

## New Zealand Society of Animal Production online archive

This paper is from the New Zealand Society for Animal Production online archive. NZSAP holds a regular annual conference in June or July each year for the presentation of technical and applied topics in animal production. NZSAP plays an important role as a forum fostering research in all areas of animal production including production systems, nutrition, meat science, animal welfare, wool science, animal breeding and genetics.

An invitation is extended to all those involved in the field of animal production to apply for membership of the New Zealand Society of Animal Production at our website [www.nzsap.org.nz](http://www.nzsap.org.nz)

[View All Proceedings](#)

[Next Conference](#)

[Join NZSAP](#)

The New Zealand Society of Animal Production in publishing the conference proceedings is engaged in disseminating information, not rendering professional advice or services. The views expressed herein do not necessarily represent the views of the New Zealand Society of Animal Production and the New Zealand Society of Animal Production expressly disclaims any form of liability with respect to anything done or omitted to be done in reliance upon the contents of these proceedings.

This work is licensed under a [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](http://creativecommons.org/licenses/by-nc-nd/4.0/).



You are free to:

**Share**— copy and redistribute the material in any medium or format

Under the following terms:

**Attribution** — You must give [appropriate credit](#), provide a link to the license, and [indicate if changes were made](#). You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

**NonCommercial** — You may not use the material for [commercial purposes](#).

**NoDerivatives** — If you [remix, transform, or build upon](#) the material, you may not distribute the modified material.

<http://creativecommons.org.nz/licences/licences-explained/>

## The Sheep Improvement Limited (SIL) genetic engine

S.A. NEWMAN, K.G. DODDS<sup>1</sup>, J.N. CLARKE<sup>2</sup>, D.J. GARRICK<sup>3</sup> AND J.C. MCEWAN<sup>1</sup>

### ABSTRACT

The Sheep Improvement Limited (SIL) genetic engine (GE) interfaces with the SIL national database. It calculates breeding values from data provided by the database and reports these back to the database along with adjusted deviations and genetic trends graphs. The SIL system is a distributed and modular design, with the database and the GE located on separate computers connected via the Internet. The individual GE components are also modular and version 2 consists of 6 separate goal trait modules: growth, meat, wool, reproduction, survival and host resistance to internal parasites. This modular design allows each goal trait module to be easily upgraded and additional modules can be added as required. The GE uses across flock and year multi-variate animal model BLUP. Evaluations can be requested as required by the user via Internet access to the database. The size of the job submitted scales transparently up to requests involving more than 200,000 animals, sufficient for the largest across flock sire referencing schemes. The intention is that the system will be continuously extended and improved. Near term developments are likely to extend to across breed analyses, the inclusion of a facial eczema module, and improved diagnostics. In the longer term it is anticipated that information from genetic markers will be included.

**Keywords:** sheep; breeding values; BLUP; computer; genetic evaluation.

### INTRODUCTION

The Sheep Improvement Limited (SIL) redevelopment of sheep performance recording in New Zealand has allowed many advances in computers and genetics to be incorporated into the new system. The core of any such system is the genetic evaluation of the information collected to enable estimated breeding values to be calculated. A major advance for New Zealand performance recording ram breeders is that SIL uses best linear unbiased prediction technology (BLUP). Previously, this was only available via specialised analyses and the number of traits was also limited. BLUP evaluations have been available for a number of years to both beef and dairy farmers and offer several advantages. The ability to rank animals across sexes or age classes allow better decisions to be made about what breeding replacements to use and this directly translates into faster genetic progress. It also allows the genetic progress of the flock to be easily estimated. BLUP also allows the genetic progress of a flock to be easily estimated and depicted as "genetic trends" graphs, which are a major aid to ram purchasers selecting a ram breeder with goals appropriate to their situation. The availability of across flock estimated breeding values, also further aids genetic progress because ram breeders can make better informed decisions about what outside sires to use in their flock. However, the primary benefit is that the estimated breeding values are more accurate as information from all related ancestors and descendants contribute to individual estimates. Previous systems only used measurements from the individual and close relatives. The value of this additional information over existing methods is low for highly inherited traits that can be measured on the individual itself, but are high for low heritability and sex-limited traits, such as number of lambs born. The potential increases in genetic progress mean the introduction of this technology is essential for the long-term survival of the New Zealand sheep industry.

