# PIGBLUP Version 6.00

A.L. Henzell

*Animal Genetics and Breeding Unit (AGBU), University of New England, Armidale, NSW 2351. AGBU is a joint venture between New South Wales State Department of Primary Industries and University of New England.*

## Overview

PIGBLUP has been significantly enhanced since V5.20 (see Henzell, 2006) – particularly with the expansion in the number of Production and Carcase traits we are now able to analyse.

PIGBLUP now runs faster than before due to some more efficient data structures being used. However, the large speed-up I had hoped to introduce with V6.00 by using full dynamic memory allocation could not be completed in the time available for this version's release. Preliminary steps such as pre-scanning the data file to glean sizing information are complete – and the information may be viewed in the *Reports/Summaries* section. I hope to complete this modification with the next release after V6.00.

These changes to PIGBLUP necessitated an overhaul and further expansion of PigCheck's error checking capabilities in addition to allowing checking of the new traits. Also, more selective viewing of errors is possible now.

With V6.00's greatly expanded traits, Migrate is essential when upgrading from earlier versions. Migrate is able now to correctly migrate all re-defined traits.

A new module, PigSched, is introduced with V6.00. This is a standalone scheduling module which allows the user to request analyses at any time of the day conditional on the availability of data files. At the scheduled analysis time, PigSched will start PIGBLUP, perform the requested analysis, log the outcome of the analysis, close PIGBLUP and, if requested, re-schedule another run at a nominated time or day.

FileMerger is altered little from V5.20 other than the standard PIGBLUP data file formats having been expanded to include the new traits.

In keeping with the past, AGBU does not support PIGBLUP on the initial release of new Windows operating systems. While several users are running PIGBLUP on the *Vista* operating system, we do not currently support it on Vista. AGBU has deferred supporting PIGBLUP on Vista at least until after the release of V6.00.

# PIGBLUP V6.00

Version 6.00's main innovations will be described first followed by a brief description of minor modifications and bug fixes.

PIGBLUP is now able to analyse up to 11 Production traits in addition to the standard *Average Daily Gain (*ADG), *Test Daily Gain* (TDG), *Back Fat* (BF) and *Feed Conversion Ratio* (FCR) traits. Further, 13 Carcase traits in addition to *Carcase Fat* (CFat), *Carcase Muscle (*CMus), and *Lean meat Yield* (LMY) can be analysed now. All but ADG and TDG may be re-defined – previous hard-coded special handling of these traits has been removed to the traits definition file; now all traits other than ADG and TDG are equal. New traits are generic traits: users have to define them, setup appropriate variance components, set suitable limits and index weights. All Production and Carcase traits are evaluated in a single multivariate analysis.

All PIGBLUP data formats have been updated to accommodate the expanded traits. A new data format 3.0 has been added to increase the capacity and consistency of some fields (especially user-defined management groups), tidy-up record structures and make future development simpler. Data format 3.0 is now the recommended data format.

In the *Set-Up/Parameters* section a new *User-Defined Base* page has been defined. The purpose of this option is to allow EBVs to be reported relative to a fixed or moving set of *base animals*. These base animals may be identified on this page by sex, breeds, and a date-of-birth window. After an analysis is completed, if a *User-Defined Base* has been defined, PIGBLUP scans the EBVs file calculating the average EBVs of the base animals. The EBVs record of every animal is then re-written with the base animals' EBV averages subtracted. The base animals' average EBV values may be viewed in the Reports/Summaries section. The *raw* EBVs from the analysis are retained in a separate file for use in the Trend Graphs section.

Users who process a variety of data files with a range of identifier (ID) widths for different herds and breeds no longer need to change the ID width in the configuration file; ID width is now defined for each breed in the FARM1.INP parameter file. PIGBLUP is better able now to determine a file's ID width and the data format being used.

Previously, sows had to have a parity one record for any of their reproduction data to be analysed. All reproduction records for consecutive parities up to the final or first missing parity were then analysed; later parities were ignored. A new *Enforce Initial Parities* option has been added to the *Set-Up/Parameters/Model* screen. If this is unchecked, parity one is not required and PIGBLUP will analyse the first sequence of consecutive parities.

The following minor changes have been made:

- A selected reproduction trait without data no longer terminates an analysis of reproduction traits; a single reproduction trait with data is sufficient for the multi-trait analysis to proceed.

