A History of the GLIM Statistical Package

Murray Aitkin

Mathematics and Statistics, University of Melbourne, 300 Swanston Street, Melboune, Victoria 3000, Australia
E-mail: murray.aitkin@unimelb.edu.au

Summary

This article documents the history of the GLIM statistical package, from Royal Statistical Society documents, GLIM Working Party documents and recollections of the developers of GLIM and its supporting papers, manuals and books.

Key words: GLIM, GLM, Nelder, Wedderburn, Baker.

1 Introduction and Acknowledgements

I began this history because I feared that the great value of this package, and its popularity in the 1970s and 80s, would be forgotten with the passage of time and the plethora of packages that developed before and after its demise. It is more than 30 years since my 1-year term as Chairman of the GLIM Working Party of the Royal Statistical Society, and records and memories are both fading.

There is another important reason for my concern with this history: the GLIM story is to me another sad example of a brilliant British innovation, which failed because of the narrow limitations placed on its development.

I am grateful to many people for information about the development progress of GLIM, especially Bob Baker, Brian Francis and Bob Gilchrist, to Janet Foster, the archivist of the Royal Statistical Society (RSS) for assistance with the minutes of meetings of the Society, and to Gail Austin for information from NAG. The early development of GLIM is also well-documented in the papers by Richardson & Baker (1980) and Payne (2004).

As far as possible, I have relied on Society documents for this history, but some meetings have not been minuted or are incompletely minuted. The views expressed in the article are my own and are not endorsed by the Royal Statistical Society. I use the third person in references to my own contributions.

2 Background to the GLM Paper

John Nelder was fundamental to the early development of GLIM through his work on generalized linear models (Section 3). He was Head of the Statistics Department at Rothamsted Experimental Station and Chairman of the RSS Working Party for Statistical Computing (later renamed the GLIM Working Party), which was formally constituted as a working party of the Society on 23/6/71. He was Chairman of the RSS Executive and Policy Committee from 21/12/85 and became President of the RSS in 1985. He also held a Visiting Professorship of Statistics at the Imperial College Department of Mathematics from the early 1970s.

Nelder's contributions to statistical computing at Rothamsted were described at length by Roger Payne and John Gower, who both gave (Payne, 2004 pp. 110-111; Gower, 2015 pp. 366-367) brief descriptions of GLIM development. Nelder's obituary in J.R.Statist.Soc A (2013) 174, part 2, pp. 499-504 gives additional information about his life.

The history given here is taken from the minutes of meetings of the RSS Working Party for Statistical Computing (WPSC), the GLIM Working Party (GWP), the GLIM sub-committee of the RSS Council (GSC), the Executive and Policy Commitee (EPC) of the RSS Council, and the full RSS Council (COU). Bob Baker and Bob Gilchrist have provided additional information on the early development.

3 Generalized Linear Modelling and the Start of GLIM

Theoretical and applied statistics were both convulsed by the publication of the GLM paper by Nelder & Wedderburn (1972). It showed that regression models in the one-parameter exponential family distributions could be fitted by maximum likelihood using an iteratively weighted least squares (IWLS or IRLS) algorithm.

Special cases of this result were already in use for the probit and logit regression models (Finney, 1947 and Lord, 1952, respectively), but the generality of the result meant that a concise algorithm could be constructed for all the one-parameter exponential family distributions and could be extended to the two-parameter family (including the normal and gamma distributions) by a simple (non-maximum likelihood) estimation of the second scale parameter.

Nelder and Wedderburn at Rothamsted began from early 1972 to develop code for the algorithm. Members of the WPSC became participants in this development, especially Mike Clarke at Queen Mary College. Release 0 of GLIM was distributed in September 1973 and release 1 in 1974.

4 Baker's Appointment

Bob Baker took a first degree in Leicester and enrolled in a Master's degree in statistics at Newcastle. Robin Plackett, the department head, suggested that as GLMs were an interesting new topic, Baker might like to write a system for fitting these models in APL, the then-current statistical computing system in Newcastle and elsewhere. Baker developed the GRAPL (Generalized Regression in APL) package and graduated with his MSc in October 1975.

Wedderburn died in a climbing accident in June 1975; Baker was hired in October 1975 to replace him in the GLIM development. He found that the early versions of GLIM were very rough, and his earliest work was to improve the current version. He became responsible for GLIM programming and had other responsibilities as well in the Rothamsted statistics department, as was standard at the time: he was a consultant to the agricultural research workers on data analysis (as was Wedderburn previously), and over many years, developed in FORTRAN a simulation model for the leaching of pesticides through soil into groundwater.

At this time, GLIM was a minor development at Rothamsted. Genstat development had the major effort, but Baker did not work on it: he was the 'backroom boy' working on GLIM. Release 2 was announced in the Society's journal Applied Statistics (Nelder, 1975) and was reported on by Council in their 1976 report:

> The second release of GLIM, a package sponsored by the Working Party, was made by the Numerical Algorithms Group (NAG) at Oxford on 1 October 1975. It contained numerous improvements, including the provision of macros, line-printer plots and sorting. At 31 March 1976, a total of 56 copies of the package had been distributed to 12 countries, for use on six machine ranges.

Conversions to three more ranges have been made independently by interested users. Considerable work has been done on a third release (due in 1977), including the provision of branching and looping instructions.

5 Early Difficulties

The early program syntax was little help, and the program manual was cryptic. As Lindsey put it later (1997 p. vi):

> The introduction of the idea of generalized linear models in the early 1970s had a major impact on the way applied statistics is carried out. In the beginning, their use was primarily restricted to fairly advanced statisticians because the only explanatory material and software available were addressed to them. Anyone who used the first versions of GLIM will never forget the manual which began with pages of statistical formulae, before actually showing what the program was meant to do or how to use it.

The power and flexibility of the early package were impressive once you understood how to use it, though the actual data handling tools in the package were quite restricted. If these could be made more flexible, as they were in later releases, the package could be very powerful.

GLIM had several unusual features, reflecting Nelder's prejudices. Parameter estimates and standard errors were given by default to only four figures, though more could be requested. This was a welcome change from the 12 figures given in many large packages, most of which were irrelevant (or wrong).

A less welcome feature was the estimation of the scale parameters in the normal and gamma distributions. These estimates were not ML, though that for the normal was the restricted maximum likelihood (REML) estimate. For the gamma distribution the GLIM estimate—the *scale factor* based on the *mean deviance*—was ML only asymptotically, and in small samples could be far away from the MLE. This did not affect the (regression) parameter estimates, but it did affect their standard errors, which were scaled by the scale factor.

A similar feature was the calculation of the *deviance*, a term invented by Nelder. In models other than the normal, the deviance was twice the negative log maximized likelihood of the fitted model, relative to twice the negative log maximized likelihood of the *saturated* model, with a parameter for each observation giving a perfect fit. So the deviance was the likelihood ratio test statistic for the fitted model relative to the saturated model. For the normal model, however, the deviance was the residual sum of squares from the fitted model. This inconsistency arose from the (assumed) *asymptotic* χ^2 distribution of the deviance in the non-normal models, and the *exact* $\sigma^2 \chi^2$ distribution of the residual sum of squares in the normal model. This cause initial confusion, which was increased by the assumption that the deviance for non-normal models would have its usual asymptotic χ^2 distribution and so could be used as a *goodness-of-fit* test for the fitted model.

This assumption was incorrect, because as the sample size increased, so did the number of parameters in the saturated model, at the same rate. The likelihood in the saturated model parameters never approached its asymptotic form: there was always one observation per parameter. It was only for deviance *differences* between two fitted (and nested) models that the asymptotic distribution might be assumed.

6 The EM Algorithm

A major stimulus to the extensions of GLMs came from the publication of the EM algorithm (Dempster, Laird & Rubin, 1977). It was presented in November 1976 as a read paper to a

Research Section meeting of the Royal Statistical Society. It was attended by a huge audience of statisticians from everywhere.

The idea was very powerful: any data structure with missing or incomplete data could be expressed in terms of a complete data model and a missingness process for the incompleteness. If missingness was at random, then standard ML algorithms for complete data could be adapted to give the ML estimates for the observed incomplete data using a simple 'EM' algorithm, which could be easily programmed in GLIM.

However, the computation of the covariance matrix of the ML estimates required several matrix operations at convergence of the algorithm, which were not provided in the GLIM computational tools. This applied to essentially *all* of the applications of the EM algorithm, so it was a severe limitation of the computational language for this class of incomplete data models. Ad hoc methods could be developed: one was to equate the squared Wald test statistic for a parameter to the likelihood ratio test statistic (lrts) for a zero value of it:

$$\frac{\widehat{\theta}^2}{\text{SE}(\widehat{\theta})^2} = \text{lrts}$$

$$\text{SE}(\widehat{\theta}) = \frac{\widehat{\theta}}{\sqrt{\text{lrts}}}.$$

This was valid only asymptotically in general, when the likelihood was normal.