### IMPLEMENTATION

#### Communication

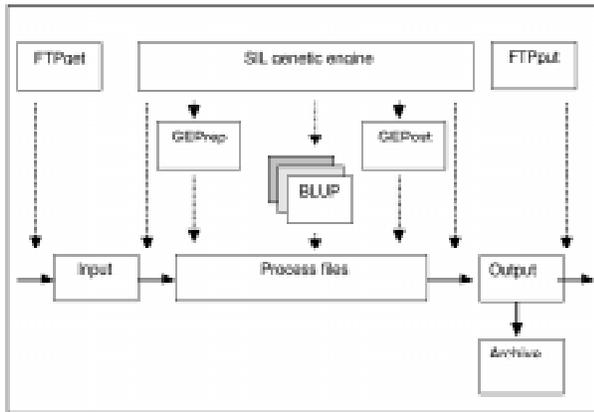
In order to use BLUP effectively for across flock

analyses the data has to be stored in a common format and a central location, hence the essential need for the SIL national database. Access to this database by many users located throughout the country in turn means that communication can only be via the Internet. However, the BLUP algorithm requires considerably more computer processing than previous sheep genetic evaluation systems, particularly when it is implemented as an animal model. Given the large interactive user demands on the computer hosting the database it was obvious that the breeding value estimation should be undertaken on a separate machine. Fortunately, the use of the Internet by the database computer provides a simple interface for the database and genetic engine computers to communicate with each other. An added bonus is that the interface is robust and available on a wide variety of operating systems. This in turn means few restrictions on hardware and the possibility of many genetic engine developments to be totally independent of the database. The amount of data required to be transferred between the database and the genetic engine means these computers could be widely separated geographically given a high speed Internet link. Such an implementation method also allows the processing of the genetic evaluations to be undertaken in parallel if implemented appropriately. Parallel processing would provide a rapid and transparent expansion in both the number and size of the genetic evaluations at minimal cost. However, in order for such a communication procedure to be effective it requires a simple, comprehensive data format to the genetic engine that can easily be modified to include new measurements, or new analyses. Currently, data is transferred by comma separated variable files with header records that identify the traits involved and the breeding values required. The actual data transfer to and from the database is controlled by two programs independent of the genetic engine itself called FTPput and FTPget (Figure 1), which are located on the genetic engine computer. These simply transfer files to and from specific directories after polling the directories at specified intervals to see if any files are present for transfer.

## Genetic engine

The genetic engine needs to undertake a number of tasks including: breeding value estimation, pre- and post processing of files, the generation of genetic trends graphs and phenotypic trait deviations, and various operating system, logging and file archiving functions. Undertaking all these using one programming language would be expensive and difficult to maintain and modify. Traditionally, breeding value estimation has been undertaken by programs written in Fortran, which are extremely fast, but difficult or impossible to use for graphing or complex operating system commands. To overcome these issues the genetic engine was highly modularised and the most appropriate language for each task used (Figure 1). This modularization in turn has the decided advantage that creation, maintenance and future upgrades of the components can largely be undertaken without rewriting of other modules.

