National Pig Improvement Program

Ron Crump and Susanne Hermesch

Animal Genetics and Breeding Unit, a joint institute of NSW Agriculture and the University of New England, University of New England, Armidale, NSW, 2351

Introduction

The National Pig Improvement Program (NPIP) is an across-herd genetic evaluation system for Australian pig breeders. Established in 1995 by the Queensland Department of Primary Industries (QDPI) as a means of encouraging genetic improvement in seedstock herds and dissemination of this genetic improvement to the commercial tier, the NPIP was the responsibility of Michael Macbeth with technical support from the Animal Genetics and Breeding Unit (AGBU).

Originally, the NPIP used the PEST genetic evaluation program to calculate across-herd estimated breeding values (EBVs) alongside other programs to manage the data and calculate other required information (eg accuracies). In 2001, responsibility for all aspects of the NPIP passed to AGBU. As a result, the genetic evaluations now use the PIGBLUP analysis modules and opportunities exist for joint future development of the two systems.

The genetic evaluations in the NPIP include lifetime daily gain, ultrasonic backfat depth and number born alive defined as different traits in the first parity and later parities. The EBVs for these traits are combined into two selection indexes, one targeted at terminal sire lines and one at maternal lines (see page 35). Currently, the NPIP includes three breeds (Large White, Landrace and Duroc).

The NPIP web site can be found at `http://npip.une.edu.au`. On this site you will find general information pertaining to the NPIP along with the publicly available parts of the NPIP results (eg EBVs for AI boars, EBV distributions and genetic trends). The evaluation results on the web site are automatically updated every time the system is run.

For information or general assistance on the NPIP contact Susanne Hermesch at AGBU (`susanne.hermesch@une.edu.au`).

Procedure

Participation in the NPIP is open to any Australian seedstock producer. However, they must undertake to create and maintain genetic links with other member herds. Registration is required before any data can be included in the analyses.

Members submit data via e-mail to `npip@une.edu.au` and receive their results back in the form of a ZIP file attached to an e-mail. The ZIP file contains the results in a number of HTML format files, which can be viewed using a web browser and, if required, imported into word processing or spreadsheet software.

What information does the system require?

1. Data on NPIP members

When a member joins the NPIP certain things are setup before data can be analysed. The breeds for which data will be submitted, e-mail addresses from which data will be sent and basic information pertaining to the data files (breed codes, sex codes etc). this information can be modified later, eg if a new breed needs to be added or e-mail addresses updated, but requires a request to the NPIP Administrators.

2. Data on pigs

There are then three types of information pertaining to pigs, the data file, a list of identifiers of imported/AI animals and a list of animals currently in the AI stations.

A routine run of the NPIP system results from the receipt of a PIGBLUP data file, so this is updating the data for that herd-breed combination. Information on across-herd identification and animals in the AI stations is not included in the PIGBLUP data format, and so this has to be collected separately. Currently this entails a process of prompting members for new or updated information followed by manual input of the data, which may be returned by e-mail, fax or via the phone.

The maintenance of identification and AI animal information is very important. Failure to identify animals across herds results in genetic connections not being exploited and improper use of the information. In the most extreme case, ie where no connecting animals were identified across herds, a series of within-herd BLUPs would be running simultaneously rather than a single across-herd evaluation. The AI animal information is important in reporting results on the current publicly available AI animals, both to members and at `http://npip.une.edu.au`. Out of date information reflects badly on both the system and its members.

To facilitate the exchange of identification information, Susanne Hermesch has recommended an across-herd numbering scheme to the NPIP members, which includes the three letter herd code and the original animal identification. Successful, complete uptake of this numbering system will enable the processing of link animals automatically *at some point in the future*, rather than the manual updating of this information that currently goes on. A number of breeders have already adopted this identification system which has made the identification of boars across herds easier.

Data submission

PIGBLUP v1.0 format data files are submitted via e-mail to `npip@une.edu.au`. This e-mail address is not used for any purpose other than receipt of NPIP data. If the e-mail does not contain a recognisable PIGBLUP-format data file or does not originate from a registered NPIP client e-mail address the NPIP Administrators are notified and no further action is taken.

How often should data be submitted? The quality of the evaluations and therefore the value of the system is dependent upon the data submitted. In the past, data submission to the NPIP could best be described as ‘patchy’, with breeders only submitting data if they wanted NPIP results for a specific purpose (eg for a sale). Attempts are being made to encourage more regular and frequent data submission. The first goal to be achieved here will be regular, monthly submission by all members and then ways to encourage and assist in more frequent submission will be investigated and implemented.

