Final report

Development of accuracy metrics for genomic products

Project code: L.GEN.2501

Prepared by: Li Li, Phillip Gurman, Andrew Swan, Daniel Brown
Animal Genetics and Breeding Unit

Date published: June 2025

PUBLISHED BY
Meat & Livestock Australia Limited
PO Box 1961
NORTH SYDNEY NSW 2059

Meat & Livestock Australia acknowledges the matching funds provided by the Australian Government to support the research and development detailed in this publication.

Abstract

This project addressed a key limitation in the current MERINOSELECT genomic evaluation pipeline: the absence of accuracy estimates for genomic-only breeding values (GEBVs) and Flock Profiles. These values are important tools in the Australian Merino sheep industry, especially for animals without recorded pedigree or phenotypes. However, their usefulness is reduced without an indication of accuracy, particularly for flocks with weak genetic links to the reference population.

The project aimed to develop and implement new methods to estimate the accuracy of GEBVs and Flock Profiles. After a review of existing methodologies, new algorithms based on the prediction error covariance (PEC) of SNP effects were developed and integrated into the existing accuracy program accounting for genomic information - snpEPN software which is using in both Sheep Genetics and BREEDPLAN evaluations. These algorithms were tested and validated using simulations, SNP accuracy measures, and real-world industry datasets.

Results showed that the accuracy of GEBVs and Flock Profiles is strongly influenced by the reference population size and trait heritability, and genetic connectedness to the reference population. The new algorithms for GEBVs and Flock Profiles were computationally efficient and demonstrated reliable performance across different datasets. A four-level reporting system for accuracy was proposed to support routine use in Sheep Genetics evaluations.

These outcomes improve confidence in genomic evaluations for the Australian Merino sheep industry and support the broader adoption across other sheep breeds and species.

Executive summary

Background

This project aimed to address a key limitation in the Sheep Genetics evaluation system: the absence of accuracy estimates for genomic-only breeding values (GEBVs) and Flock Profiles. These genomic tools, derived through post-analysis of single-step genomic BLUP (SS-GBLUP), are increasingly used in the Merino sheep industry to inform selection decisions for animals without phenotype or pedigree data. However, without accuracy estimates, users face challenges interpreting the reliability of GEBVs—particularly in commercial flocks that are weakly linked to the reference population. This gap affects the confidence and comparability of these genomic predictions.

Objectives

- Develop and validate new methods to estimate accuracy for genomic-only breeding values.
- Develop and validate new methods to estimate accuracy for Flock Profiles.

Both of the above objectives were successfully achieved. A comprehensive literature review was conducted, formulas for estimating the accuracy of GEBVs and Flock Profiles were derived. These formulas were validated using a small testing dataset by comparing the results with those obtained from an independent program which calculated full empirical accuracies, confirming that the estimated accuracies were consistent with expectations.

- Implement new methodology into the routine evaluations for Sheep Genetics.

These new algorithms were integrated into the existing genomic accuracy program (snpEPN) and applied to a latest Merino industry dataset to generate accuracy estimates for GEBVs and Flock Profiles across all traits to confirming these new algorithms were computationally efficient.

- Document and publish the new methodology.

The literature review, algorithm derivations, and validations using a small testing dataset and the latest industry dataset were documented in the Milestone 1 and 2. Some of the results were also published at the 2025 AAABG conference (Li et al. 2025), and the full findings are planned for publication in a peer-reviewed journal or presentation at an international conference.

Methodology

- A comprehensive literature review was conducted and the new algorithms to estimate prediction error variance (PEV) and prediction error covariance (PEC) of SNP effects were derived.
- Four validation approaches were used: SNP effect accuracy, genetic connectedness measures, random SNP simulation, and comparison with SS-GBLUP-derived accuracy.
- Accuracy estimates were tested using small-scale and full-scale MERINOSELECT datasets, covering over 61 traits.

Key Findings

- These algorithms were implemented into the existing *snpEPN* software, and a python script was provided for post-run analysis.

- New accuracy estimation methods closely aligned with SS-GBLUP benchmarks and were computationally efficient.
- The accuracy of GEBVs and Flock Profiles is strongly influenced by the size of the reference population, trait heritability, and genetic linkage to the reference.
- Accuracy levels varied widely across traits: traits with high heritability and large reference datasets (e.g., WWT, YFD) showed high accuracy, while traits with limited data (e.g., IMF, BWT) showed lower accuracy.
- A new four-category reporting system for accuracy was proposed for reporting to breeders.

Benefits to Industry

This project provides a validated, efficient method to estimate accuracy values for genomic predictions, enabling breeders and commercial producers to make more informed decisions. These accuracy estimates enhance confidence when using GEBVs and Flock Profiles, particularly for animals lacking pedigree and phenotype records.

Future Research and Recommendations

- Explore multi-trait extensions to improve accuracy for traits with limited reference data.
- Extend methodology for broader use across other species like beef cattle.

Table of contents

1. Background

The application of genomic information into national breeding programs has transformed the way to predict breeding values. In particular, genomic-only breeding values (GEBVs) allow genetic merit to be predicted for animals without phenotypic or pedigree records. This is especially valuable in commercial Merino flocks, where genotyping is often the only source of information available. MERINOSELECT now has a very large reference population for most of the key traits and thus provides and excellent platform to prediction GEBVs for commercial flocks.

In the MERINOSELECT genetic evaluation system, GEBVs are derived post-analysis via back-solving SNP effects estimated from single-step genomic BLUP (SS-GBLUP). These GEBVs are used to generate Flock Profiles—summarised breeding values for a representative group of animals within a commercial flock. However, unlike standard Australian Sheep Breeding Values (ASBVs), since they are calculated outside the main analysis neither GEBVs nor Flock Profiles currently include an estimate of accuracy.

The absence of accuracy information presents a major limitation. Industry experience and past research have shown that commercial flocks vary widely in their genetic linkage to the reference population used in MERINOSELECT. This affects the reliability of genomic predictions. Without accuracy estimates, breeders and commercial producers cannot distinguish between high-confidence and low-confidence genomic predictions, which may lead to misinformed selection or benchmarking decisions.