In later releases of the package, this was ameliorated by the provision of the $PASS directive, which allowed structures to be passed out of GLIM into the operating system, where the matrix operations could be performed in a FORTRAN routine and then passed back into GLIM for further analysis. While this could extend substantially the range of models, which could be fitted in GLIM/FORTRAN (Rigby and Stasinopoulos 2001 and 2005 used this for their Generalized Additive Models extension), it was cumbersome for simple models, for which a FORTRAN routine would have to be written, and was no substitute for matrix tools in the package.

7 The Lancaster Centre for Applied Statistics

Murray Aitkin was appointed Professor of Applied Statistics and Director of the new Centre for Applied Statistics (CAS) at the University of Lancaster in 1979, after a 3-year SSRC Professorial Fellowship in the Department of Mathematics there. Aitkin was awarded an SSRC large programme grant to develop applications of the EM algorithm for the statistical modelling of complex social data: mixtures, multilevel structures and missing data among others, implemented in GLIM and Genstat. The programme also included the development of a user-friendly manual for GLIM, supported two research associates and provided funds to support annual distinguished visitors for periods of 1 to 2 months at Lancaster. This was an essential inducement, given Lancaster's small size and remoteness.

This boost in computational applied statistics gave a spur to the further development of GLIM: coincidentally release 3.12, a substantal upgrade to release 2, came out in 1979. The CAS became an academic centre for applied GLIM development, as was the Department of Mathematics, Statistics and Computing at the Poytechnic of North London, with Bob Gilchrist playing a major role. This evolved into the StORM research centre (Statistics, OR, Mathematics) at the renamed University of North London, now London Metropolitan University. Programming was done initially at Rothamsted.

However, in the 1980s, the Thatcher Government contractions in public spending had a drastic effect on GLIM development, as they did on University finances.

8 Financial Clouds

Financial cuts to the UK Agricultural Research stations' budgets affected statistical computing at Rothamsted. Nelder was directed by the Rothamsted administration to stop working on the second Rothamsted statistical package. He arranged the transfer of the copyright in the package from Rothamsted to the Royal Statistical Society.

Nelder enlisted members of the WPSC in GLIM's further development with Baker, particularly Mike Clarke at Queen Mary College. He remained Chairman of the WPSC and continued to supervise GLIM development. The facilities in GLIM continued to be developed, though there was no timetable or administrative framework for a further release. In 1980, the Annual Report of Council reported on the Working Party:

> Mr R.J. Baker and Mr M.G. Richardson joined the Working Party during the year. Four meetings were held, with three main lines of work continuing. ... The distribution of release 3 of GLIM had reached 273 copies by 4 March 1980, and the first issue, in December 1979, of a GLIM newsletter, edited by Mr C.D. Payne and Mr M. Richardson, has been well received. Discussions have been held on the contents of GLIM release 4, and coding of this version has now started. ...

In this development process, major contributions (unpaid) were made by (alphabetically, in addition to Baker):

- Mike Clarke—numerical analysis;
- Mick Green—programming;
- Allan Reese—programming ideas;
- Mel Slater—advanced graphics.

Richardson & Baker (1980) referred to the next planned release of GLIM, GLIM 4, with new features. Coding of GLIM 4 was said to be already well under way, with distribution of the package scheduled for 1981.

9 External Support for GLIM

9.1 Why was GLIM so successful?

There were several reasons for GLIM's very rapid success:

- because the theory was new, GLIM had no competitors in computational statistical packages at the time, in several critical areas like binomial and Poisson regression;
- because it was *general* over the exponenial family: only one set of directives was needed for the entire family;
- the language was parsimonious (a characteristic of Nelder).
- the package was small but very fast.

Aitkin called it "the Mini-Minor of statistical packages: fun to drive and fast but rough and uncomfortable".

An example will illustrate its power and simplicity. A canonical "difficult" data set was the Quine data on absence from school of 146 Australian Aboriginal and white children (Aitkin 1978, Aitkin, Anderson, Francis and Hinde 1989 pp. 152–164, 223–225). The response DAYS was a non-negative integer, with four categorical variables (age A, sex S, culture C and learner L (slow or normal) in a severely unbalanced cross-classification with several empty cells. Aitkin (1978) used the normal model in the main text analysis:

$factor C 2 S 2 A 4 L 2 $
$yvar DAYS $error n $link i $fit C*S*A*L $disp e $

The first line declares the four variables as factors, with the appropriate number of levels. The second specifies the response variable, the probability ('error') distribution (a term disliked by many statistician, especially for discrete distributions), the transformed scale of the mean on which the regression model is fitted (log), the regression model to be fitted and the "display" directive, which controlled which parts of the output are to be printed to the screen, here just the estimates and their standard errors. The deviance and its degrees of freedom were printed automatically as well, and the number of iterations needed for convergence of the scoring algorithm. The model is of deficient rank, but this causes no difficulty: non-estimable high-order interactions are aliased: set to zero.

For the normal model, the error and link (identity) specifications are the defaults and do not need to be given explicitly. In the paper discussion, several discussants suggested the log transformation, and in his reply, Aitkin gave this analysis, using D1 = log(DAYS+1) to avoid the log of zero:

$calc d1 = %log(DAYS+1)$
$yvar d1 $ $fit C*S*A*L $disp e $.

An alternative model is the Poisson, though the variability of absence within individual cells of the four-way table was much greater than for a Poisson distribution. To fit the Poisson model on the log mean scale with the full four-way cross-classification model (corresponding to the lognormal model), the directives are

$yvar DAYS $error p $link l $fit C*S*A*L $disp e $.

This format for the model specification was followed very closely in S and R with the *glm* function: Venables and Ripley (2002 pp.169–176) gave a very detailed discussion of the Quine data using the same *glm* approach.

9.2 Conferences and short courses

Soon after GLIM 3 was distributed, it was apparent that GLMs, GLIM and their possibilities deserved major conferences, associated with training courses in the uses of the package. The first of these was held very successfully 13–15 September 1982 at the Department of Mathematics, Statistics and Computing of North London Polytechnic. It attracted more than 40 statisticians from the UK, Europe, Australia and North America. The full conference proceedings were published in the Springer Lecture Notes in Statistics series (number 14, Gilchrist 1982).

The CAS ran annual training courses at Lancaster in GLM applications using GLIM from 1980 to 1985. The 1985 course preceded the second GLIM conference at Lancaster in September 1985, at which attendance increased. The conference proceedings were again published in the Springer Lecture Notes in Statistics series (number 32, Gilchrist, Francis & Whittaker 1985). This form of conference, combining theoretical and computational papers with applications papers in GLIM, was found very valuable by the attenders, and several new developments resulted from it.

There were two further GLIM conferences: at Trento in 1989 and Munich in 1992, with proceedings again published by Springer in the Lecture Notes series. Conferences not involving GLIM developments explicitly were also held every year after 1985, rotating around the countries and Universities that supported them and were called the International Workshops in Statistical Modelling (IWSM). These explored broader statistical modelling both within and beyond the GLM and GLIM frameworks. The first was held in 1986 in Innsbruck.

9.3 Books

The need for a simple and clear exposition of the power of GLMs, and of GLIM, increased steadily with GLIM's increasing success. Several early manuals were written to help the novice into GLIM, notably Gilchrist and Green's 1980 book.

The short courses run from the CAS became the basis of the GLIM manual Aitkin had promised to the SSRC, but it also became clear that it was necessary to cover the theory of GLMs and whatever extensions of them could be developed in the book publishing time horizon.

McCullagh and Nelder's book appeared in 1983 (second edition 1989). It sold over 25 000 copies. It covered the theory expertly, as expected, but there was a curious feature of the examples. There were many of them, some now classics of the field, but there were no details of how they were fitted, no GLIM (or Genstat) code, and very few (4) mentions of GLIM except in the Software Appendix E (which also announced the new package PRISM, 'now under development'). This was curious, as McCullagh had used GLIM extensively in his published papers. It was as though GLMs could be fitted by hand-waving.

Some lower-level books, for example by Annette Dobson (1990), followed the same line: it was sufficient to express the data structure as a GLM and give the parameter estimates and standard errors for an example.

Adena & Wilson (1982) went in the opposite direction, re-expressing a large number of epidemiological examples in the major book of Breslow & Day (1980) as particular examples of GLMs and then showing how these could be fitted in GLIM. Healy's (1988) introductory book arrived later, as did Lindsey's more advanced books (1989, 1992) and the CAS GLIM book (Aitkin, Anderson, Francis & Hinde, 1989).

