# Cross-validation of single step BLUP applied to terminal sire sheep in Australia

*P.M. Gurman[1], A.A. Swan[1], V. Boerner[1], D.J. Brown[1]*

*[1] Animal Genetics and Breeding Unit, University of New England, Armidale, 2351, Australia*
*PGurman@une.edu.au (Corresponding Author)*

## Summary

To inform decision making regarding the implementation of single step for routine genetic evaluations in sheep in Australia, forward cross-validation and 5-fold cross-validation schemes were used to compare single-step genomic BLUP to pedigree BLUP (hereby referred to as SS-GBLUP and ABLUP, respectively). Accuracies and slopes of EBVs regressed against phenotypes were used to validate the two methods used. Accuracies for genotyped animals from forward cross-validation were, on average across traits, 0.073 higher for SS-GBLUP compared to ABLUP and were 0.092 higher calculated using 5-fold. Slopes for genotyped animals from forward validation were, on average across all traits for reference flock animals 0.92 for SS-GBLUP EBVs and 0.80 for ABLUP EBVs, whereas for 5-fold cross-validation the respective slopes where 0.93 and 0.89. Greater increases in accuracies were observed in genotyped sheep with these increases also observed in hard to measure traits, including worm egg counts and carcase traits.

*Keywords: cross-validation, single-step, sheep, Australia*

## Introduction

Single-step genomic BLUP (Legarra *et al.*, 2014) is increasingly being adopted for routine genomic evaluation (e.g. Australian Brahman BREEDPLAN) because it allows animals with pedigree only and animals with pedigree and genotypes to be evaluated in a single BLUP analysis. The Australian sheep industry transitioned to single-step genomic BLUP, herein SS-GBLUP, evaluation in 2016 and 2017 for routine genomic evaluation (Brown *et al.*, 2018). To inform decision-making around this change, accuracy and bias metrics were required to compare SS-GBLUP and pedigree-only BLUP, herein ABLUP. Previous research has calculated these metrics for Australian sheep using combinations of ABLUP, genomic BLUP, BayesA and BayesR (Daetwyler *et al.*, 2012; Moghaddar *et al.*, 2014; Swan *et al.*, 2014), but not for SS-GBLUP. The aim of this work was to determine the accuracy and bias of EBVs from SS-GBLUP relative to ABLUP across a range of traits in terminal sire sheep breeds.

## Methods

Forward cross-validation and 5-fold cross-validation schemes were used to compare the ability of SS-GBLUP and ABLUP to estimate breeding values for records on 2.7 million sheep across 24 traits. These traits covered body weight at different age stages from birth to adult, carcase lean meat yield, eating quality traits including intramuscular fat and shear force, eye muscle and fat depths in both live animals and carcases, and worm egg count. For forward cross-validation, the validation set of animals were those born during or after 2014,

including animals with and without genotypes, though results are presented here only for genotyped animals. These animals had their phenotypes removed before breeding values were estimated for all sheep through either SS-GBLUP or ABLUP using multi-trait models. For the 5-fold cross-validation, sets of animals were constructed based on a genomic reference population comprising of animals from information nucleus flock animals and MLA reference animals designed primarily for carcase and eating quality traits (van der Werf *et al.*, 2010), hereby referred to as reference flock animals. Animals within the population were assigned to one of five validation groups based on sire family, with progeny of a sire always assigned to the same validation group. Five datasets were then constructed based on these groups, where phenotypes were removed for animals in that validation group. Breeding values were then estimated for each dataset through either multi-trait SS-GBLUP or ABLUP and the EBVs for the validation animals were collated from their respective runs. Only 14 traits were included in the 5-fold analyses because of the extra computational requirements (note that worm egg count was not included in the 5-fold analyses). All analyses were multi-trait and the SS-GBLUP analyses were performed using the equations of Aguilar *et al.* (2010), with a lambda value of 0.5 based on McMillan & Swan (2017), resulting in equal weighting being placed on the pedigree and genomic relationships for animals with both recorded.

Data used in the analyses were pre-corrected for standard fixed effects such as age of dam, birth-rearing type, and age of measurement, and a fixed effect for contemporary group was included directly. Trait values used in validation statistics were corrected for contemporary group. For each trait, the pre-corrected records were regressed against the contemporary group, with the residuals from this analysis retained as the adjusted trait values ($y^*$).

Metrics of predictive performance were then computed for animals in the validation dataset. Accuracies of the EBV predictions for each trait were calculated following the method described by Legarra *et al.*, (2008), as the correlation between $y^*$ and EBV divided by the square root of the trait heritability. Regression slopes were calculated by predicting the trait value from the EBV, with slopes of approximately one showing zero bias in the EBVs, slopes lower than one indicating overestimation of the EBV, i.e. increased variance in EBVs from expectation, and slopes greater than one indicating underestimation of the EBV i.e. decreased variance in EBVs from expectation.

