Application of a single-step SNP BLUP random regression model to test-day yields and somatic cell scores 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

Test-day yields or somatic cell scores have been evaluated with a multi-lactation random regression test-day model for dairy cattle breeds in Germany. The Liu-Goddard single-step SNP BLUP model directly estimates the effects of SNP markers together with all other model effects and has been demonstrated to be most efficient in handing large genotype data among all variants of the single-step model. The aims of this study were to test the implementation of the single-step SNP BLUP model to the test-day traits in German Holsteins and to investigate accuracy and bias of the genomic prediction. Approximately one million genotyped Holstein animals were jointly evaluated with c.a. 12 million dairy cows having test-day data. Pseudo-phenotype data of more than 138,000 Holstein bulls were integrated as a correlated trait to the national test-day yields or somatic cell scores. A genomic validation was conducted by removing test-day records in last four years for the national cows and truncating the youngest four birth years of the integrated bulls. The single-step model gave higher correlation of the SNP effect estimates between the full and truncated evaluation than the current multi-step model. In addition, regression coefficient of the SNP effect estimates from the full on the truncated evaluation was closer to 1 for the single-step model. Based on the results for the validation bulls we can draw a conclusion that the single-step model leads to neither an inflation nor a deflation of genomic prediction for the four test-day traits. No post-processing of GEBV of young animals would be needed for the genomic prediction in German Holsteins. The impact of selecting bulls with foreign daughters was investigated. We have found that removing genotype records of older bulls led to average Mendelian sample effects closer to zero for genotyped female animals.

Key words: single-step SNP BLUP model, random regression model, test-day traits, dairy cattle

Introduction

A single-step SNP BLUP model (ssSNPBLUP, Liu et al. 2014) directly estimates the effects of SNP markers together with all other effects of the single-step model (SSM). The ssSNPBLUP model was successfully applied to conformation traits of German Holsteins (Alkhoder et al. 2022) and promising validation results were obtained for all the conformation traits.

Test-day milk, fat and protein yields and somatic cell scores have been evaluated for German dairy cattle breeds using a random regression test-day model (RRTDM, Liu et al. 2004). Combined lactation breeding values on 305-day basis, defined as a linear function of random regression coefficients (RRC), are published routinely for the four test-day traits and submitted to Interbull bull MACE evaluations as the German official EBV of bulls. MACE EBV of all bulls on the German scale have been deregressed using an iterative method by Jairath et al. (1998). The deregressed EBV (DRP) of bulls have been used as pseudo-phenotype for e.g. foreign bulls in the current German multi-step genomic evaluation. Likewise, the DRP can be used to integrate foreign phenotype information of bulls into the national single-step evaluation.

The objectives of this study were 1) to implement the ssSNPBLUP model to the four test-day traits, 2) to conduct a genomic

validation for assessing the accuracy and bias of the single-step evaluation, and 3) to compare predicative ability of the ssSNPBLUP model to the current multi-step genomic model (MSM).

Materials and Methods

A single-step SNP BLUP random regression test-day model

For German dairy breeds like Holsteins, test-day traits in first three lactations were analyzed as correlated traits and separately for the four traits: milk yield (MIL), fat yield (FAT), protein yield (PRO) or somatic cell scores (SCS) (Liu et al. 2004). Legendre polynomials with three parameters were used to model genetic or permanent environmental lactation curves. Test-day yields or SCS of a genotyped cow were evaluated with a single-step SNP BLUP RRTDM model:

$$\mathbf{y} = \mathbf{X}_h\mathbf{h} + \mathbf{X}_f\mathbf{f} + \mathbf{p} + \mathbf{Zg} + \mathbf{a} + \mathbf{e} \quad [1]$$

