Comparison of a single-step with a multistep single nucleotide polymorphism best linear unbiased predictor model for genomic evaluation of conformation traits in German Holsteins

H. Alkhoder, Z. Liu,* D. Segelke, and R. Reents

IT-Solutions for Animal Production (vit), Heinrich-Schroeder-Weg 1, D-27283 Verden, Germany

ABSTRACT

Genomic evaluation based on a single-step model uses all available data of phenotype, genotype, and pedigree; therefore, it should provide unbiased genomic breeding values with a higher correlation of prediction than the current multistep genomic model. Since 2019, a mixed reference population of cows and bulls has been applied to the routine multistep genomic evaluation in German Holsteins. For a fair comparison between the single-step and multistep genomic models, the same phenotype, genotype, and pedigree data were used. Because of its simple structure of the standard multitrait animal model used for German Holstein conventional evaluation, conformation traits were chosen as the first trait group to test a single-step SNP BLUP model for the large, genotyped population of German Holsteins. Genotype, phenotype, and pedigree data were taken from the official August 2020 conventional and genomic evaluation. Because of the same trait definition in national and multiple across-country evaluation for the conformation traits, deregressed multiple across-country evaluation estimated breeding value (EBV) of foreign bulls were treated as a new source of data for the same trait in the genomic evaluations. Due to a short history of female genotyping in Germany, the last 3 yr of youngest cows and bulls were deleted, instead of 4 yr, to perform a genomic validation. In comparison to the multistep genomic model, the single-step SNP BLUP model resulted in a higher correlation and greater variance of genomic EBV according to 798 national validation bulls. The regression of genomic prediction of the current, full evaluation on the earlier, truncated evaluation was slightly closer to 1 than the multistep model. For the validation bulls or youngest genomic artificial insemination bulls, correlation of genomic EBV between the 2 models was, on average, 0.95 across all

the conformation traits. We did not find overprediction of young animals by the single-step SNP BLUP model for the conformation traits in German Holsteins.

Key words: single-step model, single nucleotide polymorphism best linear unbiased predictor model, conformation trait, dairy cattle

INTRODUCTION

Genomic prediction (VanRaden, 2008) based on the genomic model (Meuwissen et al., 2001) has revolutionized animal breeding and selection, particularly in Holstein dairy cattle. Single-step genomic models (Aguilar et al., 2010; Christensen and Lund 2010; Legarra and Ducrocq 2012; Liu et al., 2014; Mäntysaari and Strandén 2016; Gao et al., 2018; Misztal et al., 2020) use all available information on phenotype, genotype, and pedigree and should thus provide unbiased genomic prediction. For routine genomic evaluation in German Holsteins, a multistep SNP BLUP model including a residual polygenic model (**RPG**; Liu et al., 2011) has been applied to a mixed reference population of bulls and cows (Alkhoder et al., 2017). With genomic selection implemented for German Holstein in 2010, higher genetic progress has been achieved (e.g., due to shorter generation intervals). Because the current multistep genomic model (**MSM**) based on a conventional evaluation cannot account for the genomic pre-selection properly and completely, prediction bias in the conventional and the subsequent genomic evaluation has become evident with an underestimation for youngest animals.

In comparison to the single-step genomic BLUP model (**ssGBLUP**; Aguilar et al., 2010; Gao et al., 2018), the single-step SNP BLUP model (**ssSNPBLUP**; Liu et al., 2014) does not need to set up the inverse of potentially huge genomic relationship matrix $\mathbf{H}^{-1}$, thus the ssSNPBLUP can analyze millions of genotyped animals without making approximation on the genomic relationships among animals. Vandenplas et al. (2019, 2020) compared alternative ssSNPBLUP models and confirmed the single-step model (Liu

Received August 12, 2021.

Accepted December 22, 2021.

*Corresponding author: zengting.liu@vit.de

et al., 2014) was most efficient. Therefore, the variant of single-step model, the ssSNPBLUP model by Liu et al. (2014), was considered to be feasible and chosen for the German Holstein genomic evaluation with a very large genotyped population.

Routine conventional evaluations for German Holsteins cover a wide range of trait groups all evaluated with multitrait animal models. The statistical model for genetic evaluation of conformation traits is a standard multitrait animal model (Reents, 1993) and represents the simplest form of the multitrait animal model. This form of the multitrait model does not involve complex modeling of additive genetic effects such as a random regression test-day model (Schaeffer et al., 2000). Furthermore, our multitrait animal model for the conformation traits did not have correlated genetic effects, such as for calving traits, or other nongenetic random effects, such as permanent environmental effects. Thus, we decided to use the conformation traits, as the first of 13 trait groups, for the purpose of testing the ssSNPBLUP model in German Holsteins. Using national phenotype data of 18 linear type traits, Tsuruta et al. (2021) investigated the causes of bias in single-step genomic evaluation for the US Holstein population with more than 2.3 million genotyped animals.

The objectives of this study were (1) to compare genomic predictions using the single-step and the current multistep SNP BLUP model for the conformation traits of German Holsteins, and (2) to investigate the bias and consistency of prediction of the 2 genomic models via a truncated validation data set for the conformation traits.

MATERIALS AND METHODS

Phenotype, genotype, and pedigree data for German Holsteins were obtained from routine evaluation in August 2020. A total number of 23 conformation traits were evaluated in the conventional evaluation (Reents, 1993) as well as in the subsequent multistep genomic evaluation. The conformation traits recorded on first-lactation cows in Germany included all 21 conformation traits covered in Interbull multiple across-country evaluation (**MACE**), except overall conformation score, and 3 national-only type traits. These 23 conformation traits were evaluated in 3 independent groups. The body and rump group included the following 9 traits: stature (**STA**), chest width (**CWI**), body depth (**BDE**), angularity, rump angle (**RAN**), rump width, BCS, dairy type (**MTY**), and body score (**KOE**). The udder group included the following 8 traits: fore udder attachment, rear udder height, udder support (**USU**), udder depth (**UDE**), front teat placement, front teat

length, rear teat placement (**RTP**), overall udder score (**OUS**). The feet and legs group included the following 6 traits: rear leg set, rear leg rear view, foot angle (**FAN**), overall feet and leg score (**OFL**), locomotion (**LOC**), and hock quality (**SPR**). Each of the 3 trait groups was evaluated independently using a multitrait animal model (Reents, 1993). A full list of the conformation traits with their abbreviations is given in Table 1.