Early versions of the CAS text were supplied to participants in the CAS short courses. The book was nearly ready in 1985: the final chapter on survival analysis was being written just before the CAS training course in this area in 1985, preceding the 1985 GLIM conference. The release of GLIM 3.77 in October 1985 required revisions of the macro use, and Aitkin's personal crisis and departure from Lancaster delayed considerably the completion of the book. The first edition was very successful (though not on the McCullagh and Nelder scale), selling 5 543 copies and being widely adopted as a teaching text.

9.4 The Statistical Modelling Society

The Society idea developed at the 1985 GLIM conference at Lancaster, where participants felt that both annual Workshops (the IWSMs) and a formal journal (Statistical Modelling) would build on the impetus provided by the two GLIM conferences. The Society website
`http://www.statmod.org`
gives the Workshop history:

> The workshop grew out of two GLIM conferences held in London (1982) and Lancaster (1985), and a number of short courses organised by Murray Aitkin at Lancaster University in the early 1980s. The focus then was on generalised linear modelling, and many European statisticians were attracted to these activities. At this time, a group of Austrian, Italian and British statisticians saw both the opportunity and the need for a continuation of regular meetings of similar spirit. These meetings should focus on various aspects of statistical modelling in an informal workshop environment, specifically aimed at applied statistics, but also including theoretical developments and computational methods.
>
> The inaugural workshop took place in Innsbruck in 1986 and brought together an enthusiastic group of about 30 European statisticians interested in statistical modelling. The programme focused

on GLMs and was characterised by a number of features—a friendly and supportive academic atmosphere, tutorial sessions and invited speakers presenting new developments in statistical modelling. The academic programme allowed plenty of time for presentations and for discussions, and made available copies of all papers beforehand. A very well organised social programme complemented the activities.

GLIM was not an essential element of papers or for participants, and the breadth of applied modelling increased steadily over time. A special 25th Anniversary Workshop was held in Glasgow in 2010. The Web site also gives the journal history:

Statistical Modelling: An International Journal is published by SAGE Publications on behalf of the Statistical Modelling Society. It publishes original and high-quality articles that recognize statistical modelling as the general framework for the application of statistical ideas. Submissions must reflect important developments, extensions, and applications in statistical modelling.

The journal also encourages submissions that describe scientifically interesting, complex or novel statistical modelling aspects from a wide diversity of disciplines, and submissions that embrace the diversity of applied statistical modelling. Furthermore, we would like to get self-contained tutorials on topics that have been spread into several papers. Interesting practical problems using an existing or novel adaptation of a modelling technique are strongly encouraged.

10 1982 PRISM

A major surprise (to non-members of the WPSC) was the announcement at the 1982 GLIM conference (by Nelder in the Proceedings) of a new package, PRISM (PRogram for Interactive Statistical Modelling), based on GLIM. Following his introduction, the next five papers in the Proceedings gave a very detailed description of the new system.

PRISM was to be a modular development with four modules that could be bought or rented separately:

- KERNEL: a set of housekeeping routines for interpreting the user's input, allocating space, formatting output, handling faults and messages; also general facilities for tabulation, array manipulation, sophisticated program-control and data manipulation.
- GLIM 4: the new release of GLIM; new facilities include certain conditional and marginal distributions, composite and other link functions, output of predicted tables and margins, a more general structure formula specification, simplified user-defined models.
- AOV: an analysis-of-variance module, using the algorithm as in the ANOVA directive of Genstat; handles any 'generally-balanced' design; covariance analysis; full output of fitted tables and margins.
- GRAPH: a new module for flexible, device-independent graphics; standard statistical capabilities include histograms, bar-charts, scatter plots, stem-and-leaf etc; other standard features are surface views, contour plots, line-charts; low-level features such as axes, lines, labelling, colouring, area-infill, etc.

Also an implementation of GKS, the new ISO graphics standard.

This was a remarkable step-up from GLIM 3, merging facilities of both GLIM and Genstat.

It was unclear to those outside the GWP how much of this actually existed and how long it would take to deliver the system. Scepticism was fed by the repeated postponements of release dates. An article by Bob Baker in the March 1983 Newsletter promised PRISM in Autumn 1983. In the Editorial in Newsletter 7 (June 1983), PRISM is scheduled for release in June 1984.

In the 1983 Annual Council Report, more delay was evident:

The package GLIM, sponsored by the Working Party, has now been distributed to more than 600 sites in 41 countries. New countries for the year include Chile, Ethiopia, Jamaica, Fiji, Lebanon, New Caledonia, Syria, Taiwan and Thailand. Conversions for six more machine ranges bring the total of ranges for which GLIM is now available to 35. Work on a successor, provisionally called PRISM, has continued, and the aim is to provide a flexible and powerful interactive language suitable for driving modules concerned with fitting generalized linear models, the analysis of designed experiments, array and table manipulation, and graphical facilities as defined in the GKS specification.

In Newsletter 8 (June 1984), it was announced that GLIM 4 was a long way off and that it had been decided to release an updated version of GLIM 3 (provisionally called GLIM 3A). There was no Working Party report in 1984.

In the 1985 Working Party report, PRISM was dismissed in one line:

The Working Party has been very active in the past year and has met seven times. A revised version of GLIM 3 (3.77) has been developed and is now available for many machine ranges. The source code has been rewritten to the Fortran 77 standard and several new facilities have been added to enhance both the modelling capabilities and the command language. Work is in progress to make the program available on low-cost single-user computers. The Manual has been completely revised. The full Manual and a Primer will be published in 1985. Over 900 copies of GLIM 3 have now been distributed worldwide. The work involved in distribution is carried out by NAG Ltd, who also publish the GLIM Newsletter. The contract between NAG Ltd and the RSS is currently being re-negotiated. The Working Party has submitted new draft terms of reference to Council for approval. These will not alter the type of work undertaken but will formalize the position of Working Party members. Further software developments are currently being planned. *The name PRISM, referred to in previous reports, has had to be dropped as it is already the protected name of a computer program.* (emphasis added) The name GLIM has been registered as a trademark in the name of the Society to avoid any similar problems in the future. The members of the Working Party are listed below. Dr J.A. Nelder (Chairman until March 1985), Professor M. Aitkin (Chairman from April 1985), Mr R A. Reese (Hon. Secretary), Mr R.J. Baker, Mr M.R.B. Clarke, Mr B. Francis, Dr M. Green, Mr M. Slater, Dr R. Gilchrist, Mr C.D. Payne, Mr A.V. Swan, Mr R. White, Mr M.G. Richardson and Ms J. Webb (representing NAG Ltd).

This major proposal did not materialise as a system, for reasons, which may have been known to the GWP members, but which were not explained publicly. Clive Payne (p. 111) attributed it to

...the name PRISM had to be abandoned because of trademark clashes, the project itself failed to gel, and the collaboration between Rothamsted and the Working Party on Statistical Computing came to a close in 1984 when John retired as Head of the Rothamsted Statistics Department and Chairman of the Working Party. So instead GLIM 3.77 appeared in 1985 ...

The issue of a reserved name would not have prevented the program's development and distribution. The 'failure to gel' was an ominous sign for GLIM's future.

GLIM development was already straining the GWP members. The Thatcher attack on University funding was placing heavier working loads on University staff everywhere. Department heads were now hoping for funding to come in from their staffs' outside work, rather than this being contributed free to another organisation. At this time, the RSS was earning little from royalties from GLIM sales: in 1980–1982 these were £3 490.78; of this, £2 481 was outlaid by the RSS on a sandwich-course student working on PRISM under Nelder's guidance.

Planning GLIM's further development into something different was bedevilled by financial isssues. These were generally the same or of the same intractable form:

- Where would funding for the development come from? The RSS did not have funds to support such developments.
- Since these developments would have, or wanted to have, the imprimatur of the Society, what proportion of the revenue would the Society receive from sales or leasing of the developed system?
- If the Society did receive such revenue, how much of this, if any, should be used to support the imprimatured development?

GLIM was sold outright, not licensed annually, unlike Genstat. GLIM and Genstat were distributed by NAG, the Numerical Algorithms Group company. The support of the program development by the RSS led to a 'gentleman's agreement' that the RSS would receive from the distributors a proportion of the royalties from GLIM sales. This proportion was not fixed but did not exceed 10%. The program sold very successfully. By December 1981, GLIM 3 was selling 11 copies a month, and 450 copies had been sold since its release. The RSS Council was not satisfied with the 10% royalty deal it had with NAG, and this was increased considerably from the GLIM 3.77 release in 1985. Royalties were calculated on the NAG gross sales, less costs of overseas distribution, at 60% for the first 160 copies and 40% on sales over 160. By 1987, cumulative RSS income from GLIM was more than £69 000. Requests to the RSS Council by the GWP for financial support for GLIM issues were sometimes approved for increased publicity or sales but rarely if ever for actual development of the package.

