Application of a single-step SNP BLUP model to conformation traits of German Holsteins

H. Alkhoder and Z. Liu

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

Abstract

Genomic evaluation based on a single-step model utilizes all available data of phenotype, genotype and pedigree and therefore provides unbiased genomic prediction with a higher accuracy than the current multi-step genomic model. Until today, a mixed reference population of cows and bulls has been applied to the routine multi-step genomic evaluation in German Holsteins. For a fair comparison between the single-step and multi-step genomic models, the same phenotype, genotype and pedigree data were used. Because of the standard multi-trait 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 (Liu-Goddard) for the large, genotyped population of German Holstein. 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 MACE evaluation for the conformation traits, deregressed MACE EBV of foreign bulls were treated as a new source of data for the same trait in the single-step evaluation. Due to a short history of female genotyping, last three years of youngest cows and bulls were deleted, instead of four years, for performing a genomic validation. In comparison to the multi-step genomic model, the single-step SNP BLUP model resulted in a higher prediction accuracy and greater GEBV variance 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 multi-step model. For the validation bulls or youngest genomic AI bulls, correlation of GEBV between the two models was, on average, 0.95 across all the conformation traits. We found no major concern about a possible over-prediction of young animals by the single-step SNP BLUP model for the conformation traits in German Holsteins.

Key words: single-step model, SNP BLUP model, conformation traits, dairy cattle

Introduction

Genomic evaluation (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; Liu et al. 2014) utilizes all available information on phenotype, genotype and pedigree and therefore provides unbiased genomic prediction.

For routine genomic evaluation in German Holsteins, a multi-step SNP BLUP model with a residual polygenic model (RPG, Liu et al. 2011) is 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 multi-step 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 evaluations have become evident with largest underestimation for youngest animals.

In comparison to single-step genomic BLUP model (ssGBLUP, Aguilar et al. 2010), 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 several alternative ssSNPBLUP models and confirmed the Liu-Goddard single-step model (Liu et al. 2014) being most efficient. Therefore, the variant of single-step model, Liu-Goddard ssSNPBLUP model, has been chosen for German Holstein genomic evaluation.

Routine conventional and genomic evaluations for German Holsteins cover a wide range of trait groups all evaluated with multi-trait animal models. The statistical model for genetic evaluation of conformation traits is a standard multi-trait animal model (Reents, 1993) and represents the simplest form of the multi-trait animal model. Thus, we decided to use the conformation traits for testing the ssSNPBLUP model as the first of 13 trait groups in German Holsteins.

The objectives of this study were 1) to test the implementation of Liu-Goddard single-step model for conformation traits, 2) to compare genomic evaluations using the single-step model and the current multi-step SNP BLUP model, and 3) to investigate prediction accuracy and bias via a truncated validation data set for the two genomic models.

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 the subsequent multi-step genomic evaluation. The conformation traits recorded on first lactation cows in Germany included all 21 conformation traits covered in Interbull MACE evaluation, except overall conformation score (OCS), and three national only type traits. These 23 conformation traits were evaluated in three independent groups: body and rump with 9 traits: stature (STA),

chest width (CWI), body depth (BDE), angularity (ANG), rump angle (RAN), rump width (RWI), body condition score (BCS), dairy type (MTY), and body score (KOE); udder with 8 traits: fore udder attachment (FUA), rear udder height (RUH), udder support (USU), udder depth (UDE), front teat placement (FTP), front teat length (FTL), rear teat placement (RTP), overall udder score (OUS); and feet and legs with 6 traits: rear leg set (RLS), rear leg rear view (RLR), foot angle (FAN), overall feet and leg score (OFL), locomotion (LOC), and hock quality (SPR). Each of the three trait groups was evaluated using a multi-trait animal model (Reents, 1993). Two conformation traits with strong genetic trend, STA and UDE, and two traits with weak or flat genetic trend, 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*year effects was performed prior to solving all effects of the models. The Liu-Goddard ssSNPBLUP model was applied to the conformation traits:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{u} + \mathbf{e} \qquad [1]$$

where $\mathbf{y}$ is a vector of conformation trait values adjusted for the heterogeneous classifier*year variances, $\mathbf{b}$ is a vector of all fixed effects including the major fixed effect of herd-classification-date, age at calving x lactation stage, classifier x year of classification x daughter type, $\mathbf{X}$ is the incidence matrix for all the fixed effects, $\mathbf{u}$ is a vector of GEBV 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 all MACE bulls were used here as pseudo-phenotype (Liu, 2011). 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. Please note that the deregression of the bull MACE EBV was done one trait at a time. Because the DRP were free of the fixed effects of model [1], a single pseudo-ID was assigned to each of the fixed effects for the DRP records.