The current MSM for German Holsteins (Liu et al., 2011) is a single-trait model; a trait-specific proportion of residual polygenic variance (**RPV**) in additive genetic variance was applied to each trait. Table 1 shows the RPV in percentage of additive genetic variance for each of the 23 conformation traits for German Holsteins. The assumed RPV percentage was 10% for the following 7 traits: STA, RAN, rump width, UDE, front teat placement, front teat length, and RTP. The assumed RPV percentage was 20% for the following 7 traits: CWI, BDE, angularity, rear leg set, rear leg rear view, fore udder attachment, rear udder height. The assumed RPV percentage was 30% for the following 9 traits: FAN, USU, OUS, OFL, LOC, BCS, MTY, KOE, and SPR. For genotyped animals, parental averages of conventional evaluation or conventional EBV were combined with direct genomic values (**DGV**) from the MSM to obtain genomic breeding values (**GBV**) based on the selection index theory (Liu et al., 2011, 2019).

Two conformation traits with strong genetic trends, STA and UDE, and 2 traits with weak or flat genetic trends, CWI and USU, were chosen as examples of the conformation traits for detailed analysis. In the conventional as well as the single-step evaluation, a pre-adjustment for heterogeneous variance in classifier $\times$ year effects was performed before solving all effects of the models (Reents, 1993). The ssSNPBLUP model (Liu et al., 2014) was applied to the conformation traits as follows:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{u} + \mathbf{e}, \quad [1]$$

where $\mathbf{y}$ is a vector of conformation trait values adjusted for the heterogeneous classifier $\times$ year variances; $\mathbf{b}$ is a vector of all fixed effects including the major fixed effect of herd $\times$ classification-date, age at calving $\times$ lactation stage, and classifier $\times$ year of classification; $\mathbf{X}$ is the incidence matrix for all the fixed effects; $\mathbf{u}$ is a vector of GBV for the cows with the type traits; and $\mathbf{e}$ is a vector of residuals.

The single-step evaluation of conformation traits was amended with a new source of data from Interbull bull MACE evaluation. Deregressed EBV (**DRP**) of

Table 1. A list of conformation traits evaluated routinely for German Holsteins

Item	Trait name	Trait group	RPV ¹ (%)	Abbreviation
International traits	Stature	Body	10	STA
	Chest width		20	CWI
	Body depth		20	BDE
	Angularity		20	ANG
	Rump angle		10	RAN
	Rump width		10	RWI
	Body condition score	Feet and legs	30	BCS
	Rear leg set		20	RLS
	Rear leg rear view		20	RLR
	Foot angle		30	FAN
	Overall feet and legs		30	OFL
	Locomotion		30	LOC
	Fore udder attachment	Udder	20	FUA
	Rear udder height		20	RUH
	Udder support		30	USU
	Udder depth		10	UDE
	Front teat placement		10	FTP
	Front teat length		10	FTL
	Rear teat placement		10	RTP
	Overall udder score		30	OUS
National traits	Dairy type	Body	30	MTY
	Body score	Body	30	KOE
	Hock quality	Feet and legs	30	SPR

¹Assumed proportion of residual polygenic variance in the total additive genetic variance of the multistep genomic model for German Holstein conformation evaluation.

all MACE bulls were used here as pseudophenotype. It was assumed that the MACE trait and national conformation traits were genetically identical and treated as the same trait. Effective daughter contribution (**EDC**) on the animal-model basis were used as weights for their DRP for all the bulls in MACE evaluation. If a bull had daughters in Germany as well as in foreign countries, his weight was the difference in the animal-model EDC between all daughters worldwide from MACE and domestic daughters in Germany. When a bull had only daughters in Germany, then his weight would be zero and he would not have the additional DRP record. However, if a bull had only daughters in foreign countries, his MACE EDC would be the weight for his DRP based on the MACE evaluation. Because the DRP were free of the fixed effects of model [1], a single pseudo-identification was assigned to each of the fixed effects for the DRP records.

The ssSNPBLUP by Liu et al. (2014) allows fitting an RPG effect in the single-step SNP BLUP model. We assumed that the RPG explained 30% additive genetic variance for all the 23 conformation traits. The choice of the 30% RPV was consistent with the MSM model where some type traits already had 30% RPV in the current MSM model (Table 1). Andersen et al. (2021) demonstrated that the 30% RPV was optimal for conformation traits. In contrast to the ssSNPBLUP model, the MSM assigned a trait-specific RPV percentage to each of the 23 traits due to the single-trait MSM model. To make a fair comparison between the 2 models, the

conformation trait LOC with the same 30% RPV was additionally selected for analyzing the genomic evaluation results.

A fixed effect, the so-called J factor (Hsu et al., 2017), was included per trait in ssSNPBLUP model [1] to account for different genetic levels of the genotyped and nongenotyped animals. The J factor considers the fact that genotypes may be only available on selected animals. For the estimation of the SNP effects of the model [1], genetic correlations between traits within each of the 3 trait groups were considered due to the multitrait SSM model, whereas the MSM ignored the genetic correlations between traits in the SNP effect estimation because of the single-trait MSM model (Liu et al., 2011).

Ancestors of the genotyped animals (including genotyped young animals) and ancestors of the cows with type trait records or bulls with DRP were traced back in pedigree as far as possible. Unknown parent groups (**UPG**) were fitted for all animals as well as for the genotyped animals using the concept of Quaas-Pollak transformation (Vandenplas et al., 2021). In contrast to fitting genetic group contributions as regression effects (Misztal et al., 2013), Vandenplas et al. (2021) applied the Quaas-Pollak transformation, by treating genetic groups as phantom parents, to mixed model equations of the ssSNPBLUP model (Liu et al., 2014) and successfully derived a new set of equations including the UPG effects. The authors (Vandenplas et al., 2021) demonstrated that their way of modeling UPG

via the Quaas-Pollak transformation was more efficient than the way of fitting UPG as covariates (Misztal et al., 2013) for the ssSNPBLUP model (Liu et al., 2014).

The software MiXBUP (Ten Napel et al., 2020) was used for the single-step evaluation based on the ssSNPBLUP model (Liu et al., 2014). The ssSNPBLUP model (Liu et al., 2014) was also implemented in software MiX99 (Strandén and Lisdauer, 1999) but in a special way (Mäntysaari, personal communication). The SNP markers were treated in the special implementation as if they were animals with neither known parents nor progeny. Solutions of the SNP markers, as fake animals, from the special implementation must be divided by a constant to obtain original effects of the SNP markers. Both software packages MiXBUP and MiX99 implemented an efficient algorithm for multiplying a vector with the inverse of pedigree relationship matrix for all genotyped animals, $\mathbf{A}_{22}^{-1}$ (Strandén and Mäntysaari 2014; Masuda et al., 2017). The sparse presentation of $\mathbf{A}_{22}^{-1}$ allows processing genotype data of tens of millions of animals at a relatively low cost.

Our current MSM model (Liu et al., 2011) was applied to DRP of the reference cows and MACE DRP of the bulls (Alkhoder et al., 2017) for estimating effects of SNP markers, which were then used for calculating DGV of genotyped animals. We used the selection index method to combine DGV and conventional parental averages or EBV for calculating GEBV of the genotyped animals (Liu et al., 2019). Nongenotyped cows with phenotype data were not considered in the SNP effect estimation of the MSM model. However, their phenotype data were considered in the conventional parental averages if they are linked to the genotyped animals. The DRP of all bulls from the MACE evaluation were also considered in the parental average calculation for all genotyped animals in the multistep genomic evaluation. In contrast to the MSM model, the ssSNPBLUP model was able to evaluate all animals with phenotypes or all genotyped animals simultaneously.

