Single-step genomic prediction for metabolic body weight of Nordic Red dairy cattle

R.M. Kempe[1*], T. Mehtiö[1], A.-M. Leino[1], M. Koivula[1], T. Pitkänen[1], R.B. Stephansen[2], E.A. Mäntysaari[1], E. Negussie[1], J. Pösö[3], U.S. Nielsen[4], W.F. Fikse[5], G.P. Aamand[6] and M.H. Lidauer[1]

[1] Natural Resources Institute Finland (Luke), Myllytie 1, FI-31600 Jokioinen, Finland; [2]Aarhus University, QGG, Blichers Alle 20, 8830-Tjele, Denmark; [3] Faba co-op, P.O. BOX 40, FI-01301 Vantaa, Finland; [4]SEGES Cattle, Agro Food Park 15, DK-8200 Aarhus N, Denmark; [5]Växa, Box 288, SE-75105 Uppsala, Sweden; [6] Nordic Cattle Genetic Evaluation, Agro Food Park 15, DK-8200 Aarhus N, Denmark; [*]riitta.kempe@luke.fi

Abstract

The aim of this study was to predict genomic breeding values for metabolic body weight (MBW) by a single-step model and to compare obtained breeding values to those from a multi-step genomic prediction model and a pedigree-based BLUP model. Data contained over 3 million observations from Danish, Finnish and Swedish Red dairy cattle born in the years 1990 to 2017. The applied multiple-trait animal models included six traits: 1st, 2nd, and 3rd parity MBW, and first parity stature, chest width and body depth. Breeding values from the single-step model were in good agreement with those from the multi-step model (correlations were 0.98-0.99), but BLUP breeding values differed from genomic breeding values slightly for bulls and clearly for cows. The single-step model gave higher validation reliabilities for bulls and cows than the multi-step model. The results support using a single-step approach for the genomic evaluation of MBW.

Introduction

In 2020, Nordic Cattle Genetic Evaluation (NAV) included the Saved Feed Index, which contains the metabolic body weight (MBW) and residual feed intake in the Nordic Total Merit index for the Danish, Finnish, and Swedish Holstein (HOL), Red dairy cattle (RDC) and Jersey populations. MBW is the core trait for the genetic evaluation of maintenance efficiency as the feed requirement for maintenance is assumed to have a linear relationship with MBW. The MBW evaluation (Lidauer et al., 2019) is based on three core traits; MBW in 1st, 2nd, and 3rd parity, and three correlated conformation traits; stature (ST), chest width (CW), and body depth (BD). Currently, genomic breeding values for MBW are predicted by a multiple-step approach but a single-step genomic prediction model (Aguilar et al., 2010; Christensen and Lund, 2010) is expected to improve predictions. The aim of this study was to develop a single-step genomic prediction model for MBW in Nordic RDC. Cross-validation statistics and regression of recent breeding values on earlier breeding values (Legarra and Reverter, 2018) were applied to compare the developed single-step model with the currently used multiple-step model, and with the pedigree-based BLUP model.

Materials & Methods

Phenotypic data and pedigree. The data contained all available body weight (BW) and conformation traits observations of Danish, Finnish and Swedish RDC cows born in the years 1990 to 2017. This included BW observations from Finnish cows, which were measured by tape (heart girth), BW observations from Danish cows, which were measured by scale, and first parity conformation observations from cows of all countries. The raw BW observations were pre-processed as explained by Lidauer et al. (2019) to obtain for each lactation, that had BW

records, one MBW observation ($BW^{0.75}$) with an associated weighting coefficient that accounted for the type and number of the raw BW observations. The first parity conformation observations were taken from the NAV routine conformation evaluation (NAV, 2019) with the following trait definition: ST (linear, cm), CW (linear score from 1 to 9) and BD (linear score from 1 to 9). After editing, the data consisted of 1.28 million MBW observations and 2.94 million conformation observations. The pedigree of the cows with observations included 2.99 million animals and the base population was described with 182 unknown parent groups (UPG).