The Liu-Goddard ssSNPBLUP allows fitting a RPG effect in the single-step SNP BLUP model. We assumed that the RPG explained 30% additive genetic variance for each of the 23 conformation traits. A J factor per trait was included in ssSNPBLUP model [1] to account for different genetic levels of the genotyped and non-genotyped animals (Hsu et al. 2017). For SNP effects of the model [1] genetic correlations between traits within each of the three trait groups were considered, whereas the multi-step genomic model (MSM) on a single-trait basis (Liu et al. 2011) ignored the genetic correlations between traits in the SNP effect estimation.

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 Quaas-Pollak transformation (Vandenplas et al. 2021). The software MiXBLUP (Ten Napel et al. 2020) was used for the single-step evaluation based on the Liu-Goddard model. The Liu-Goddard ssSNPBLUP model was also implemented in software MiX99 (Strandén and Lidauer, 1999) in a special way (Mäntysaari 2018, personal

communication). SNP markers were treated as if they were animals with neither known parents nor progeny in this special implementation. 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.

To compare with the ssSNPBLUP model, the current multi-step genomic model (Liu et al. 2011) was applied to DRP of the reference cows and MACE DRP of the bulls (Alkhoder et al. 2017). Both the single-step ssSNPBLUP and the multi-step genomic model MSM started from the same phenotype and genotype data. Non-genotyped cows with phenotypes were not able to be considered in the multi-step genomic evaluation if those animals were not linked to the genotyped animals. However, DRP of all bulls from the MACE evaluation were considered in the parental average calculation for GEBV calculation of all genotyped animals in the multi-step genomic evaluation. In contrast, the single-step model was able to include all animals with phenotypes or genotypes in genomic evaluation.

A validation study for the single-step model

Due to a relatively short history of female animal genotyping in German Holsteins (Liu et al. 2019), only the youngest three birth years of cows with type records 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 validation bulls were selected. Since daughters of the validation bulls may be included in the truncated reference population for 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 1 shows the numbers of animals with phenotype data of conformation traits for the full and truncated evaluations. A total number of 875,252 genotyped Holstein animals were considered, including culled male candidates. Pedigree file for the genotyped and phenotyped animals contained 9,012,965 animals for the ssSNPBLUP full evaluation. Unknown parent groups (UPG) were defined according to breeds and country origins, four selection paths and birth years of animals. The number of UPG was 138 for the full evaluation. 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.

Table 1. Description of phenotype data sets for a full and a truncated evaluation of all 23 type traits in German Holstein

	Full data set	Truncated set
Cows with type records	2,715,550	2,454,801 (-260,749)
§Bulls with MACE data	115,552	112,605 (-5,945)

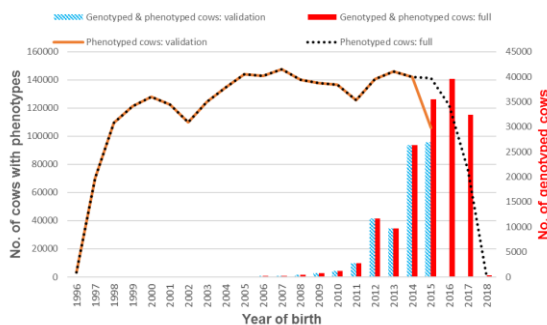§ Bulls must have daughters outside Germany

Figure 1. Numbers of phenotyped or genotyped cows by birth year for the full and truncated evaluation

