech), and we collaborated in the preparation of a report. The survey results showed that anthelmintics are widely used in circumstances where they are unlikely to preserve helminth-free pastures or significantly to improve livestock productivity. (Church and Leech, with Dr J. F. Michel, Central Veterinary Laboratory, and Mr J. O. Latham, ADAS Winchester)

Data on the isolates of the National Milk Quality Survey were collated and more than 100 summary tables produced. (Leech) Consulting work continued on the British Sugar Corporation survey of representative sugarbeet fields (Church) and information from this survey on the incidence of weedbeet in 1980 was summarised. (Leech)

Routine analysis

New equipment. The evaluation of data-entry and word processing systems using magnetic media was completed. The Jacquard J 100 system was chosen and installed. It has the advantage that the terminals can be used for either data-entry or word processing. The Computer Department will also use one of the terminals for word processing. A direct link to the System 4 computer is available, allowing the transfer of files in either direction. (Alvey, Dunwoody and Todd, with Gledhill, Computer Department) The J100 has good word-processing software, but for data verification we have had to devise our own programs using the Basic language provided by the system. (Todd) Further programs have been required to make the link to the System 4 as simple to use as possible. (Todd, with Higgins, Computer Department) The system has been in use since early October and has generally proved satisfactory. Operators like the new equipment, which has improved their working conditions.

Data processing. The service we provide for the punching and verification of data and their subsequent analysis has again been widely used with at least 124 workers from other departments and 15 from outside stations submitting jobs. Data comprising 1·75 million items were involved, an increase of 17% on last year. There was an increase of 40% in the number of items processed for ADAS (1·02 million compared with 0·73 million in 1979). By contrast, work for Rothamsted users decreased by 4% and that for Overseas Development Administration (ODA) by 15%. The average turn-round time was 7 days, the same as in 1979. (Dunwoody, Dyer, Smith and Todd)

Theory

Ordination between and within groups. Canonical variate analysis allows between- and within-group variation to be properly combined, but equivalent forms of analysis using distances other than Mahalanobis-D^2 have not been developed. A significant step forward has been made by working out formulae for inter-group distances given the distances (however defined) or similarities between all pairs of individuals. Thus group-centroids can be analysed by any selected ordination method and the within-group variation superimposed by the add-a-point technique (Gower, J. C. (1968). Adding a

https://doi.org/10.23637/ERADOC-1-137

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

point to vector diagrams in multivariate analysis. *Biometrika* **55**, 582–585). A preliminary paper is being prepared. The method is capable of extensive further development. (Gower and Digby).

Asymmetry. Models have been explored for asymmetric matrices describing immigration–emigration in animal populations. A paper has been prepared illustrating the use of previously derived theory to suggest an appropriate form of model for such data. The basic least-squares approach can indicate the terms required in a model but is, however, unsatisfactory for a proper statistical analysis. This can be remedied by accepting the model terms but not the least-squares parameter estimates, which are replaced by estimates using various forms of generalised linear model. The analysis highlights the relationships between over-parameterisation and unidentifiability, showing how the inclusion of an extra parameter leads to inseparability of other parameters. (Gower, with Dr A. G. Constantine, CSIRO Adelaide)

Multivariate methods. Three established methods of multivariate analysis are principal-components analysis, the bi-plot and correspondence analysis, and these have so far been regarded as basically distinct. However all three methods can be described in terms of the singular-value decomposition of a certain matrix, and thus can be unified by expressing them as special instances of a more general procedure. The methods can also be associated with implied underlying parametric models. (Digby)

Non-linear inference. The definition of likelihood-based residuals (*Rothamsted Report for 1977*, Part 1, 304) serves to unify the method of maximum likelihood and the method of least squares. Differentiation or differencing of the vectors of residuals allows methods for minimising sums of squares to be applied to the non-linear parameters only in a model with separable linear parameters, and to models with non-Normal errors, correlated errors, and even to robust error distributions. When appropriate stable parameters are used the speed and accuracy of fitting may be dramatically improved. The method is based on the idea that the solution locus for separable models when fitting the non-linear parameters consists only of those points corresponding to fitted values of the linear parameters, and that for non-Normal errors the data space may be transformed so that the likelihood-ratio criterion behaves like a squared distance.