GLIM 4 was temporarily forgotten in the GLIM 3 update, GLIM 3.77, released in October 1985, using the new FORTRAN 77 compiler. Some of the changes previously announced for GLIM 4 were incorporated in this release. These did not include the array calculation functions, the ANOVA procedure from Genstat or the greatly improved graphics. Some minor changes were also made to the syntax, which caused the Lancaster GLIM book authors some extra work, and delay in their book's completion. Meanwhile, the use and success of GLIM 3 continued to increase.

11 1985

1985 was a turbulent year. Early in the year, Nelder retired from Rothamsted and from the GWP Chairmanship. His retirement was quickly followed by John Gower's appointment as head and Baker's resignation. Rothamsted staffing was being cut, and positions were being lost. Gavin Ross was being asked to resign. Gower told Baker that he wanted GLIM work at Rothamsted to end and that Baker should think of leaving.

Baker had already decided that

- there was no scope in the Government cuts for any future at Rothamsted;
- he was not contributing enough to the GLIM development;
- statistics was not the area he was going to be good at; programming seemed more attractive.

He left Rothamsted for a commercial programming job. Gavin Ross was moved to his Rothamsted slot. Gower (2015 p. 367) commented

In 1985, because of lack of resources, work at Rothamsted associated with the GLIM project was reduced to a residual level.

Aitkin had written to Clive Payne, a GWP member from Oxford, about what he saw as necessary developments for GLIM. Payne passed on his comments to Nelder, who wrote a long letter to Aitkin (3/1/85). This established clearly his view for GLIM development:

Dear Murray,

...The critical question to me is about the directions in which GLIM ought to be expanded; I don't think we can answer this without considering carefully its ecological niche in the software market generally. The most characteristic feature of GLIM at present is its power in relation to its size. I think we should try to preserve this. The new version of GLIM 3 still fits on one IBM-PC-sized floppy disc. This is where it belongs, I contend.

We had long discussions with NAG about the relative positions of GLIM and Genstat, which they need to establish to sell them properly. Genstat is the large system, including e.g. matrix arithmetic, multi-variate analysis procedures, classification procedures and so on. Genstat includes the GLM extension to regression, *of course* (emphasis added). GLIM is built around GLMs and is intended to be much smaller.

The point of making these distinctions is to underline that what I think you are pressing for is a new and better version of Genstat. Now before I left Rothamsted we were well along the way to a complete redesign of the Genstat language to make it modern, powerful, and more systematic, while preserving existing algorithmic code as far as possible. John Gower, who succeeded me, is continuing to direct this work, and he is very keen to obtain collaboration (with associated external funding) for future developments in this area. I would like to encourage you very strongly to make contact with him and to see what can be done towards making Genstat V the sort of package you have in mind. I am not saying this, I should stress, in any way to divert your interest in GLIM, but because I think your ideas are much more naturally compatible with the plans for Genstat V. When you have had time to look at the spec. for Genstat V, I should be very glad to have your reactions to these thoughts. ...

Yours sincerely, John Nelder

Aitkin replied to him on 15/1/85:

Dear John,

Many thanks for your letter and the consideration of my comments to Clive. As you know I have always felt that our position with respect to GLIM was rather anomalous, and our book perhaps underlines the anomaly. As we become more closely involved with GLIM, its limitations become increasingly frustrating.

Of course I am well aware of the GLIM/Genstat philosophy, and I will certainly find out from John Gower what the Genstat plans are. Perhaps this will change my view: unlike you and John Hinde and Dorothy Anderson, I am only able to master one system (if that!), and therefore feel that GLIM is so good it's a shame not to extend it in those areas of *model fitting* (emphasis in original) where it is currently lacking. ...

With kind regards,
Yours sincerely,
Murray Aitkin

Aitkin wrote the same day to Gower:

Dear John,

Belated congratulations on your appointment, and best wishes for the future!

I have been discussing the future of GLIM with John Nelder and other members of the working party, and John has suggested I contact you to find out the plans for Genstat V and how these might relate to my views of GLIM development. Could you send me specifications for the proposed system, or working papers if the specifications are not yet finalised?

With kind regards.
Yours sincerely,
Murray Aitkin

Gower replied (21/1/85):

> Dear Murray,
>
>   I can't be much help on Genstat 5. We have bulky files full of paper but no consolidated document. The project is now well under way and we hope to demonstrate a preliminary version at the Genstat conference in York in September.
>
>   The main features of Genstat 5 are:
>
> 1 It is fully interactive
> 2 The syntax is consistent and simplified
> 3 It will have a powerful procedure mechanism (rather than macros)
> 4 Graphics will be much improved
> 5 It will be more user-friendly (menus, etc.).
>
>   I don't know what plans you have for GLIM but Bob Baker tells me that one idea is to add an algorithm for generalised variance component estimation. That at least should be easily incorporated into Genstat 5 as a procedure. Certainly one thing that should be avoided is duplication between the facilities of Genstat and GLIM; apart from being silly, it would be very wasteful of resources.
>
>   Perhaps you would like to call in when you are next down this way; you could then find out more about Genstat 5.
>
>   Yours sincerely,
>   John

Aitkin interpreted Gower's letter as Gower's being offered possible GLIM facilities for Genstat, which should then not be implemented in GLIM—this would be wasteful of resources. The GLM code from GLIM had already been implemented in Genstat; this was evidently *not* a waste of resources . . . Aitkin did not take up this suggestion, or look further into how GLIM facilities might be incorporated into Genstat, or Genstat facilities might be incorporated into GLIM.

12 President of the RSS

Nelder's appointment as the next President of the RSS followed soon after his retirement. He was asked to submit to the next Council meeting (17/4/85) a paper which would

- name his suggested successor as Chairman of the GLIM Working Party;
- suggest new and more precise terms of reference for the GLIM Working Party;
- outline proposals for the future development of GLIM and the necessary arrangements to deal with them.

Aitkin was approached by a senior member of the GWP (not Nelder), who asked if he would be willing to be nominated to Council as the next Chairman of the GWP. Aitkin said that he would, on one condition: that his role as Chairman would be to develop GLIM into an internationally successful general statistical package. This was reported to the GWP, a majority of whose members were happy with the condition.

At the Council meeing of 17/4/85, Nelder was overseas. Gower introduced his report. Nelder reported that the GWP recommended Aitkin's appointment as the next Chairman. Baker had left Rothamsted for a commercial position. Nelder suggested that, given his and Baker's departures, Rothamsted no longer be the base for GLIM development and that the base should be moved to a suitable site that could be tendered for and a development licence issued by the RSS to the chosen site. The development licence would define an executive body charged with implementing a specification agreed with the GWP. The form of the executive body was then being discussed by the GWP, which would put a firm proposal to Council in due course.

Nelder proposed (following a discussion between the RSS's officers and himself) that the GWP should in future act as an advisory body. The actual implementation of further versions of GLIM should become the responsibility of the executive body. He wrote that the GWP accepted this view of their future activities.

Council agreed:

- that the Working Party's recommendation that Professor Murray Aitkin should succeed Dr Nelder as Chairman for a 3-year period be accepted; [Allan Reese became Secretary]
- that Dr Nelder be urged to continue to serve on the Working Party to ensure continuity;
- that a sub-committee of Council be set up to study the further development of GLIM;
- the members of the committee were Nelder, Gower, Ewan Page and David Hill (Ian Blenkinsop ex officio),

which should report as soon as possible on

- ...the development licence;
- the appointment of members to the GWP;
- ...the terms of reference of the GWP.

It is notable that the new Chairman of the GWP was *not* to be a member of the sub-committee. Aitkin was unaware of the existence of this sub-committee, and his non-appointment to it, until much later.

Council minutes 15/5/85: A meeting of the Council sub-committee for GLIM future development was arranged for 30/5/85.

Council minutes 26/6/85: The terms of reference proposed by the GLIM sub-committee for the GWP were accepted.

The Council minutes do not give details of the new terms of reference, as they were proposed in the minutes of their 17/4/85 meeting. In particular, there would be an Executive body, which would implement any new GLIM version; the GWP would act only as an avisory body.

Following his appointment and for the GLIM conference at Lancaster in August, Aitkin wrote a paper describing the areas that needed to be developed (Aitkin, 1985). These areas were well-known and had been increasingly criticised by many otherwise enthusiastic GLIM users. They were, briefly

- the poor-quality line-printer graphics;
- the complete absence of any matrix operations, or even matrix structural handling;
- the very limited use of arrays, which could not be operated on *as* arrays, but had to be treated as sets of vectors and looped through;
- the inability to handle multi-level structures and models.

GLIM did provide the $PASS facility for exporting a GLIM structure to FORTRAN, operating on it there and returning it to GLIM for further analysis. This was particularly complex to use and was no substitute for matrix and array tools within GLIM (already proposed for PRISM).