Results & Discussion

The single-step genomic evaluation was performed using the full and truncated data sets with the software MiXBLUP. The total number of estimated effects or equations for the full evaluation was 217,423,347. A total of 3387 rounds of iteration were needed 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 hours was required for the full evaluation and the memory usage was 65Gb VmPeak and 39 Gb VmHWM.
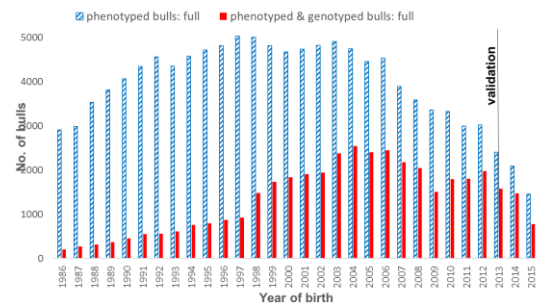

Figure 2. Numbers of bulls with phenotype or genotype data by birth year for the full and truncated evaluation

Comparison of the two software packages

Both MiXBLUP and MiX99 were tested for the ssSNPBLUP model using the same data of the conformation traits. The two software packages differed in computational efficiency for the Liu-Goddard ssSNPBLUP model, because the implementations of the single-step model were different. A second-level pre-conditioner (Vandenplas et al. 2019) was used in both software programs. With a 2-bit representation of SNP genotypes, MiXBLUP showed some advantage in memory usage as well as computing time for both pre-processing and solving steps.

Despite the differences in computational efficiency, the two software packages gave identical estimates of all the model effects. Correlation of SNP effects between the two 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 can conclude that both MiXBLUP and MiX99 resulted in equal effect estimates for the conformation traits of German Holsteins. It is worth noting that the 2-bit representation of SNP genotypes can enable processing possibly tens of millions of genotyped animals with the Liu-Goddard ssSNPBLUP model.

As the two software packages led to identical evaluation results, we present here only results from the software MiXBLUP.

SNP effect estimates

The Liu-Goddard ssSNPBLUP model directly estimated SNP effects. Figure 3 shows the observed correlations between SNP effect estimates of the two datasets or between the two genomic models. For the ssSNPBLUP model, SNP effect correlations (SS_SS-VAL, the solid line in red) ranged from 0.91 to 0.95 between the full and truncated evaluations 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 (MS_MS-VAL, the dotted line in green) were much lower, between 0.80 and 0.91 with an average of 0.86, between the full and truncated evaluations. The much lower SNP effect correlations can be explained by the fact that many non-genotyped cows or bulls with own phenotype data were not considered by the MSM but by the ssSNPBLUP model for the estimation of the SNP effects in both the full and truncated data sets. For the full data set, SNP effect estimates were correlated between the two models (MS_SS, dashed line with dots in black) with an average of 0.82, ranging from 0.77 to 0.88. The three national type traits, MTY, KOE, and SPR, had the lowest correlations between the two models. Similar SNP effect correlations were also observed for the truncated validation data set between the
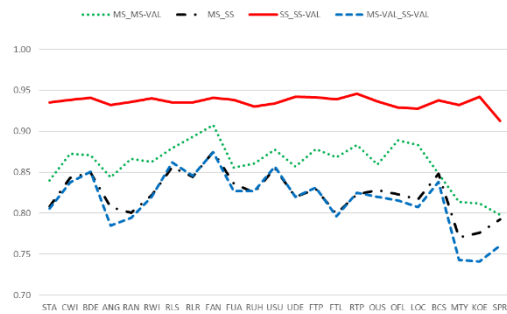
two models (MS-VAL_SS-VAL, dashed line in blue).

Figure 3. Observed correlations between SNP effect estimates from the full and truncated evaluations

Figure 4 shows the regression coefficients of SNP effect estimates of the full evaluation on the truncated evaluation for each of the two 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 has 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 is 0.927. The regression coefficients indicate that the MSM slightly over-predicted 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.