Estimates of the defining parameters of curves are often so highly correlated and asymmetrically distributed that they are of little use, while stable parameters may not convey the essential algebraic form. As a compromise partially stable parameters may sometimes be defined with interpretable forms and low correlations.

A study of the various roles of parameter transformations in non-linear modelling has been published (Paper 28). The main new results relate to automatic correction for asymmetry in a line search for an optimum, and transformations to straighten curved valleys in a two-dimensional subspace. (Ross)

Fitting differential equations. A fast method of fitting first-order differential equations of the form $y' = f(y,\theta)$ was developed. The model is fitted in stages, first using observations instead of expectations to compute y', and to obtain rough estimates of parameters, then repeating the process using the fitted values previously obtained. The method compares very favourably with the more obvious method of solving the equation for each value of the parameters, which can be very unstable in optimisation. Linear parameters, including the initial value or boundary conditions, may be treated as separable. (Ross)

https://doi.org/10.23637/ERADOC-1-137

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

Dual polyclaves. A polyclave is a set of punched cards (or its equivelent) used to identify specimens from a known set of taxa. It has the advantage that the user is not forced to examine a fixed sequence of characters—as in the identification key—but is free to examine any convenient character. However, the polyclave gives no guidance about which character to examine next during an identification. This deficiency can be overcome by using a *dual polyclave*, which indicates how each character separates the taxa not eliminated by the characters already observed (Paper 22). (Payne and White with Mr R. R. Lamacraft, CSIRO Adelaide)

Estimation methods for generalised linear models (GLMs). The standard method of estimation, as used currently in Genstat and GLIM, may require prohibitively much computer space for large problems with many parameters. Investigation has begun of non-matrix-based methods for maximising the likelihood of GLMs. This includes the method of cyclic ascent. A class of such methods has been defined and although, for some methods in the class, convergence results have been obtained for the solution of standard linear equations, few results have been obtained for the case where the weights alter at each cycle. Nevertheless, the possibilities seem encouraging and a paper is being prepared. (Baker, with Mr M. R. B. Clarke, Queen Mary College, London, and Dr D. Ratcliffe, CSIRO Brisbane)

Statistical programming

Genstat. The number of licensees has risen from 103 to 136 sites, now in 25 countries. The new licensees were organisations in the UK (19), the rest of Europe (6), Australasia (4), Africa (2) and N. America (2). The handover of the distribution of the program to the Numerical Algorithms Group Limited (NAG) was completed, and the new scheme is now working well. (Alvey and Watson)

Releases. Versions of release 4.02 were completed for the CII Iris 80 (Mile A. Bouvier, Jouy-en-Josas, France) and the Honeywell (under GCOS). (Mr J. S. Lemon, University of Aberdeen)

Release 4.03 was made available on the 4-70—later than hoped—in July; it contains a major extension for time-series analysis (see below). Versions for the IBM 370 and similar machines (Mr N. McLaren, University of Cambridge), Prime (Simpson), CDC 6000 series (Dr P. A. Baghurst, CSIRO Adelaide), DEC Vax (Mr A. Nairn, University of Stirling) and Honeywell (GCOS) (Mr J. S. Lemon, University of Aberdeen) are now being distributed. That for the CII Iris 80 is now complete, and it is expected that versions for CDC 7600, ICL 2980 and the DEC-10 will be ready for distribution early in 1981. Conversions are in progress on the Modcomp Classic 7870, Telefunken TR 440, Honeywell (Multics), Xerox Sigma 6, Harris and HP 3000. If all these conversions are completed Genstat will be available on 22 different machines from 18 manufacturers.