where **y** is a vector of the cow's test-day yields or SCS in first three lactations that were pre-adjusted for heterogeneous variances by herd-test-date-parity-milking-frequency, **h** is a vector of fixed effects of herd-test-date-parity-milking-frequency associated with her test-day records, **f** is a vector of fixed lactation curve effects, modelled as regressions on days in milk (DIM) using Wilmink function (Liu et. al. 2004), **p** is a vector of permanent environmental effects that are expressed as RRC, **g** is a vector of all SNP marker effects also in form of RRC, **a** is a vector of residual polygenic effects (RPG) expressed in RRC, **e** is a vector of random error effects. The random effects **p** and **a** have 3 RRC for each of the three lactations, leading to a total of 9 RRC for each of the effects. For every SNP marker, there are 9 RRC to be estimated for every trait. $\mathbf{X}_h$ is the incidence matrix for the fixed herd-test-date-parity-milking-frequency effects associated with this cow's test-day records, $\mathbf{X}_f$ is the incidence matrix for the fixed lactation curve effects, **Z** is the design matrix containing genotype data of the cow. RPG variance was assumed to be 30% of the total additive genetic variance for any of the test-day traits.

If the cow is not genotyped, then model 1 is simplified to a regular RRTDM model:

$$\mathbf{y} = \mathbf{X}_h\mathbf{h} + \mathbf{X}_f\mathbf{f} + \mathbf{p} + \mathbf{u} + \mathbf{e} \quad [2]$$

where **u** is a vector of EBV for the cows with test-day data, also expressed in 9 RRC.

According to the ssSNPBLUP model (Liu et al. 2014), GEBV of a genotyped animal has two components:

$$\mathbf{u} = \mathbf{Zg} + \mathbf{a}\,. \quad [3]$$

Integration of foreign phenotype of bulls

In the German dairy cattle evaluation, breeding values on 305-day basis for the three yield traits or on an average daily basis for SCS were weighted across three lactations and were defined as the official breeding values, which were submitted to Interbull's bull MACE evaluation. Deregressed MACE EBV of a bull from the MACE evaluation (Liu, 2011) was analyzed with a single-trait animal model:

$$\boldsymbol{y} = \boldsymbol{\mu} + \mathbf{z}'\mathbf{g}_{\boldsymbol{m}} + \boldsymbol{a} + \boldsymbol{e} \quad [4]$$

where **y** represents DRP of the bull for a given MACE trait, $\mathbf{g}_m$ is a vector of SNP effects for the MACE trait, $\boldsymbol{a}$ is RPG effect of the bull, and $\boldsymbol{e}$ is the residual effect with

$$\boldsymbol{var}(\boldsymbol{e}) = \boldsymbol{\sigma}^2_{\boldsymbol{e\_m}}/\boldsymbol{n} \quad [5]$$

where $\boldsymbol{\sigma}^2_{\boldsymbol{e\_m}}$ is error variance of the MACE trait, $\boldsymbol{n}$ is effective daughter contribution (EDC) of the bull converted to an animal-model basis.

If the bull is not genotyped, then model 4 is simplified to a regular animal model:

$$\boldsymbol{y} = \boldsymbol{\mu} + \boldsymbol{u_m} + \boldsymbol{e} \quad [6]$$

where $\boldsymbol{u_m}$ is breeding value of the bull for the MACE trait.

Deregression of bull MACE EBV was followed by the method by Liu (2011), using all bulls included in MACE evaluation and complete pedigree data of the bulls. To calculate EDC of all the domestic or foreign bulls on German country scale, national EDC from all participating countries and genetic

correlations among the countries were considered (Liu, 2011). A validation study or reversibility test of the deregression method for bull MACE EBV was successfully performed for all traits on the German scale (Liu and Masuda, 2021).

Let $\boldsymbol{y_n}$ and $\boldsymbol{y_m}$ represent deregressed EBV of a bull from the national conventional and MACE evaluation, respectively, with corresponding national EDC $\boldsymbol{n_n}$ and MACE EDC $\boldsymbol{n_m}$. Pseudo-phenotype of this bull for model 4 or 6 was adjusted for the contribution of national daughters:

$$\boldsymbol{y} = (\boldsymbol{n_m y_m} - \boldsymbol{n_n y_n})/(\boldsymbol{n_m} - \boldsymbol{n_n}) \quad [7]$$

with its corresponding weight changed to:

$$\boldsymbol{n} = \boldsymbol{n_m} - \boldsymbol{n_n} \; . \quad [8]$$

If the bull had no daughters outside Germany, then his MACE data would not need to be integrated in to the national single-step evaluation. In case that this bull had no domestic daughters at all, his deregressed MACE EBV $\boldsymbol{y_m}$ and $\boldsymbol{n_m}$ MACE EDC would be directly used for the integration without the adjustments above.

