
VOLUME 62 2002

The National Database for sheep genetic evaluation and research

D.J. GARRICK, J.N. CLARKE[1] AND C.F. CALLOW[2]

Department of Animal Science, Massey University, Palmerston North, New Zealand.

ABSTRACT

The long-term success of any recording scheme depends on access to a genetically-structured database of performance records. Ram breeders in New Zealand have indicated preference for a system providing within-flock breeding value prediction via on-farm personal computers or through commercial bureaux on a regional level. This paper describes recent developments in industry structure in relation to animal evaluation and access to field data for research.

Ram buyers, breeders and extension agents all benefit from consistent techniques for calculating and summarising the breeding worth of animals. The within-flock methodology prototyped in Animalplan has now been used to form the basis of a "genetic engine", available under licence, for incorporation into commercial programmes with database and reporting capabilities.

Across-flock genetic evaluation (e.g. sire reference analyses) may require access to information from a number of distributed databases. A national identification system which recognises animals used in flocks other than their birth flock is pre-requisite to analysis. The National Database has been initiated to ensure the integrity of data and to safeguard such pedigree and performance records for a wide variety of purposes. It will be updated annually to provide a single source for across-flock Best Linear Unbiased Prediction (BLUP) evaluation.

The provision of the National Database ensures continued access to field data for animal breeding and other applied research. Genetic development of the scheme will come from improved specification of the production model (e.g. new traits and indicators of non-genetic effects) and from improved statistical approaches (e.g. BLUP and estimates of distributional parameters), all requiring access to and analysis of large bodies of genetically-structured field data.

Keywords Database, genetic evaluation, sheep breeding, flock recording.

BACKGROUND

National Performance Recording

Performance recording schemes designed for genetic improvement on a national scale have six major components.

1. Committed Buyers.

Buyers must be committed to the use of sires which are, on average, genetically superior for the traits included in the buyer's objective. Buyers must be prepared to pay a premium for superior animals, that reflects the increased profits derived from their progeny. Buyers also have important roles to ensure breeders use selection objectives in line with commercial requirements and to keep researchers informed (through extension agents) of future needs to enable research to be done in advance of application.

2. Motivated Breeders.

Long-term dedication is required to collect and maintain accurate pedigree and performance records. Furthermore, extra costs associated with identifying superior animals have to be carried by the breeders in return for a buyer premium. Breeders need to be sensitive to future production, processing and market requirements to help identify the long-term needs of the industry. Breeders particularly, must keep abreast of relevant research developments which provide opportunities for improved genetic progress and must stimulate research according to their needs.

3. Database.

A system is required to collect, validate, input, store, correct and manipulate pedigree and performance records. The system must also be capable of reporting relevant results in a suitable form for day-to-day use by breeders and for use by prospective buyers. These functions are well suited to computers. A national system may need to incorporate database requirements at several levels, namely, breeders, collective organisations such as breed societies, as well as for (across-flock) genetic evaluation purposes. It may be an advantage to provide micro-computer systems for on-farm use linking to a larger system for national purposes.

4. Breeding Value Prediction.

The prediction of breeding values (BVs) provides a summary of the breeding worth of an animal. Furthermore, the summary combines all available information in an optimal manner, providing breeders and buyers with objective information for assisting in their decision making. Once the data has been collected and entered into the database, there is a considerable benefit relative to costs from using up-to-date methods to combine pedigree and performance records to obtain breeding values. In most cases, the method of choice (on technical and economic grounds) is one known as Best Linear Unbiased Prediction (BLUP). It facilitates the combining of all sources of information (on relatives and on correlated traits) for the estimation of breeding values. Prediction of breeding values is also an important component of animal breeding research as breeding values are a pre-requisite to the estimation of genetic parameters (heritabilities,

[1] Genetics Section, Ruakura Research Station, Hamilton, New Zealand.
[2] The New Zealand Animal Breeding Trust, c/- Massey University, Palmerston North, New Zealand.

genetic correlations) which are essential for comparing alternative selection criteria and selection policies.