Previous investigations at AGBU identified potential methods to estimate the accuracy of GEBVs and Flock Profiles based on prediction error variance (PEV) and the prediction error covariance (PEC) of SNP effects. However, these approaches had not been fully developed or integrated into the routine Sheep Genetics evaluation pipeline.

This project was designed to close this gap. It aimed to:

- Review and develop methods for calculating the accuracy of GEBVs and Flock Profiles.
- Quantify the genetic connectedness between commercial flocks and the reference population.
- Implement and validate new algorithms within the existing SNP-based post-analysis system.
- Provide an efficient framework for reporting accuracy alongside genomic predictions.

The outcomes of this project are directly relevant to genetic service providers and commercial breeders. With the growing adoption of genomic tools in the sheep industry, the ability to report accuracy estimates for genotype-only predictions is crucial for Sheep Genetics products.

2. Objectives

The aim of this project was to develop and implement robust methods to estimate the accuracy of genomic-only breeding values (GEBVs) and Flock Profiles for the Sheep Genetics evaluation system. Specifically, the project aimed to:

- Develop and validate new methodologies to estimate accuracy for genomic only breeding values.

- Develop and validate new methodologies to estimate accuracy for Flock Profiles.
- Implement new methodologies into the routine evaluations for Sheep Genetics.
- Document and publish the new methodology.

These objectives were successfully achieved. The final outputs are ready for integration into routine evaluations within the Australian sheep industry.

3. Methodology

3.1 Derivation of Accuracy Estimation Methods

The project began with a comprehensive literature review covering methodologies related to:

- Prediction error variance (PEV) and prediction error covariance (PEC) for genomic predictions,
- Accuracy estimation for genomic-only breeding values (GEBVs),
- Group-level accuracy estimation (e.g., Flock Profiles),
- Genetic connectedness across populations

This review informed the derivation of new formulas to estimate the accuracy of GEBVs and Flock Profiles using SNP effect predictions from the existing *snpEPN* software. The methods rely on calculating the PEC matrix of SNP effects, from which the PEV of GEBVs and the PEV of flock means are derived.

For flock-level accuracy, additional statistical components were incorporated to account for the genomic relationships between sampled animals and the covariance structure of their prediction errors.

3.2 Algorithm Development and Software Implementation

The derived algorithms were implemented into:

- A revised version of AGBU's existing *snpEPN* program (Fortran-based), and
- A Python-based tool (*FPA.py*) to support post-analysis flexibility.

Both tools calculate accuracy estimates for:

- Individual animal GEBVs (using SNP-based PEV approximation),
- Flock Profiles (adjusted for within-flock covariance and genomic relatedness).

3.3 Validation approach

Validation was performed using four independent approaches:

1. SNP Effect Accuracy
   Accuracy estimates for individual SNP effects were computed to ensure they aligned with expectations based on trait reference sizes and heritability.

2. Comparison with SS-GBLUP PEV
   Accuracy estimates from the updated *snpEPN* were compared with those obtained directly from single-step GBLUP (SS-GBLUP) using a small test dataset. High correlations confirmed the validity of the approximation.

3. Random Genotype Simulation
   A simulation study assessed the sensitivity of accuracy estimates to random genotype values and varying flock sizes, providing insight into the accuracy level for random genotype and how flock structure impacts accuracy.

4. Genetic Linkage Metrics
   Genetic connectedness between test flocks and the reference population was quantified using PEV-based metrics (PEVD and coefficient of determination [CD]). These metrics were evaluated for their correlation with estimated accuracy.

3.4 Dataset Application and Industry Testing

Three datasets were used in the project:

- Small Validation Dataset
  Approximately 106,000 animals with about 38,500 genotypes and three traits (e.g. WWT, YGFW, ASL), allowing focused validation of algorithms.
- Industry Example Dataset (Oct 2024)
  Approximately 6.36 million animals, with about 524,000 genotypes and 996 test flocks. This dataset enabled cross-trait evaluation and benchmarking of accuracy estimates against genetic link metrics.
- Latest Full Industry Dataset (Apr 2025)
  Approximately 6.55 million animals with 613,000 genotypes and over 1,000 flocks across 61 traits. This was used to test software efficiency, evaluate accuracy distribution across traits, and prepare for routine deployment.

3.5 Reporting Framework

Given the variability in accuracy across traits and flocks, a four-category system was proposed for industry reporting:

- Not Reportable (below threshold),
- Low (≤ 0.59),
- Medium (0.60–0.79),
- High (≥ 0.80).

This classification aims to support interpretation by breeders and industry stakeholders. Technically these thresholds should vary across traits as they have different underlying heritabilities but this would potentially add to confusion for breeders.

4. Results

4.1 Accuracy Estimation Algorithms Successfully Developed

New algorithms to estimate the accuracy of genomic-only breeding values (GEBVs) and Flock Profiles were derived based on prediction error covariance (PEC) of SNP effects. These were implemented in AGBU's *snpEPN* software and a new Python-based tool, *FPA.py*. The algorithms account for:

- The variance of SNP effects,
- The structure of genomic relationships (GRM), and
- The prediction error covariance among individuals or flock samples.

The final formula for Flock Profile accuracy includes both the genetic relatedness between animals and their prediction error covariance—ensuring more realistic and robust group-level accuracy estimates than a simple average of individual accuracies.

4.2 Validation Results from Small Dataset

A small validation dataset (approximately 106,000 animals) was used to test algorithm correctness:

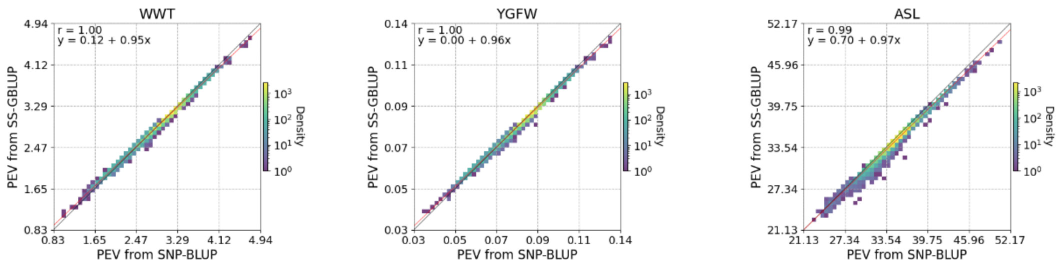
- PEV values derived from the updated *snpEPN* matched closely with those from SS-GBLUP, confirming the reliability of the new implementation (Figure 1).
- Accuracy of GEBVs was highest in Merino breeds, reflecting the size and relevance of the reference population. For less related breeds, accuracy was significantly lower.
- SNP effect accuracies were moderate to high depending on the trait (e.g. mean SNP accuracy of 0.66 for WWT; 0.08 for ASL), highlighting the role of trait heritability and reference size.
- 

Figure 1. PEV from SNP-BLUP (*snpEPN*) versus PEV from SS-GBLUP for weaning weight (WWT), yearling greasy fleece weight (YGFW), and adult staple length (ASL).