Due to a relatively short history of female animal genotyping in German Holsteins (Liu et al., 2019), only the youngest 3 birth years of cows with conformation records (2016, 2017, 2018) were removed for a genomic validation for the ssSNPBLUP as well as for the MSM model. Youngest reference bulls born from 2013 to 2015

were selected as validation bulls if they had daughters in at least 10 herds and most daughters in Germany. In total, 798 national validation bulls were selected. Because daughters of the validation bulls may be included in the truncated reference population for the MSM model or truncated phenotype data for ssSNPBLUP model, we further removed all daughters of the validation bulls from the truncated reference population or truncated phenotype data set. With the same selection criteria, 2,964 foreign validation bulls were defined, and they must not have daughters in Germany.

Table 2 shows the numbers of animals with phenotype data of conformation traits for the full and truncated evaluations. The total number of cows with own conformation records was 2,715,550, and the number of Holstein bulls with daughters outside Germany was 115,552. A total number of 875,252 genotyped Holstein animals were considered, including culled male candidates. Pedigree file for the genotyped or phenotyped animals contained 9,012,965 animals for the ssSNPBLUP full evaluation. Unknown parent groups were defined according to breeds and country origins, 4 selection paths, and birth years of animals. The number of UPG effects was 138 for the full evaluation, and the UPG effects were treated as random effect by adding 1 to the corresponding diagonals in the mixed model equations. Figure 1 shows the numbers of genotyped or phenotyped cows by birth year in the full and truncated evaluation. In Figure 2, one can see the numbers of bulls with either own phenotype DRP or daughters with records and the number of bulls with genotype data across all birth years.

RESULTS AND DISCUSSION

The single-step genomic evaluation was performed using the full and truncated data sets with the software MiXBUP. The total number of estimated effects or equations for the full evaluation was 217,423,347. A total of 3,387 rounds of iteration were required to reach a pre-defined convergence criterion. Using 15 of a total of 48 cores on a Linux server, a total clock time of 49 h was needed for the full evaluation, and the memory usage was 65 Gb peak virtual memory usage and 39 Gb peak resident set size.

Table 2. Description of phenotype data sets for a full and a truncated evaluation of all 23 conformation traits in German Holstein

Item	Full data set	Truncated data set
No. of cows with own conformation records	2,715,550	2,454,801
No. of bulls with MACE data ¹	115,552	112,605

¹Bulls must have daughters outside Germany. MACE = multiple across-country evaluation.

Figure 1. The numbers of phenotyped or genotyped cows by birth year for the full and truncated evaluations.

Comparison of the 2 Software Packages

Both MiXBUP and MiX99 were tested for the ssSNPBLUP model using the same data of the conformation traits. The 2 software packages differed in computational efficiency for the ssSNPBLUP model (Liu et al., 2014) because the implementations of the single-step model were different. A second-level pre-conditioner for the SNP effects (Vandenplas et al., 2019) was used in both software programs. Using the compressed PLINK BED format for storage of SNP genotypes, MiXBUP showed some advantage in memory usage as well as

computing time for both pre-processing and solving steps.

Despite the differences in computational efficiency and drastically different ways of implementing the ssSNPBLUP model by Liu et al. (2014), the 2 software packages gave identical estimates of all the model effects. Correlation of SNP effects between the 2 software packages exceeded 0.99 for any of the 23 type traits. GEBV correlation was above 0.995 for any group of genotyped animals. We obtained equal average and variance particularly for reference bulls or cows as well as for female or male candidates. We concluded that

Figure 2. The numbers of bulls with phenotype or genotype data in the full and truncated evaluations.

both MiXBLUP and MiX99 resulted in equal effect estimates for the conformation traits of German Holsteins. It is worth noting that the compressed PLINK BED storage format of SNP genotypes can enable processing possibly tens of millions of genotyped animals with the ssSNPBLUP model (Liu et al., 2014). As the 2 software packages led to identical evaluation results, we presented here only results from the software MiXBLUP.

SNP Effect Estimates

The ssSNPBLUP model by Liu et al. (2014) directly estimated SNP effects. Figure 3 shows the observed correlations between SNP effect estimates of the 2 data sets or between the 2 genomic models. For the ssSNPBLUP model, SNP effect correlations between the full and truncated evaluations (**SS_SS-VAL**; the solid line in red) ranged from 0.91 to 0.95 for the 23 conformation traits with an average of 0.94. The national trait SPR containing no foreign bull MACE information had the lowest correlation, 0.91. In comparison to the ssSNPBLUP model, the SNP effect correlations for the MSM model between the full and truncated evaluations (**MS_MS-VAL**; the dotted line in green) were much lower, between 0.80 and 0.91 with an average of 0.86. To explain the much lower SNP effect correlations of the MSM than the ssSNPBLUP model, we took a closer look at the reference population of the MSM model for the trait STA. In total, there were 215,705 (113,398) reference animals for the SNP effect estimation as follows: 175,208 (76,734) reference cows and 40,497 (36,664) bulls, for the full (truncated) evaluation. A total of 751,497 (494,006) ancestors were found for the reference animals, 61,699 (46,572) male and 689,798 (447,434) female ancestors, in pedigree file of the full (truncated) evaluation of the MSM model. Of the 61,699 (46,572) male ancestors, 8,260 (6,689) bulls had daughters in MACE but no genotypes in the full (truncated) data. Among the 689,798 (447,434) female ancestors, 108,495 (62,520) cows had own records in the national phenotype data but no genotypes in the full (truncated) data. In contrast to the SSM model, the MSM model ignored the phenotype data of the 8,260 (6,689) ancestor bulls and 108,495 (62,520) ancestor cows when estimating SNP effects in the full (truncated) evaluation. As shown here, most of these nongenotyped ancestors with own phenotype data in the full evaluation were also present in the truncated data set of the SSM model. These common ancestors with own phenotype data between the full and truncated data sets of the SSM model likely led to the higher correlation of the SNP effects for the SSM than the MSM model.

For the full data set, SNP effect estimates were correlated between the 2 models (**MS_SS**, dashed line with dots in black) with an average of 0.82, ranging from 0.77 to 0.88. The 3 national type traits, MTY, KOE, and SPR, had the lowest correlations between the 2 models. Similar SNP effect correlations were also observed for the truncated validation data set between the 2 models (**MS-VAL_SS-VAL**, dashed line in blue).