While the 5-fold cross-validation comparisons focused on reference flock animals only, the forward cross-validation comparisons included a mixture of the most recent reference flock animals along with animals genotyped by breeders in industry flocks.

## Results

### Accuracies

The average SS-GBLUP accuracy calculated from forward cross-validation across industry and reference flock \ animals across all traits was 0.30, with the mean increase from ABLUP to SS-GBLUP, 0.073. The average SS-GBLUP accuracy calculated from 5-fold cross-validation across all traits was 0.48, with the increase in accuracy from ABLUP to SS-GBLUP, 0.092.

Accuracies for animals split by validation method (forward vs 5-fold) are presented in

Table 1. In this table, accuracies are reported, averaged across trait group. 'Body Weight' included weights recorded at a range of production stages; 'LMY and Eating Quality' included lean meat yield, shear force, dress and intramuscular fat; 'Eye Muscle and Fat' includes eye muscle depth and fat recorded at a range of production stages; and 'WEC' includes worm egg counts recorded at multiple stages of production. For the forward cross-validation, the mean accuracy calculated from SS-GBLUP EBVs for reference flock animals across all traits was 0.41. These mean accuracies are greater than those estimated from ABLUP EBVs, with the mean increase across all traits for the reference flock animals 0.097. For the 5-fold cross-validation method, the mean accuracy calculated for reference flock animals across all traits from SS-GBLUP EBVs was 0.48, which is 0.07 higher than the accuracy estimated using the forward cross-validation method across all traits. The mean increase in accuracy from ABLUP to SS-GBLUP across all traits was 0.092. Validation was also performed on industry animals, though low accuracies were estimated for industry animals.

*Table 1: Estimated accuracies for reference flock animals calculated from EBVs calculated using SS-GBLUP and ABLUP averaged for each trait group. Results are presented using forward validation and 5-fold cross-validation. The mean number of animals in the validation group with trait observations across all traits in the trait group are included as 'Mean Sheep'.*

| Validation Method | Trait Group | Mean Sheep | SS-GBLUP Accuracies | | | Accuracy increases from ABLUP | | |
|---|---|---|---|---|---|---|---|---|
| | | | Mean | Min | Max | Mean | Min | Max |
| **Forward** | Body weight | 1257 | 0.55 | 0.34 | 0.92 | 0.16 | 0.07 | 0.39 |
| | EMD & Fat | 1407 | 0.40 | 0.32 | 0.54 | 0.04 | 0.02 | 0.07 |
| | LMY & EQ | 807 | 0.37 | 0.26 | 0.49 | 0.08 | -0.02 | 0.18 |
| | WEC | 411 | 0.19 | 0.09 | 0.29 | 0.12 | 0.02 | 0.22 |
| **5 Fold** | Body weight | 6019 | 0.66 | 0.41 | 0.88 | 0.17 | 0.07 | 0.34 |
| | EMD & Fat | 4197 | 0.43 | 0.25 | 0.56 | 0.05 | 0.02 | 0.08 |
| | LMY & EQ | 5809 | 0.38 | 0.37 | 0.39 | 0.09 | 0.07 | 0.11 |

**Regression Slopes**

Slopes estimated for SS-GBLUP or ABLUP EBVs of reference animals, split by validation method are presented in Table 2, with the trait groupings the same as in Table 1. Generally, slopes were closer to one for EBVs calculated using SS-GBLUP compared to ABLUP. For forward validation the average slope across all traits estimated for reference flock animals from SS-GBLUP EBVs was 0.92 and was 0.80 from ABLUP EBVs, whereas for 5-fold cross-validation the respective slopes where 0.93 and 0.89. Slopes estimated for industry animals were low, consistent with the low accuracies for the same group of animals.

*Table 2: Estimated slopes (i.e. bias) calculated from EBVs from SS-GBLUP and ABLUP for each trait group compared between forward and 5-fold cross-validation methods. The mean number of animals with trait observations across all traits in the trait group are included as 'Mean Sheep'.*

| Validation Method | Trait Group | Mean sheep | SS-GBLUP Slopes | | | ABLUP Slopes | | |
|---|---|---|---|---|---|---|---|---|
| | | | Mean | Min | Max | Mean | Min | Max |
| **Forward** | Body weight | 1257 | 1.05 | 0.44 | 2.32 | 0.84 | 0.37 | 1.67 |
| | EMD & Fat | 1407 | 0.91 | 0.60 | 1.11 | 0.98 | 0.72 | 1.16 |
| | LMY & EQ | 807 | 0.85 | 0.75 | 0.97 | 0.79 | 0.41 | 1.05 |
| | WEC | 411 | 0.73 | 0.33 | 1.13 | 0.30 | 0.28 | 0.31 |
| **5 Fold** | Body weight | 6019 | 1.17 | 0.50 | 1.97 | 1.05 | 0.51 | 1.47 |
| | EMD & Fat | 4197 | 0.84 | 0.43 | 1.07 | 0.86 | 0.49 | 1.01 |
| | LMY & EQ | 5809 | 0.81 | 0.69 | 0.92 | 0.74 | 0.57 | 0.84 |