New facilities. The principal addition to release 4.03 is a set of new directives for time-series analysis of the kind popularised by Box and Jenkins. The facilities include (i) a set of functions, such as sample and partial autocorrelations, useful in model selection, (ii) directives for filtering and preliminary estimation, (iii) univariate estimation and forecasting using ARIMA models, and (iv) multi-input estimation and forecasting using transfer-function models. There are extensive sets of options for controlling the iteration during estimation and for saving parts of the output. We are greatly indebted to Dr G. Tunnicliffe-Wilson of Lancaster University for designing this section of Genstat and for implementing the algorithms. P. W. Lane wrote the interpreters and embedded

https://doi.org/10.23637/ERADOC-1-137

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

the algorithms in the system. Careful initial planning resulted in a relatively trouble-free incorporation of a major new feature to the program.

Other new facilities include a directive for producing histograms, new functions for tables, and automatic checking for consistency where structures have other structures as attributes, (Alvey, Simpson and Watson)

Documentation and support. An introduction to Genstat on video-tape has been prepared in collaboration with the Audio-Visual Unit of London University. The first tape has been filmed and the second is about to be started. (Alvey and Lane, with Dr D. Clark of the Audio-Visual Unit)

A primer is being written to provide a simple introduction to the basic ideas in Genstat, and to serve as a way in to the main manual. (Lane and Alvey, with Mr N. Galwey, University of Cambridge)

A set of forms has been prepared to aid users with the specification of certain standard analyses. The user fills up the form according to the instructions provided and at the end has a complete Genstat program which he can present to the computer. Forms have been designed for basic statistics, canonical variate analysis, contingency tables, analysis of simple designed experiments, non-parametric statistics, principal-component analysis, probit analysis, regression, simple graphs and histograms, and tabulation. (Alvey)

Genstat was displayed by NAG as part of their stand at the COMPSTAT 80 conference in Edinburgh in August. Daily demonstrations, using the interactive version on a Prime computer at ABRO, attracted considerable interest. (Digby, Lane and Payne)

Introductory courses were given to ADAS at Guildford, also in London, Aberdeen, Reading, Bristol and Nottingham. (Alvey, Digby, Lane and Simpson, with Dr S Evans, University of Bristol)

Genkey. Release 3.02 was completed. New facilities include the incorporation of improved test selection criteria for tests with different costs; the construction of keys to identify the group to which a specimen belongs, based on data describing individual group members rather than group summaries (*Rothamsted Report for 1979*, Part 1, 250); the construction of dual polyclaves; and improvements to the printing of diagnostic tables. (Payne and White)

Maximum Likelihood Program (MLP). Interest in MLP continues to grow and by November 1980 it was being used at over 20 sites. Version 3.06, which was released with the new Manual has been or is being converted for the following machine ranges:

ICL 4-70	Rothamsted
IBM 370	Cambridge University
PRIME	Edinburgh RCC
ICL 2980	Edinburgh RCC
UNIVAC	Copenhagen
Amdahl	Leeds University
Honeywell Multics	Bristol University
Burroughs	Otago, New Zealand
CDC	Manchester RCC
Nord	Reading University

The package was demonstrated at COMPSTAT 80 in Edinburgh, and a poster and a new prospectus were prepared. Various minor amendments were made following tests on different machine ranges and exposure to a wider class of user.

Version 3.07, now being developed, provides the major enhancement of an alternative

https://doi.org/10.23637/ERADOC-1-137

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

optimisation algorithm based on the Gauss–Newton method applied to likelihood-based residuals; this will allow problems with more parameters to be fitted, space permitting. Other important additions are a set of regression functions that can be used within models; these allow, amongst other things, parallel curve analysis to be extended to general models and local smoothing functions to be fitted to irregularly spaced data. Automatic plots of general models were added, and some initial-value procedures were improved. (Ross and White)

Generalised linear interactive modelling (GLIM). Release 3 of GLIM (distributed by NAG) is now to be found in over 300 centres throughout the world, and versions are available for 25 machine ranges. Release 4 has been under development during the year, incorporating suggestions from users together with new features defined by the originators. The opportunity has been taken to modularise the code further and to remove restrictions on, e.g., directory size, by moving to a dynamic space-allocation system. In addition provision is being made for alternative kernels within the system, whereby, for example, the existing kernel for fitting generalised models (GLMs) could be replaced without difficulty by one for analysing data from designed experiments.

