719. Cross-validation of single-step genetic evaluation in U.S. Katahdin sheep

A.J. McMillan[1*], D.J. Brown[1], J.M. Burke[2], J.L. Morgan[3] and R.M. Lewis[4]

[1]AGBU, a joint venture of NSW Department of Primary Industries and University of New England, 2351, Armidale, Australia; [2]USDA, ARS, Dale Bumpers Small Farms Research Center, Booneville, AR 72927, USA; [3]Round Mountain Consulting, Fayetteville, AR 72701, USA; [4]Department of Animal Science, University of Nebraska – Lincoln, Lincoln, NE 68583, USA; amcmill4@une.edu.au

Abstract

Genomic information is used in genetic evaluation to improve prediction accuracy in most livestock species. Such is less so in sheep, particularly in the U.S. In this study, the impact of implementing single-step genomic BLUP as compared to pedigree BLUP for weight and faecal egg count traits was evaluated in U.S. Katahdin sheep. Two methods of cross validation were utilised to compare the predictive ability and bias of estimated breeding values of these methods. Genomic information improved predictive ability for both traits, and reduced bias in the evaluation of faecal egg counts. Accuracies of estimated breeding values improved appreciably in genotyped animals. The benefit from including genomic information based on cross validation was minimal but is expected to improve as the reference population grows. Single-step genomic BLUP procedures developed will be used to update those applied in the routine genetic evaluation offered through the U.S. National Sheep Improvement Program.

Introduction

Many global genetic evaluation systems now utilise genomic information routinely. To date, this has not been possible for the National Sheep Improvement Program (NSIP) evaluation system, which provides genetic evaluation services to the U.S. sheep industry (Notter, 1998). Through a U.S. Department of Agriculture grant, and in collaboration with NSIP Katahdin breeders, a large number of genotypes and phenotypes were collected on key traits. Katahdin are an increasingly popular hair sheep breed in the U.S., recognized for tolerance to parasite resistance and twinning. These data provided a foundation to test and implement genomically assisted genetic evaluation. As a result of this research and development, and as a first for the U.S. sheep industry, NSIP recently implemented single step genomic BLUP (SS-GBLUP) (Legarra *et al*., 2014) within their genetic evaluation for Katahdin sheep. SS-GBLUP has been implemented and is successfully running in Australian evaluations since 2016 (Brown *et al.,* 2018), each using the same evaluation platform (OVIS; Brown *et al.,* 2007*)*. The aim of this study was to use cross validation methods with forward prediction to compare pedigree based BLUP (PBLUP) with SS-GBLUP in the Katahdin evaluation for weaning weight (wwt), post-weaning weight (pwt), weaning faecal egg count (wfec), and post-weaning faecal egg count (pfec).

Materials & methods

Data for this paper was based on the June 2021 Katahdin NSIP genetic evaluation. The analysis has close to 100,000 sheep and is multi-trait with 12 traits across age stages including body weights, carcase scans, faecal egg counts and reproductive traits. Approximately 5,000 animals have been genotyped using the GeneSeek Genomic Profiler Ovine 50k array (Neogen Corp., Lincoln, NE, USA) with most genotyped animals being born within the last 5 years.

Forward cross validation was undertaken using both PBLUP and SS-GBLUP procedures. These were full multi-trait analyses. The SS-GBLUP evaluation used the equations of Aguilar *et al.* (2010). A lambda value

R.F. Veerkamp and Y. de Haas (eds)

DOI: 10.3920/978-90-8686-940-4_719, © A.J. McMillan *et al.* 2022

of 0.5 was used (McMillan and Swan, 2017), which placed equal emphasis on the pedigree and genomic relationships for animals when both recorded.

Phenotypic data were pre-corrected for standard fixed effects such as age of dam, birth-rearing type, and age of measurement; a fixed effect for contemporary group was included directly. Phenotypes for animals born after 2018 were removed from the data and the BLUP analyses used to estimate breeding values (EBV) from this partial dataset. In Table 1, phenotypic data for wwt, pwt, wfec and pfec are shown. Accuracy of prediction was assessed by examining the ability for EBV to predict phenotypes pre-corrected using the OVIS evaluation procedures (Brown *et al.*, 2007). The expectation of these regressions is a slope of 1.0. The accuracy of prediction was also calculated using the correlation of adjusted phenotypes with EBV and dividing by the square root of the heritability. As an additional method of validation, accuracy and bias metrics were calculated using the LR method (Legarra and Reverter, 2018) as the correlation and regression slopes between the EBV from the full analysis and the partial analysis for the animals in the validation group with phenotypes (Legarra *et al.,* 2008). The EBV accuracy was estimated for each model using OVIS and compared as a further test of the improvements in predictive ability from each model.