Figure 4. Regression coefficients of SNP effect estimates of the full on the truncated evaluations

Accuracy of genomic prediction

GEBV correlation of validation animals between an early, truncated and a later, full evaluation is a measure of accuracy of genomic prediction. GEBV of all the 798 validation bulls were compared between the full and truncated evaluations and between the two genomic models. It can be seen clearly in Figure 5 that the ssSNPBLUP model resulted in significantly 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 two 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 two models for the truncated data set (MS-VAL_SS-VAL, dashed line in blue).
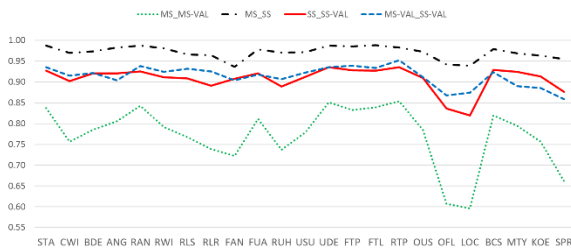

Figure 5. Observed correlations of GEBV of validation bulls between the full and the truncated evaluations

We also investigated the correlation of GEBV of the validation bulls with their deregressed EBV. Please keep in mind that DRP of these validation bulls were calculated from the conventional evaluation, not from the single-step evaluation. DRP of the validation bulls should be less auto-correlated 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). Overall, the GEBV of the truncated evaluation are slightly lower correlated with their DRP than with their GEBV of the full evaluation in Figure 5.
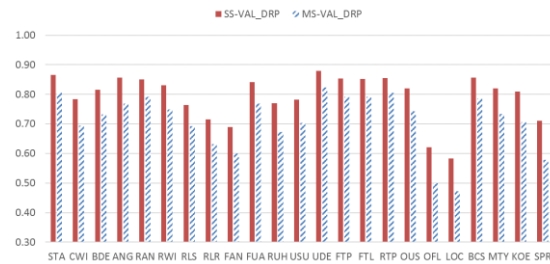

Figure 6. Observed correlations of GEBV of validation bulls with their deregressed EBV of the full conventional evaluation

If foreign validation bulls had no daughters in Germany, their deregressed MACE EBV from the current MACE evaluation should have lower reliability than the German national validation bulls. For all conformation traits included in MACE evaluation, except the total overall conformation OCS, we compared the accuracy of genomic prediction between the foreign and domestic validation bulls 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 DEU (dotted line in green) vs SS-VAL_DRP FOR (solid line in red). 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 showed a higher correlation of GEBV with DRP for any of the traits than the foreign validation bulls: MS-VAL_DRP DEU (dashed line with dots in black) vs MS-VAL_DRP FOR (dashed line in red). But the average difference in the correlation is 0.05, being less than the average difference for the ssSNPBLUP model.
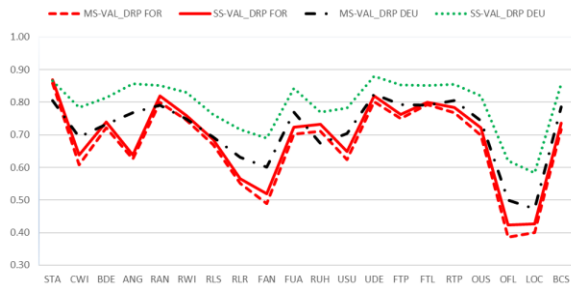
Figure 7. Accuracy of genomic prediction for the foreign and domestic validation bulls

Dispersion of genomic prediction

Regression of GEBV of the validation bulls from the full evaluation on GEBV of the early, truncated evaluation gives an indication whether the genomic prediction is 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) are close to 1, in fact 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) is slightly lower, 0.98, also being close to 1. We can conclude that neither the ssSNPBLUP nor the MSM model results in an over-prediction or underestimation of genomic evaluation.
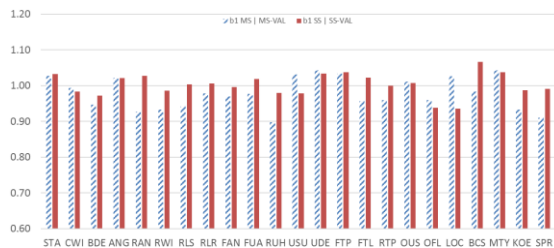

Figure 8. Regression coefficients of GEBV of validation bulls from the full on the truncated evaluations