**FIGURE 1:** Diagram of the major components of the SIL genetic engine and its structure. Data flow and locations are in black arrows and boxes at the bottom of the diagram. Various program modules and their control are shown above and their actions by dashed arrows. Note this diagram does not detail inter-program communication



The pre-processing step (Geprep) undertakes: various data checks, creates appropriate contemporary groups, and covariates, rescales the data, calculates phenotypic deviations, and creates appropriate files for subsequent BLUP evaluation. Its actions are controlled by user requests contained in the header of the data file. The BLUP modules consist of six goal trait groups of growth, meat, wool, reproduction, survival and disease (internal parasites), each using a variety of predictor measurements, where available, to estimate breeding values specific to that goal trait group (Table 1). The BLUP modules are implemented as a multivariate animal model and two, reproduction and disease, can also utilise repeated measurements. These modules transparently handle across flock analyses. While the range of breeding values that can be calculated is already extensive, further modules are being planned including: facial eczema, dagginess, additional carcass traits and additional wool traits. In addition, across breed evaluations are under development. The post processing step parses and combines the output from the separate BLUP modules, creates genetic trends graphs and adjusts the breeding values to a constant base year. At each step extensive log files are created so that any problems can be quickly investigated, as and when required. These log files also allow overall

summaries of the activity of the genetic engine to be regularly produced. Core files are archived so that any specific job can be recreated in response to future queries.

As currently configured, the genetic engine can transparently handle jobs ranging from 100 to more than 200,000 animals. The latter being the size of the largest across flock runs previously undertaken. Using commercially available hardware and recompiling the programs would allow this to be increased eight-fold, i.e. to approximately 1.6 million animals. Thus the system should be able to handle all requirements in the foreseeable future as long as care is taken to maintain the existing scalability in future enhancements.

The genetic engine is entirely automatic and requires no manual intervention under normal operating conditions. Checking the status of the system and, if required, manual intervention can be done remotely. To date over 1000 separate jobs have been processed in the five months since version two was implemented, a typical job requires approximately seven minutes of processing time and computer utilisation is currently less than 20%. The engine has been run continuously for periods in excess of a month at a time, albeit checked daily, without requiring any maintenance.

**TABLE 1:** SIL genetic engine BLUP goal trait modules (version 2.0 Nov 1999)<sup>1</sup>

Goal trait group	Predictor traits	Breeding values estimated
Growth	Weaning weight, autumn live weight, yearling live weight, carcass weight <sup>2</sup> , adult ewe live weight <sup>2</sup>	Weaning weight, autumn live weight, yearling live weight, carcass weight, adult ewe live weight
Meat	Weaning weight, autumn live weight, ultrasonic or assessed fat depth, ultrasonic EMD, ultrasonic EMW	Fat depth, EMD, EMW, lean weight, fat weight, EMA
Wool	Weaning weight, autumn live weight, yearling live weight, hogget fleece weight, fibre diameter, lamb fleece weight, two tooth fleece weight.	Fibre diameter, lamb fleece weight, hogget fleece weight, ewe fleece weight
Reproduction <sup>3</sup>	Autumn live weight, number of lambs born	Number of lambs born
Survival	Lamb survival birth to weaning	Lamb survival
Disease <sup>3</sup>	Weaning weight, autumn live weight, hogget fleece weight, summer strongyle egg count, summer <i>Nematodirus</i> egg count, autumn strongyle egg count, autumn <i>Nematodirus</i> egg count, host resistance antibody test	Summer strongyle egg count, autumn strongyle egg count, adult ewe faecal egg count

<sup>1</sup> Implemented as multivariate BLUP animal models for each goal trait group

<sup>2</sup> To be implemented

<sup>3</sup> Repeated trait model

## **CONCLUSIONS**

The SIL genetic engine can be easily modified and expanded as a consequence of its modular design. Also, by being located on a separate machine from the SIL database, it is independent of its processing and operating system requirements. A further advantage is that it can easily be modified to allow individual jobs to be processed in parallel on separate remote computers. Currently, up to 200,000 animals can be analysed for all goal trait groups in a single analysis, but this can be expanded transparently eight-fold by purchase of available hardware. The intention is that the genetic engine will continue to be developed, to provide increased functionality, and incorporate advances in genetic evaluation techniques.