The first meeting of a sub-committee of the GLIM Working Party, following Aitkin's appointment as Chairman, was held on 14/6/85 at the RSS to produce a draft specification of GLIM 4. Present were Aitkin (Chairman), Mike Clarke, Mick Green, Nelder (first half of meeting), Mel Slater and Tony Swan. It is not clear when Nelder left the meeting, but it is clear from the minutes that GLIM 4 was to be a substantial extension of GLIM 3; in particular:

General matrix and array handling facilities will be provided, with Cholesky factorization and singular value decompostions of matrices. ...

It is expected that GLIM 4 will offer much more sophisticated graphics facilities, probably as an extra-cost option, but these are still under discussion and will be developed separately by the Working Party.

At this meeting, presumably in Nelder's absence, alarm was expressed by members who knew of the proposal for the GWP to become only an advisory body for the Executive that would implement GLIM 4. Aitkin subsequently wrote to the Council expressing their concern.

Council minutes 11/7/85:

Concern was expressed by some Council members about the possibility of confusion between the GLIM Working Party and the GLIM sub-committee. It was pointed out that the latter was an ad hoc sub-committee of Council set up to define the terms of reference of the GLIM Working Party and the method of appointing members to it. The minutes of the only meeting of the sub-committee had been considered at the June Council meeting. Subsequently a letter had been received from Professor Aitkin, the new Chairman of the GLIM Working Party, expressing disquiet over some of the decisions reached by the sub-committee. Agreed that:

- a further meeting of the GLIM sub-committee should be held to consider the points raised in Professor Aitkin's letter;
- a report of the meeting, together with copies of Professor Aitkin's letter and a paper by the Hon. Officers, be submitted to the next Council meeting.

13 1985 IKBS Funding

The difficulty for *any* development of GLIM was the funding of the GWP members to do the development, whatever it was. Members were unpaid for their development work; this had in the past been regarded by their Universities as a part of their research work. Increasing pressure on University finances now made this position difficult to support. External funding was essential.

A new initiative of the EEC in Intelligent Knowledge-Based Systems (IKBS) made funding available for *products* of research groups, beyond computer hardware. Transforming GLIM looked very suitable for this, especially as IKBS funding had to be for participating groups from at least three European countries, one of which had to be France. GLIM had such a solid base of European statisticians that it was quite easy to formulate an international team to divide up the development.

Aitkin discussed this possibility for development with many of the European statisticians attending the 1985 GLIM conference in Lancaster, as well as with the GWP members. There was general enthusiasm for the idea. Aitkin wrote up a rough draft of the proposed development, for a package with the tentative name of ESP—the European Statistical Package—and circulated it to the European statistics groups most likely to participate. His aim was to produce 'the Mini-Cooper of statistical packages: even faster but much better equipped, more powerful and smooth'.

The draft version submitted to the EEC for comment was from eight collaborating laboratories in six countries, in the 'Stimulation Action' programme: for 'Stimulation of European Cooperation and Scientific and Technical Interchange: Application for an "Operation" ' .

The laboratories and their collaborating members were

- University of Lancaster: Murray Aitkin, Brian Francis, Mick Green
- Queen Mary College: Mike Clarke, Mel Slater

- Polytechnic of North London: Bob Gilchrist
- University of Wuppertal: Gerhard Arminger
- Odense University: Bent Jorgensen
- Istituto Applicatione del Calculo, Rome: Piergiorgio Gherardini
- Université Paul Sabatier, Toulouse: Henri Caussinus
- University of Leiden: Jan de Leeuw.

The scientific case for the development was a rewrite of Aitkin's 1985 GLIM conference paper. The roles of the laboratories and their contributions to the new program were spelled out in some detail.

The EEC evaluation was very positive, and they asked for the full proposal within 6 months, in the range £500–1 000 K. This was very encouraging, but before proceeding further, Aitkin needed to discuss the proposal with the RSS Council.

14 RSS Council Reaction

The following developments are documented in minutes of the relevant committees. At this time, Nelder was RSS President, John Gower was Secretary and Ewan Page (Vice-Chancellor of the University of Reading and former Professor of Computing, University of Newcastle) was Treasurer, of the RSS. They were all members of the Council sub-committee for GLIM, the existence of which Aitkin was still unaware. He did not consult other members of the Council: the three most senior members were all well versed in the RSS's computing isssues. He asked for a meeting with them and sent a copy of the draft EEC proposal.

The meeting was at the RSS (15/10/85) with Nelder, Gower, Page, David Hill and Brian Ford from NAG in attendance, minutes by Ian Blenkinsop. Gower reported the distribution of the GLIM 3.77 contract had been signed with NAG, and there was discussion of the wording of the GLIM 4 development licence.

Aitkin spoke to the draft EEC proposal, as the way of solving the funding problem for GLIM 4 development. Because the funding would be from the EEC for a package that was not GLIM 4, but a development from GLIM, of course this would have financial implications for the RSS, yet to be determined in the final submission and the EEC discussion.

Nelder was the main sub-committee commentator on this. His views were already clear from the history of the GLIM development, but he spelled them out:

- The statistical package market was already crowded, and GLIM would not be successful if it developed in this way.
- The GLIM code was already in Genstat, and so GLIM users who wanted more matrix facilities could use Genstat.
- There was no room in the British statistical package market for another general package. GLIM should remain a small specialised package for fitting GLMs.

Page added: if this development *were* to go ahead, then

- The name GLIM and the logo or name of the Royal Statistical Society must not be used.
- The code of GLIM was copyright and could not be used in any such development.

In response, Aitkin pointed out that

- there was no proposal to use the Genstat Wilkinson–Nelder algorithm for generally balanced experimental designs (as proposed for PRISM), because this could not handle generally unbalanced designs of the kind in social surveys and

- users of GLIM were working in quite different applied areas from users of Genstat, so there was no real issue of competition.

The minutes of this sub-committee meeting (15/10/85) do not mention Nelder's projection of GLIM's future. On the GLIM 4 development issue, the minutes report:

Professor Aitkin reported that GLIM 4 would not be a development of the source code of GLIM 3.77. He believed that it would be entirely different, and it was not clear that the intellectual property would be vested in the Society to the same extent as for GLIM 3.77. Initial financial support might be sought from the EEC and thereafter from ESRC. In the long run, however, it would be necessary to ensure that the development was financially self-supporting.

The Hon. Treasurer (Prof Page) pointed out that the use of a different code did not of itself diminish the rights of the Society to the intellectual property. However it was generally agreed that the Society would wish to participate in and encourage further development and that in any agreement covering such development it might be thought proper that the proportion of royalties accruing to other participating organisations and individuals should be increased. Any such agreement should also contain an explicit statement that the Society would not unreasonably seek to prevent exploitation by others in the event of it not wishing to do so itself. This would not, however, preclude the society from negotiating a royalty related to the extent to which the new development was dependent on previous versions of GLIM.

Page's tortuous comment clearly showed the degree of complexity of the Society's GLIM financial arrangements, and the difficulties facing any potential developer.

The sub-committee then agreed:

- that Professor Aitkin should prepare outline proposals for the development of GLIM 4 by the end of December for submission to the EEC;
- that the same proposal should be submitted to the Council meeting on 16 January 1986 with some indication of the expected degree of Society participation;
- that the Society would not wish to express a view about any formal structure, which might be set up at the University of Lancaster.

Nelder's position was clearly influenced by the pressure from the Government on agricultural research spending, including at Rothamsted, and the inappropriateness (to the Government) of agricultural research money being spent on computing software development.

This pressure was historically totally unfair, given the value of computing developments by Nelder and Yates to progress in agricultural experimental and survey data analysis, let alone the design of agricultural experiments. Genstat license fees went to Rothamsted, and GLIM's expanding success was seen at Rothamsted as a serious threat to its future Genstat funding. The GWP/EEC proposal for an expansion of GLIM was therefore even more threatening.

Aitkin sympathised with the Rothamsted people over their funding threat. As with the Universities, government funding cuts were based on an economic philosophy, not on reality. However, his responsibility was to the RSS for the future of GLIM, not to Rothamsted for the future of Genstat. He did not accept Nelder's prediction of GLIM's failure if it was developed: on the contrary, its failure was clearly coming if it was *not* developed. The question was how this could proceed.

However, he had in the meantime to inform the full Council of the details of the EEC proposal. This was substantially delayed by a personal crisis, as he and his wife separated and were subsequently divorced. He was unable to perform any serious academic task for several months. As a consequence, the proposal he sent to the full RSS Council was only a minor rewrite of the original EEC document.

14.1 Whither or wither, GLIM?

Following the meeting with the GLIM sub-committee, Aitkin met an immediately available subset of the GWP to describe the meeting and to consider the future of the package. There was a great deal of anger expressed, along the lines of

'We don't *need* the GLIM code! We can write a new version ourselves! We should go ahead with the EEC package anyway!'