Phenotype, genotype and pedigree data were obtained from the official April 2021 evaluation for Germany dairy breeds. Test-day data from year 2000 were included in routine conventional evaluation. A total of 242,121,126 test-day records from 12,432,940 cows of the breeds Holsteins, Red Dairy Cattle and Jersey were analyzed together with 138,770 MACE Holstein bulls that had daughters outside Germany. The total number of national cows with test-day data and integrated bulls with foreign daughters amounted to 12,571,710. All genotyped Holstein animals, including culled animals, were jointly evaluated with those animals with phenotype data, and the number of genotyped animals was 949,636 for the April 2021 evaluation. A maximum number of 20 generations was used to trace ancestors of the genotyped or phenotyped animals. Additionally, the oldest bulls with daughters or cows with records were guaranteed to have at least three generations of ancestors. The pedigree file for the single-step evaluation contained 20,461,400 animals and 177 phantom parent groups that were defined according to breed, country of origin, four selection paths and birth year of animals with missing parents.

A validation study for the single-step model

We followed the rules of Interbull trend validation test III and GEBV test (Mäntysaari et al. 2010) to assess the predictive ability of the ssSNPBLUP model for the four traits. Validation bulls were defined as youngest bulls with daughters born in the last four years 2013 through 2016, they must have daughters in at least 10 herds in Germany with a minimum EDC of 20. Last four years of test-day records of the national cows were removed to simulate a genomic evaluation four years ago. Due to a lack of MACE EBV from a truncated MACE evaluation, bull MACE evaluation from April 2021 were used for the validation study. The youngest four birth years of the bulls were deleted from the MACE data. As an extra step, national daughters of the validation bulls were deleted from the test-day data, if there were any daughters left in the truncated test-day data set.

All genotyped animals used in the full evaluation were also included in the truncated genomic evaluation. After the data truncation, 222,634,210 test-day records from 10,903,891 cows and 128,504 bulls with foreign daughters remained in the phenotype data for the validation study.

Table 1 shows the numbers of cows with test-day records and bulls with integrated MACE data for the full and truncated single-step evaluations. Figure 1 shows the numbers of cows with test-day records, genotyped or non-genotyped, by birth year, used in the full evaluation and the truncated validation run. Figure 2 displays the number of integrated MACE bulls with or without genotypes for the full single-step evaluation.

Table 1. Description of phenotype data sets for a full and a truncated evaluation of the four test-day traits in German dairy cattle breeds

	Full data set	Truncated set
Cows with test-day records	12,432,940	10,903,891 (-1,529,049)
§Bulls with MACE data	138,770	128,504 (-10,266)

§ Bulls must have daughters outside Germany

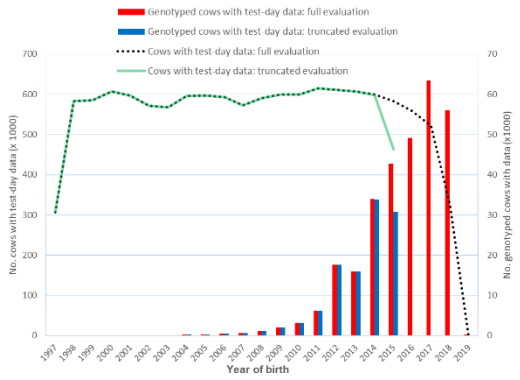

Figure 1. Numbers of cows with test-day records or with genotype data by birth year for the full and truncated evaluations