The 'housekeeping' routines for release 4 are virtually complete; these include an item recogniser, routines for space control and garbage collection, for input and output control, and for the compilation and execution of macros. (Baker) The code for GLM model specification and fitting is complete, and awaits incorporation. (Mr M. R. B. Clarke, Queen Mary College, London) New facilities include declaration of standard order for units, user-defined contrast for indexing factors, the inverse Gaussian distribution for errors, the use of a grouping factor to produce within-group analyses, and the Givens algorithm as an alternative to the existing Gauss–Jordan one for fitting GLMs. (Baker and Nelder, with Mr M. R. B. Clarke)

Work on an alternative kernel to provide the Genstat algorithm for the analysis of designed experiments has resulted in a common basic set of routines usable in either package. In addition interpreters and other linking routines have been written to imbed the algorithm in the GLIM framework. (Baker and Payne, with Dr J. D. Henstridge, University of Western Australia)

Contributions have been made to the GLIM newsletter and a set of conversational examples for teaching the use of the package has been developed and successfully used in courses. (Baker and Nelder)

Macros have been developed to handle control mortality in probit analysis, for Wadley's problem, and for using Wedderburn's solution to the problem of a model specified in terms of constraints. (Baker)

Rothamsted General Survey Program (RGSP). We helped in the preparation of a new manual for Part 1 of the Rothamsted General Survey Program (RGSP, Mark 2). The Mark 2 release includes improved facilities for data input and validation, also table-manipulation subroutines which aid in calculation of sampling errors, especially for hierarchical data. (Church, with Beasley and Yates, Computer Department)

Future developments. So far our programs have been written in Fortran, a language, which though still widespread, increasingly looks its age. There is little doubt that with modern high-level languages. statistical programs would be easier to write, and to maintain. Unfortunately the Babel of languages is as great as ever, and availability on many machines of efficient compilers remain both a necessity for us and a condition that is hard to fulfil.

In conjunction with the Division of Mathematics and Statistics, CSIRO (Australia),

https://doi.org/10.23637/ERADOC-1-137

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

a project is under way to evaluate several high-level languages with respect to our needs, some of which are discussed in Paper 20. Attention will be paid to ease of definition of complex data structures, the amount of code generated by the compiler and efficiency of execution. (Nelder, Payne and Simpson)

Commonwealth and Overseas

The task of our ODA-sponsored unit remains the provision of a biometric service for agricultural research workers, primarily in countries where ODA supports projects.

Overseas visit. One visit was made in 1980, by D. A. Preece to Tanzania for 16 days in May–June. He visited the Oilseeds Research Project, Naliendele, and the Cotton Research Project, Ukiriguru. Discussions with the Project Leader at Naliendele helped to clear up statistical matters that had led to difficulties in postal correspondence between Naliendele and Rothamsted. Preece also met several of the British workers involved in the Lindi/Mtwara Regional Integrated Development Programme (RIDEP). At Ukiriguru, he had detailed discussions about the layout and statistical analysis of the balanced-lattice experiments in the cotton-breeding programme. He also discussed other statistical matters with three of the British ODA-funded staff at Ukiriguru and assessed the computing needs of the cotton programme.

Work on ODA projects. Most of this work concerned the design and analysis of agricultural experiments. The experiments were on many crops (including maize, rice, wheat, triticale, millet, sorghum, cowpea, groundnuts, sesame, sunflower, cotton, tobacco, sugar, tomatoes, potatoes) and were in many lands (Belize, Bolivia, Botswana, the Falkland Islands, the Gambia, Ghana, Kenya, Malawi, Nepal, Paraguay, Sierra Leone, Sudan, Tanzania, Uganda, Zambia).