Figure 4 shows the regression coefficients of SNP effect estimates of the full evaluation on the truncated evaluation for each of the 2 genomic models. For the ssSNPBLUP model (**b1: SS | SS-VAL**; in red), the regression coefficients were all close to 1, varying from 0.987 for OFL to 1.045 for STA with an average of 1.018. In comparison, the MSM model had regression coefficients (**b1: MS | MS-VAL**; in blue) all lower than 1, ranging from 0.899 for the national trait SPR to 0.952 for BCS, and the average regression coefficient was 0.927. The regression coefficients indicated that the MSM slightly overpredicted the variance of SNP effects, and the ssSNPBLUP model resulted in neither over- nor under-prediction of SNP effects for any of the conformation traits.

Correlation of Genomic Prediction

The GEBV correlation of validation animals between an early-truncated and a later-full evaluation may be regarded as statistics closely related to the accuracy of genomic prediction, although it does not exactly follow the definition of accuracy by Legarra and Reverter (2018). The GEBV of all the 798 validation bulls were compared between the full and truncated evaluations and between the 2 genomic models. It can be seen clearly in Figure 5 that the ssSNPBLUP model resulted in higher GEBV correlations, with an average of 0.91, between the full and the truncated evaluation (**SS_SS-VAL**, solid line in red) for any of the traits than the MSM model (**MS_MS-VAL**, dotted line in green), with an average of 0.77. For the full data set, GEBV of the 2 genomic models are highly correlated (**MS_SS**, dashed line with dots in black), ranging from 0.94 to 0.99. These GEBV correlations were lower between the 2 models for the truncated data set (**MS-VAL_SS-VAL**, dashed line in blue) than for the full data set.

We also investigated the correlation of GEBV of the validation bulls with their deregressed EBV. Please keep in mind that DRP of the validation bulls were calculated from the conventional evaluation, not from the single-step evaluation. The DRP of the validation bulls should be less autocorrelated with their GEBV from the truncated evaluation than their GEBV from the full evaluation. It can be seen in Figure 6 that the ssSNPBLUP model led to a higher correlation between

GEBV of the truncated evaluation and DRP of the full conventional evaluation (SS-VAL_DRP, in red) for any of the traits than the MSM (MS-VAL_DRP, in blue), even though the bull DRP were calculated from the conventional evaluation. Overall, GEBV of the truncated evaluation were less correlated with their DRP than with their GEBV of the full evaluation in Figure 5. Comparing to the validation correlations in Table 3 of the US study (Tsuruta et al., 2021), we could see our validation correlations were equal or slightly higher for the common type traits. Our higher correlations of GEBV of the validation bulls with their DRP may be caused by use of the bull MACE data in our study. In contrast, Tsuruta et al. (2021) analyzed only the US national phenotype data of their 18 linear type traits. If foreign validation bulls had no daughters in Germany, their deregressed MACE EBV from the current MACE should have lower reliability than the German national validation bulls. For all conformation traits included in

MACE, except the total overall conformation score, we compared the GEBV correlation of genomic prediction between the foreign and domestic validation bulls shown in Figure 7. For the ssSNPBLUP model, foreign validation bulls have consistently much lower correlation of GEBV from the validation evaluation with their DRP than the national validation bulls: SS-VAL_DRP FOR (solid line in red) versus SS-VAL_DRP DEU (dotted line in green). The average difference in the correlation across the 20 MACE conformation traits was 0.10, with a mean of 0.79 for the national and 0.69 for the foreign validation bulls, respectively. For the MSM model, the national validation bulls also showed a higher correlation of GEBV with DRP for any of the traits than the foreign validation bulls as follows: MS-VAL_DRP DEU (dashed line with dots in black) versus MS-VAL_DRP FOR (dashed line in red). However, the average difference in the correlation was 0.05, which was less than the average difference for the ssSNPBLUP model.

Figure 3. Observed correlations between SNP effect estimates from the full and truncated evaluations (MS_MS-VAL = SNP effect correlations between the multistep full and validation evaluation; SS_SS-VAL = the correlations between the single-step evaluation with the full and truncated validation data set; MS_SS = the SNP effect correlations between the multistep and single-step evaluation with the full data set; MS-VAL_SS-VAL = the correlations between the 2 genomic models with the truncated validation data set). STA = stature; CWI = chest width; BDE = body depth; ANG = angularity; RAN = rump angle; RWI = rump width; RLS = rear leg set; RLR = rear leg rear view; FAN = foot angle; FUA = fore udder attachment; RUH = rear udder height; USU = udder support; UDE = upper depth; FTP = front teat placement; FTL = front teat length; RTP = rear teat placement; OUS = overall udder score; OFL = overall feet and legs; LOC = locomotion; MTY = dairy type; KOE = body score; SPR = hock quality.

Figure 4. Regression coefficients of SNP effect estimates of the full on the truncated evaluations (b1: MS | MS-VAL represents the regression of SNP effect estimates of the multistep full valuation on the multistep validation evaluation; b1: SS | SS-VAL represents the regression of SNP effect estimates of single-step full evaluation on the single-step validation evaluation). STA = stature; CWI = chest width; BDE = body depth; ANG = angularity; RAN = rump angle; RWI = rump width; RLS = rear leg set; RLR = rear leg rear view; FAN = foot angle; FUA = fore udder attachment; RUH = rear udder height; USU = udder support; UDE = upper depth; FTP = front teat placement; FTL = front teat length; RTP = rear teat placement; OUS = overall udder score; OFL = overall feet and legs; LOC = locomotion; MTY = dairy type; KOE = body score; SPR = hock quality.

Dispersion of Genomic Prediction

Regression of GEBV of the validation bulls from the full evaluation on GEBV of the early, truncated evaluation indicated whether the genomic prediction was inflated or underestimated. It can be seen clearly in Figure 8 that the regression coefficients for the ssSNPBLUP model (b1 SS | SS-VAL, in red) were close to 1, with an average of 1.00, ranging from 0.935 for LOC to 1.066 for BCS. The average regression coefficient for the MSM model (b1 MS | MS-VAL, in blue) was slightly lower, 0.98, also close to 1. We concluded that neither the ssSNPBLUP nor the MSM model resulted in a notable overprediction or underestimation of genomic evaluation.

Using the pseudophenotype of the validation bulls, DRP, we calculated the regressions of DRP on their early GEBV from the truncated evaluation, shown in Figure 9. In contrast to the regression coefficients in Figure 8, there was a greater variation among the traits. The average of regression coefficients was 1.03 or 1.09 for the ssSNPBLUP (b1 DRP | SS-VAL, in red) or the MSM model (b1 DRP | MS-VAL, in blue), respectively. The US study (Tsuruta et al., 2021) found similar regression coefficients also close to 1. However, more regression coefficients deviating from 1 were ob-

tained in their study (e.g., 0.74 for OFL and 1.25 for RTP). Our regression coefficients had the lowest value (0.96) for BDE and the highest value (1.20) for FAN.