- Several 21-day litter weight bugs have been fixed: If *Number at weighing* was non-zero and *Weighing Date* was valid but *Weight at weighing* was

blank, the trait for the animal was analysed with a zero weight. Also, periods without data now have correctly calculated fixed effects. Adjustment regressions are now reported with respect to the actual data means rather than the *Age at weighing* limits mid-point and *Minimum number of piglets fostered*. This has no effect on the EBV calculated.

- Fixed effects for Weaning to Conception Interval (WCI) were incorrect for periods without data. Also, to make WCI consistent with other traits, its data file statistics in the RUN.LOG file for WCI now report the *raw* data statistics for the trait rather than the WCI statistics for the animals to be analysed.

- *Defaults Editing/Printer Set-Up* now updates the printer setup and saves the settings for use in later sessions.

- A *DfltMaxAnimalsViewed* field has been added to the configuration file. This value is used to limit the number of animals displayed when viewing EBVs to improve the response time when handling very large data files. A copy of this value may be edited in the *View EBVs* options. This limit is applied *after* sorting. Thus, a change to the sort field being used will display a different subset of animals.

- A *FixedHeader* parameter has been added to the configuration file to enable the user to omit header records being written to the fixed-format EBVs file (FXCUSTOM.INX).

- Limits editing has been re-vamped to handle the expanded set of traits and to simplify editing.

- Flags editing has been made easier to use by enabling options only as their pre-conditions for use are met. Also, screen resizing now correctly re-locates screen components rather than them being lost off screen as was happening previously.

- The Production/Carcase and Reproduction convergence graphs are now saved to file and may be viewed at any time from the *Reports* screen.

- A Genetic Audit bug was fixed whereby *selection intensity* statistics were incorrect if no reproduction traits were being analysed.

- PBSAMA displays average inbreeding levels and average *$Indexes* of the litter resulting from a specific mate.

- When *Analyse* is clicked, PIGBLUP checks whether current scan statistics exist in the directory containing the data file. If no such statistics file exists or it pre-dates the data file, a fast new scan is carried out. These statistics may have been generated by a PigCheck scan and both PigCheck and PIGBLUP allow the statistics to be viewed. (In the next version, these statistics will be used for full dynamic memory allocation and should result in a further significant speed-up of PIGBLUP.)

- A more detailed summary of how genetic groups are formed is now written to RUN.LOG and made viewable in *Reports/Summaries*.

PIGBLUP may be protected now with a USB Ultra-Pro dongle rather than the old parallel port style dongle. There is a charge for the new USB dongle (and a partial refund on the old). Contact AGBU in writing should you require this change. Note that there are Sentinel Pro and Ultra Pro parallel port dongles in use. If you are using a Sentinel Pro dongle, you will need a new copy of the PIGBLUP software; otherwise, it is just an exchange of dongles.

## PigCheck

PigCheck is a standalone program for identifying data errors and inconsistencies and correctly sorting animals and their records in a PIGBLUP data file.

The main enhancement to PigCheck other than catering for the expanded set of traits is in the selective viewing of validation errors.

Previously, users had to scan through an unordered error log. The new *Sieve* option permits errors to be viewed

- By individual animal,

- By groups/lists of animals,

- By record type, and

- By Severity level range.

Furthermore, lists of animals for viewing may be built easily and used to generate detailed displays of the errors discovered for each animal in the list.

A new PigCheck chapter has been written for the PIGBLUP manual describing each option in the new PigCheck. The relevant section in this chapter should be read before trying to use the *Sieve* option.

*Pedigree Checking* has new options for displaying <u>possible</u> boar and sow mis-identifications as well as more complete litter inconsistency detection. Mis-identifications are sensed using an *edit distance* measure between two animal IDs. This can be thought of as the minimum number of keystrokes needed to change one ID into another. For sows, only sows with similar farrowing dates are compared. Such errors are not found as litter inconsistencies as they cause what should have been a single litter to be split into two or more independent litters. Typing errors usually are the cause of such errors. It is important to realise that these are only <u>possible</u> errors.

A Data Summary Scan available in PIGBLUP as outlined above may also be requested in PigCheck and the summary statistics and some errors identified during the scan may be viewed.

## FileMerger

FileMerger is a very general program which may be used to combine recorded animal data held in formatted or un-formatted non-PIGBLUP text files and convert it into a correctly sorted PIGBLUP data file.

Several minor bugs have been fixed and all standard PIGBLUP data file templates have been modified for the expanded trait set.

It is now possible to use FileMerger to convert a PIGBLUP data extract from one format to another.

Several new functions have been added to aid construction of calculated fields.