Using the pseudo-phenotype 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 is a greater variation in the regression coefficients. The average of regression coefficients is 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.
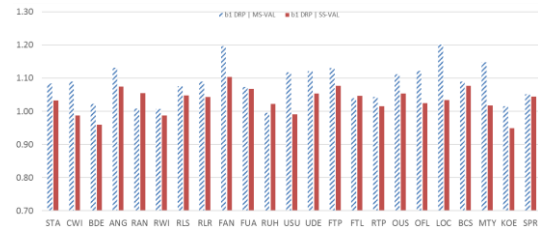

Figure 9. Regression coefficients of deregressed conventional EBV (DRP) from the full evaluation on GEBV of the truncated evaluation for the validation bulls

Averages and variances of GEBV

Genotyped German Holstein animals were chosen for comparing trends and variances of GEBV between the ssSNPBLUP and MSM models. Table 2 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 can be considered an unselected sample, thanks to the whole-herd genotyping project KuhVision in Germany (Liu et al. 2019). Two conformation traits, stature STA and udder depth UDE, were chosen representing conformation traits with a high genetic trend, whereas the other two selected traits, chest width CWI and rump angle RAN represented the type traits with a flat genetic trend.

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 stature STA. The secondary Y axis shows the number of genotyped female animals in blue bars. The truncated evaluation

Table 2. Numbers of genotyped AI bulls, 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	1012		
2004	1079		
2005	882		
2006	950		
2007	907		
2008	875		
2009	495		
2010	484	5119	
2011	486	8151	
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		4375	56,145
Total	12,119	119,002	459,300

of the ssSNPBLUP model (dotted line in green) gives slightly higher GEBV than the full ssSNPBLUP evaluation (solid line in red), though the difference being only 5% genetic standard deviation. Similarly, the truncated evaluation of the MSM model (dashed line with dots in blue) has a little bit higher genetic trend than the full evaluation (dashed line in black). The MSM model gave slightly higher trend than the ssSNPBLUP model. For the four selected type traits, differences in the model or data had rather limited impact on genetic trends in the genotyped female animals. Particularly for the traits CWI and RAN with little or no trend, the model and data differences had nearly no impact on the trends for the female animals.

Genotyped male candidates, excluding AI bulls, showed higher genetic trend than the genotyped female candidates shown in Figure

10, because only male candidates with higher genomic parental average usually got genotyped. Among the three genotyped groups of animals, AI bulls have the highest genetic trend in STA or UDE. Figure 11 shows genetic trends of UDE in the AI bulls. The truncated evaluations seem to have lower trends in youngest AI bulls without own daughters: SS (solid line in red) vs SS-VAL (dotted line in green) or MS (dashed line in black) vs MS-VAL (dashed line with dots in blue). It is worth noting that the youngest AI bull of birth year 2019 may be three generations or more away from the reference population in the truncated evaluation.
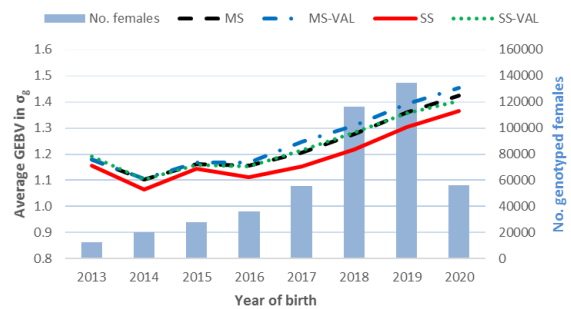

Figure 10. Genetic trends of stature (STA) in genotyped German Holstein female candidates

For the 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 two models or the two data sets for the two traits CWI and RAN for the genotyped Holstein AI bulls.
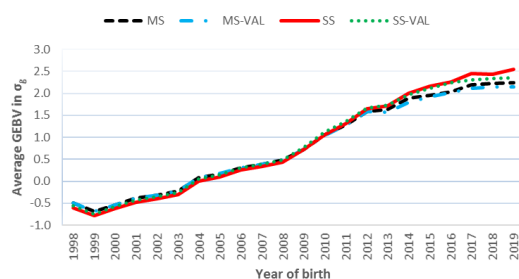

Figure 11. Genetic trends of udder depth (UDE) in genotyped German Holstein AI bulls

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 four 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 is clearly smaller than the older AI bulls with daughters of birth years up to 2015 for the trait UDE. Because the ssSNPBLUP model utilizes 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 set. The trend in GEBV variance by birth year of the AI bulls seems to be logical and meets the expectation.