Aitkin pointed out the obvious difficulty.

Without RSS support for this development, we won't be successful, because certainly Nelder, and possibly the others, will be asked to give opinions on its success. Nelder will kill it. We won't be able to argue against that.

This increased the feelings of frustration in the room. Aitkin had already decided what he should do. He had made it a point of his acceptance of the Chair that GLIM would be developed. If this was not going to happen, there was no point in his continuing as the Chair.

Executive and Policy Committee minutes (15/10/85, Nelder chair):

Proposals for the development of GLIM 4 would be submitted by Professor Aitkin to the Executive and Policy committee on 16 January and thereafter to the Council meeting in February.

Council minutes 20/11/85 (Gower):

Professor Aitkin was preparing a specification for GLIM 4, details of which would be submitted to Council in due course.

Council minutes 12/1/86:

Considered a draft document, submitted by Professor Aitkin covering the suggested procedure for the development of GLIM 4. Agreed

- that the GLIM sub-committee should examine the proposals more carefully and report back to Council in due course with particular reference to

  – the degree of RSS involvement expected;
  – the general question of interface with other systems;
  – proposals for marketing;
  – the proportion of royalties to be expected by the Society.

- that a more detailed proposal be submitted in due course
- that the final proposals be submitted to Council when available.

Nelder wrote following the Council meeting (21/2/86):

Dear Murray,

Council considered your draft proposals for your software development on February 12. There was a general view, I think, that your project is very different in scale, scope and method from that of GLIM 3, and in consequence we were asked to reconvene the GLIM Sub-Committee to gain some further background information. In particular Council would like to know:

1 How you see the Society's role in this project.
2 In return for the Society's backing and permission to use the name GLIM, the Society would expect some financial return. What arrangement do you have in mind?

3 Whether you intend to submit this proposal to a grant-giving body as Chairman of the GLIM Working Party or in some other capacity.

If you could supply some preliminary comments on these questions your reply could then form the basis for a discussion by the Sub-Committee. In the meantime Council asks that if you make any application to EEC or elsewhere, you should not for the time being refer to the product as GLIM 4, although I see no objection to mentioning that this new project is motivated by, or represents an extension of, earlier GLIM developments. I hope we can arrange a meeting of the Sub-Committee quite soon.
Yours sincerely,
J.A. Nelder
President

Aitkin was now aware that there was a GLIM Sub-Committee, though its membership was obscure. But nothing in the letter changed the situation, which had indeed not changed. He judged that whatever royalty and naming compromises might be negotiated, the EEC development that he had proposed (for which the final submission would still require an immense amount of work) was not going to be successful. Nelder had made that plain.

Aitkin wrote a formal resignation letter (24/2/86) to the RSS as GWP Chairman, without giving any explanation of his action, other than that he was going overseas for 3 years. He wrote a more personal letter to the GWP members in which he did mention the personal crisis, which most knew about already. Aitkin gave considerable thought to a full explanation and description of the GLIM sub-committee meeting to the full RSS Council but decided that this would only sour his relations with the Society for years to come, without any benefit to anyone.

Nelder called him a day later to ask whether his resignation was connected to their discussions. He said only that he had a number of personal difficulties (which was certainly true) and could not continue as Chair under the circumstances. This seemed to reassure Nelder.

Executive and Policy Committee minutes 20/3/86.

The Hon. Secretary reported that Professor Aitkin had tendered his resignation as Chairman of the GLIM Working Party as he was going abroad for three years. A meeting of the GLIM sub-committee had been arranged for 24/3/86 to review the situation. Agreed that the recommendations of the GLIM sub-committee be awaited but that an expression of the Executive and Policy Committee's concern about the present situation be conveyed to the sub-committee.

Council minutes 20/3/86.

The Hon. Secretary reported that Professor Aitkin's proposals for the development of a "GLIM 4" had been presented at the last Council meeting as a result of which the GLIM sub-committee had been asked to give detailed consideration to the proposals. Subsequently however Professor Aitkin had tendered his resignation as Chairman of the GLIM Working party. The GLIM sub-committee had recommended that the GLIM Working Party should limit its responsibility to the maintenance and modest improvement of GLIM 3.77. Although Professor Aitkin's proposals for the development of new software are in abeyance, the sub-committee nevertheless recognised that the Society would wish to encourage the development of statistical software. It was therefore recomended that an announcement should be published in News and Notes to the effect that Fellows were encouraged to submit proposals for new software development for consideration by Council.
Council would, when appropriate, endorse such proposals in the hope that this would be of help in obtaining development funds from grant giving bodies.

This seemed to Aitkin to stand logic on its head. The GLIM Working Party was not to submit proposals for new software, but the Society would wish to encourage such proposals from

someone else and would endorse them to support applications for development funds. But had not he already sought the development funds, with good prospects for success, without this endorsement? What was the difficulty?

15 Further Developments

The GLIM 4 development matter however was not dead. On April 4, Nelder wrote to the members of the GWP, following a meeting of the GLIM Sub-Committee of Council, conveying its recommendations following Aitkin's resignation. These were that

- the GWP should propose a new Chairman;
- the GWP should in future act as an advisory body on maintaining the current GLIM;
- the actual implementation of further versions of GLIM should become the responsibility of the executive body,

as previously proposed by the Sub-Committee.

This produced a strong reaction and began a struggle. Allan Reese wrote (1/5/86) to Nelder as Secretary of the GWP, objecting strongly to two of the recommendations:

- It was not appropriate to propose a Chairperson until it was determined whether the GWP would continue to exist, under its terms of reference.
- While the maintenance of the current GLIM was an important role for the GWP, it would be unhappy at any attempt to veto or curtail its planned activities on extensions to GLIM.

> ... if it is Council's intention to relegate the WP to a mere service role on the existing program, then the WP see no honourable alternative but reluctantly all to resign, and explain publicly why they had done so....
>
> For the WP to continue in a meaningful role ... it must be permitted and should be encouraged to take a lead in the definition and provision of statistical computing. ...
>
> Meanwhile, the WP sees no point in holding back Professor Aitkin from submitting his current appliction to the EEC, as a cooperative enterprise, not necessarily under the auspices of the RSS. Otherwise, the opportunity may be missed. This is entirely consistent with your suggestion that a development group might be autonomous.

This produced a rapid reaction from Council, reported in a letter of 19/5/86 from Nelder to Gilchrist. Nelder quoted the Council minutes:

1. Council recognises the need for a group to continue to look after GLIM 3-77, as a contractual obligation by the Society in its relations with NAG.
2. It would welcome initiatives by that group for propagating interest in GLIM 3-77, and money could be made available to support such activities.
3. Council was opposed to the Society's sponsoring new statistical software directly.
4. Council was willing to set up a procedure for evaluating proposals for new software and would be prepared to endorse any proposal for which the evaluation was satisfactory, in the hope that such endorsement would be useful to the originators in seeking funds for implementation.

   In view of (3), a proposal would have to come from a named set of individuals, and could not come from any official subgroup of the Society.

This was an even stronger restatement of the GLIM Sub-Committee's position: the GWP was to work *only* on GLIM 3.77 *support*, meaning that 3.77 was the *final* version of GLIM. Any further development would have to come from someone else.

A follow-up meeting of the Sub-Committee and members of the GWP was held on 28/5/86. As a consequence, Council changed its view. At the subsequent Council meeting (18/6/86), it was agreed that the GLIM Working Party should be responsible to Council for

- the maintenance of existing versions of GLIM including GLIM 3.77;
- the specification of extensions to GLIM 3.77 and new versions of GLIM;
- advising on further developments of GLIM or related software.

Council re-affirmed that the Working Party, in pursuing these objectives, must continue to report its proposals to Council, as and when necessary, and that the final decisions on the acceptability of any such proposals must remain with Council.

In the course of further discussions it was also agreed that:

1 Council should have a formal representative on the GLIM Working Party; Dr P.J. Green was suggested;
2 in considering applications from the GLIM Working Party for funds, Council will take into account income accruing from royalties;
3 the GLIM Sub-committee be wound up and the responsibility for decisions on matters of urgency be delegated to the Hon. Officers.
4 Council's appreciation of the work done by the members of the GLIM Working Party be recorded.

So, the GWP received a vote of confidence from Council and retained its responsibility for new developments, and the Council now had an ear and a voice directly at the GWP, without the filtering of the wound-up GLIM Sub-committee. Development of GLIM 4, or something else, could be considered and pursued by the GWP without interference: this was not the end of GLIM 4.

Tony Swan was appointed Chairman of the GWP with Clive Payne as Secretary (Swan retired as Chairman in March 1997, and Gilchrist took over as Chair, with Brian Francis as Secretary). The GWP began a period of consideration of how GLIM 4 should be developed, which took several years to formulate. Meanwhile, GLIM 3.77 sales continued robustly, as did the GLIM royalties.