4.3 Simulation with Random Genotypes

Simulated datasets tested how flock size and structure affect Flock Profile accuracy under random SNP values:

- Accuracy ranged from 0.28 (flock size = 1) to 0.38 (flock size = 1,000), with minimal gain beyond ~15 animals.

- Cloned flocks yielded no additional accuracy benefit, confirming that genetic diversity within the flock sample is essential for meaningful accuracy.

4.4 Accuracy Across Industry Dataset

Using the October 2024 MERINOSELECT dataset (~6.36 million animals):

- High accuracy was observed for traits with large reference populations and moderate to high heritability (e.g. WWT, YWT, YFD) (Figure 2).
- Lower accuracy traits (e.g. IMF, BWT) showed greater variability, with GEBV accuracy in some flocks as low as ~0.4.
- Flock Profiles generally exhibited higher accuracy than individual GEBVs due to averaging, especially when genomic relationships and covariance were properly accounted for.

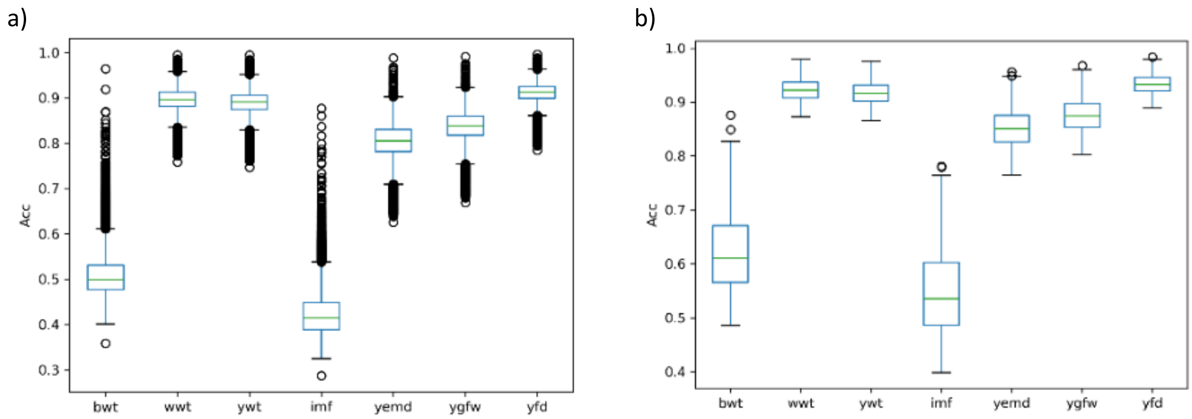


Figure 2. Accuracy distributions across traits for a) individual GBV; b) Flock Profiles.

4.5 Relationship Between Accuracy and Genetic Connectedness

- A strong correlation was observed between flock-level accuracy and genetic connectedness to the reference population.
    - Coefficient of determination (CD) between flocks and reference population had an r = 0.92–1.00 correlation with flock accuracy.
    - Flocks with poor genetic links to the reference showed lower accuracy even for traits with high heritability.
- This confirms that genetic connectedness to reference population is a key driver of accuracy for flock-level accuracy.

4.6 Implementation and Computational Performance

- The updated *snpEPN* software was applied to the April 2025 full industry dataset (~6.55 million animals, 61 traits).
- Runtime increased only marginally (by approximately 1 to 2 hours), demonstrating high computational efficiency and suitability for routine evaluations.
- Core values such as effective progeny number (EPN) remained unchanged, confirming that the update did not affect the existing pipelines.

4.7 Accuracy Reporting System

A four-category system was proposed to classify GEBV and Flock Profile accuracy in routine outputs:

- Not Reportable: Below trait-specific threshold or fixed threshold (e.g. <0.4),
- Low: ≤ 0.59,
- Medium: 0.60–0.79,
- High: ≥ 0.80.

This system provides a clear and interpretable framework for users, supporting transparency and informed decision-making. These thresholds might need to be review across all traits to ensure their suitability for lowly heritable traits like reproduction.

5. Conclusion

This project successfully addressed a major gap in the MERINOSELECT evaluation system by developing and validating new methodologies for estimating the accuracy of genomic-only breeding values (GEBVs) and Flock Profiles. The outcomes represent a significant advancement in the reliability and interpretability of genomic predictions for animals without phenotype or pedigree data.

By deriving accuracy estimates directly from the prediction error covariance (PEC) of SNP effects, and accounting for both genetic relationships and prediction error covariance among animals, the project has enabled the calculation of robust and meaningful accuracy metrics. These estimates were shown to align closely with established benchmarks from SS-GBLUP models and were implemented with minimal computational overhead.

Comprehensive validation using simulations, small test datasets, and full industry datasets confirmed that accuracy is strongly influenced by reference population size, trait heritability, and the degree of genetic connectedness between flocks and the reference. The proposed four-level accuracy reporting system provides a clear and interpretable tool for breeders to assess the reliability of genomic products.

The new methodology has been integrated into the snpEPN software and is ready for routine implementation in Sheep Genetics evaluations. This program is also used by BREEDPLAN and is also being further tested for use in BREEDPLAN Select.

5.1 Key Findings

- New algorithms were developed to estimate GEBV and Flock Profile accuracy based on SNP effect PEC.
- Accuracy of genomic predictions is strongly correlated with the size and quality of the reference population and the degree of genetic connectedness.
- The updated methods produce results consistent with SS-GBLUP benchmarks and add minimal computational time.
- A four-level accuracy classification system was developed to support industry interpretation and reporting.