We have seen the same pattern of GEBV variances of the genotyped AI bulls for the other three 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 two traits CWI and RAN.
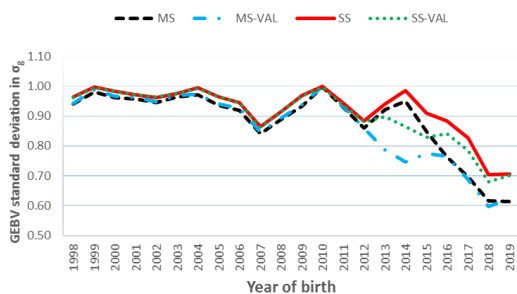

Figure 12. Standard deviations of GEBV of udder depth (UDE) in genotyped German Holstein AI bulls

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) are quite highly correlated with correlation being 1 for the years 1998 to 2012 in which the reference bulls are common to the two evaluations. For the youngest birth years of 2016 through 2019, the GEBV correlation is c.a. 0.96. For genotyped AI bulls born between 2013 and 2015, the GEBV between the two evaluations are least correlated, ranging from 0.92 to 0.95, because this group of AI bulls have no phenotype data in the truncated but in the full evaluation. In comparison to the single-step model, the MSM model has very similar GEBV correlations between the two 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 are significantly lower than the ssSNPBLUP model. This can be explained by the fact that the multi-step model MSM ignores genotypes of the candidates for SNP effect estimation, whereas the ssSNPBLUP considers the reference and candidate animals jointly. Based on the observed correlations of GEBV, we expect more stable genomic prediction over time using the ssSNPBLUP than using the MSM model. For the full data set, GEBV of the two models are also highly correlated, above 0.98 up to birth year 2015, and 0.96 or 0.97 for birth years 2016 to 2019. This indicates less reranking of top animals or bulls when the MSM model is replaced with the ssSNPBLUP model for routine genomic evaluation in future.
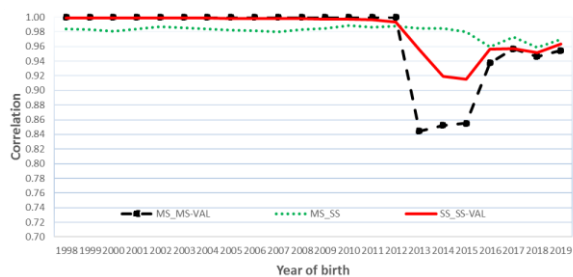
Figure 13. Correlations of GEBV of stature (STA) between two evaluations for genotyped German Holstein AI bulls

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. Genomic evaluation is said to be inflated or underestimated if the regression coefficient is less 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 nearly 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.
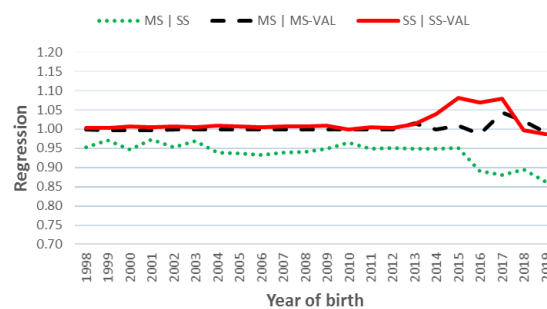

Figure 14. Regression coefficients of stature GEBV of the full evaluation on the truncated evaluation for the genotyped Holstein AI bulls

As for the trait STA, we also found no inflation or underestimation of genomic prediction by either of the genomic models for the other three 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 between 0.98 and 1.05, also suggesting no over- or underestimation of GEBV by the single-step Liu-Goddard ssSNPBLUP model.