Averages and Variances of GEBV

Genotyped German Holstein animals were chosen for comparing trends and variances of GEBV between the ssSNPBLUP and MSM models. Unlike the validation bulls shown in the previous sections, there were 3 groups of genotyped Holstein animals to be considered in subsequent sections to show their statistics such as genetic trends. Table 3 shows the number of genotyped AI bulls that have been highly selected, the number of genotyped male candidates without own phenotype data, and the number of genotyped female candidates with no own phenotype data. The genotyped male candidates were, to some degree, pre-selected for genotyping usually based on their genomic parental average. The genotyped female candidates could be considered an unselected sample, thanks to the whole-herd genotyping project KuhVision in Germany (Liu et al., 2019). As stated before, traits STA and UDE, representing conformation traits with high genetic trends, and traits CWI and RAN, representing the type traits with flat genetic trends, were chosen for further analysis. With

Figure 5. Observed correlations of genomic EBV (GEBV) of validation bulls between the full and the truncated evaluations (MS_MS-VAL = the GEBV correlations between the multistep genomic evaluation with the full and truncated validation data set; MS_SS = the correlations between the multistep and single-step evaluation using the full data set; SS_SS-VAL = the GEBV correlations between the full and truncated validation evaluation using the single-step model; MS-VAL_SS-VAL = the GEBV correlations between the 2 genomic models with the truncated validation data set). STA = stature; CWI = chest width; BDE = body depth; ANG = angularity; RAN = rump angle; RWI = rump width; RLS = rear leg set; RLR = rear leg rear view; FAN = foot angle; FUA = fore udder attachment; RUH = rear udder height; USU = udder support; UDE = upper depth; FTP = front teat placement; FTL = front teat length; RTP = rear teat placement; OUS = overall udder score; OFL = overall feet and legs; LOC = locomotion; MTY = dairy type; KOE = body score; SPR = hock quality.

an equal RPV of 30% in the 2 genomic models SSM and MSM, the conformation trait LOC was additionally selected to investigate the effect of the genomic models on genetic trends.

Genetic Trends in the Genotyped Animals

As the group of animals with lowest selection intensity, the genotyped Holstein female candidates were shown in Figure 10 with their genetic trends in STA and LOC. As described before, the trait STA was chosen as a trait with high genetic trend. For trait STA in the upper graph of Figure 10, the truncated evaluation of the ssSNPBLUP model (dotted line in green) gave slightly higher GEBV than the full ssSNPBLUP evaluation (solid line in red), though the difference was only 5% genetic standard deviation. Similarly, the truncated evaluation of the MSM model (dashed line with dots in blue) had a little bit higher genetic trend than the full evaluation (dashed line in black). The MSM model gave a slightly higher trend than the ssSNPBLUP model for trait STA. It is important to keep in mind when comparing the trends of the 2 models that STA has different RPV, 10% for MSM and 30% for SSM, between the 2 genomic models. Because of an equal 30% RPV for trait LOC between the SSM and MSM models, genetic trends of the 2 models for LOC can be directly compared. The lower graph of Figure 10 clearly shows that the SSM led to a higher genetic trend in the

genotyped female candidates than the MSM model. For the 5 selected type traits, differences in the model or data had rather limited effect on genetic trends in the genotyped female animals. In particular, for the traits CWI and RAN that had little or no trend, the model and data differences showed approximately no influence on the trends for the female animals (data not shown).

Genotyped male candidates, excluding AI bulls, showed higher genetic trend than the genotyped female candidates (data not shown) because male candidates with higher genomic parental average usually tend to get genotyped. Among the 3 genotyped groups of animals, AI bulls had the highest genetic trend in STA or UDE. Figure 11 illustrates genetic trends of UDE (upper graph) and LOC (lower graph) in the AI bulls. The truncated evaluations of trait UDE seem to have slightly lower trends in youngest AI bulls without own daughters yet [SS (solid line in red) versus SS-VAL (dotted line in green) or MS (dashed line in black) versus MS-VAL (dashed line with dots in blue)]. The difference in genetic trend between the full and truncated genomic evaluation became larger from older to younger birth years. It is worth noting that the youngest AI bull of birth year 2019 with the largest trend difference between the full and truncated evaluations may be 2 generations or more away from the reference population in the truncated evaluation. The genetic trends of trait UDE in the upper graph of Figure 11 may be a result of the following 2 factors: the difference between the 2 ge-

Figure 6. Observed correlations of GEBV of validation bulls with their deregressed EBV of the full conventional evaluations (SS-VAL_DRP = correlations of the single-step model genomic EBV with the validation data set with deregressed EBV from the conventional evaluation for the validation bulls; MS-VAL_DRP = the correlations of the multistep model GEBV with the validation data set with the deregressed EBV). STA = stature; CWI = chest width; BDE = body depth; ANG = angularity; RAN = rump angle; RWI = rump width; RLS = rear leg set; RLR = rear leg rear view; FAN = foot angle; FUA = fore udder attachment; RUH = rear udder height; USU = udder support; UDE = upper depth; FTP = front teat placement; FTL = front teat length; RTP = rear teat placement; OUS = overall udder score; OFL = overall feet and legs; LOC = locomotion; MTY = dairy type; KOE = body score; SPR = hock quality.

nomic models and unequal RPV, 10% for the MSM and 30% for the SSM model. Due to the equal RPV of 30% for trait LOC, the trend differences in the lower graph

Table 3. Numbers of genotyped AI bulls and male and female candidates of German Holstein

Year of birth	AI bulls	Male candidates	Female candidates
1998	412		
1999	477		
2000	524		
2001	512		
2002	534		
2003	1,012		
2004	1,079		
2005	882		
2006	950		
2007	907		
2008	875		
2009	495		
2010	484	5,119	
2011	486	8,151	
2012	450	10,517	
2013	396	12,235	12,607
2014	350	14,934	20,261
2015	267	14,979	28,038
2016	288	14,219	35,969
2017	310	12,382	55,608
2018	244	11,042	116,278
2019	185	11,049	134,394
2020		4,375	56,145
Total	12,119	119,002	459,300

of Figure 11 clearly show a higher genetic trend for the SSM than the MSM model. In fact, genetic trends of all the 9 conformation traits with 30% RPV between the 2 models, FAN, USU, OUS, OFL, LOC, BCS, MTY, KOE, and SPR, were higher for the SSM than for the MSM model (data not shown).

For trait STA, the differences in average GEBV in the youngest birth years were somewhat less between the truncated and full evaluation than for the trait UDE. Practically, no differences were found in genetic trends of the 2 models or the 2 data sets for the 2 traits CWI and RAN for the genotyped Holstein AI bulls (data not shown).