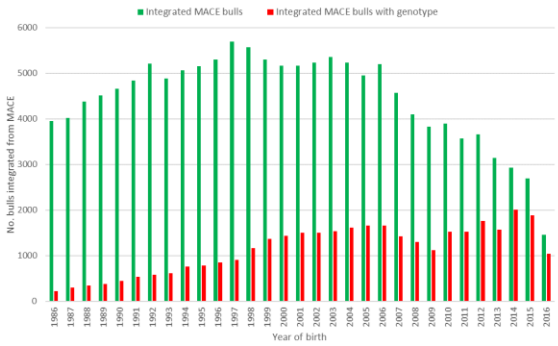

Figure 2. Numbers of integrated MACE bulls with or without genotype records for the full single-step evaluation

The current multi-step genomic model

To compare the accuracy and bias of the single-step model with the current multi-step genomic model, genomic evaluations using the MSM model were also performed with a full and truncated data set. Since 2019 a mixed bull and cow reference population has been used for genomic prediction in German Holsteins (Alkhoder et al. 2017). A single-trait SNP BLUP model was used for estimating the effects of SNP markers (Liu et al., 2011). For the three yield traits, GEBV were expressed as the official 305-day breeding values which were a linear function of 305-day breeding values of the first three lactations. In comparison, GEBV of SCS were expressed on an average daily basis.

Phenotype, genotype and pedigree data for the multi-step evaluations were taken from the official German Holstein evaluation in April 2021. The mixed reference population contained 249,363 reference cows and 43,699 reference bulls for the three yield traits. As routinely conducted for a genomic validation of the MSM model, current evaluation of April 2021 of the reference cows or bulls were used, instead of a conventional evaluation corresponding to four years ago. To consider the relative short history of female genotyping in German Holsteins, the youngest three birth years of national bulls, born in 2014 through 2016, were chosen as validation bulls. A total of 991 Holstein validation bulls were found with daughters in at least 10 herds in Germany. The youngest two birth years of reference cows were removed from the full reference population for the validation study. Furthermore, all daughters of the validation bulls were deleted from the truncated reference population as well to ensure that no daughters of the validation bulls were kept in the truncated data set.

In contrast to the validation study of the single-step model above, the genomic validation of the MSM model did not re-estimate breeding values of the reference cows using a truncated test-day data set. Additionally, the number of truncated data differed between the validation studies of the two models SSM and MSM.

For each of the four test-day traits, a total of four genomic evaluations were conducted: single-step model with the full data set (SSM_Full) and the truncated data set (SSM_Val), and multi-step model with the full data set (MSM_Full) and the truncated data set (MSM_Val).

Results & Discussion

Diverse test runs of the single-step evaluation were performed with the software package MiX99 (Strandén and Lidauer, 1999), in which the ssSNPBLUP by Liu et al. (2014) was implemented in a special way (Mäntysaari, personal communication). In addition, we applied the software package MiXBLUP (Ten Napel et al. 2020) to the same data sets. Approximately the same execution time was used with either of the software. However, MiXBLUP needed a maximum RAM of 135Gb for the full data evaluation, and MiX99 required 358Gb for the same data set. Because identical effect estimates were obtained with the two software packages, we used only the solutions from MiX99 for further analyses.

For the single-step evaluation of trait protein yield PRO, the total number of all estimated effects was 338,616,497 with the full data set. Random access memory of 347 Gb was needed with the software MiX99 for the full single-step evaluation. The total clock time per round of iteration was 1.63 minutes on a Linux server with 2x10 cores of Intel® Xeon® CPU E5-2690 v2 @3.00Ghz. Figure 3 shows convergence rates of the single-step full evaluation with 17 cores used. Based on the convergence criteria CR or CD (Strandén and Lidauer, 1999), reasonably accurate solutions could be obtained with approximately 1000 rounds of iteration.

A special conventional evaluation with only national cow test-day records was conducted with MiX99 software and all effect estimates were compared to the official conventional evaluation based on vit's own software package for the RRTDM model (Liu et al. 2004). Identical effect estimates were obtained for the herd-test-date-parity-milking-frequency, fixed lactation curves, permanent environmental and genetic effects for all the cows with own data, bulls with daughters and all their ancestors. Judging from the identical solutions of the two different software packages, we could conclude that the implementation of the MiX99 software for the test-day model was done in a correct way.