Aitkin left Lancaster in August 1986 and did not return to a UK position until 1996. This ended his connection with the CAS, and his immediate connection with GLIM, though he continued sporadically to attend the IWSMs. Brian Francis remained to keep the CAS alive, became the director of the RSS's GLIM Development Centre at Lancaster (which was assisted by Aitkin's appointment of Mick Green from PNL to a new additional tenured post in the CAS), to which the RSS contributed the modest sum of £21,000 and was in time promoted to Professor, and Director of the CAS until its eventual merger with the Department of Mathematics and Statistics in August 2005.

15.1 GLIM 4 Again

At the 1989 joint GLIM/Statistical Modelling conference in Trento, a new release of GLIM 4 was announced, and the conference proceedings contain a detailed paper (Green, Francis & Bradley, 1989) on its new features, which had some, though not all, of the developments Aitkin had proposed. There would be new presentation-quality graphics, and an extension of probability functions, including all the exponential family cdfs and quantile functions. Matrix operations and handling tools would *not* be a part of this extension. This announcement was updated in the 1992 joint GLIM/Statistical Modelling conference in Munich (Fahrmeir *et al.*,

1992), where papers on the new release were given by Gilchrist and Payne; Francis, Green and Clarke; and Swan and Francis.

The new facilities were certainly welcome and worthwhile, but in Aitkin's view, GLIM would certainly die without the matrix and array tools, even as a simple GLM fitting package. Other packages like Stata already had, or were developing, GLM procedures at least as general as those in GLIM, within a full matrix and array computational framework. There would be no reason to use GLIM.

GLIM 4 was released 1/4/93. The extensions from GLIM 3 were major:

- High resolution graphics, an extended range of models including Inverse Gaussian, inverse power link and orthogonal polynomials.
- Elimination of nuisance parameters for more efficient model fitting.
- Analysis of case control studies.
- More mathematical functions including percentiles and quantiles for χ^2, t, F, beta, binomial and Poisson distributions.
- An extended macro library; macro debugging and editing.
- Access to the fitting process and non-linear modelling.
- The PASS routine for incorporating FORTRAN routines.
- An online manual (and the comprehensive GLIM 4 manual).

However, the surface views and contour plots that had been proposed for the GRAPH module in PRISM were not in GLIM 4.

Its initial sales were astounding, with royalties in 1994 alone of more than £50 000. GLIM 4 had been developed by a large team (listed alphabetically):

- overall project direction by all the GWP members;
- programming development by Malcolm Bradley, Mike Clarke, Brian Francis, Mick Green;
- Windows implementation by Andrew Stalewski;
- manual by Malcolm Bradley, Mike Clarke, Brian Francis, Bob Gilchrist, John Hinde, Mick Green, Peter Green, Clive Payne, Allan Reese, Andrew Stalewski and Tony Swan.

There was a UNIX version paralleling the UNIX version of GLIM 3-77. The manual (Francis *et al.*, 1993) was enormous (and became widely cited)—821 pages, and nearly 3 kilos, since it contained a 323-page book—Chapter 12—by Tony Swan on GLM applications.

The release of GLIM 4 was followed by a spike in activity of the GWP members in giving courses on the new system. There were many GLIM courses held in Europe in 94/5 and some in the USA, under the auspices of the GWP, taught mainly by GWP members but with support from local statisticians who also helped organise the courses. These were part funded by the EU COMETT programme and part by NAG. Courses were given in Milan, London, Utrecht, Berlin, Leuven, Barcelona, Graz and Athens. There were two courses in the USA, one at the American Statistical Association annual conference. By the end of 1995, GLIM 4 sales had reached 730.

The new release clearly called for a new edition of the CAS book: the new facilities, especially the new graphics, made possible an expanded edition with new chapters on mixtures and random effect models, which could be fitted by nonparametric maximum likelihood, adding to the parametric modelling facilities. Oxford had little enthusiasm for a new edition: they considered GLIM already passé. Publication required considerable negotiation and was delayed until 2005. It involved a promise for an R version (published 2009). The GLIM 4 edition sold slowly; NAG's withdrawal of the package in 2004 ended sales.

The large GLIM 4 development came at a price. In the 1992 WP report, Swan wrote

Despite substantial funding by the Society in the first two years this major undertaking took an enormous amount of personal effort and commitment particularly from Brian Francis and Mick Green at Lancaster. The implementation of the high resolution graphics facilities which were not included in the basic specification was due almost entirely to their efforts. This is an example of the benefits of getting something produced by a non-commercial group driven by public-spirited motives.

...

Several of the full and co-opted members of the Working Party are leaving due to pressure of other work or change in circumstances. I would particularly like to thank Malcolm Bradley who did a large amount of programming for GLIM 4, Peter Green who has long been a very helpful source of good advice, Roger White who provided a useful link with the earlier work of Rothamsted, Carl O'Brien who brought considerable expertise from the GLIMPSE project but is now working in Brazil and David Clayton who provided much practical help in testing GLIM 4 on case-control and survival data problems.

Tony Scallan joined the Working Party to take on the responsibility of editing the Macro library. John Hinde, with a long-standing interest in linear models, statistical computing and GLIM, has also joined the Working Party.

In the 1993 WP report, Swan wrote

Finally I must report the sad news that one of our most influential and productive members, Mike Clarke, died recently after a lengthy illness. The high quality of the numerical algorithms at the heart of GLIM, and which did much to earn the package the high regard in which it is held round the world, is almost entirely due to his work and we shall miss him sorely.

16 GLIM Sales and RSS Royalties

GLIM's success is most evident in its annual sales (inconsistently reported) and the royalties on sales earned by the RSS (Table 1). The (cumulative) sales were reported in some of the WP annual reports and GLIM newsletters, and the royalties in the annual Council reports. Royalties reported (in JRSS series A) were of two kinds: journal sales (outside of subscriptions) and GLIM. GLIM royalties were not reported before 1983 but may be implicit as part of 'miscellaneous' income. This entry in the RSS income statements began in 1972, with an amount of £491. In the early years, the RSS was earning little from royalties from GLIM sales: in 1980–1982 (not shown in Table 1) these totalled £3 490; of this £2 481 was outlaid by the RSS on a sandwich-course student working on PRISM under Nelder's guidance.

GLIM royalties were reported separately from 1983 to 1987, and in later years were bundled with journal royalties, reported from 1988 to 1996. The latter were never more than £1 600 in any year, and since the GLIM component cannot be separately identified, only the combined total is shown in the table. After 1996, Council reports were published in JRSSD until 2003. They reverted to JRSSA from 2004.

The first panel in Table 1 is for GLIM releases up to 3.12, the second for GLIM 3.77 and the third for GLIM 4. The column 'RSS input' is the support provided by the Council to the GWP for GLIM development. The RSS negotiated a new contract with NAG for GLIM 3.77 with a substantial increase in royalties. GLIM 3.77 continued to sell well until the release of GLIM 4, which shows a large spike in royalties and number sold. Royalties from GLIM 4 dropped very substantially in 1995, though the WP report was optimistic:

GLIM 4 continues to sell well, with total licence sales of 730 in 47 countries at the end of October 1995. New licences continue to be added at the rate of about 15 a month.

GLIM royalties 1983–1987 were £69 004, and over the period 1988–1996 are estimated at more than £200 000, assuming the journal royalties maintained the 1974–1985 average of

Table 1. *Journal and GLIM royalties, RSS input to the GWP and GLIM sales.*

Year	Journal income	GLIM income	Comb. income	RSS input	Cum sales	Monthly sales	Coun-tries	Sites
1974	2 978							
1975	617							
1976	739				56		12	
1977	413					21	26	
1978	722							169
1979	2 349				241			
1980	1 268				350			
1981	1 454				450	11	35	500
1982	1 532						41	600
1983	1 535	2 024						
1984	1 131	2 889			887			
1985	423	7 620			921			
1986	1 409	36 356						
1987	1 403	21 518				16		701
1988			17 885	10 166			51	1 003
1989			25 655	7 909				
1990			19 796	836				
1991			27 383	326	>2 000		57	1 302
1992			26 826	365			53	1 416
1993			20 936	2 046			55	1 600
1994			54 575	3 576		20	42	600
1995			9 412	1 040	730		47	
1996			13 776			15		
1997			10 781					
1998			5 662					
1999			8 024					
2000			6 088					
2001			4 325					
2002			1 548					
2003			2 944					
2004			166					
2005			0					

£1 500 annually. Several Treasurer's reports over this period welcomed the boost to RSS income from GLIM, while warning that it could not continue indefinitely. Income declined steadily after 1994 (which Aitkin attributes to the absence of matrix facilities), and precipitately after 2000.