Variances of GEBV in the Genotyped Animals

It is expected that GEBV variances of young candidate animals without own phenotype data should be lower than genotyped animals with phenotype data. This expected trend in GEBV variance can be confirmed in Figure 12 for the genotyped German Holstein AI bulls in the following 4 scenarios: the ssSNPBLUP or MSM model with the full or truncated data sets. For AI bulls without phenotype data in birth years 2016 through 2019, GEBV variance was clearly smaller than the older daughter-proven AI bulls of birth years up to 2015 for the trait UDE. Because the ssSNPBLUP

Figure 7. Observed correlations of genomic EBV (GEBV) from the truncated evaluation for the foreign and domestic validation bulls (MS-VAL_DRP FOR and MS-VAL_DRP DEU represent the correlations of GEBV of the multistep validation evaluation with deregressed EBV for foreign and national validation bulls, respectively; SS-VAL_DRP FOR and SS-VAL_DRP DEU represent the correlations of GEBV of the single-step validation evaluation with the deregressed EBV for the foreign and national validation bulls, respectively). STA = stature; CWI = chest width; BDE = body depth; ANG = angularity; RAN = rump angle; RWI = rump width; RLS = rear leg set; RLR = rear leg rear view; FAN = foot angle; FUA = fore udder attachment; RUH = rear udder height; USU = udder support; UDE = upper depth; FTP = front teat placement; FTL = front teat length; RTP = rear teat placement; OUS = overall udder score; OFL = overall feet and legs; LOC = locomotion; MTY = dairy type; KOE = body score; SPR = hock quality.

model can use all available genotype and phenotype data, its GEBV should have larger variance than GEBV of the MSM model. This is especially evident for the youngest birth years with either the full or the truncated data sets. The trend in GEBV variance by birth year of the AI bulls seemed to be logical and to meet the expectation.

We have seen the same pattern of GEBV variances of the genotyped AI bulls for the other 3 traits and all the other remaining conformation traits. For the male or female candidates, the trend in GEBV variance by birth year was flatter, as expected, particularly for the female candidates and for the 2 traits CWI and RAN (data not shown).

Correlations of GEBV Between Evaluations

If GEBV of young candidates from an early evaluation with less complete data are highly correlated with their GEBV from a later evaluation based on more complete data, then genomic prediction is expected to be stable over time. In Figure 13, one can see that GEBV of the ssSNPBLUP model between the truncated and the full evaluation (SS_SS-VAL, solid line in red) were quite highly correlated with correlation of 1 for the years 1998 to 2012, where the reference bulls were common to the 2 evaluations. For genotyped AI bulls born between 2013 and 2015, the GEBV between

the 2 evaluations were least correlated, ranging from 0.92 to 0.95, because this group of AI bulls had no phenotype data in the truncated evaluation but had phenotype data in the full evaluation. For the youngest birth years of 2016 through 2019, the GEBV correlation was around 0.96. These genotyped young AI bulls had daughter phenotype information neither in the truncated nor in the full evaluation. Their GEBV had limited accuracy in both the full and truncated data, and both kinds of GEBV represented predictions without daughter information. Therefore, their GEBV correlations between the full and truncated data were expected to be higher than the bulls born between birth years 2013 and 2015. In comparison to the SSM model, the MSM model had very similar GEBV correlations between the 2 evaluations (MS_MS-VAL, dashed line in black) for the birth years 1998 to 2012, with common reference bulls or for the youngest birth years 2016 to 2019 with common candidates. However, the GEBV correlations for the AI bulls born in 2013 to 2015 were clearly lower than the ssSNPBLUP model. This can be explained with the same reasons as for the SNP effect correlations in Figure 3. Phenotype data of a significant number of male and female ancestors were ignored in the SNP effect estimation of the MSM model, whereas their phenotype data were accounted by the SSM model. Most of those phenotyped ancestors without genotypes were present in both the full

Figure 8. Regression coefficients of genomic EBV (GEBV) of the full on the truncated evaluations for the validation bulls (b1 MS | MS-VAL represents the regressions of the multistep GEBV of the full on the truncated validation evaluation; b1 SS | SS-VAL represents the regressions of the single-step GEBV of the full on the truncated validation evaluation). STA = stature; CWI = chest width; BDE = body depth; ANG = angularity; RAN = rump angle; RWI = rump width; RLS = rear leg set; RLR = rear leg rear view; FAN = foot angle; FUA = fore udder attachment; RUH = rear udder height; USU = udder support; UDE = upper depth; FTP = front teat placement; FTL = front teat length; RTP = rear teat placement; OUS = overall udder score; OFL = overall feet and legs; LOC = locomotion; MTY = dairy type; KOE = body score; SPR = hock quality.

and truncated phenotype data sets of the SSM model. These phenotype data of the common ancestors in the full and truncated genomic evaluations resulted in higher GEBV correlations of the SSM model for the birth years 2013 to 2015 of the genotyped AI bulls.

Based on the observed correlations of GEBV, we expected more stable genomic prediction over time using the SSM than using the MSM model. For the full data set, GEBV of the 2 models were also highly correlated, above 0.98 up to birth year 2015, and 0.96 or 0.97 for

Figure 9. Regression coefficients of deregressed conventional EBV (DRP) of the full evaluation on genomic EBV (GEBV) of the truncated evaluation for the validation bulls (b1 DRP | MS-VAL represents the regressions of deregressed EBV of the conventional evaluation on the multistep GEBV of the truncated validation evaluation; b1 DRP | SS-VAL represents the regressions of the deregressed EBV of the conventional evaluation on the single-step GEBV of the truncated validation evaluation). STA = stature; CWI = chest width; BDE = body depth; ANG = angularity; RAN = rump angle; RWI = rump width; RLS = rear leg set; RLR = rear leg rear view; FAN = foot angle; FUA = fore udder attachment; RUH = rear udder height; USU = udder support; UDE = upper depth; FTP = front teat placement; FTL = front teat length; RTP = rear teat placement; OUS = overall udder score; OFL = overall feet and legs; LOC = locomotion; MTY = dairy type; KOE = body score; SPR = hock quality.

birth years 2016 to 2019. This indicated less reranking of top animals or bulls when the MSM model was upgraded with the SSM model for routine genomic evaluation in the future.

The pattern of GEBV correlations was very similar for the other selected traits as for trait STA shown in Figure 13. For the genotyped female candidates, all the GEBV correlations between the truncated and full evaluation were 0.96 or higher, depending on birth years. Similar GEBV correlations were also found in the genotyped male candidates with older or younger

female candidates having higher or lower GEBV correlation, respectively.