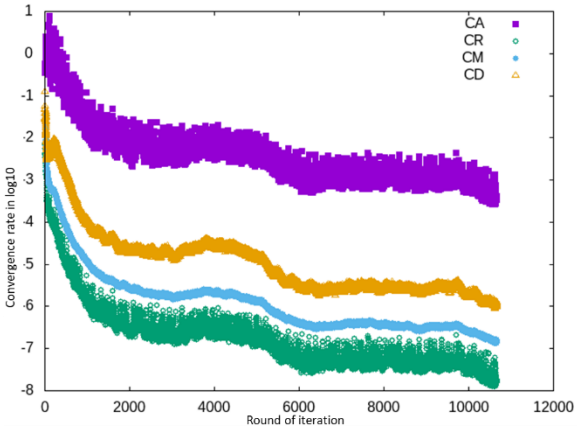

Figure 3. Convergence criteria of the single-step SNP BLUP model using the full data set for test-day protein yield

SNP effect estimates

The Liu-Goddard ssSNPBLUP model estimated SNP effects in form of RRC for the 9 national traits and as combined lactation breeding values for the MACE trait. We ignored solutions on the MACE trait, because of most animals missing this trait. SNP effects in RRC were combined to be expressed on the official breeding value scale. Figure 4 shows observed correlations of the SNP effect estimates between the full and truncated evaluations.

It can be clearly seen that the SNP effect estimates of the SSM model (black bars) are higher correlated between the full and truncated evaluation for any of the four test-day traits than the MSM model (red bars). This indicates more consistent or accurate SNP effects for the single-step model. For the full evaluation, SNP effect estimates are correlated, ranging from 0.74 for trait PRO to 0.79 for trait MIL, between the two genomic models (blue bars).

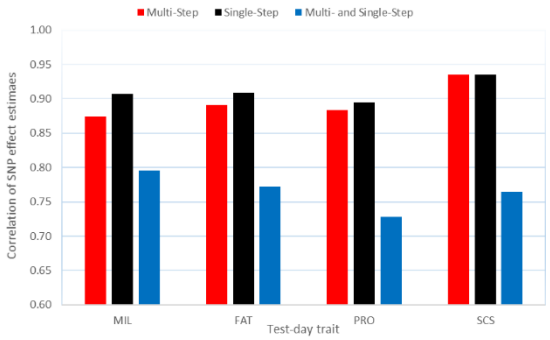

Figure 4. Correlations of SNP effect estimates between the full and truncated evaluations for the four traits

Figure 5 shows regression coefficients of the SNP effect estimates from the full on the truncated evaluation for the two models and for each trait. We can see that the SSM model (black bars) has regression coefficient being closer to 1 than the MSM model (red bars). It is desirable to have the regression coefficients closer to 1, because this indicates less inflation or deflation of the SNP effect estimates from the SSM model.

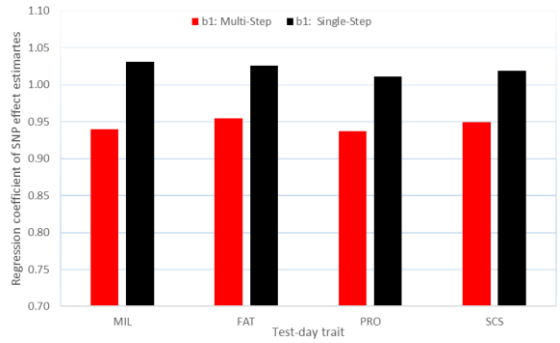

Figure 5. Regression coefficients of SNP effect estimates of the full on the truncated evaluations