5.2 Benefits to Industry

- Enables breeders and commercial producers to make informed decisions by providing accuracy metrics for GEBVs and Flock Profiles.
- Provides an efficient tool for implementation into routine Sheep Genetics pipelines.

6. Future Research and Recommendations

Several opportunities remain for further development and refinement.

- Extend the current single-trait methodology to account for genetic correlations among traits. This would improve accuracy predictions for traits with limited data or lower heritability by leveraging information from related traits.
- Extend methodology for broader use across other species like beef cattle.
- The proposed classification (Not Reportable, Low, Medium, High) should be incorporated into Sheep Genetics outputs to improve clarity and decision-making for users, subject to review across more traits.

7. References

Li L, Gurman PM, Swan AA, Brown DJ (2025). Estimates of Accuracy for Genomic-only Breeding Values and Flock Profiles in Australian Merino Sheep. Proceedings of the Association for the Advancement of Animal Breeding and Genetics 26, 205–208.

8. Appendix

8.1 Genomic accuracy for MERINOSELET analyses

The tables below summarise the estimated accuracies for genomic-only breeding values and Flock Profiles from the MERINOSELECT analysis conducted on 19 April 2026. The dataset included a total of 6,845,734 animals in the main run (MAIN). The genomic reference population comprised 739,670 genotyped animals and 61,238 SNP markers. The numbers of back-solved genotypes and flocks were 81,065 and 1288, respectively.

Table 1. Mean, standard deviation (Std), minimum (Min) and maximum (Max) accuracies for 81,065 genomic-only genotypes and 1288 flocks by traits, based on the MERINOSELECT main run. Results are presented for both additive genetic (ADD) and maternal genetic (MAT) effects.

			Genomic-only genotypes				Flock Profiles			
Ord	**Trait**	**Effect**	**Mean**	**Std**	**Min**	**Max**	**Mean**	**Std**	**Min**	**Max**
1	bwt	ADD	0.56	0.04	0.40	0.98	0.66	0.07	0.53	0.89
2	wwt	ADD	0.90	0.02	0.77	0.99	0.93	0.02	0.88	0.98
3	pwt	ADD	0.90	0.02	0.77	0.99	0.93	0.02	0.87	0.98
4	ywt	ADD	0.89	0.02	0.76	0.99	0.93	0.02	0.87	0.98
5	hwt	ADD	0.81	0.03	0.64	0.99	0.86	0.03	0.78	0.95
6	awt	ADD	0.62	0.04	0.44	0.97	0.72	0.06	0.59	0.88
7	cwt	ADD	0.42	0.05	0.29	0.88	0.55	0.08	0.39	0.78
8	wcf	ADD	0.13	0.03	0.08	0.53	0.23	0.07	0.11	0.59
9	pcf	ADD	0.70	0.05	0.52	0.97	0.77	0.05	0.66	0.95
10	ycf	ADD	0.79	0.04	0.61	0.99	0.84	0.04	0.75	0.96
11	hcf	ADD	0.56	0.04	0.39	0.95	0.66	0.06	0.52	0.83
12	cfat	ADD	0.37	0.04	0.25	0.85	0.50	0.09	0.33	0.76
13	ccfat	ADD	0.33	0.05	0.22	0.78	0.46	0.09	0.29	0.72
14	imf	ADD	0.46	0.04	0.32	0.91	0.58	0.08	0.43	0.81
15	sf5	ADD	0.31	0.05	0.20	0.76	0.44	0.09	0.27	0.70
16	wemd	ADD	0.14	0.03	0.08	0.56	0.24	0.07	0.12	0.60
17	pemd	ADD	0.72	0.05	0.54	0.97	0.79	0.05	0.68	0.95
18	yemd	ADD	0.82	0.03	0.66	0.99	0.87	0.03	0.79	0.97
19	hemd	ADD	0.54	0.04	0.37	0.95	0.64	0.06	0.50	0.82
20	cemd	ADD	0.29	0.04	0.19	0.73	0.43	0.09	0.26	0.69
21	pgfw	ADD	0.80	0.03	0.64	0.98	0.86	0.03	0.77	0.96
22	ygfw	ADD	0.86	0.03	0.71	0.99	0.90	0.03	0.83	0.98
23	hgfw	ADD	0.79	0.04	0.62	0.99	0.84	0.04	0.75	0.96
24	agfw	ADD	0.67	0.04	0.49	0.97	0.75	0.06	0.64	0.93
25	pcfw	ADD	0.70	0.04	0.52	0.97	0.78	0.05	0.66	0.93
26	ycfw	ADD	0.80	0.04	0.63	0.99	0.85	0.04	0.77	0.96
27	hcfw	ADD	0.67	0.04	0.50	0.98	0.75	0.05	0.64	0.94
28	acfw	ADD	0.49	0.05	0.34	0.97	0.61	0.07	0.45	0.88
29	pfd	ADD	0.86	0.03	0.71	0.99	0.90	0.03	0.84	0.98
30	yfd	ADD	0.93	0.02	0.83	0.99	0.95	0.01	0.91	0.99
31	hfd	ADD	0.83	0.03	0.68	0.99	0.88	0.03	0.79	0.98