In the first instance, an automated reminder system has been set up. Just before midnight on the last day of each month a reminder is sent to each NPIP member that has not successfully submitted data in that calendar month. If the member only responds to these e-mails, submission will be effectively bi-monthly. For example, if a member receives a reminder on March 1 and submits data in response they will not be sent a reminder March 31, but would at the end of April. This method appears to have been successful with data submission being stimulated by the reminders. Having run this system for a few months with no negative feedback, it will be adapted to encourage a monthly distribution of data submissions. Overall, data submissions have been going through smoothly for all members since 2002.

Notes:

- The data file submitted can be compressed using any ZIP utility (eg WinZip, ZipGenius) in order to minimise the length of time spent on line.
- The data file does not need to span all years of data. It must go back far enough to include all modified data, including changes to historical identifiers and data as well as to currently active animals.
- Change of e-mail address or additional e-mail addresses must be notified to Susanne Hermesch to allow us to update the registered list. Data submitted from non-registered e-mail addresses will not be run automatically.
- Submission of data via floppy disks and Australia Post is not encouraged, but can be carried out if circumstances dictate.

Data preparation

The NPIP evaluations are carried out using the PIGBLUP analysis modules (compiled as standalone programs to run on the Linux-based PC that functions as the NPIP machine). As with a PIGBLUP analysis, only a single data file is supported and identifiers must be unique within that file. It is therefore necessary to process the data files received from each NPIP member to combine the old and new data within a herd and to combine the files across herds ensuring that herd effects will be included (user-defined management group fields in the data file are used to nest the user-defined management groups found in the data files within herds) and that identifiers are unique across the system.

On receipt of a PIGBLUP data file for a herd-breed, this file is compared to and combined with the pre-existing file for this herd and breed. Assuming there are not too many differences between the information common to the old and new files, then the combined file is kept and placed in a queue for analysis.

When a run starts all of the data files for the breed being analysed are combined into one, large file and a file which maps unique across-herd identifiers to within-herd identifiers is created. It is at this stage that the information on link animal identification and AI animals is used and amalgamated into the PIGBLUP data file. The across-herd identifier contains information on the herd of origin and whether the animal is a link or AI animal.

PIGBLUP allows users to flag animals that must receive an evaluation, ie animals that are neither parents nor part of the recent cohort. In many cases, the data received by the NPIP has excessive numbers of flagged animals. That is, animals are still flagged which if still alive would have to be parents by now. To reduce this, and maintain the appropriateness of the Sow, Gilt, Old Boar, Young Boar categorisation in the results active flags are stripped from animals over two years old. Boars in AI studs are flagged to ensure that they get an updated EBV.

Notes:

- A large number of changes to historical data between data submissions will result in a delay. The automatic system will ask the NPIP Administrators to check the problem. It is then up to them to manually accept the data and start the run, or reject the data. This has been observed when herd-recording systems have been upgraded such that different rounding was applied (ie the last digit of a number has varied) or previously blank fields were replaced with zeroes (still effectively blank for many variables in PIGBLUP).

Output

The outputs from the NPIP system consist of EBVs and accuracies for individual traits, and both terminal sire and maternal $INDEXs (more of which later) plus summaries of these (trends, graphs) and of the data structure.

The results from a routine NPIP run updates the web pages and sends results via e-mail to the member that had submitted data to that run. That is, no data submission, no updated/new results.

The results sent out to members are attached to an e-mail message as a compressed archive. The files, other than some graphics files, are in HTML format. The use of this format means that the members will all have access to viewing software (a web browser), but can also import the data into other software for further processing or reporting.

Originally, results were sent out in the body of an e-mail message. This resulted in a very unwieldy message that was hard for members to navigate, understand or use further. The use of HTML allows easy navigation to be built in, via hyperlinks and frames, easy access to definitions of abbreviations and easy importation of the results into other software packages.

The results archive contains information in the following areas:

1. News

Any information on modifications etc is visible to the member immediately upon opening the index.html file.

2. Results

For the members own animals and their parents. As in PIGBLUP, these results are divided into Old Boar, Young Boar, Gilt and Sow categories. Estimated breeding values and accuracies for each trait are reported along with information on how many herds, litters, and recorded offspring an animal has. Rankings are presented for both the terminal sire and maternal $INDEXs. For Young Boars and Gilts, deciles are presented for each EBV. That is, which ten percent of the population of Young Boars or Gilts the animal is in for that EBV.

Results (EBVs and accuracies) for AI boars are presented separately, within AI centre.

3. Breeding value summaries

- Genetic Trends

Graphs of genetic trends are available for each trait and both $INDEXs. In each genetic trend graph the trend for the member's own herd is presented with the average trend in all other NPIP herds.

- Distributions

Distributions of Young Boar and Gilt EBVs have been plotted such that the member's herd distribution can be seen within the overall, across-herd distribution of EBVs.

4. Information to assist in maintaining and identifying genetic links

- Base Parents

Animal that appear only as parents in the member's herd and have no known parents are reported. The aim of this is to try to provide a prompt for member's to update identification information on any imports of genetic material.

