

BREEDPLAN SINGLE-STEP GENOMIC EVALUATIONS DELIVERS INCREASED ACCURACIES ACROSS ALL BREEDS AND EBVS

D.J. Johnston, M.H. Ferdosi, N.K. Connors, J. Cook, C.J. Girard and A.A. Swan

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

SUMMARY

Forward cross-validation analyses were used to quantify the changes in BREEDPLAN EBVs from single-step genetic evaluations compared to traditional pedigree-based evaluations for Angus, Brahman, Hereford, Santa Gertrudis and Wagyu breeds. EBVs were generated from full multi-trait evaluations for each breed and compared to EBVs from an evaluation where all the phenotypic records were removed from the last four year drops of animals (termed Validation). Results for the sub-set of validation animals that were SNP genotyped showed the population-based accuracy of single-step EBVs were higher than pedigree-based accuracies for all breeds and traits. However, the magnitudes of the accuracy increases differed across breeds and traits, and generally reflected differences in the size of the training populations for each trait. The largest increase in accuracy, averaged across all traits in a breed, was observed for Angus (24%) and the smallest for Santa Gertrudis (5%). Across breeds, the largest increases in accuracy occurred for the growth trait EBVs compared to smaller increases for abattoir carcase, female reproduction and NFI EBVs. This study has shown the benefits of single-step genomic evaluations, and the opportunity to increase rates of genetic progress, through the increased accuracy generated. The study also highlighted breeds and traits which could benefit from additional recording to increase accuracies from single-step.

INTRODUCTION

Inclusion of SNP-based data is now routine in most livestock genetic evaluations worldwide. BREEDPLAN has included DNA marker data since 2010 and in 2017 implemented single-step SNP-based genomic evaluations (Johnston *et al.* 2018). Increasing rates of industry genotyping, coupled with the development of breed reference populations has resulted in increased accuracy of EBVs, especially for genotyped animals (Jeyaruban *et al.* 2019). Quantifying the benefits from single-step evaluations is not straightforward from large evaluations however Legarra and Reverter (2018) proposed using forward validation and semi-parametric estimates (called Method LR) as a relatively simple method to quantify the changes in accuracy, bias, and dispersion between two evaluations. The aim of this study was to use forward cross-validation and the Method LR to assess the changes in accuracies of single-step versus pedigree-based evaluations for a range of breeds that differ in the numbers of genotyped animals and the size and structure of their genomic reference populations.

MATERIAL AND METHODS

Data. The dataset used in this study included performance, pedigrees and genotypic data extracts from each of the five breed's databases from December 2022. Full BREEDPLAN multi-trait evaluations including maternal effects and genetic groups were run for each breed. Single-step evaluations were performed according to the procedures of Connors *et al.* (2017) and Johnston *et al.* (2018). Traits included: birth weight (BW), gestation length (GL), 200d weight (WW), 400d weight (YW), 600d weight (FW), cow weight (MCW), bull ultrasound rib fat (BRF), P8 fat (BP8), eye muscle area (BEMA), intramuscular fat percent (BIMF), heifer ultrasound rib fat (HRF), P8 fat (HP8), eye muscle area (HEMA), intramuscular fat percent (HIMF), days to calving (DC), abattoir

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

carcass weight (CWT), rib fat (CRIB), P8 fat (CP8), eye muscle area (CEMA), intramuscular fat (CIMF), marble score (CMS), marble fineness score (CMF), retail beef yield (CMY), shear force (CSF), net feed intake-postweaning (NFIP), net feed intake-finishing (NFIF), scrotal circumference (SC), heifer age at puberty (AP), flight time (FT) and percent normal sperm (PNS).

Forward validation. To perform the cross-validation animals were defined as “validation” animals or ‘training’ based on their year of birth. Animals born after 2018 (except Wagyu, where 2019 was a better split in the data of genotyped animals) comprised the validation sub-set. BREEDPLAN evaluations were performed using pedigree-based BLUP analyses (PED) and single-step analyses (S-S). These runs included all phenotypes, with the resulting EBVs for the validation subset of animals denoted as $\hat{\mathbf{u}}_w$. Phenotypes for the validation animals were then removed and the analyses repeated, with the resulting EBVs denoted as $\hat{\mathbf{u}}_p$. The subscripts “w” and “p” refer to “whole” and “part” analyses respectively, with the part EBVs of validation animals ($\hat{\mathbf{u}}_p$) informed by their pedigree and genomic relationships with the training animals, respectively. A series of metrics were computed using the Method LR (Legarra and Reverter 2018) to compare EBVs from part versus whole subsets from each of the PED and S-S evaluations. A population-based accuracy (acc) of the evaluations was computed using the approximation below:

$$acc = \sqrt{\frac{cov(\hat{\mathbf{u}}_w, \hat{\mathbf{u}}_p)}{(diag(\mathbf{K}) - \bar{\mathbf{K}})\sigma_{u,\infty}^2}}$$

where, \mathbf{K} is the relationship matrix (i.e. NRM for PED and GRM for S-S) for the validation animals with phenotypes for each trait and $\sigma_{u,\infty}^2$ is assumed to be the genetic variance. The dispersion (i.e. slope) was estimated by $disp = cov(\hat{\mathbf{u}}_w, \hat{\mathbf{u}}_p)/var(\hat{\mathbf{u}}_p)$ and the bias was estimated as $bias = (\bar{\hat{\mathbf{u}}}_p - \bar{\hat{\mathbf{u}}}_w)$. While the validation animals included both genotyped and pedigree-only animals, metrics reported in this paper are for genotyped animals only. Bias and dispersion metrics were computed but only summary results are reported here.