Around 1996, Francis and other GWP members worked on the feasibility of a full Windows version, tentatively called GLIM 5, retaining the current GLIM 4 FORTRAN code, recompiling it as a Windows Dynamic Link Library (DLL) and providing an interface between this and a Windows application.

Gilchrist continued this with a small amount of support from the RSS, and a beta/near-alpha version was produced and distributed to the WP in the early 2000s. Development stopped over funding disagreements between the RSS and the WP. The WP enthusiasm waned for this project (and for carrying on the support needed for a marketed version), and it was not finalised and distributed. WP members were well aware that S was on the rise. This ended GLIM development. The GLIM Working Party was wound up in 2005, though there was no public (or private to the WP members) announcement or report of this. The GLIM code also disappeared, rather than being made publicly available.

GLIM was put by NAG into a state they referred to as 'limbo' (i.e. not actively promoted or marketed, but still available) by the Product Committee shortly before NAG's internal structure

was reorganised to replace the Production and Services Division 2001–2004. GLIM support stopped in 2004.

Nelder continued to develop software based on GLIM and Genstat. GLIMPSE was an early attempt at an expert statistical system (O'Brien, 1989). Payne summarised it (p. 112–113):

> GLIMPSE was developed by John, with Carl O'Brien at NAG and David Wolstenholme at Imperial College, to provide a knowledge-based front-end for GLIM. ...
>
> In practice, GLIMPSE became more a tool for experts than a tool to provide expert help to novices. Its long-term effect on statistical computing has been perhaps more through the influence of its ideas of model strategy and model checking ... than through its own wide-spread use.

Nelder continued to insist that GLIM should not compete with Genstat. In his unpublished paper of 28/4/92 *The RSS and statistical software*, he wrote in the final paragraph:

> I believe that neither GLIM nor Genstat has adequate resources to compete successfully in the market of the 1990s. If both attempt to continue separately there is a strong possibility that within 5 years neither will survive. Already the S plus group in Seattle has at least three times as much programming effort available as GLIM and Genstat combined, and is supported by hundreds of thousands of NASA dollars. If we wish to see a continuing presence for a UK package then I believe we have no alternative but to merge the efforts on Genstat and GLIM in a single co-ordinated system. Further, I think that the RSS has an important role to play in this. The Society has gained appreciable income from the royalties from GLIM, and these will continue, at least in the short term. A new framework will be needed if a single product is agreed on, but I see no difficulty in principle in the Society's continuing to gain benefit from royalties for, say, new modules for Genstat, sponsored by the Society and sold separately. The exact framework will need discussion, and the time to start such discussion is now.

17 Retrospect

GLIM's early success showed clearly the importance of GLMs as a major tool of both theoretical and applied statistics. For applied statistics, it changed the focus of analysis from *statistical methods* to *statistical models and modelling*. GLIM provided, for a time, the computational tool of choice for GLMs, as well as providing a continuing focus for modelling statisticians through the IWSMs. Its remarkable sales success in 1994 was not maintained, and it is clear that computational statisticians were already moving to other packages like S or Stata, and later R, to obtain features that GLIM did not provide. Despite Nelder's efforts to promote Genstat as an alternative, statisticians needing a more powerful tool than GLIM did not move to Genstat, but to Stata, S or even MATLAB.

US statisticians who participated in the GLIM conferences and IWSMs were agreed that to give the development of GLIM to a professional society would lead to its death. While small University groups could develop great algorithms, a properly funded commercial development was essential to ensure the success of statistical software. The success of SAS, R and BUGS illustrated different aspects of this argument.

The continued obstruction of GLIM's development into a more general and powerful package contributed to its demise. Even with such development, it might have died eventually anyway, but to me, it is a great pity that it did not have the opportunity.

References

Adena, M. & Wilson, S. (1982). *Generalized Linear Models in Epidemiological Research: Case-Control Studies* Sydney. Instat Foundation.

Aitkin, M. (1978). The analysis of unbalanced cross-classifications. *J. R. Stat. Soc. Ser. A*, **141**, 195–223.

Aitkin, M. (1985). GLIM 4 – Directions for development. In *Generalized Linear Models: Proceedings, Lancaster 1985*, Eds. Gilchrist, R., Francis, B. & Whittaker, J. Berlin: Springer-Verlag, pp. 6–13.

Aitkin, M., Anderson, D., Francis, B. & Hinde, J. (1989). *Statistical Modelling in GLIM,* (second GLIM 4 edition 2005). Oxford University Press.

Breslow, N. & Day, N. (1980). *Statistical Methods in Cancer Research. Volume 1—The Analysis of Case-Control Studies.* Lyon: International Agency for Research on Cancer.

Dempster, A., Laird, N. & Rubin, D. (1977). Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. Ser. B*, **39**, 1–38.

Dobson, A. (1990). *An Introduction to Generalized Linear Models.* Boca Raton, FL: CRC Press.

Fahrmeir, L., Francis, B., Gilchrist, R. & Tutz, G. (eds). (1992). Advances in GLIM and Statistical Modelling: Proceedings of the GLIM92 Conference and the 7th International Workshop on Statistical Modelling, Munich. Springer-Verlag, New York.

Finney, D. (1947). *Probit Analysis: A Statistical Treatment of the Sigmoid Response Curve.* Cambridge: Cambridge University Press.

Francis, B., Green, M. & Payne, C. (1993). *GLIM 4: The Statistical System for Generalized Linear Modelling. Release 4 Manual.* Oxford: Oxford Science Publications.

Gilchrist, R., Francis, B. & Whittaker, J. (eds). (1985). *Generalized Linear Models: Proceedings, Lancaster 1985.* Berlin: Springer-Verlag.

Gilchrist, R. (ed). (1982). *GLIM 82: Proceedings of the International Conference on Generalised Linear Models.* New York: Springer-Verlag.

Gilchrist, R., Francis, B. & Whittaker, J. (eds). (1985). *Generalized Linear Models: Proceedings of the GLIM 85 Conference.* Berlin: Springer-Verlag.

Gilchrist, R. & Green, M. (1980). *GLIM: A Primer.* Polytechnic of North London, Depatment of Mathematics.

Gower, J. (2015). The development of statistical computing at Rothamsted. *Int. Stat. Rev.*, **83**, 357–370.

Green, M., Francis, B. & Bradley, M. (1989). GLIM 4—structure and development. In *Statistical Modelling: Proceedings of GLIM 89 and the 4th International Workshop on Statistical Modelling held in Trento, Italy*, Eds. A. DeCarli, B. Francis, R. Gilchrist & G. Seeber, pp. 18–25. New York: Springer-Verlag.

Healy, M. (1988). *GLIM: An Introduction.* Oxford: Clarendon Press.

Lindsey, J. (1989). *The Analysis of Categorical Data Using GLIM.* Berlin: Springer-Verlag.

Lindsey, J. (1992). *The Analysis of Stochastic Processes Using GLIM.* Berlin: Springer-Verlag.

Lindsey, J. (1997). *Applying Generalized Linear. Models.* New York: Springer-Verlag.

Lord, F. (1952). *A Theory of Test Scores: Psychometric Monograph no. 7.* Psychometric Corporation, Richmond Virginia.

Nelder, J. (1975). Announcement by the Working Party on Statistical Computing. *Appl. Stat.*, **24**, 259–261.

Nelder, J. & Wedderburn, R. (1972). Generalized linear models. *J. R. Stat. Soc. Ser. A*, **135**, 370–384.

McCullagh, P. & Nelder, J. (1983). *Generalized Linear Models,* (second edition 1989). London: Chapman and Hall.

O'Brien, C. (1989). The GLIMPSE system. In *Statistical Modelling: Proceedings of GLIM 89 and the 4th International Workshop on Statistical Modelling held in Trento, Italy*, Eds. A. DeCarli, B. Francis, R. Gilchrist & G. Seeber, pp. 18–25. New York: Springer-Verlag.

Payne, R. (2004). Algorithms, data structures and languages—the computational ingredients for innovative analysis. In *Methods and Models in Statistics: In Honour of Professor John Nelder*, pp. 95–118. London: Imperial College Press.

Richardson, M. & Baker, R. (1980). The development of the GLIM system up to GLIM-4. In *COMPSTAT 1980: Proceedings in Computational Statistics*, pp. 523–529. Vienna: Physica-Verlag.

Rigby, R. & Stasinopoulos, D. (2001). The GAMLSS project: A flexible approach to statistical modelling. In *New Trends in Statistical Modelling: Proceedings of the 16th International Workshop on Statistical Modelling*, Eds. B. Klein & L. Korsholm, pp. 337–345. Odense, Denmark.

Rigby, R. & Stasinopoulos, D. (2005). Generalized Additive Models for Location, Scale and Shape (with discussion). *Appl. Stat.*, **54**, 507–554.

Venables, W. & Ripley, B. (2002). *Modern Applied Statistics with S.* New York: Springer.