Regressions of GEBV of the Genotyped Animals

Regression of GEBV of a later evaluation with complete data on those of an early evaluation with less complete data indicates an unbiased or an inflation or underestimation of GEBV variance (Mäntysaari et al., 2010; Legarra and Reverter 2018). Genomic evaluation is said to be inflated or underestimated if the regression

Figure 10. Genetic trends of stature (upper graph) and locomotion (lower graph) in genotyped German Holstein female candidates; the y-axis shows average genomic EBV (GEBV) in genetic standard deviations. MS or MS-VAL represents the average GEBV of the multistep genomic model with the full or truncated validation data; SS or SS-VAL represents the average GEBV of the single-step model with the full or truncated validation data). σ_g = SD.

coefficient is less than or greater than 1, respectively. Figure 14 shows the regression coefficients of stature GEBV from the full on the truncated evaluation for the genotyped AI bulls. The regression coefficients of the ssSNPBLUP model (SS | SS-VAL, solid line in red) are approximately 1 for the older bulls and range from 0.99 to 1.08 for the youngest birth years, suggesting neither severe inflation nor significant underestimation of genomic prediction. The same is also true for the MSM model (MS | MS-VAL, dashed line in black) with

regression coefficients varying between 0.99 and 1.04. For the full data set, GEBV regression coefficients of the MSM on the ssSNPBLUP model are all lower than 1, particularly for the youngest AI bulls.

As for the trait STA, we also found no inflation or underestimation of genomic prediction by either of the genomic models for the other 3 conformation traits, based on the regression coefficients of the genotyped AI bulls. For the genotyped male or female candidates, the regression coefficients of the ssSNPBLUP model varied

Figure 11. Genetic trends of udder depth (upper graph) and locomotion (lower graph) in genotyped German Holstein AI bulls; the y-axis shows average genomic EBV (GEBV) in genetic standard deviations; MS or MS-VAL represents average GEBV of the multistep genomic model with the full or truncated validation data; SS or SS-VAL represents average GEBV of the single-step model with the full or truncated validation data). σ_g = SD.

Figure 12. Standard deviations of genomic EBV (GEBV) of udder depth in genotyped German Holstein AI bulls; the y-axis shows standard deviations of GEBV in genetic standard deviations; MS or MS-VAL represents standard deviations of GEBV of the multistep genomic model with the full or truncated validation data; SS or SS-VAL represents standard deviations of GEBV of the single-step model with the full or truncated validation data). σ_g = SD.

between 0.98 and 1.05, also suggesting no significant over- or underestimation of GEBV by the ssSNPBLUP model (Liu et al., 2014).

Although the main investigation here was focused on the genotyped animals, we also compared GEBV of nongenotyped animals (e.g., domestic cows with own conformation records) between the MSM and SSM models. In our current MSM genomic evaluation, nongenotyped cows that were included in the preceding

conventional evaluation were not postprocessed with foreign bull MACE data. However, the nongenotyped cows were included in the SSM evaluation combining the MACE phenotype data of foreign bulls with phenotype data of national cows. For the nongenotyped cows, variance of GEBV of the single-step model was, as expected, higher than the multistep model. In fact, this was also observed for the genotyped animals. Another cause for the higher GEBV variance for the

Figure 13. Correlations of genomic EBV (GEBV) of stature between 2 evaluations for genotyped German Holstein AI bulls. MS_MS-VAL represents the GEBV correlations of the multistep model between the full and truncated validation evaluation; MS_SS represents the GEBV correlations between the multistep and single-step evaluation with full data set; SS_SS-VAL represents the correlations of the single-step model between the full and truncated validation evaluation).

Figure 14. Regression coefficients of stature genomic EBV (GEBV) of the full evaluation on the truncated evaluation for the genotyped Holstein AI bulls. MS | SS represents the GEBV regressions of the multistep on single-step evaluation with full data set; MS | MS-VAL represents the GEBV regressions of the full on truncated validation evaluation using the multistep model; SS | SS-VAL represents the GEBV regressions of the full on truncated validation evaluation using the single-step model.

nongenotyped cows was the inclusion of deregressed MACE EBV of all foreign bulls in the single-step evaluation. In contrast, the MACE phenotype data of all the foreign bulls were considered only in the calculation of conventional parental average or EBV and in the SNP effect estimation or DGV calculation in the multistep genomic evaluation based on the conventional evaluations. Similar to the nongenotyped cows, ancestors of the genotyped or phenotyped animals in the conformation genomic evaluation had also higher variance of GEBV in the single-step than the multistep evaluation.

A total of 23 conformation traits were evaluated for German Holsteins using the single-step SNP BLUP model (Liu et al. 2014), with 875,252 genotyped animals and more than 2.7 million cows with phenotype records. Deregressed MACE EBV of foreign bulls were integrated as a new source of data of the same trait as the national conformation data. Both software packages MiX99 and MiXBUP were shown to give identical solutions for the ssSNPBLUP model. We conducted a detailed investigation on GEBV of the genotyped AI bulls and genotyped male and female candidates. Genetic trends, GEBV variances, and correlations and regressions between the 2 evaluations were provided for each of the 3 animal groups. For the youngest genomic AI bulls or validation bulls, GEBV correlation between the 2 models ssSNPBLUP and MSM was, on average, 0.95 for all the conformation traits. We have shown that the ssSNPBLUP model resulted in more accurate genomic prediction than the current MSM model. As found in our study, the ssSNPBLUP can efficiently

process genotype data of tens of millions without making approximation in genotype information by using the compressed PLINK BED storage format of SNP genotypes.

CONCLUSIONS

We demonstrated, based on the GEBV of validation bulls, that the single-step model had a higher prediction correlation and a greater GEBV variance than the current multistep genomic model. For the national validation bulls, regression coefficient of GEBV from the full evaluation on the truncated evaluation was close to 1 for any of the conformation traits. In addition, the regression coefficient of the deregressed EBV on GEBV of the truncated evaluation did not differ notably from its expected value of 1. Genetic trends of the single-step model were shown to be higher than the multistep model for the current multistep genomic model, based on the 7 conformation traits with an equal RPV between the 2 models. The single-step model resulted in higher GEBV correlations between the full and truncated evaluation for the genotyped animals than the multistep genomic model. Regression coefficients indicated no over- or underestimation in genomic prediction for these animals. It is assuring that we found no overprediction of young animals by the single-step genomic model for the conformation traits. The current multistep genomic model was based on conventional evaluation that ignored the genomic selection; therefore, a biased prediction could not be avoided in the current genomic evaluation. We

have clearly demonstrated that an implementation of the ssSNPBLUP model in routine genomic evaluation of dairy cattle is computationally feasible, and more accurate genomic prediction is expected with the ssSNPBLUP model.

ACKNOWLEDGMENTS

Jeremie Vandenplas and Jan ten Napel from University of Wageningen (Wageningen, Netherlands) and Esa Mäntysaari and Ismo Strandén from the Finnish research institute LUKE (Helsinki, Finland) are greatly thanked for their help in using the software packages MiXBLUP and MiX99, respectively. The authors have not stated any conflicts of interest.