Results

Changes in phenotypic predictive ability and bias are shown in Table 2. For all traits the SS-GBLUP approach increased predictive ability based on the variance explained by the model and accuracy of prediction. The improvements ranged from 2% for wwt up to 10% for pfec across both metrics. For the weight traits, there was a tendency for genomic information to overestimate the phenotypic differences compared to PBLUP. Although, the fec traits showed less bias with the SS-GBLUP model, the phenotypes still were under-predicted.

Result for the EBV regression are shown in Table 3. Based on the correlation of EBV, over all for animals in the validation set, the predictive ability slightly increased for all traits with the incorporation of genomic information. For animals with a genotype, the improvement was more substantial and ranged from 2-3% across all traits.

The bias was either not significantly different or improved when comparing SS-GBLUP to PBLUP. In genotyped animals the bias was closer to the expectation of 1 across traits with SS-GBLUP. The trait showing the largest bias under SS-GBLUP evaluation for genotyped animals was wfec although the bias was much improved compared to PBLUP evaluation.

Table 1. Summary of pre-2018 (training) and post-2018 (validation) datasets.

Trait[1]	Pre-2018			Post-2018		
	Records	Mean	Std	Records	Mean	Std
wwt	57,791	20.52	3.99	14,038	20.52	4.22
pwt	44,525	31.50	4.82	12,254	31.70	5.08
wfec	6,910	9.41	5.96	3,522	8.89	5.53
pfec	9,779	8.78	5.63	4,058	9.63	5.37

[1] wwt = weaning weights (kg); pwt = post-weaning weight (kg); wfec = weaning faecal egg count (egg/g, cube root transformed); pfec = post-weaning faecal egg count (egg/g, cube root transformed).

Table 2. Results from regression of adjusted phenotypes and accuracy of prediction and EBV using each evaluation method.

Analysis[1]	Trait[2]	Slope (s.e.)	R^2	Accuracy
PBLUP	wwt	1.08±0.08	0.39	0.64
PBLUP	pwt	0.97±0.06	0.38	0.57
PBLUP	wfec	0.57±0.08	0.27	0.28
PBLUP	pfec	0.58±0.07	0.24	0.13
SS-GBLUP	wwt	1.20±0.07	0.41	0.70
SS-GBLUP	pwt	1.12±0.06	0.44	0.63
SS-GBLUP	wfec	0.63±0.08	0.33	0.35
SS-GBLUP	pfec	0.70±0.07	0.34	0.19

[1] PBLUP = pedigree BLUP; SS-GBLUP = single-step genomic BLUP.

[2] Traits defined as in Table 1.

Table 3. Results from LR method for regression of full EBV with part EBV.

Analysis[1]	Trait[2]	All animals			Genotyped animals		
		Records	Slope (s.e.)	Corr	Records	Slope (s.e.)	Corr
PBLUP	wwt	13,768	1.00±0.01	0.77	2,076	0.94±0.02	0.79
PBLUP	pwt	11,784	0.92±0.01	0.71	2,529	0.95±0.02	0.76
PBLUP	wfec	3,394	0.84±0.02	0.62	1,691	0.85±0.03	0.60
PBLUP	pfec	3,906	0.88±0.02	0.65	1,761	0.97±0.02	0.68
SS-GBLUP	wwt	13,768	0.98±0.01	0.77	2,076	1.00±0.02	0.81
SS-GBLUP	pwt	11,784	0.93±0.01	0.72	2,529	1.04±0.02	0.79
SS-GBLUP	wfec	3,394	0.87±0.02	0.65	1,691	0.90±0.03	0.64
SS-GBLUP	pfec	3,906	0.94±0.02	0.67	1,761	1.02±0.02	0.71

[1] PBLUP = pedigree BLUP; SS-GBLUP = single-step genomic BLUP.

[2] Traits defined as in Table 1.