Best linear unbiased prediction model. The BLUP model developed for the NAV routine MBW evaluation was the basis for building the single-step genomic predication model. The BLUP model is a multiple-trait animal model with six traits, which are MBW in 1st, 2nd, and 3rd parity, and first parity ST, CW, and BD. The applied heritabilities for these traits were 0.46, 0.51, 0.56, 0.60, 0.18, and 0.26, respectively. A detailed description of the model and the applied variance components is given in Lidauer et al. (2019). The correlated information from the conformation traits is important because there is no BW information available from Swedish cows. Moreover, recording of BW by tape is reducing in Finland and measuring BW by scale is slowly becoming more common. Genetic correlations between the three MBW traits and the conformation traits were on average 0.67, 0.55, and 0.49 with ST, CW, and BD, respectively. Furthermore, to improve predictions of Finnish herds the observations from HOL cows were also included to increase contemporary group sizes. The estimated breeding values (EBV) for MBW were used to form for each animal one combined MBW EBV by weighting 1st, 2nd, and 3rd parity EBVs with the weights 0.30, 0.25, and 0.45, respectively, and changing the sign of the index values so that a higher index value refers to a lower MBW.

Marker data. Genotype information from RDC animals only was included for developing the single-step model. Bulls were genotyped using the Illumina Bovine SNP50 Bead Chip (Illumina, San Diego, CA) and the cows using the lower-density EuroG MD chip (http://www.eurogenomics.com/). The unobserved genotypes with the lower-density chip were imputed to the 50K density. After editing there were 46,914 single nucleotide polymorphisms (SNP) markers available for 51,417 animals.

Multi-step genomic prediction model. For building the multi-step genomic prediction model, which was based on SNP-BLUP (msSNPBLUP), all genotyped bulls which had at least 20 daughters and a MBW EBV reliability of at least 0.5 and all genotyped cows that had observations were included. This resulted in 5,554 bulls and 43,276 cows and in total 48,830 genotyped animals. De-regressed proofs (DRP) based on the combined EBV for MBW of 1st, 2nd, and 3rd parity were used as observations for the SNP-BLUP evaluation and each animal's observation was weighted by the animal's effective record contribution (ERC). In addition to marker effects, the SNP-BLUP model included a polygenic effect explaining 10% of the genetic variation.

Single-step genomic prediction model. Relationships between animals in the single-step genomic prediction model (ssGBLUP) were described by the **H** matrix (Aguilar et al., 2010; Christensen and Lund, 2010) which was implemented as described by Koivula et al. (2021). The VanRaden method I (VanRaden, 2008) was used for building the genomic relationship matrix by blending the **G** matrix with the residual polygenic effect that accounts for 10% of the genetic variation. The QP transformation was applied to account for the UPG. A combined genomic enhanced breeding value (GEBV) for MBW was formed in the same way as the combined EBV for MBW.

Model validation. The ssGBLUP model was validated by forward prediction cross-validation, which was done for all the three models. For the cross-validation, we selected 1,721 cows and 354 bulls that had no own information in the reduced data set. For the evaluation with the reduced data, observations from most recent four years of bulls and their progenies were excluded. The same pedigree and genomic information were used as for the full data set evaluation to obtain breeding values for candidates (BV_c). Thus, in the SNP-BLUP evaluation for the reduced data all candidate animals were included but had no own data. The cross-validation reliability (r^2_{cv}) was calculated as:

$$r^2_{cv} = \text{corr}(DRP, BV_c)^2 / r^2_{DRP}, \quad (1)$$

where DRP were calculated separately for both the bull validation candidate, and the cow validation candidate group, to ensure proper calculation of ERC weights used for the de-regression. The second statistic we applied was the regression of full data breeding values on reduced data breeding values (Legarra and Reverter, 2018), which has an expectation of b_1=1.0.

Results

Correlations between ssGBLUP and msSNPBLUP breeding values of genotyped animals were 0.98-0.99, indicating that the models gave almost the same genomic breeding values (Table 1). The correlation between BLUP and ssGBLUP or msSNPBLUP breeding values was 0.97 in bulls and 0.84 to 0.86 in cows, respectively. Thus, BLUP breeding values were slightly different in bulls and clearly different in cows compared with genomic breeding values. The ssGBLUP GEBV had 5 to 12% higher standard deviation compared with msSNPBLUP or BLUP breeding values in cows and 0.2 to 5.5% higher in bulls.

The correlations between candidates' BV_c and their future DRP are presented in Table 2. Correlations were the highest when BV_c were estimated with ssGBLUP for both bulls and cows. This resulted also the highest r^2_{cv} with ssGBLUP being as high as 0.91 for the bull candidates. The r^2_{cv} of BV_c obtained with the msSNPBLUP model was higher than that when using BLUP in candidate bulls, but not in cows. For all three models, r^2_{cv} was higher for bulls compared with cows. Mean DRP reliability was for bulls 0.71 and for cows 0.61.