Correlation and dispersion of genomic prediction

We would like to know how accurately we predicted genomic breeding values of animals, e.g. the validation bulls, in the later-full evaluation based on their early-truncated evaluation as young candidates. Regressing GEBV of the validation bulls from the later-full evaluation on the early-truncated evaluation (Legarra and Reverte, 2018) gave us an indication on both the accuracy and dispersion of the genomic prediction. Table 2 shows the results of genomic validation for the two genomic models SSM and MSM based on the national validation bulls. Even though the validation data sets for the two genomic models SSM and MSM were not identical and the validation bulls differed, we appended the validation results of the MSM model here for information. Model R^2 value or squared correlation of GEBV between the full and truncated evaluations clearly show higher consistency of GEBV of the SSM than MSM, because the MSM did not utilize information such as phenotype data of non-genotyped relatives of the reference animals. Regression intercept b_0 range from -0.002 to 0.17 genetic standard deviations, suggesting no severe over- or underestimation of the level of GEBV of the validation bulls. For the SSM regression slope b_1 value was close to 1 with lowest value 0.96 for protein yield and highest value 1.02 for milk yield. The regression slope b_1 of being near 1 indicates that the variance of the early-truncated evaluation was neither too high nor too low.

Table 2. Linear regression of GEBV of the full on the truncated evaluation of validation bulls

	b_0 in genetic std dev	Slope b_1	Model R^2 value
Single-step genomic model			
Milk yield	-0.16	1.02	0.81
Fat yield	-0.08	1.00	0.80
Protein yield	0.03	0.96	0.71
SCS	0.04	0.99	0.78
Multi-step genomic model			
Milk yield	0.17	0.98	0.70
Fat yield	-0.20	1.12	0.76
Protein yield	-0.002	1.07	0.70
SCS	-0.02	1.07	0.68

Genetic trends of the test-day traits

Genotyped German Holstein AI bulls, as a highly selected group of animals, were investigated in their averages and variances of GEBV of the full and truncated evaluations for the two genomic models. The same AI bulls were selected as in the study by Alkhoder et al. (2022) with some new AI bulls born in 2019 and 2020 added. In total, 12,249 genotyped Holstein AI bulls, belonging to German AI

studs, were born between 1998 and 2020. Figure 6 shows genetic trends of protein yield in the German AI bulls of the four evaluations. Single-step model has higher genetic trends in recent birth years than the multi-step model, with largest difference found in the youngest birth year. The trend differences between the truncated and full evaluation are marginal for any of the two models. Similar trends are seen also for traits MIL and FAT. For trait SCS the trend differences between the models or between the data sets are much smaller than the yield traits.

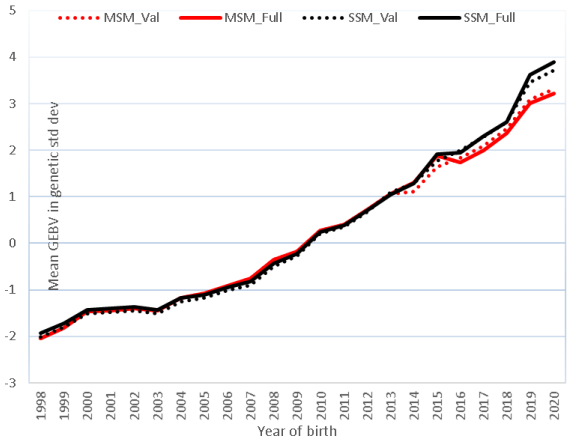

Figure 6. Genetic trends in the genotyped German AI bulls for protein yield of the two genomic models with two data sets

GEBV variances within birth years

A major concern of the single-step evaluation was too high variance of GEBV of young animals. Figure 7 shows within-year standard deviations of GEBV in the genotyped AI bulls for test-day trait fat yield. It is evident that the young AI bulls without daughters, born in 2018 through 2020, have clearly lower GEBV variances than those AI bulls with own daughters. The genomic evaluations with the truncated data set have lower GEBV variance than those with the full data set. For traits FAT and MIL, older genotyped AI bulls with daughters seem to have slightly higher GEBV variance for the SSM than the MSM model. But there is no difference in GEBV variance for the AI bulls with daughters for the traits PRO and SCS.