The comparison of EBV accuracy from each model is shown in Figure 1. Clear improvements in EBV accuracy are seen for genotyped animals from the SS-GBLUP model compared to PBLUP. Such was particularly the case where the accuracies based on PBLUP were low.

Discussion

Accuracy of prediction of phenotypes was significantly improved with the use of SS-GBLUP for all traits. The benefit was more obvious for fec traits. This is likely a function of the difference in heritabilities, and the number of phenotypes available on fec as compared to weight traits on the validation animals as weights are more commonly recorded. Thus, the weight traits had a larger amount of data supporting them compared to fec, and therefore a higher starting accuracy. The benefit from genomic information therefore becomes smaller. These differences were more obvious for genotyped animals.

The prediction of phenotypes from EBV were more biased for weight traits when comparing SS-GBLUP and PBLUP. In contrast, the predictions of phenotypes for fec traits were much less under predicted with genomic information. The benefit of the SS-GBLUP model can also be observed by comparing the improvement in estimated EBV accuracy from this model over PBLUP. This benefit is greater for animals with lower PBLUP accuracy.

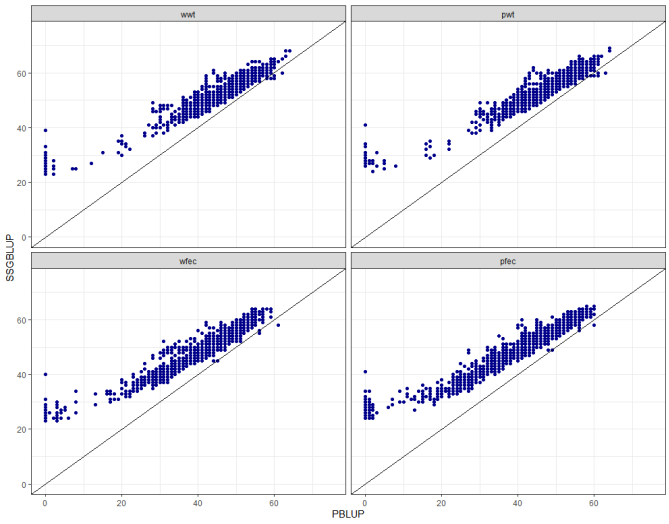

Figure 1. Comparison of EBV accuracy for genotyped animals from pedigree BLUP (PBLUP) and single-step genomic BLUP (SS-GBLUP) models for each trait.

In most cases, though, a small improvement in accuracy of prediction was observed. The benefit from including genomic information based on cross validation is minimal currently but is expected to improve as the size of the reference population grows. The new genetic evaluation procedures and models developed are now being used in the routine genetic evaluation of U.S. Katahdin sheep through NSIP. They will be the foundation for applying SS-GBLUP to other U.S. breeds once sufficient genomic information becomes available.

Acknowledgements

This project is supported by the Agriculture and Food Research Initiative Grant number 2016-51300-25723 from the USDA National Institute of Food and Agriculture, Organic Agriculture Research and Extension Initiative program.

References

Aguilar I., Misztal I., Johnson D.L., Legarra A., Tsuruta S. *et al.* (2010) J. Dairy Sci. 93:743-752. https://doi.org/10.3168/jds.2009-2730

Brown D.J., Huismann A.E., Swan A.A., Graser H.-U, Woolaston R.J. *et al.* (2007) Proc. Of the 18th AAABG, Armidale, Australia.

Brown D.J., Swan A.A., Boerner V., Li l., Gurman P.M., *et al.* (2018) Proc. of the 11th WCGALP, Auckland, New Zealand.

Legarra A., Christensen O.F., Aguilar I., and Misztal I. (2014) Livestock Science 166:54-65. https://doi.org/10.1016/j.livsci.2014.04.029

Legarra A, Robert-Granié C, Manfredi E, and Elsen J-M. (2008) Genetics 180(1):611-618. https://doi.org/10.1534/genetics.108.088575

Legarra A., and Reverter A. (2018) Genet. Sel. Evol. 50:53. https://doi.org/10.1186/s12711-018-0426-6

McMillan A.J., and Swan A.A (2017) Proc. of the 22nd AAABG, Townsville, Australia.

Notter, D.R. (1998) J. Anim. Sci. 76(9):2324–2330. https://doi.org/10.2527/1998.7692324x