## Migrate

Migrate is a standalone program which may be used by users following their installation of a new version of PIGBLUP. For all herds and breeds in use in the previous version, Migrate builds corresponding herds and breeds in the newly installed new version's directory. All necessary conversions are made so each migrated herd and breed is ready for use and commercial operation with the new version may proceed immediately.

By making all PIGBLUP traits equal in the sense that all their attributes are now specified in the traits definition file rather than some traits having some attributes specially coded, Migrate should now correctly convert all pre-V6.00 herds and breeds to V6.00.

Migrate will now prompt for a default animal ID width before performing the migration. This applies only to migration from any PIGBLUP version prior to V6.00; future migrations from V6.00 to later versions will not make this enquiry.

In order to minimise the number of old versions of PIGBLUP Migrate may be called on to convert, only versions V5.00 and later are catered for. As new versions are released this will be further reduced to V5.20 and later.

## PigSched

This is a new scheduling program specially designed to start-up PIGBLUP and interact with it to perform analyses at pre-set times and dates – possibly on a recurring basis.

At the scheduled analysis time, PigSched will run PIGBLUP, perform the requested analysis, log the outcome of the analysis, close PIGBLUP and, if requested, re-schedule another run at a nominated time or day.

Some users receive data extracts during the night from remote sites. PigSched allows them to schedule analyses after a scheduled start-of-analysis time and within a chosen time window subject to the availability of a named data file. When the start-time is reached, PigSched will check for the data file. If it is found and it is not running another PIGBLUP analysis, it will start-up PIGBLUP and instruct it to run a particular analysis.

If the data file is not available yet, it can check whether any other job has reached its start time and is ready to run.

Once an analysis has been completed, a new analysis may be rescheduled if the user has elected to do so. Also, if a data file does not become available in the nominated start window, PigSched can automatically reschedule the analysis for the next scheduled run time.

Other features provide for such things as automatic re-start after a power failure and to detect when a user has gone home and left PIGBLUP running idle (in which case, after a nominated time, PigSched may shutdown PIGBLUP so it can run any scheduled jobs).

In order to automatically run scheduled analyses, PIGBLUP has been modified transparently to detect the scheduler's request and simulate a user's mouse clicks to select a herd/breed, perform the analysis and finally exit PIGBLUP. In order for the two programs to interact, they need to know where to find each other. Both PIGBLUP and PigSched register themselves with the Windows registry when installed or, if no registry entry is found when started, re-register themselves. Both programs have options to de-register themselves also. This may be necessary when either program is re-installed.

## EurekaLog

The programs of PIGBLUP V6.00 will now be released with an embedded third party module called *EurekaLog*. This program has minimal overheads in terms of memory space or run time. The configuration file has several new options which allow you to switch *EurekaLog* off or on.

*EurekaLog*'s function is to catch *Access Violations*, *Illegal Memory Accesses*, *Invalid Mathematical* operations (eg division by zero) and *Memory Leaks*. Usually, such errors are caught during testing prior to release but Memory Leaks are typically a failure by the programmer to perform all housekeeping chores before exiting a program.

Should you notice a pop-up window shortly after exiting PIGBLUP or PigSched like
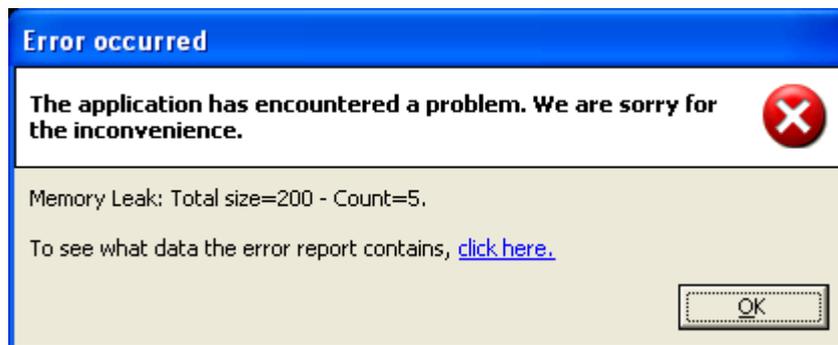


*EurekaLog* has been configured to write error details to special files and, should you agree, email them to AGBU. (Note that you may examine the generated file prior to your agreeing to its being sent. It does not contain any client data.)

*EurekaLog* has been used extensively during the development of V6.00 to locate a number of housekeeping errors. These errors did not cause wrong analyses but did cause rare unexplained Access Violations. Use of EurekaLog has helped make PIGBLUP and its associated programs even more robust than before.

## References

Henzell, A.L. (2006) PIGBLUP Update, *AGBU Pig genetics workshop notes* pp 5-7.