Unsatisfactory data-recording was again a matter for serious concern. Some of the problems encountered involved the degree of precision to which data were recorded. More alarming, however, was the presence of salient anomalies in several sets of data—anomalies that the experimenters had overlooked but which were sufficient, if not taken into account, to distort the overall results seriously.

K. Ryder produced a Genstat macro specifically for the many balanced lattices analysed during the year, and he mounted computer programs written by Professor S. C. Pearce (University of Kent at Canterbury) for the 'post mortem' analysis of field experiments.

Intercropping. Janet Riley continued compiling her bibliography of intercropping, and entered into collaboration with Dr R. Willey and Dr C. Floyd (ICRISTAT) in connection with this. Mr R. Mead (Department of Applied Statistics, University of Reading) and Riley completed a paper entitled 'A review of statistical ideas relevant to intercropping research'.

Collection of intercropping data from institutes and from published papers is being continued. This is vital to the furtherance of statistical research. (Riley)

Precision of data. This is a topic on which most statistical textbooks and courses are silent, but on which agricultural experimenters need clear and well-founded advice. Relevant theory in the statistical literature seemingly relates only to data recorded with a 'satisfactory' degree of precision, and to providing a rough indication of what may be taken as 'satisfactory'; no literature has been found on the possible consequences of using an insufficient degree of precision. Empirical investigations of the problem have been started. (Preece and Riley)

https://doi.org/10.23637/ERADOC-1-137

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

Visitors and training. More than 30 visitors were received, including various workers from ODA projects. These visitors were concerned with agricultural work in Bolivia, Botswana, Ecuador, the Falkland Islands, the Gambia, Ghana, Indonesia, Iran, Japan, Kenya, Malaysia, Nigeria, Paraguay, Sierra Leone, Syria, Tanzania, Trinidad and Tobago and other Caribbean countries, Turkey and Zimbabwe.

An Ethiopian student, Mr I. Bekele, and an ARC bursar, Ms Brenda Roberts, spent training periods with the Unit during the summer.

Staff and visiting workers

We were saddened by the deaths of R. H. Wimble in July after 10 years in the Department, and of Mrs. Betty H. Mayton after 5 years. J. G. Pearlman, S. A. Watson, P. J. Zemroch, Linda C. Bassill and Gillian M. Carter left during the year. J. Wood spent 3 months at the University of Western Australia (February–April).

J. A. Nelder visited the National Institute of Health (Washington, DC) and Bell Telephone Laboratories (Murray Hill, NJ) in March, and in May he gave an invited paper at the Journées de Statistique held at the Université Paul Sabatier, Toulouse. He attended the 8th World Computer Congress in Tokyo and Melbourne in October, followed by a visit to the Division of Mathematics and Statistics in CSIRO at its branches throughout Australia.

J. C. Gower was an invited lecturer at the Pyschometric Society meeting in Iowa City in May and lectured also at Yale and Princeton Universities and at Bell Telephone Laboratories (Murray Hill, NJ). With P. G. N. Digby he gave an invited lecture to the Conference on Graphical Methods in Multivariate Analysis held in Sheffield in March. Jill F. B. Altman also attended.

The Department was strongly represented at the ARS Statisticians Conference at Warwick University in July, J. A. Nelder and D. A. Preece giving talks, and also at the COMPSTAT 80 meeting in Edinburgh in August. There Baker, Payne and Ross gave papers (Papers 12, 20 and 28) and all helped in demonstrating our programs.