RESULTS AND DISCUSSION

Table 1 presents population estimates of accuracies for genotyped validation animals for the five breeds across the range of BREEDPLAN EBVs. Results show accuracies from S-S were higher than PED across all breeds and all trait EBVs. This demonstrates that the validation animal’s EBVs were on average more highly correlated from the S-S evaluation compared to the PED evaluation from the part versus whole runs. Across breeds S-S accuracies were generally highest for the growth traits, with an extreme value for Brahman birth weight (also evident in PED accuracy) suggesting the additive variance assumed is smaller than the true variance or is being influenced by maternal effects.

On average the increase in S-S accuracy was 0.05, 0.12, 0.16, 0.18, 0.23, for Santa Gertrudis, Brahman, Hereford, Wagyu and Angus, respectively (Figure 1). For Angus and Hereford the increased accuracies were in general agreement with earlier analyses of Jeyaruban *et al.* (2019) that used a subset of these data from previous years. Recently, Moore *et al.* (2023) presented an alternative approach applied to the Brahman data, and while their results were based on different edits and data subsets, the changes in accuracies from single-step across the traits were generally in agreement but not always, suggesting the different methods are possibly sensitive to assumptions and need further scrutiny.

The magnitude of the increases in accuracies generally reflected the size of the training populations. Figure 2 illustrates that pooled across traits and breeds, accuracies observed in validation animals tended to increase with the size of the training set. The plots plateau for S-S at 0.80 accuracy with more than 10,000 animals with records and genotypes in the training dataset

compared to about 0.50 for pedigree evaluations of the same size. In general, for traits with greater than 5,000 animals in the training populations the S-S accuracies were above 0.60.

Estimates of dispersion (not shown here) were generally close to the expected value of unity, indicating little evidence of under- or over prediction of S-S EBVs. Bias estimates (not shown) were mostly small but further analyses are required to compare ungenotyped and genotyped contemporaries.

Table 1. Population accuracy estimates for genotyped validation animals from pedigree (PED) and Single-step (S-S) BREEDPLAN evaluations for five breeds across all EBVs

EBV*	Angus		Brahman		Hereford		Santa Gertrudis		Wagyu	
	PED	S-S	PED	S-S	PED	S-S	PED	S-S	PED	S-S
BW	0.44	0.81	0.79	0.99	0.63	0.85	0.33	0.38	0.49	0.77
WW	0.53	0.81	0.54	0.70	0.57	0.76	0.37	0.41	0.50	0.76
YW	0.53	0.81	0.43	0.57	0.64	0.83	0.41	0.45	0.53	0.85
FW	0.54	0.82	0.46	0.61	0.56	0.79	0.41	0.46	0.51	0.81
MCW	0.52	0.84	0.42	0.64	0.58	0.79	0.50	0.53	.	.
BEMA	0.58	0.81	0.41	0.49	0.54	0.66	0.43	0.52	0.38	0.56
HEMA	0.58	0.79	0.39	0.48	0.63	0.74	0.50	0.58	0.34	0.55
BIMF	0.67	0.84	.	.	0.65	0.78	0.34	0.39	0.33	0.42
HIMF	0.71	0.89	.	.	0.71	0.85	0.33	0.39	0.35	0.44
BP8	0.46	0.75	0.35	0.49	0.52	0.70	0.44	0.51	0.33	0.48
HP8	0.45	0.76	0.43	0.55	0.63	0.81	0.38	0.47	0.36	0.52
BRF	0.50	0.78	0.34	0.48	0.54	0.70	0.45	0.52	0.34	0.48
HRF	0.50	0.78	0.50	0.61	0.69	0.87	0.38	0.47	0.35	0.48
CWT	0.47	0.66	0.35	0.43	0.32	0.55	0.33	0.37	0.52	0.77
CEMA	0.39	0.69	.	.	0.34	0.46	.	.	0.33	0.44
CIMF	0.50	0.73	0.40	0.46	0.35	0.51	0.31	0.34	0.45	0.61
CMY	0.44	0.60
CP8	0.35	0.64	0.35	0.47	0.24	0.47	0.35	0.38	0.40	0.61
CRF	0.39	0.66	0.37	0.43	0.29	0.45	0.38	0.41	.	.
CMF	0.27	0.37
CMS	0.40	0.62	0.53	0.63
CSF	.	.	0.36	0.42	.	.	0.28	0.32	.	.
DTC	0.47	0.54	0.36	0.60	0.25	0.31	0.55	0.67	.	.
AP	.	.	0.41	0.47
PNS	.	.	0.21	0.26	.	.	0.26	0.29	.	.
SC	0.44	0.72	0.59	0.76	0.49	0.61	0.40	0.46	.	.
GL	0.42	0.67	0.32	0.33	0.60	0.68	0.24	0.25	0.30	0.41
FT	.	.	0.41	0.46	.	.	0.51	0.57	.	.
NFIF	0.22	0.40	.	.	0.17	0.17
NFIP	0.24	0.36