## Discussion

Accuracies and bias estimates calculated from SS-GBLUP EBVs were better than those estimated from ABLUP EBVs. Increases in accuracy between SS-GBLUP and ABLUP were observed for genotyped animals for all trait groups, including hard to measure traits such as eating quality, carcase eye muscle depth and fat, and worm egg count. This indicates that SS-GBLUP can be used to provide more accurate breeding values for animals for traits that cannot be readily recorded or where the recording of that trait is destructive. SS-GBLUP can also be used to increase accuracy for "easy to measure" traits.

Accuracies estimated in Table 1 using forward and 5-fold cross-validation for the eye muscle depth and fat trait group are generally similar to, or greater than those reported previously for eye muscle depth by GBLUP (Daetwyler *et al.*, 2012). In comparison to previous results presented by Swan *et al.*, (2014), body weight trait accuracies are higher, eye muscle depth and fat traits have similar accuracies, while WEC accuracies estimated here are lower than the 'pwec' reported by Swan *et al.*, (2014). In this analysis, multiple WEC traits are included, measured at different stages of production, unlike the study of Swan *et al.*, (2014), which included only a single stage of production.

Difficulties were observed in validation of EBVs for genotyped animals in industry flocks, largely because of a strong tendency by breeders to selectively genotype and phenotype these animals. It is expected that as the number of genotypes increases for industry animals, validation for these animals will become more accurate.

Higher accuracies were observed when calculated using the 5-fold method. This may have been caused by the use of reference flock animals only in the 5-fold validation sets. This may also have been as a result of the splitting the validation set between five different analyses. When less animals are dropped in each analysis, the impact on the estimated accuracy will be lower. Genomic linkage between the validation and training sets will also be different between forward and 5-fold cross-validation, with fewer animals removed from the analysis.

## Conclusions
Single-step provides improved predictive ability compared to pedigree-alone methods for estimating breeding values. These results provide industry decision-makers with evidence of improved accuracy through the use of single step compared to pedigree methods alone and highlights that the benefits for genotyped animals are applicable across all traits.

## Acknowledgements

## List of References

Aguilar, I., I. Misztal, D.L. Johnson, A. Legarra, S. Tsuruta, & T.J. Lawlor. 2010. Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. J. Dairy Sci., 93(2): 743–752.

Brown, D.J., A.A. Swan, V. Boerner, L. Li, P.M. Gurman, A. McMillan, & J.H.J. van der Werf. 2018. Single Step Genetic Evaluations in the Australian Sheep Industry. 11th World Congr. Genet. Appl. Livest. Prod.

Daetwyler, H.D., A.A. Swan, J.H. van der Werf, & B.J. Hayes. 2012. Accuracy of pedigree and genomic predictions of carcass and novel meat quality traits in multi-breed sheep data assessed by cross-validation. Genet. Sel. Evol., 44: 33.

Legarra, A., O.F. Christensen, I. Aguilar, & I. Misztal. 2014. Single Step, a general approach for genomic selection. Livest. Sci., Genomics Applied to Livestock Production, 166: 54–65.

Legarra, A., C. Robert-Granié, E. Manfredi, & J.-M. Elsen. 2008. Performance of Genomic Selection in Mice. Genetics, 180(1): 611–618.

McMillan, A.J., & A.A. Swan. 2017. Weighting Of Genomic And Pedigree Relationships In Single Step Evaluation Of Carcass Traits In Australian Sheep. Proc. Assoc. Adv. Anim. Breed. Genet. 22: 557–560.

Moghaddar, N., A.A. Swan, & J.H.J. van der Werf. 2014. Genomic prediction of weight and wool traits in a multi-breed sheep population. Anim. Prod. Sci., 54(5): 544–549.

Swan, A., D. Brown, H.D. Daetwyler, B. J Hayes, M. Kelly, N. Moghaddar, & J. H J Van Der Werf. 2014. Genomic Evaluations in the Australian Sheep Industry. 10th World Congr. Genet. Appl. Livest. Prod. Vancouver, BC, Canada.

van der Werf, J.H.J., B.P. Kinghorn, & R.G. Banks. 2010. Design and role of an information nucleus in sheep breeding programs. Anim. Prod. Sci., 50(12): 998–1003.