32	afd	ADD	0.64	0.05	0.46	0.98	0.72	0.06	0.61	0.94
33	pdcv	ADD	0.76	0.04	0.58	0.97	0.82	0.04	0.72	0.95
34	ydcv	ADD	0.86	0.03	0.71	0.99	0.90	0.03	0.83	0.98
35	hdcv	ADD	0.70	0.04	0.52	0.98	0.78	0.05	0.67	0.94
36	adcv	ADD	0.51	0.05	0.35	0.95	0.62	0.07	0.47	0.90
37	pcuv	ADD	0.81	0.03	0.65	0.98	0.86	0.03	0.77	0.96
38	ycuv	ADD	0.88	0.03	0.75	0.99	0.92	0.02	0.85	0.98
39	hcuv	ADD	0.75	0.04	0.57	0.99	0.81	0.04	0.71	0.96
40	acuv	ADD	0.57	0.05	0.40	0.97	0.67	0.06	0.54	0.91
41	psl	ADD	0.75	0.04	0.58	0.98	0.81	0.04	0.72	0.96
42	ysl	ADD	0.82	0.03	0.66	0.99	0.87	0.03	0.79	0.97
43	hsl	ADD	0.69	0.04	0.51	0.99	0.77	0.05	0.66	0.92
44	asl	ADD	0.49	0.05	0.33	0.95	0.60	0.07	0.44	0.84
45	pss	ADD	0.49	0.05	0.33	0.92	0.60	0.07	0.45	0.87
46	yss	ADD	0.60	0.05	0.43	0.98	0.69	0.06	0.56	0.90
47	hss	ADD	0.46	0.05	0.30	0.91	0.57	0.08	0.41	0.83
48	ass	ADD	0.38	0.05	0.25	0.93	0.51	0.08	0.33	0.78
49	psc	ADD	0.40	0.05	0.27	0.90	0.51	0.09	0.36	0.86
50	ysc	ADD	0.58	0.05	0.40	0.95	0.67	0.07	0.53	0.90
51	hsc	ADD	0.24	0.06	0.16	0.85	0.34	0.09	0.21	0.75
52	dress	ADD	0.30	0.04	0.20	0.72	0.44	0.09	0.26	0.71
53	lmy	ADD	0.22	0.05	0.14	0.65	0.34	0.09	0.19	0.58
54	bwt	MAT	0.47	0.05	0.32	0.93	0.56	0.07	0.42	0.82
55	wwt	MAT	0.76	0.04	0.58	0.98	0.81	0.04	0.71	0.94
56	cwt	MAT	0.06	0.02	0.03	0.40	0.09	0.03	0.04	0.28
57	fw	MAT	0.41	0.05	0.27	0.81	0.51	0.07	0.37	0.78

Table 2. Mean, standard deviation (Std), minimum (Min) and maximum (Max) accuracies for 81,065 genomic-only genotypes and 1288 flocks by traits, based on the MERINOSELECT VISUAL sub-run. Results are presented for both additive genetic (ADD) and maternal genetic (MAT) effects.

			Genomic-only genotypes				**Flock Profiles**			
Ord	**Trait**	**Effect**	**Mean**	**Std**	**Min**	**Max**	**Mean**	**Std**	**Min**	**Max**
1	ebwr	ADD	0.90	0.02	0.77	0.99	0.93	0.02	0.86	0.98
2	lbwr	ADD	0.75	0.04	0.57	0.98	0.81	0.04	0.72	0.96
3	ebdwr	ADD	0.45	0.04	0.31	0.84	0.57	0.07	0.42	0.82
4	lbdwr	ADD	0.71	0.04	0.53	0.98	0.78	0.05	0.67	0.94
5	ebcov	ADD	0.85	0.03	0.71	0.99	0.89	0.03	0.82	0.97
6	lbcov	ADD	0.64	0.05	0.47	0.94	0.73	0.06	0.61	0.94
7	edag	ADD	0.51	0.05	0.35	0.96	0.62	0.07	0.47	0.88
8	ldag	ADD	0.77	0.04	0.60	0.99	0.83	0.04	0.74	0.95
9	ecol	ADD	0.75	0.04	0.58	0.99	0.82	0.04	0.72	0.95
10	lcol	ADD	0.56	0.05	0.40	0.95	0.66	0.07	0.53	0.86
11	efrot	ADD	0.69	0.04	0.51	0.99	0.77	0.05	0.66	0.92
12	lfrot	ADD	0.50	0.05	0.34	0.93	0.60	0.07	0.46	0.84
13	echar	ADD	0.73	0.04	0.56	0.99	0.80	0.04	0.70	0.94
14	lchar	ADD	0.56	0.05	0.39	0.95	0.65	0.07	0.52	0.87

15	ebwr	MAT	0.51	0.05	0.34	0.91	0.61	0.07	0.47	0.85
16	lbwr	MAT	0.27	0.04	0.17	0.63	0.37	0.08	0.23	0.72
17	ebdwr	MAT	0.13	0.03	0.08	0.52	0.19	0.06	0.11	0.43
18	lbdwr	MAT	0.22	0.04	0.14	0.57	0.31	0.07	0.19	0.55
19	ebcov	MAT	0.47	0.05	0.31	0.85	0.58	0.07	0.43	0.83
20	edag	MAT	0.31	0.05	0.19	0.76	0.41	0.08	0.27	0.77
21	efrot	MAT	0.30	0.05	0.19	0.75	0.40	0.08	0.27	0.73
22	lfrot	MAT	0.10	0.03	0.06	0.40	0.16	0.05	0.08	0.36

Table 3. Mean, standard deviation (Std), minimum (Min) and maximum (Max) accuracies for 81,065 genomic-only genotypes and 1288 flocks by traits, based on the MERINOSELECT COMPREPRO sub-run. Results are presented for both additive genetic (ADD) effect.

			Genomic-only genotypes				Flock Profiles			
Ord	**Trait**	**Effect**	**Mean**	**Std**	**Min**	**Max**	**Mean**	**Std**	**Min**	**Max**
1	awt	ADD	0.75	0.04	0.58	0.99	0.82	0.04	0.72	0.94
2	pscanwt	ADD	0.72	0.05	0.55	0.98	0.79	0.05	0.68	0.95
3	pcf	ADD	0.67	0.05	0.49	0.97	0.75	0.06	0.63	0.94
4	pemd	ADD	0.69	0.05	0.51	0.97	0.76	0.05	0.65	0.94
5	yscanwt	ADD	0.82	0.03	0.66	0.99	0.87	0.03	0.79	0.97
6	ycf	ADD	0.77	0.04	0.59	0.99	0.83	0.04	0.73	0.96
7	yemd	ADD	0.81	0.04	0.64	0.99	0.86	0.03	0.78	0.97
8	hscanwt	ADD	0.58	0.05	0.40	0.95	0.66	0.06	0.51	0.84
9	hcf	ADD	0.53	0.05	0.36	0.95	0.63	0.06	0.48	0.82
10	hemd	ADD	0.54	0.05	0.38	0.95	0.64	0.06	0.50	0.83
11	psc	ADD	0.41	0.05	0.28	0.91	0.52	0.09	0.37	0.86
12	ysc	ADD	0.57	0.05	0.39	0.95	0.66	0.07	0.52	0.89
13	mbs	ADD	0.26	0.05	0.17	0.82	0.38	0.09	0.20	0.78
14	cs	ADD	0.52	0.05	0.35	0.94	0.63	0.07	0.49	0.88
15	ycon	ADD	0.46	0.06	0.31	0.91	0.56	0.08	0.42	0.86
16	con	ADD	0.47	0.05	0.31	0.89	0.60	0.08	0.43	0.86
17	yls	ADD	0.24	0.05	0.16	0.74	0.34	0.08	0.21	0.71
18	ls	ADD	0.53	0.05	0.36	0.90	0.65	0.07	0.49	0.87
19	era	ADD	0.41	0.05	0.26	0.82	0.55	0.08	0.37	0.82