5. Animal Breeding Research.

The long-term success of performance recording requires a system which can move with the times and allow breeders to take advantage of new market opportunities and scientific discoveries relating to identification of animals which produce superior offspring. The research needs to be an integral part of the programme to allow results to be immediately available to industry and to enable breeders to guide research direction. However, from a financial viewpoint and to be progressive in exploiting advances in new genetic, reproductive and molecular technologies, the research should attract public good funding and not be seen as an activity to be funded entirely by users. Indeed, performance recording schemes provide a critical channel for the transfer of new technology to industry.

6. Extension.

It is vital that extension organisations have a high profile to provide two-way communication between the various components in the system. They must encourage buyers to make informed decisions and demonstrate the financial advantage of using superior animals. They are important to assist breeders in the use and understanding of the programme. Furthermore, they must help stimulate short- and long-term improvements in the system by helping researchers keep in touch with industry requirements.

An important emerging opportunity is the development of computer-based decision-making software which, based on producers' own policies, cost and return expectations, can assist breeders to set breeding objectives, calculate appropriate relative economic values and decide among different recording options.

The major factor determining the effectiveness of a national scheme is the way that each of the components is matched to industry needs and mutually co-ordinated. The database and research components are linked through BV prediction as this technology arises from animal breeding research and can be improved only through research access to the database. The extension organisation also needs to be in close association with the database operation, and have good and up-to-date knowledge relating to the BV prediction and research components. Integration of on-farm recording, database functions and BV prediction followed by the use of this information in selection decisions has and will continue to play the major role in achieving genetic improvement in New Zealand livestock species (Clarke *et al.*, 1987).

RECENT DEVELOPMENTS

National Performance Recording for Sheep in New Zealand

In the early days of national performance recording in New Zealand, four of these components were under the umbrella of the Ministry of Agriculture & Fisheries/National Flock Recording Scheme (MAF/NFRS) and then MAF/Sheeplan, with collaborative research provided by Massey University, primarily through the work of Professor A.L. Rae. Operationally, these systems worked as well as could be expected from a central bureau system, with funding by users for the database operation and by government for research and extension components.

The development of user pays philosophies in MAF occurred at about the same time as Sheeplan was being upgraded and improved to form a more flexible recording system known as Animalplan. Most would recognise that it is reasonable to charge the users (breeders) for the operation of the database component. However an equitable appropriation of funds for research and extension (including promotion) is more difficult. As a result, there was a considerable erosion of funding for animal breeding research and extension related to Animalplan. Finally, MAF decided to close down the Animalplan operation rather than try to separate the commercial and public good elements.

At the time that the demise of MAF's animal recording services was announced, a number of commercial companies stepped in to offer (regional) computer services for data collection and reporting. These companies cannot themselves be a complete replacement for Animalplan as this would require specialised knowledge of animal breeding theory and practice, in addition to having capability and funding for co-ordinated extension and research components. However, the continuation of an Animalplan-like total recording service has been assured by the co-ordination of private companies through The New Zealand Animal Breeding Trust (NZABT). This non-profit organisation was formed in 1991 to service the application of genetic technology to the breeding industry. Its first priority is to ensure that the Animalplan technology for calculating BVs is made available to industry. This has involved detailed documentation and reprogramming of BV calculations (Johnson *et al.*, 1989) and their development into what has become known as the Animalplan Genetic Engine. The Engine and associated information have been made available, under licence, to three companies to offer bureau services. Clients of MAF's Animalplan will be progressively transferred to these bureaux during 1992. The development of a PC-package incorporating the Animalplan Genetic Engine has also been initiated, and will shortly be put out to tender. Thus, genetic evaluations of animals akin to the previous services offered by MAF's Animalplan will now be available through the use of the Trust Genetic Engine via either licensed bureaux or a leased PC-package.

Other priorities of the NZABT include information services for extension, and the formation of a single National Database for animal breeding research and higher level processing. Higher level processing (HLP) refers to those genetic evaluation services that are beyond the current capability of the Genetic Engine/Bureaux, including numerically-intensive processing such as across-year BLUP analyses or analyses of data sourced from more than a single bureau and/or PC-database (e.g. for across-flock analysis of sire reference data; Garrick 1991). Other examples of HLP involve traits under development (e.g. faecal egg counts, antibody measures, body composition) or procedures under development (e.g. analysis of crossbred records, separation of major gene effects) that require preliminary research to be carried out in conjunction with the processing service. Results from HLP will usually be breeder-appropriable and would therefore be funded on a user-pays basis. The remainder of this paper will focus on the public-good value of the National Database.