REFERENCES

- Aguilar, I., I. Misztal, D. L. Johnson, A. Legarra, S. Tsuruta, and 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.* 93:743–752. <https://doi.org/10.3168/jds.2009-2730>.
- Alkhoder, H., Z. Liu, F. Reinhardt, and R. Reents. 2017. Impact of adding cows to the current EuroGenomics bull reference population on genomic prediction. *Interbull Bull.* 51:86–90.
- Andersen, T., U. S. Nielsen, G. P. Aamand, E. Mäntysaari, J. Pösö, and E. Rius-Vilarrasa. 2021. Using single-step genetic evaluation for type traits in the Nordic countries. *Interbull Bull.* 56:90–93.
- Christensen, O. F., and M. S. Lund. 2010. Genomic prediction when some animals are not genotyped. *Genet. Sel. Evol.* 42:2. <https://doi.org/10.1186/1297-9686-42-2>.
- Gao, H., M. Koivula, J. Jensen, I. Strandén, P. Madsen, T. Pitkanen, G. P. Aamand, and E. Mäntysaari. 2018. Genomic prediction using different single-step methods in the Finnish red dairy cattle population. *J. Dairy Sci.* 101:10082–10088. <https://doi.org/10.3168/jds.2018-14913>.
- Hsu, W. L., D. J. Garrick, and R. L. Fernando. 2017. The accuracy and bias of single-step genomic prediction for populations under selection. *G3 (Bethesda)* 7:2685–2694. <https://doi.org/10.1534/g3.117.043596>.
- Legarra, A., and V. Ducrocq. 2012. Computational strategies for national integration of phenotypic, genomic, and pedigree data in a single-step best linear unbiased prediction. *J. Dairy Sci.* 95:4629–4645. <https://doi.org/10.3168/jds.2011-4982>.
- Legarra, A., and A. Reverter. 2018. Semi-parametric estimates of population accuracy and bias of predictions of breeding values and future phenotypes using the LR method. *Genet. Sel. Evol.* 50:53. <https://doi.org/10.1186/s12711-018-0426-6>.
- Liu, Z., H. Alkhoder, Ch. Kubitz, K. Stock, J. Heise, D. Segelke, F. Reinhardt, and R. Reents. 2019. Genomic prediction of health traits using a mixed reference bull and cow reference population for German Holsteins. *Interbull Bull.* 55:35–45.
- Liu, Z., M. E. Goddard, F. Reinhardt, and R. Reents. 2014. A single-step genomic model with direct estimation of marker effects. *J. Dairy Sci.* 97:5833–5850. <https://doi.org/10.3168/jds.2014-7924>.
- Liu, Z., F. R. Seefried, R. Reinhardt, S. Rensing, G. Thaller, and R. Reents. 2011. Impacts of both reference population size and inclusion of a residual polygenic effect on the accuracy of genomic prediction. *Genet. Sel. Evol.* 43:19. <https://doi.org/10.1186/1297-9686-43-19>.
- Mäntysaari, E. A., Z. Liu, and P. M. VanRaden. 2010. Interbull validation test for genomic evaluations. *Interbull Bull.* 41:10–14.
- Mäntysaari, E. A., and I. Strandén. 2016. Single-step genomic evaluation with many more genotyped animals. Page 366 in *Proceedings of the 67th annual meeting of the European Association for Animal Production*. EIP-AGRI. Wageningen Academic Publishers.
- Masuda, Y., I. Misztal, A. Legarra, S. Tsuruta, D. A. L. Lourenco, B. O. Fragomeni, and I. Aguilar. 2017. Technical note: Avoiding the direct inversion of the numerator relationship matrix for genotyped animals in single-step genomic best linear unbiased prediction solved with the preconditioned conjugate gradient. *J. Anim. Sci.* 95:49–52.
- Meuwissen, T. H. E., B. J. Hayes, and M. E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157:1819–1829. <https://doi.org/10.1093/genetics/157.4.1819>.
- Misztal, I., D. Lourenco, and A. Legarra. 2020. Current status of genomic evaluation. *J. Anim. Sci.* 98:ska101. <https://doi.org/10.1093/jas/skaa101>.
- Misztal, I., Z. Vitezica, A. Legarra, I. Aguilar, and A. Swan. 2013. Unknown-parent groups in single-step genomic evaluation. *J. Anim. Breed. Genet.* 130:252–258. <https://doi.org/10.1111/jbg.12025>.
- Reents, R. 1993. Estimation of breeding values for type traits in Germany. *Interbull Bull.* 8:16.
- Schaeffer, L. R., J. Jamrozik, G. J. Kistemaker, and B. J. Van Doormaal. 2000. Experience with a test day model. *J. Dairy Sci.* 83:1135–1144. [https://doi.org/10.3168/jds.S0022-0302\(00\)74979-4](https://doi.org/10.3168/jds.S0022-0302(00)74979-4).
- Strandén, I., and M. Lidauer. 1999. Solving large mixed linear models using preconditioned conjugate gradient iteration. *J. Dairy Sci.* 82:2779–2787. [https://doi.org/10.3168/jds.S0022-0302\(99\)75535-9](https://doi.org/10.3168/jds.S0022-0302(99)75535-9).
- Strandén, I., and E. A. Mäntysaari. 2014. Comparison of some equivalent equations to solve single-step GBLUP. *Proc. 10th World Congr. Genet. Appl. Livest. Prod.*
- Ten Napel, J., J. Vandenplas, M. Lidauer, I. Strandén, M. Taskinen, E. Mäntysaari, M. P. Calus, and R. F. Veerkamp. 2020. MiXBLUP user's guide.
- Tsuruta, S., T. J. Lawlor, D. A. L. Lourenco, and I. Misztal. 2021. Bias in genomic predictions by mating practices for linear type traits in a large-scale genomic evaluation. *J. Dairy Sci.* 104:662–677. <https://doi.org/10.3168/jds.2020-18668>.
- Vandenplas, J., M. P. L. Calus, H. Eding, and C. Vuik. 2019. A second-level diagonal preconditioner for single-step SNPBLUP. *Genet. Sel. Evol.* 51:30. <https://doi.org/10.1186/s12711-019-0472-8>.
- Vandenplas, J., H. Eding, M. Bosmans, and M. P. L. Calus. 2020. Computational strategies for the preconditioned conjugate gradient method applied to ssSNPBLUP, with an application to a multivariate maternal model. *Genet. Sel. Evol.* 52:24. <https://doi.org/10.1186/s12711-020-00543-9>.
- Vandenplas, J., H. Eding, and M. P. L. Calus. 2021. Technical note: Genetic groups in single-step single nucleotide polymorphism best linear unbiased predictor. *J. Dairy Sci.* 104:3298–3303. <https://doi.org/10.3168/jds.2020-19460>.
- VanRaden, P. M. 2008. Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91:4414–4423. <https://doi.org/10.3168/jds.2007-0980>.

ORCIDS

- H. Alkhoder <https://orcid.org/0000-0001-5956-9355>
 Z. Liu <https://orcid.org/0000-0002-7112-0320>
 D. Segelke <https://orcid.org/0000-0002-7351-530X>
 R. Reents <https://orcid.org/0000-0003-1481-7436>