Table 4. Mean, standard deviation (Std), minimum (Min) and maximum (Max) accuracies for 81,065 genomic-only genotypes and 1288 flocks by traits, based on the MERINOSELECT WEC sub-run. Results are presented for both additive genetic (ADD) effect.

			Genomic-only genotypes				Flock Profiles			
Ord	**Trait**	**Effect**	**Mean**	**Std**	**Min**	**Max**	**Mean**	**Std**	**Min**	**Max**
1	Wwec	ADD	0.49	0.05	0.35	0.89	0.61	0.07	0.47	0.85
2	Pwec	ADD	0.58	0.05	0.41	0.95	0.68	0.07	0.54	0.89
3	Ywec	ADD	0.53	0.05	0.36	0.96	0.63	0.07	0.48	0.90
4	Hwec	ADD	0.34	0.05	0.22	0.89	0.47	0.09	0.29	0.76

8.2 AAABG Paper

ESTIMATES OF ACCURACY FOR GENOMIC-ONLY BREEDING VALUES AND FLOCK PROFILES IN AUSTRALIAN MERINO SHEEP

L. Li, P.M. Gurman, A.A. Swan and D.J. Brown

Animal Genetics Breeding Unit*, University of New England, Armidale, 2351, Australia

SUMMARY

As uptake of genomic-only estimated breeding values (GEBVs) and Flock Profiles – defined as the average GEBV from a random sample of approximately 20 animals from a commercial flock – expands within the Australian sheep industry, the need for accurate and reliable results becomes increasingly critical. Currently, GEBVs are presented without accuracy estimates, making it difficult to judge the genomic connectedness of the genotype to the reference population. This study aims to develop and implement a robust methodology for estimating the accuracy of GEBVs and Flock Profiles. Building on existing genetic evaluation accuracy frameworks, we propose a new method that can be integrated into routine evaluations, providing users with confidence in the reliability of their results.

INTRODUCTION

Single-step genomic BLUP (SS-GBLUP) combines pedigree, phenotypic, and genomic information and is routinely used in genetic evaluations for the Australian sheep and beef cattle industries (Brown *et al.* 2018; Johnston *et al.* 2018). One of the key advantages of genomic selection is its ability to calculate genomic-only estimated breeding values (GEBVs) for individuals without phenotypes and pedigree information. This application has been implemented in the Australian Merino sheep industry by back-solving SNP effects from SS-GBLUP EBVs, and then calculating GEBVs for animals with only genotype data (Swan *et al.* 2018). Currently, the resulting GEBVs for the genotype-only animals are then averaged to establish a benchmark for commercial flocks, referred to as Flock Profiles, using a random sample of approximately 20 animals from a flock.

However, this process does not estimate the associated accuracy for these predictions. Some flocks are genetically more different from the genotypes of the reference population in MERINOSELECT, which is expected to result in GEBVs of lower accuracy. It is crucial for the interpretation of GEBVs that predictions are sufficiently accurate and, like ASBVs, are only reported if they reach reporting thresholds for accuracy. Additionally, having an accuracy estimate attached to Flock Profiles and individual animal GEBVs is vital as these products are further adopted. This study aimed to derive a methodology for approximating the prediction error variance (PEV) and accuracy for both GEBVs and Flock Profiles. Initially, the methods were validated using a small dataset by comparing the approximated PEV with true PEV. Subsequently, these methodologies were applied to recent data from the Merino sheep industry to assess the accuracy of both GEBVs and Flock Profiles for a subset of traits.

MATERIALS AND METHODS

Derivation of Accuracy for GEBVs. Single-trait SNP-BLUP models can approximate accuracies for the genomic component of the single-step BLUP model (Li *et al.* 2023). In short, consider a single trait SNP-BLUP model with

* A joint venture of NSW Department of Primary Industries and Regional Development and the University of New England

a mixed model equation (MME) $\begin{bmatrix} \mathbf{X'X} & \mathbf{X'ZW} \\ \boldsymbol{W'}\mathbf{Z'X} & \boldsymbol{W'Z'}\mathbf{ZW}+\mathbf{I}\lambda \end{bmatrix}\begin{bmatrix} \boldsymbol{b} \\ \boldsymbol{s} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X'}y \\ \boldsymbol{W'Z'}y \end{bmatrix}$, where **y** is the vector of observations, **b** is the vector of fixed effects, **s** is the vector of estimated SNP effects, **X** and **Z** are incidence matrices which map observations to fixed effects and individual animals' breeding values, respectively, $\boldsymbol{W}$ is the animals by markers matrix of centred and scaled marker genotypes and $\lambda = \sigma_e^2/\sigma_s^2$, i.e. the ratio of the residual variance and the marker variance. The prediction error covariance (PEC) for SNP effect, $\boldsymbol{C^{22}}$, can be obtained by inverting the left-hand side of MME, which by block matrix inversion rules can be written as $\mathbf{C^{22}} = [\boldsymbol{W'}\mathbf{Z'}(\mathbf{I}-\mathbf{P})\mathbf{Z}\boldsymbol{W} + \mathbf{I}\boldsymbol{\lambda}]^{-1}$, where $\mathbf{P} = \mathbf{X}(\mathbf{X'X})^{-1}\mathbf{X'}$.