- Link animal usage

Counts of recorded animals born to sires with recorded animals in other NPIP member herds are created. These are summarised by quarter to give a picture of the pattern of link animal usage over time. The aim is to maintain 10% of births by link animals. For this to provide useful information in the evaluation, these animals must be recorded.

- Connectedness

Originally, connectedness calculations in the NPIP were only performed intermittently. The methodology used by the beef cattle evaluation system BREEDPLAN has now been adopted and these much quicker calculations are performed during each NPIP run. The calculations identify valid statistical comparisons between recent groups of animals based on pedigree. At present, all member herds within the NPIP are connected, however the degree of connectedness (that is, how confident you can be that any comparison between two herds is going to be correct) varies. The results are presented visually, with colours representing the degree of connectedness. Currently, the results are on a relative scale (that is, the scale goes from least connected to best connected within the evaluation) however this is going to be modified to an absolute scale in order that between evaluation improvements in connectedness may be observable.

Other Areas

1. Indexes

In order to make selection decisions, breeders need to be able to weigh individual EBVs by their economic importance and combine them into an index. The PIGBLUP system provides breeders with the tool of the $INDEX which uses information on economic, production and marketing parameters to combine EBVs. Two indexes have been set up in the NPIP to accommodate different selection strategies by breeders. The Terminal Sire Line index places emphasis on backfat and growth rate and has no significant weight on litter size. The Maternal line index has a higher weight on litter size and growth and places less weight on backfat. The input parameters used for the two indexes in the NPIP are shown in the Appendix.

Both indexes express the economic worth of differences in EBVs on a per litter basis. The approximate economic weights of each index are shown in Table 1. The economic weight quantifies the economic benefit of improving each trait by one unit; what is the economic worth of improving growth rate by one gram per litter?

Table 1 Approximate economic weights for growth rate (ADG), backfat (BF) and litter size (NBA) for the Terminal Sire line index and Maternal line index

	ADG* (per g)	BF (per mm)	NBA (per piglet)
Terminal sire line	0.63	- 19.7	+ 7.00
Maternal line	0.68	- 4.5	+ 24.0

* Weights are expressed on a per litter basis.

Information on economic weights and EBVs is then combined into one index by taking the sum of all EBVs each multiplied by its corresponding economic weight (ec. Wt.).

$$\$Index = ec.\ Wt._{(trait1)} * EBV_{(trait1)} + ec.\ Wt._{(trait2)} * EBV_{(trait2)} + \ldots + ec.\ Wt._{(traitn)} * EBV_{(traitn)}$$

Cameron and Crump (2001) derived economic weights reflecting average input parameters of the Australian pig industry which are shown in Table 2 as a point of reference.

Table 2 Economic weights of pig production traits in Australian dollars (Cameron and Crump, 2001).

Trait	Increment	Per farrowing (litter)
Average daily gain	10 g / day	3.97
Carcase backfat (mean: 18mm)	-1 mm	-16.55
Feed conversion ratio	-0.1 kg / kg	-17.04
Number born alive	1 piglet	31.72
Number born alive (parity=1)	1 piglet	10.19
Number born alive (parity > 1)	1 piglet	21.53
Dressing percentage	1 %	11.25

2. Connectedness

The presence of genetic links between herds is essential in any across-herd evaluation of livestock and there is the question of 'how much connectedness is enough?' In the original system, connectedness was reported as good, fair, poor or bad for each pair of herds and for each trait. The accuracy of comparisons between two herds will always be lower for a trait with low heritability than for a trait with high heritability. However, this also holds for within-herd comparison of animals, and as long as we are aware of this, we should be able to use a method that is not trait (heritability) specific.

An approach based on that used by the beef cattle evaluation system BREEDPLAN has been implemented in the NPIP system. However, in BREEDPLAN a single cut-off point is used to determine whether two herds are connected or not whereas in the NPIP we use a range of cut-off points to assign the connectedness between two herds to a point on a scale. These points are then represented visually, by means of shade of colour, in the results.

3. Sire Reference Scheme

In order to improve genetic links between herds a sire reference scheme was in place until recently. Initially, the sire reference scheme required breeders to nominate boars as reference sires and then vote for individual sires. In 2002, six boars were nominated by Susanne Hermesch after consultation with breeders as reference sires. However, breeders preferred to use sires from other herds that suited their individual breeding programs, which reduced the use of the proposed reference sires across herds. It is believed that a reference sire scheme is too restrictive for breeders and the responsibility has been placed with the breeders to use boars from other NPIP members for 10% of their matings.

4. Heterogeneity of variance

The observed data, and EBVs, for any herd can be described in terms of a mean and a variance (the 'spread' of the values about the mean). The inclusion of within-herd contemporary groups in the model for the BLUP evaluation removes differences in mean performance between herds and between time periods within herds. However, there may be differences in the level of variation across herds.