The b_1 estimates were closer to the expectation for ssGBLUP in candidate bulls than those for the BLUP model and msSNPBLUP model (Table 2). In cows, the *b1* estimate for BLUP was slightly closer to expectation than for the ssGBLUP or msSNPBLUP models. The b_1 estimates for msSNPBLUP were larger than 1.0 indicating that BV_c underpredicted the future breeding values, whereas the b_1 estimates for BLUP were below 1.0 indicating that BV_c overpredicted future breeding values.

Table 1. Standard deviation (Std) and correlation between MBW breeding values and those with de-regressed proofs (DRP) for reference population animals.

			Correlations		
	Prediction model[1]	Std	msSNPBLUP	ssGBLUP	DRP
Cows	BLUP	4.11	0.86	0.84	0.88
n=35,075	msSNPBLUP	4.36		0.98	0.75
	ssGBLUP	4.59			0.74
Bulls	BLUP	4.56	0.97	0.97	0.95
n=5,539	msSNPBLUP	4.33		0.99	0.90
	ssGBLUP	4.57			0.91

[1]BLUP=best linear unbiased prediction; msSNPBLUP=multi-step genomic BLUP, ssGBLUP=single-step genomic BLUP

Table 2. Cross-validation estimates correlation ($r_{(DRP,BVc)}$), validation reliability (r^2_{cv}), regression coefficient (b_1) for the bull and cow candidate groups with different models.

		Cross-validation reliability		Legarra-Reverter	
	Prediction model[1]	$r_{(DRP,BVc)}$	r^2_{cv}	b_1	R^2
Cows	BLUP	0.67	0.74	0.96	0.51
n=1,721	msSNPBLUP	0.65	0.69	1.09	0.70
	ssGBLUP	0.73	0.86	1.05	0.84
Bulls	BLUP	0.74	0.77	0.96	0.56
n=354	msSNPBLUP	0.75	0.80	1.10	0.74
	ssGBLUP	0.81	0.91	1.01	0.85

[1] BLUP=best linear unbiased prediction; msSNPBLUP=multi-step genomic BLUP, ssGBLUP=single-step genomic BLUP

Discussion

Breeding values from the single-step model were in good agreement with those from the multi-step model and reasonably different from those of the BLUP model. Correlations between traditional BLUP and genomic models were clearly lower than correlations among genomic models. Also, country-wise (Denmark, Finland, Sweden) investigation of the breeding values confirmed that the single-step model is modelling the data appropriately. The obtained cross-validation reliabilities were higher than those for RDC yield traits ssGBLUP GEBVs (Koivula et al. 2018). This can be explained by the high heritability of the MBW traits and ST, and the high genetic correlation between MBW and ST. The reason why we obtained a lower r^2_{cv} value for the msSNPBLUP evaluation compared with the BLUP evaluation for the cow candidate group was unclear and requires further assessments.

The regression of candidates' full data GEBV on candidates reduced data BV_c were close to the expectation when applying ssGBLUP, and therefore ssGBLUP is preferable over msSNPBLUP for obtaining GEBV for candidates. The results of this study support the implementation of ssGBLUP for the genomic evaluation of MBW in Nordic RDC and to develop ssGBLUP also for the Nordic HOL and Jersey MBW evaluation.

References

Aguilar I., Misztal I., Johnson D.L., Legarra A., Tsuruta S., *et al.* (2010) J. Dairy Sci. 93(2):743–752. https://doi.org/10.3168/jds.2009-2730

Christensen O.F., and Lund M.S. (2010) Genet. Sel. Evol. 42:2. https://doi.org/10.1186/1297-9686-42-2

Koivula M., Strandén I., Aamand G.P. and Mäntysaari, E.A. (2018) J. Anim. Breed. Genet. 135(2):107-115. https://doi.org/10.1111/jbg.12318

Koivula M., Strandén I. Aamand G.P. and Mäntysaari E.A. (2021) J. Dairy Sci. 104(9):10049–10058. https://doi.org/10.3168/jds.2020-19821

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

Lidauer M.H., Leino A.-M., Stephansen R.S., Pösö J., Nielsen U.S., *et al.* (2019) Interbull Bulletin no. 55:21-25. https://journal.interbull.org/index.php/ib/issue/view/78

NAV (2021) NAV routine genetic evaluation of dairy cattle - data and genetic models (12th ed). Available on: https://nordicebv.info/wp-content/uploads/2021/10/NAV-routine-genetic-evaluation_EDITYSS-08102021.pdf

VanRaden P. M. (2008) J. Dairy Sci. 91: 4414–4423. https://doi.org/10.3168/jds.2007-0980