Conclusions
Single-step genomic model uses all available data on phenotype, genotype, and pedigree for genomic prediction. The Liu-Goddard single-step SNP BLUP model directly estimates the effects of SNP markers and was shown to be most efficient among the variants of the single-step model in literature and feasible for processing up to millions of genotyped animals using the 2-bit representation of SNP genotypes. A total of 23 conformation traits were evaluated for German Holstein using the ssSNPBLUP model, with 875,252 genotyped animals and more than 2.7 millions of cows with phenotype records. The multi-trait model included a J factor to account for different genetic levels in the genotyped and non-genotyped animals. Deregressed MACE EBV of foreign bulls were integrated as a new source of data of the same trait as the national conformation data. Thanks to the efficient memory usage of the Liu-Goddard model, it

was feasible to conduct the single-step evaluation with the very large genotype data within reasonable time. Both software packages MiX99 and MiXBLUP were shown to give identical solutions for the ssSNPBLUP model. Genotype and phenotype data were truncated to simulate an early evaluation for assessing the accuracy and bias of the ssSNPBLUP model. Based on the GEBV of validation bulls, we showed that the single-step model had a higher accuracy and a greater GEBV variance than the current multi-step genomic model. For the national validation bulls, regression coefficient of GEBV from the full on the truncated evaluation was close to 1 for any of the conformation traits. In addition, the regression coefficient of their deregressed EBV on GEBV of the truncated evaluation did not differ significantly from its expected value of 1. We conducted a detailed investigation on GEBV of the genotyped AI bulls, genotyped male and female candidates. Genetic trends, GEBV variances, and correlations and regressions between the two evaluations were provided for each of the three animal groups. The single-step model resulted in higher GEBV correlations between the full and truncated evaluation for the genotyped animals than the multi-step genomic model. Regression coefficients indicated no over- or underestimation in genomic prediction for these animals. For the youngest genomic AI bulls or validation bulls, GEBV correlation between the two models ssSNPBLUP and MSM was, on average, 0.95 for all the conformation traits. It is assuring that we found in our study no over-prediction of young animals by the single-step genomic model for the conformation traits.

References

Alkhoder, H., Z. Liu, F. Reinhardt, and R. Reents. 2017. *Interbull Bulletin* 51:86-90.

Aguilar, I., I. Misztal, D. L. Johnson, A. Legarra, S. Tsuruta, and T. J. Lawlor. 2010. *J. Dairy Sci.* 93:743–752.

Hsu, W. L., D. J. Garrick, and R. L. Fernando. 2017. *G3* (Bethesda) 7:2685-2694.

Liu, Z. 2011. Use of MACE Results as Input for Genomic Models. *Interbull Bulletin* 43.

Liu, Z., F. R. Seefried, R. Reinhardt, S. Rensing, G. Thaller, and R. Reents 2011. *Genet. Sel. Evol.* 43:19.

Liu, Z., M. E. Goddard, F. Reinhardt, and R. Reents. 2014. *J. Dairy Sci.* 97:5833-5850.

Liu, Z., H. Alkhoder, Ch. Kubitz, K. Stock, J. Heise, D. Segelke, F. Reinhardt, and R. Reents. 2019. *Interbull Bulletin* 55:35-45.

Meuwissen, T. H. E., B. J. Hayes, and M. E. Goddard. 2001. *Genetics* 157:1819-1829.

Reents, R. 1993. *Interbull Bulletin* 8.

Strandén, I. and M. Lidauer. 1999. *J. Dairy Sci.* 82:2779-2787.

Ten Napel, J., J. Vandenplas, M. Lidauer, I. Stranden, M. Taskinen, E. Mäntysaari, M. P. Calus, and R. F. Veerkamp. 2020 MiXBLUP user's guide.

Vandenplas, J., M. P. L. Calus, H. Eding, and C. Vuik. 2019. *Genet. Sel. Evol.* 51:30

Vandenplas, J., H. Eding, M. Bosmans, and M. P. L. Calus. 2020. *Genet. Sel. Evol.* 52:24.

Vandenplas, J., Eding, H., and Calus, M. P. L. 2021. *J. Dairy Sci.* 104:3298-3303.

VanRaden, P. M. 2008. *J Dairy Sci.* 91:4414-4423.