*see text for trait names; “.” indicates trait not recorded or too few validation animals for the breed

CONCLUSIONS

This study has shown that single-step BREEDPLAN evaluations are delivering increased accuracies of EBVs across the full range of EBVs and breeds. This improvement in accuracies allow more genetic progress, particularly for economically important traits that are generally low accuracy at the time of selection. Improvements in accuracy from single-step will also benefit the commercial beef sector through better genetic description across a broader range of EBVs and allow more precise matching of genetics to specific production systems and markets. Further increases in accuracies are possible from single-step evaluations of particular breeds and traits by increasing the size of their

reference populations.

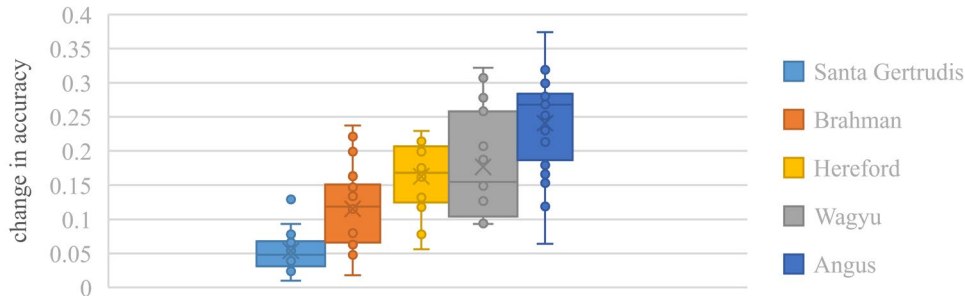


Figure 1. Average change in accuracy of validation animals across all traits by breed

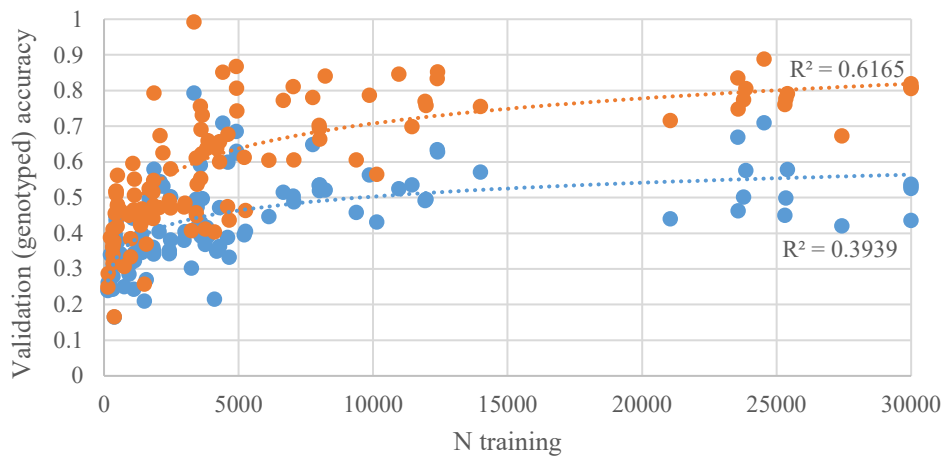


Figure 2. Accuracy of validation animals versus size of training populations for all traits and breeds for single-step (orange) and pedigree (blue) evaluations.

ACKNOWLEDGEMENT

This research was funded by Meat and Livestock Australia project L.GEN.2204. The authors acknowledge each of the five breed societies and their members for access to their data, and to the Agricultural Business Research Institute for providing phenotypic and genotypic data extracts.

REFERENCES

- Connors N.K., Cook J, Girard C., Tier B, Gore K, Johnston D.J. and Ferdosi M.H. (2017) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **22**: 317.
- Legarra A. and Reverter A. (2018) *Genet. Sel. Evol.* **50**: 53.
- Jeyaruban M.G., Gurman P.M., Johnston D.J., Swan A.A., Banks R.G. and Girard C.J. (2019) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **23**: 79.
- Johnston D.J, Ferdosi M., Connors N., Boerner V., Cook J., Girard C., Swan A. and Tier B. (2018). *Proc. 11th World Cong. Genet. Appl. Livest. Prod.* **11**: 269.
- Moore K.J., Gurman P.M. and Johnston D.J. (2023). *Proc. Assoc. Advmt. Anim. Breed. Genet.* **25**: *These proceedings.*