The PEC for GEBVs for all animals can therefore be calculated as $\boldsymbol{PEC_{GEBVs}} = \boldsymbol{WC^{22}W'}\sigma_e^2$. In practice, only the diagonal elements of the PEC are required for individual animal calculations. After the derivation of PEC, the GEBV accuracy of individual animal can be calculated as $\mathbf{Acc} = \sqrt{\mathbf{1} - \frac{\mathbf{diag}(\boldsymbol{PEC_{GEBVs}})}{diag(\boldsymbol{WW'})\sigma_a^2}}$, where $diag(\boldsymbol{WW'})$ are the diagonal elements of the genomic relationship matrix and σ_a^2 is the additive genetic variance. It is worth noting that the PEV for GEBVs calculated by this algorithm used SNP information only. We consider these PEV as approximations for the PEV of EBV of these genotype-only animals if they are derived in the routine SS-GBLUP analysis directly. This model utilizes an H matrix with a lambda value to integrate genomic and pedigree relationship matrices to derive EBVs (Brown *et al.* 2018).

Derivation of PEV and Accuracy for Flock Profiles. The above formula for an individual animal's accuracy can be extended to describe the accuracy for a group of animals. Suppose the flock profile contains genotypes from *n* individual animals, with the centred and scaled matrix $\boldsymbol{W}_{FP}$ obtained as the subset of $\boldsymbol{W}$ for the animals in the profile. The PEC for the animals is $PEC_{FP} = \boldsymbol{W}_{FP}\boldsymbol{C}^{22}\boldsymbol{W}'_{FP}\sigma_e^2$. If we define a vector $\boldsymbol{v} = \frac{1}{n}\mathbf{1'}$ as an n-dimensional column vector with each weight as 1/n, then the accuracy for the flock profile is $Acc_{FP} = \sqrt{1 - \frac{\boldsymbol{v'PEC_{FP}}\mathbf{v}}{\boldsymbol{v'}(\boldsymbol{W_{FP}W'}_{FP})\mathbf{v}\sigma_a^2}}$. These calculations have been incorporated into the current snpEPN genomic accuracy software (Li *et al.* 2023).

Validation. A small test dataset from the MERINOSELECT analysis was extracted to validate the methodology as a subset of flocks and genotypes. The validation aimed to confirm that our PEV estimates are close to true values and achieve higher accuracy for Merino genotypes versus lower accuracy for non-Merino genotypes. This dataset comprised 106,000 animals, with 38,492 having genotypes for 61,146 SNPs. Approximately half of the genotyped animals were extracted from Merino flocks designated as the reference population, while the remaining genotypes were added from non-Merino flocks, which were expected to be genetically highly different from these reference genotypes for testing purposes. Three traits with different reference sizes were investigated: weaning weight (WWT, 15,279), yearling greasy fleece weight (YGFW, 8,379), and adult staple length (ASL, 1,186). Genetic parameters used in this validation were extracted from the MERINOSELECT with heritability of 0.29, 0.34 and 0.34 for WWT, YGFW and ASL, respectively. In this validation, the PEV of GEBVs from snpEPN were first compared with the PEV based on SS-GBLUP model (SS-PEV) with assumption that these genotype-only animals were included in the routine analysis. SS-PEV were calculated directly from the inverse of mixed model equations using an H relationship matrix with a lambda value of 0.95 to combine genomic and pedigree relationship matrices. Then, the accuracy of GEBVs was calculated and compared by breeds to demonstrate the influence of reference breeds on predictive accuracy.

Example data. The described methods were also applied to a MERINOSELECT analysis extracted from the October 1, 2024 run, including 6,359,892 animals, of which 523,845 were genotyped for 61,260 SNPs. Additionally, 78,852 genotypes were back-solved, including 27,670 genotypes from 996 test flocks. Traits investigated included birth weight (BWT), weaning weight (WWT), yearling weight (YWT), intramuscular fat (IMF), yearling eye muscle depth (YEMD), yearling greasy fleece weight (YGFW), and yearling fleece diameter (YFD). Heritability values used for accuracy approximation, along with the count of animals having both genotypes and phenotypes (Nref) ranging from 7,357 for IMF to 396,307 for WWT for each trait, are detailed in Table 1.

Table 1. Heritability (h^2) and number of animals with both genotypes and phenotypes (Nref) for live weight, carcass and fleece traits extracted from the MERINOSELECT analysis

Trait	BWT	WWT	YWT	IMF	YEMD	YGFW	YFD
h^2	0.22	0.34	0.41	0.52	0.31	0.34	0.58
Nref	29,775	396,307	273,915	7,357	165,088	197,280	212,179