A difference in variance between herds is termed heterogeneity of variance. The result of heterogeneity of variance is that the best, and worst, animals in the evaluation will appear to be from the most variable herds. It depends on the cause of the heterogeneity whether this is a problem.

For example, one herd could be using a more genetically diverse (variable) group of boars than another herd and this would be reflected in the variability of the performance of the offspring. This would not be a problem. However, if the difference in variation is related to management, or failure to record differences in management between groups of animals within herd, then this should be addressed.

A preliminary simple investigation of heterogeneity of variance in the NPIP data indicates that there is heterogeneity present. This heterogeneity is both between herd and within-herd across time periods. Management or environmental causes cannot be identified since the information is not available from the PIGBLUP data files. In a second study we hope to identify whether the heterogeneity has arisen in an acceptable way, and if not we will consider how to deal with it.

If it is deemed necessary to address the issue in the evaluation system, there are a number of approaches which boil down to two main groups; pre-adjustment of the variance prior to the evaluation (quick, approximate), adjustment of the variance during the evaluation (slows down the evaluation but statistically preferable, eg Pujol *et al.*, 1998).

Future

1. New Traits

There will be new traits included in the NPIP evaluations, but only as their uptake by members becomes widespread. In the first instance, the analysis of juvenile IGF-1 as an indicator of FCR will be implemented as most current NPIP members are recording this trait.

Where a trait is only recorded by a subset of members, results for that trait will only be reported to members that are currently recording that trait (ie have records taken within a set recent time frame). This is in order to maintain some advantage for people that are investing in recording.

2. New Tools

A list of the top 5% or *n* (whichever is greater for the breed) Young Boars for each EBV will be published on the web pages and sent out to members. This will provide an indication of where to shop if you are looking for the best animals for a particular trait or $INDEX.

Diagnostic tools, for example to try to explain changes in EBVs for an animal between NPIP runs, are being considered. However, the NPIP does not currently back up solutions and data between runs, and this procedure would also have to be in place.

Software to assist with data submission by the members is a possibility. If the comparison of old and new data files took place on the members computer, far smaller files would have to be uploaded for analysis. In addition, any such software could be used to store link animal identification information and hence automate the updating of this also.

Ultimately, it is easily possible to imagine an on-line database system where members upload their new data, modify their own settings, view and download the set of results they want and possibly answer some of their own questions through a set of data queries.

References

Crump, R. (2001). "The National Pig Improvement Program." *2001 Pig Genetics Workshop, Armidale.* pp 6-10.

Pujol, M.R., J.Piedrafita, R. Quintanilla, A. Reverter and J. Tibau (1998) *Proc. 6th World Congr. Genet. Appl. Livest. Prod.* **25**:645-648.

Appendix: Input parameters used for Terminal Sire and Maternal line indices used for the National Pig Improvement Program.

	Unit	Terminal sire line	Maternal line
3. **Economic parameters**			
Average Carcase Market Price	$/kg	2.45	2.45
Premium for Aus-meat grid fat class cypher 0	$	.20	.05
Premium for Aus-meat grid fat class cypher 1	$	.20	.05
Premium for Aus-meat grid fat class cypher 2	$	.00	.00
Premium for Aus-meat grid fat class cypher 3	$	-.15	-.10
Premium for Aus-meat grid fat class cypher 4	$	-.30	-0.15
Premium for Aus-meat grid fat class cypher 5	$	-.50	-.15
Cost per kg Feed in Feeder	$/kg	0.30	0.24
Non-feed Costs per Day (eg.,Labor&Housing)	$pig/day	.30	.40
4. **Production parameters**			
Number of Pigs Born Alive per litter,	n	10.35	10.35
Pre-weaning Mortality	%	17.00	17.00
Post-weaning Mortality	%	3.00	3.00
Average Daily Gain to Market Weight	gm/day	589	585
Mean p2 Fat Depth	mm	14.00	14.00
Mean Muscle depth		55.0	55.0
Is feed restricted in grow/finish to reduce P2 fat?	(Y/N)	N	N
Feed Conversion, (kg feed per kg live wt.)	kg/kg	3.11	3.11
Target Market Carcase Weight	kg	75.00	75.0
Dressing Percent	%	78.00	78.0
5. **Marketing Parameters**			
Percent of Boars sold (or used) as Terminal Sires	%	40.00	0.00
Percent of Boars sold (or used) as Maternal Sires	%	0.00	40.00
Percent of Boars sold (or used) as Slaughter Boars	%	60.00	60.00
Percent of Gilts sold (or used) as Replacement Gilts	%	20.00	70.00
Percent of Gilts sold (or used) as Slaughter Gilts	%	80.00	30.00
Payment grid fit - use a weighted quadratic fit		On	On