Requirements of the National Database

Four important components determine the genetic effectiveness of the National Database.

1. Unique Animal Identifiers

A system of uniquely identifying animals independently of their field tag numbers is essential to the operation of the National Database. Field tags are not suitable because animals can be transferred from flock to flock (e.g. sires purchased from outside which may have the same tag number as animals already in the flock). Reference sires, particularly those used via artificial insemination, can sire progeny in many flocks at one time and may not even have a field tag. The animal identifier must also allow records to be related when they have been collected over the lifetime of an animal that may have changed location during that time.

2. Pedigree Relationships

Where available the national identifier of the sire and the dam of every individual should be recorded. Furthermore, the system must allow recording of special circumstances (e.g. foster parents, particularly when a dam is used to foster an animal soon after its birth). Surrogate dams may be used in the case of embryo transfer, when the pre-natal maternal environment is then provided by a dam other than the genetic dam. The system must also distinguish clones (animals that are genetically identical such as those created by embryo splitting) from fullsibs which share parents but are not genetically identical because of the chance effects of Mendelian sampling. Syndicate matings may also need recording when a group of ewes has been mated to a team of sires and individual parentage is not known. Finally, genetic grouping may be required. In its simplest form this would require relating animals to breed groups.

3. Performance Traits

These may be physical entities which can be counted or measured; alternatively they may be subjective traits such as visual scores. Performance traits must have a consistent definition over years (including units of measurement). Well-defined validation rules are required for excluding outliers. These rules must be consistent regardless of the origin of the data. Well-defined categories are required for non-genetic effects, including factors such as age of dam, sex and date of birth. The database must be capable of storing repeated traits that are recorded more than once in the animal's lifetime. Finally, the traits which are recorded must be open-ended in the sense that there is flexibility to add new traits as opportunities arrive.

4. The National Database must be accessible

Records from the database may need to be accessed in a number of ways. For example, it might be of interest to relate traits which have been recorded on the same animal at different times in its life, or to obtain records on animals in the same management group. Records need to be related among animals in different generations, for example when parent-offspring regressions are to be calculated. Often records on animals in the same genetic group such as progeny of one sire need to be withdrawn. The database also needs to be mutually accessible to other databases, for example, health databases, meat works' databases from slaughtered animals and perhaps consultants' databases. Other issues of accessibility include time and cost.

PUBLIC GOOD VALUE OF THE NATIONAL DATABASE

Specially structured databases are important to public good science in animal genetics and a National Database has particular advantages (Wickham, 1984). First, a National Database is required to enable studies making use of records which have been obtained from different flocks. Examples of between-flock studies include investigation of genotype-by-environment interaction and analysis of sire reference schemes. Second, a database needs to be national to be commercially realistic in the sense that the records are obtained from animals which are commercially available and are managed according to normal farm practices. Results from research flocks have frequently been criticised because the animals are not typical of those available to the industry and their management is not in accord with "normal farm practice". Thirdly, the National Database needs to have considerable size to allow for meaningful estimation of genetic parameters. Reliable estimates of heritabilities and genetic correlations require records from many sires obtained across a range of flock and year environments. Finally, the database needs to be national because increased reliance is placed on commercial records as reduced Crown funding results in the discontinuation of many of our research flocks.

The public good value of the National Database is best demonstrated by considering the definition of public good according to the Foundation for Research Science and Technology (FRST; see FRST 1991).

1. Public good science outputs increase our knowledge and understanding of the physical, biological and sociological environment. The study of quantitative genetics certainly increases our knowledge of the biological environment. In the process of genetic evaluation of animals one routinely calculates estimates of various non-genetic factors that are responsible for differences in performance among animals which are managed as a group. Estimation of variance components can explain the relative importance of a number of factors determining animal productivity. Many physiological studies have been aimed at identifying the mechanisms involved. In the first instance, these studies rely upon divergent animals that have been identified as having high or low levels of performance for the trait under investigation. Characterisation of industry practices, determining the structure and dynamics of the sire breeding industry, also relies on access to a national breeder database.