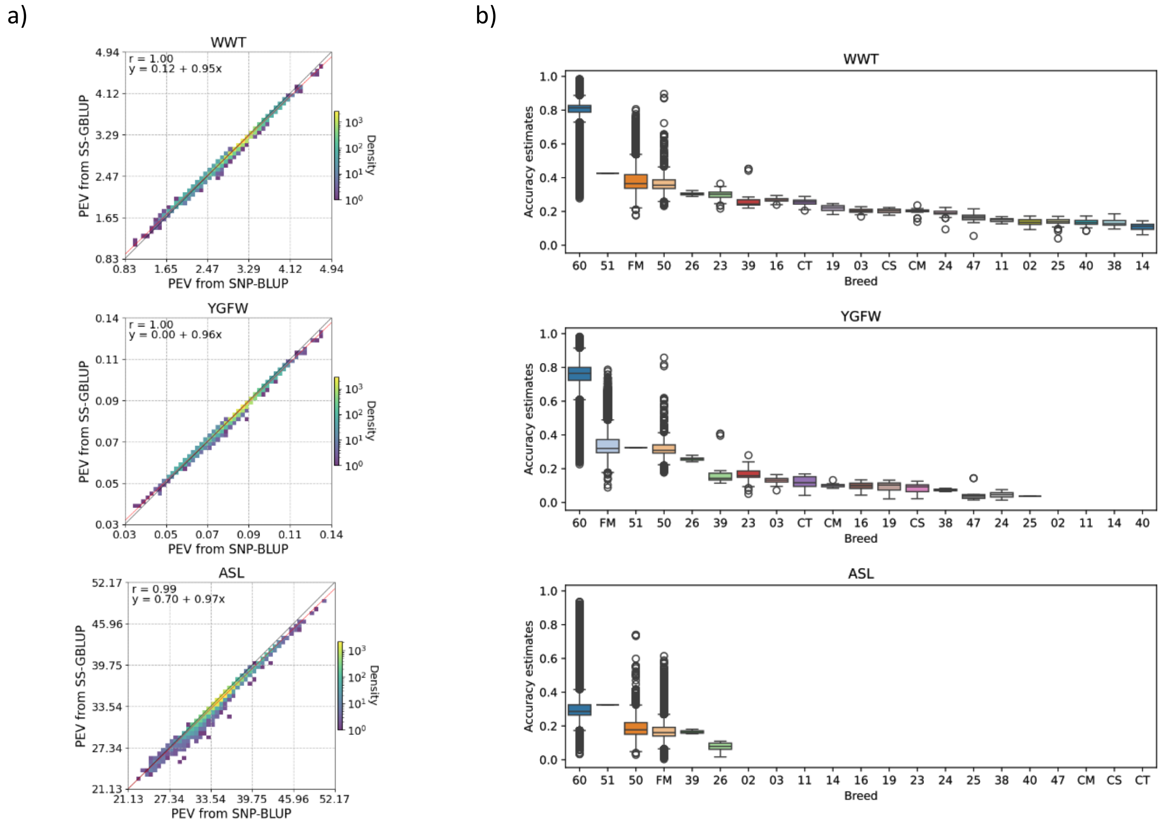


Figure 1. a) PEV from SNP-BLUP versus PEV from SS-GBLUP; b) Accuracy estimates for all animals by breeds for weaning weight (WWT), yearling greasy fleece weight (YGFW), and adult staple length (ASL). Note: Corresponding breed names for the codes listed in b) can be found at: https://www.sheepgenetics.org.au/resources/breeders-guide/animal-identification/.

RESULTS AND DISCUSSION

Validation results. The PEV of GEBVs for genotyped animals estimated from SNP-BLUP was very close to the PEV based on SS-GBLUP model, with correlations near 1 for all traits (Fig. 1a), indicating that the algorithms implemented were functioning as expected. The accuracy of EBVs for all animals across traits by breeds are shown in Fig. 1b. The Merino breed (60) consistently displayed higher accuracy for all traits, with a wider spread indicating variability within the breed. Other breeds less related to the reference flocks generally presented much lower accuracies, especially for YGFW and ASL, where most breeds had accuracy less than 0.2 (YGFW) and close to zero (ASL). This may be due to the reference animals predominantly being from Merino breeds, with the reference size reducing from 15,279 for WWT to 1,186 for ASL.

Example data results. Figure 2 shows two boxplots illustrating the accuracy of GEBVs and Flock Profiles across all traits. The GEBVs show a wide range of accuracies, with BWT and IMF showing much lower accuracy (~0.4-0.5) and greater variability. In contrast, other traits display higher accuracies (>0.8) with narrower spreads, indicating more accurate and consistent predictions. These results correspond to the size of the reference population, as detailed in Table 1. Not surprisingly, the accuracy of Flock Profiles (Figure 2b) presents a pattern like GEBVs but with generally higher and more uniform accuracies. These high accuracy levels for both GEBVs of genotype-only animals and Flock Profiles are highly promising for the implementation of genomic selection in the Australia Merino sheep industry. However, these high accuracy estimates require further validation. This could be achieved, for instance, by comparing the accuracy of testing flocks with the linkage metrices between genomic reference populations and testing flocks.

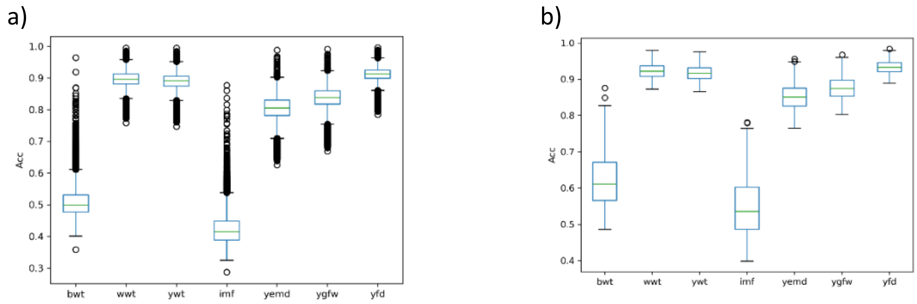


Figure 2. Accuracy by traits for a) GEBVs; b) Flock Profiles

CONCLUSIONS

A methodology for estimating the accuracy of GEBVs and Flock Profiles has been developed and incorporated into the current snpEPN program. This methodology was validated using a small dataset showing high correlations between PEV derived from SNP-BLUP and PEV from SS-GBLUP. Very high accuracy (>0.8) was observed for both GEBVs and Flock Profiles within a real Merino sheep dataset for traits with substantial-sized reference populations, such as WWT, YWT, YEMD, YGFW and YFD. However, lower accuracy (0.4-0.5) was noted for traits with smaller reference populations, such as BWT and IMF. By providing accuracy estimates for GEBVs and Flock Profiles, the Merino sheep industry can benefit from a more transparent and reliable genetic evaluation system. This methodology can be implemented in other sheep breeds and beef cattle.

ACKNOWLEDGEMENTS

This research was funded by Meat and Livestock Australia project L.GEN.2501. The authors acknowledge the contribution of data from Australian sheep breeders.

REFERENCES

Brown D.J., Swan A.A., Boerner V., Li L., Gurman P.M., McMillan A.J., van der Werf J.H.J., Chandler H.R., Tier B. and Banks R.G. (2018) *Proc. of the 11th World Cong. on Genet. App. to Livest. Prod.* **11:** 460.

Johnston D.J., Ferdosi M.H., Connors N.K., Boerner V., Cook J., Girard C.J., Swan A.A. and Tier B. (2018) *Proc. of the 11th World Cong. on Genet. App. to Livest. Prod.* **11:** 269.

Li L., Gurman P., Swan A.A. and Tier B. (2023) *Anim. Prod. Sci*. **63:** 1086.

Swan A.A., Gurman P.M., Boerner V., Brown D.J., Clark S., Gore K., Granleese, T. and van der Werf J.H.J. (2018) *Proc. of the 11th World Cong. on Genet. App. to Livest. Prod.* **11:** 1-6.