G. J. S. Ross gave talks to local groups of the Royal Statistical Society in Exeter and Newcastle, and presented a paper at the ARC Crop Modellers Conference in Littlehampton (Paper 29). J. N. Perry addressed the Mathematical Ecology Workshop of the Biometrics Society at Cambridge, and was a discussant at the April meeting of the Royal Entomological Society on mathematical modelling. J. Wood gave a paper at the 15th Colloquium of the International Potash Institute at Wageningen, Holland.

Dr A. G. Constantine returned to CSIRO Adelaide in April after a year in the Department. Dr J. D. Henstridge from the University of Western Australia joined the Department in September for 4 months, and in December Mr J. E. Chiria from Namulonge Research Station, Uganda, joined the ODA Section for 3 months.

Publications

General Papers

1 Altman, J. F. B. (1979) Statistical perspectives in the use of analytical results. In: *Nutrition and Analytical Chemistry Joint Training Conference Papers*. London: ADAS, pp. 17–22.

2 Altman, J. F. B. (1980) Experimentation. In: *Feeding strategies for dairy cows*. Ed. W. H. Broster, C. L. Johnson & J. C. Taylor. London: ARC, pp. 15.1–15.14.

https://doi.org/10.23637/ERADOC-1-137

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

3 BEASLEY, J. D., CHURCH, B. M. & YATES, F. (1980) *The Rothamsted General Survey Program (Part I) Program Manual.* Rothamsted Experimental Station (RGSP21).

4 CHURCH, B. M. (1980) *Fertiliser practice on cereals, 1978* (SS/C/711), 10 pp.

5 CHURCH, B. M. (1980) *The pH and lime requirements of agricultural soils in England and Wales, 1969–78: Information from the Representative Soil Sample Survey* (SS/CH/2), 10 pp.

6 CHURCH, B. M. & LEECH, P. K. (1980) Fertiliser use on farm crops in England and Wales, 1979, with comparative metric data from 1970. London: Ministry of Agriculture, Fisheries and Food (SAFC/39), 25 pp.

7 (HARDY, R.) & ALTMAN, J. F. B. (1981) A comparison of autumn-born Friesian or Hereford × Friesian calves finished out of yards at 18–21 months, or grass finished at 2 years old. *Experimental Husbandry* **37**, 83–93.

8 SIMPSON, H. R. (1980) *Genstat Installation Guide* (2nd edition). Statistics Department, Rothamsted Experimental Station, 32 pp.

9 SIMPSON, H. R. (1980) *Genstat Implementer's Manual* (2nd edition) Statistics Department, Rothamsted Experimental Station, 66 pp.

PAPER IN ROTHAMSTED REPORT, PART 2

10 CHURCH, B. M. (1981) Use of fertilisers in England and Wales, 1980. *Rothamsted Experimental Station. Report for 1980*, Part 2, 115–122.

RESEARCH PAPERS

11 BAILEY, L., BALL, B. V. & PERRY, J. N. (1980) Viruses for the honey bee, *Apis mellifera. Annals of Applied Biology* **97**, 109–118.

12 BAKER, R. J. & (RICHARDSON, M. G.) (1980) The development of the GLIM system up to GLIM-4. *Proceedings of the 4th Symposium on Computational Statistics, COMPSTAT 1980.* Wien: Physica-Verlag, pp. 523–529.

13 BANFIELD, C. F. & GOWER, J. C. (1980) A note on the graphical representation of multivariate binary data. *Applied Statistics* **29**, 238–245.

14 BENZIAN, B. & LANE, P. W. (1981) Interrelationship between N concentration in grain, grain yield and added fertiliser nitrogen in wheat experiments of South-east England. *Journal of the Science of Food and Agriculture* **32**, 35–43.

15 BROMILOW, R. H., BAKER, R. J., FREEMAN, M. A. H. & (GOROG, K.) (1980) The degradation of aldicarb and oxamyl in soil. *Pesticide Science* **11**, 371–378.

16 GOWER, J. C. (1980) A modified Leverrier-Faddeev algorithm for matrices with multiple eigenvalues. *Linear algebra and its applications* **31**, 61–70.