2. Public good science outputs can be defined as research that is likely to develop, maintain or increase research skills or scientific expertise that is, or are likely to be, of particular importance to New Zealand. Research skills in genetic modelling, in the practice of biometrics and in computational procedures rely on access to a large database. The development and maintenance of these skills will increasingly rely on the National Database as we have fewer genetically-structured Crown-funded flocks at research stations.

3. Public good science outputs may be of benefit to New Zealand but are unlikely to be funded from non-Government sources. This is certainly the case in the sheep industry where there is no single source of industry funds. A levy on all producers for the collective benefit of the industry might be considered, perhaps administered through the Wool and Meat Boards. However, the benefits of animal breeding advances accrue to the breeders, to the farmers who make use of these stock and to the various manufacturing and processing industries beyond the farm gate. The research is certainly unlikely to be funded from non-Government sources in entirety, though partial funding is likely to be available.

Operational aspects of Research using the National Database

In the first instance the National Database will hold all existing data from MAF Animalplan files. In future, licenced bureaux using the Animalplan genetic engine will export breeders' data to the National Database on an annual basis. Export and analysis of existing pedigree and performance records would be one possible use of the database for research. More commonly it is anticipated that researchers will add in data on promising new traits, for example eye muscle depth and width measured on the live animal. Records of pedigree and performance for the new and existing traits would then be exported and analysed together. These records would be used to obtain parameter estimates to determine the extent to which the trait is under genetic control and to identify non-genetic influences of performance. Parameter estimates can then be used to obtain new prediction criteria that are better able to make progress towards the selection objective. As appropriate, this research can be incorporated into the genetic engine by modifying the genetic parameter files, the genetic evaluation model and its distributional properties. Research findings that are not of a form which can be incorporated in the genetic engine will be disseminated to the industry via extension programmes co-ordinated by theTrust.

Funding to support the National Database

Once fully structured and established, the National Database will be funded on a user-pays basis. Breeders will make use of the National Database as a back-up source of their historical records and to access reference and transfer animals that have originated in outside flocks. Via the bureaux, breeders will pay a fee for access on a per-national identifier basis. The National Database is also used as a source of records for higher level processing such as sire reference analysis. Breeder charges for higher level processing will include a fee for accessing data from the National Database. Flocks that are not recording through licenced bureaux will also need to be entered via the National Database if they are to be used for higher level processing. This is to ensure that data edits, validations and national identifiers are consistent. According to FRST policy, database costs for research use will be charged as a cost against research rather than by direct funding through FRST. Accordingly, research users of the database will be charged for access using the same formula applying for higher level processing.

CONCLUSION

The National Database is one component of an integrated scheme of performance recording which spans buyers, breeders, bureaux, researchers and extension agents. The components of the scheme include both user pays and public good appropriable components. The database is important for both basic and applied research and is expected to become increasingly important as time goes on.

REFERENCES

Clarke J.N.; Wickham B.W.; Baker R.L.; Callow C.F.; Morris C.A.; Packard P.M. 1987. Development of animal recording for genetic improvement of pastoral livestock production in N.Z. *Proceedings of the IVth Animal Science Congress of the Asian-Australasian Association of Animal Production Societies,* **IV:** 43-46a.

FRST 1991. A guide to the Public Good Research Fund Assessment and Allocation Process. Foundation for Research, Science & Technology. October 1991, p10.

Garrick D.J. 1991. Best Linear Unbiased Prediction for across-flock/year breeding values. *Proceedings of the New Zealand Society of Animal Production,* **51:** 411-416.

Johnson D.L.; Rae A.L.;Clarke J.N. 1989. Technical aspects of the Animalplan system. *Proceedings of the New Zealand Society of Animal Production,* **49:** 197-202.

Wickham B.W. 1984. The research potential of shared information resources. *Proceedings of the New Zealand Society of Animal Production,* **44:** 83-86.