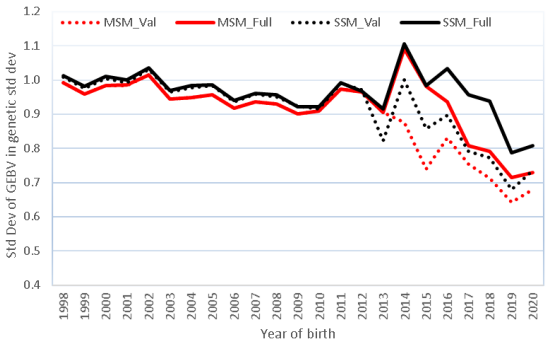

Figure 7. GEBV standard deviations of fat yield in genotyped German Holstein AI bulls for the two genomic models using the two data sets

Correlations of GEBV between evaluations

Should 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, genomic prediction is expected to be stable over time. To investigate GEBV correlation between the full and truncated evaluation with full data set, we choose all genotyped female animals, which may be regarded as a group of animals with the lowest selection intensity. In comparison to the study by Alkhoder et al. (2022), only a limited number of newly genotyped female animals were added in last birth years of 2020 and 2021. Figure 8 shows GEBV correlations of trait SCS between the truncated and full evaluations for the two genomic models. In addition, GEBV correlations are also presented between the two models using the full data set (blue line). For either of the models, the two evaluations are highly correlated for the female animals, with an average above 0.97. Because no truncated conventional evaluation was conducted for the validation of the MSM model, GEBV correlations (red line) are marginally higher than those of the SSM model. With the full data set, GEBV of both models have an average correlation of 0.96 across all birth years and 0.95 for the candidate years 2019 to 2021. Similar pattern of GEBV correlations for the female animals were also observed for the other three yield traits, except with slightly lower GEBV correlations than for SCS.

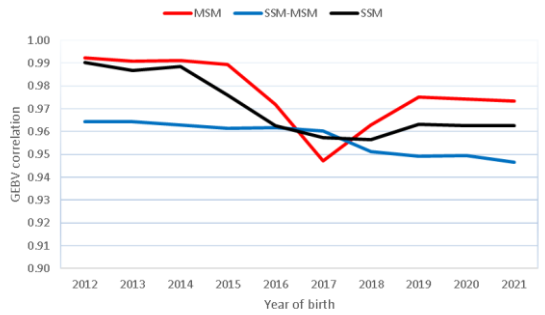

Figure 8. GEBV correlations between the truncated and full evaluations of trait SCS for the genotyped Holstein female animals

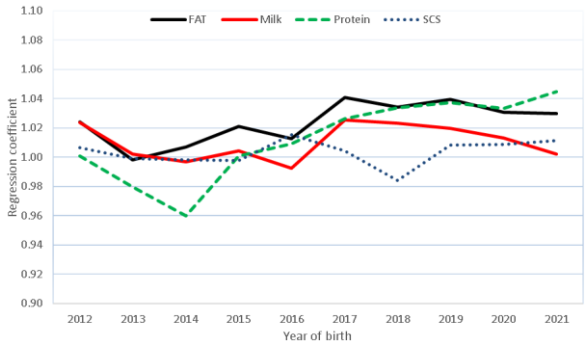

Figure 9. Regression coefficients of GEBV of the genotyped male candidates from the full on the truncated single-step evaluation

Regressions of GEBV of the full on the truncated genomic evaluation

Regression of GEBV of a later evaluation with complete data on those of an early evaluation with less complete data can show 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. We choose the third group of the genotyped animals, genotyped male candidates without own phenotype, to show the regression coefficients of the validation for the SSM model. In comparison to the number of genotyped Holstein male candidates (Alkhoder et al. 2022), only newly genotyped male animals were added to the youngest birth year 2020 and 2021. Figure 9 shows the regression slope values of the single-step full on the truncated evaluation for all the genotyped male candidates. Nearly all regression coefficients fall in the range of 0.98 and 1.04, indicating there is no evidence of an inflation or deflation of genomic prediction using the single-step model for this group of animals. We have found also similar regression slope estimates for the MSM model than the SSM model for the four test-day traits. For the genotyped female animals, the regression slope values range from 1 to 1.04. For the highly selected genomic AI bulls without own daughters, the regression slopes fall mostly in the range between 0.95 and 1.10.