17 GOWER, J. C. (1980) An application of the Leverrier–Faddeev algorithm to skew-symmetric matrix decompositions. *Utilitas Mathematica* **18**, 225–240.

18 GOWER, J. C. (1980) Problems in interpreting asymmetric chemical relationships. *Chemosystematics: principles and practice*. Eds. F. A. Bisby, J. G. Vaughan & C. A. Wright. London: Academic Press, pp. 399–409.

19 (JOHNSON, P. A.) & ZEMROCH, P. J. (1980) Manurial experiments with spring barley on non-chalk soils. *Experimental Husbandry* **36**, 34–43.

20 (LAMACRAFT, R. R.) & PAYNE, R. W. (1980) A new look at data structures for statistical languages. *Proceedings of the 4th Symposium on Computational Statistics, COMPSTAT 1980*: Wien: Physica-Verlag. pp. 436–469.

https://doi.org/10.23637/ERADOC-1-137

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

21 McEwen, J., Simpson, H. R. *et al.* (1981) The effect of irrigation, nitrogen fertilizer and the control of pests and pathogens on spring-sown field beans (*Vicia faba* L.) and residual effects on two following winter wheat crops. *Journal of Agricultural Science, Cambridge* **96**, 129–150.

22 Payne, R. W., (Lamcraft, R. R.) & White, R. P. (1981) The Dual Polyclave: an aid to more efficient identification. *New Phytologist* **87**, 121–126.

23 Payne, R. W. & Preece, D. A. (1980) Identification keys and diagnostic tables: a review (with discussion). *Journal of the Royal Statistical Society*, Series A 143, 253–292.

24 Pearlman, J. E. & Lawlor, D. W. (1981) Tracer experiments and compartmental modelling in analysis of plant metabolism. In: *Mathematics and plant physiology*. Ed. D. A. Rose & D. A. Charles-Edwards. London: Academic Press, pp. 37–52.

25 Perry, J. N., Macaulay, E. D. M. & Emmett, B. J. (1980) Phenological and geographical relationships between catches of pea moth in sex-attractant traps. *Annals of Applied Biology* **97**, 17–26.

26 Perry, J. N., Wall, C. & Greenway, A. R. (1980) Latin square designs in field experiments involving insect sex attractants. *Ecological Entomology* **5**, 385–396.

27 Preece, D. A. (1981) Distributions of final digits in data. *The Statistician* **30**, 33–62.

28 Ross, G. J. S. (1980) Uses of non-linear transformations in non-linear optimisation problems. *Proceedings of 4th International Symposium on Computational Statistics COMPSTAT 1980*. Wien: Physica-Verlag, pp. 381–388.

29 Ross, G. J. S. (1981) The use of non-linear regression methods in crop modelling. In: *Mathematics and plant physiology*. Ed. D. A. Rose & D. A. Charles-Edwards. London: Academic Press, 269–282.

30 Sawicki, R. M., Devonshire, A. L., Payne, R. W. & Petzing, S. M. (1980) Stability of insecticide resistance in the peach-potato aphid, *Myzus persicae* (Sulzer). *Pesticide Science* **11**, 33–42.

31 Taylor, L. R., Woiwod, I. P. & Perry, J. N. (1980) Variance and the large-scale spatial stability of aphids, moths and birds. *Journal of Animal Ecology* **49**, 831–854.

32 Wall, C. & Perry, J. N. (1980) Effects of spacing and trap number on interactions between pea moth pheromone traps. *Entomologia experimentalis et applicata* **28**, 313–321.

33 Wall, C. & Perry, J. N. (1980) Interactions between insect pheromone traps: a study of behaviour in the male pea moth. In: *Olfaction and taste* VII. Ed. G. van der Starre. London: Information Retrieval Ltd, p. 410.

34 Wimble, R. H. (1980) Theoretical basis of fertiliser recommendation. *Chemistry and Industry*, 680–683.

https://doi.org/10.23637/ERADOC-1-137