Impact of the selection of integrated bulls

Selection of the integrated bulls was shown to have an impact on the genomic validation results (Alkhoder and Liu, 2022). Using all the bulls with foreign daughters seem to give the least bias for the validation bulls and highest GEBV correlation. When only genotyped bulls with foreign daughters were integrated, the GEBV correlation decreased and the prediction bias increased, in comparison to the standard scenario of using all bulls with foreign daughters for the integration. When the birth years of the integrated bulls were left-truncated to 1995, which corresponded to the earliest test-day records of the national cows, little difference in the GEBV correlation or bias was found to the standard scenario.

Impact of genotype edits on older bulls

At the beginning of genomic selection in year 2008 or 2009, not all bulls with daughters in the national test-day data had semen available for genotyping, which was observed in both national and international bulls. The genotyped bulls tended to have a higher genetic level than those non-genotyped bulls. Such an involuntary selective genotyping may lead to a biased genomic prediction. To investigate the impact of the selective genotyping in the early years of genomic selection, genotype records of bulls born before 2005 were deleted from the genotype data. A genomic evaluation was conducted for trait PRO using both the full and truncated data sets (Alkhoder and Liu, 2022). In comparison

to the standard scenario given in this study, no difference in GEBV correlation was observed and the regression slope was equal. However, average Mendelian sampling effects of the genotyped female animals were closer to zero for this scenario with genotype edits than for the standard scenario of using all phenotype and genotype data of the bulls. In addition, young candidates seem to have lower average of GEBV than for the standard scenario.

Conclusions

The Liu-Goddard single-step SNP BLUP model was successfully applied to the test-day data in German dairy cattle. The three test-day yield traits and SCS of the national cows were evaluated with the ssSNPBLUP random regression test-day model. Deregressed MACE EBV of bulls with foreign daughters were integrated as a correlated trait to the national single-step test-day model. Thanks to the high efficiency of the ssSNPBLUP model, approximately one million genotyped animals were able to be evaluated jointly with all the national cows and bulls with foreign phenotypes, no approximation of genomic relationship between any pair of the genotyped animals was needed. Regarding the SNP effect estimates, the single-step model had clearly higher correlation between the truncated and full evaluation than the multi-step model. Regression of SNP effect estimates from the full on the truncated evaluation showed that the single-step model resulted in less inflated or deflated SNP effects than the current multi-step model. In comparison to the current multi-step genomic model, the single-step model resulted in higher correlation of GEBV between the full and truncated evaluations for the validation bulls. Higher genetic trends and greater GEBV variances were found in the young animals for the single-step model than the multi-step model. For the validation bulls, regression slope of GEBV of the full on the truncated evaluation was close to 1 for all the four test-day traits, ranging from 0.96 for PRO to 1.02 for MIL, indicating no severe inflation or deflation of genomic prediction with the single-step model. For the highly selected genotyped young AI bulls, GEBV correlation between the full and truncated evaluation had an average of 0.95 across the traits. Between the single-step and multi-step model GEBV correlation for this group of AI bulls was 0.93 averaged across the four traits. We investigated further the impact of selection of MACE bulls on genomic prediction. Furthermore, we conducted an additional genomic evaluation to evaluate the impact of removing genotype data of older bulls. Based on the diverse test evaluations of the single-step model, we can conclude that no post-processing of GEBV of young animals seem to be necessary. Further topics will be addressed such as genomic reliability approximation, interim genomic evaluation on a weekly basis, and integration of MACE SNP effects in to the national single-step evaluation.

Acknowledgements

Dr. Jeremie Vandenplas and Dr. Jan ten Napel from University of Wageningen and Dr. Esa Mäntysaari and Dr. Ismo Strandén from the Finish research institute LUKE are greatly thanked for their help in using the software packages MiXBLUP and MiX99, respectively. The authors have stated no conflicts of interest.

References

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

Alkhoder, H., Z. Liu, D. Segelke, and R. Reents. 2022. *J. Dairy Sci.* 105:3306-3322. https://doi.org/10.3168/jds.2021-21145.

Alkhoder, H., and Z. Liu. 2022. WCAGLP, Rotterdam, The Netherlands, July 3-8, 2022.

Jairath, L., Dekkers, J.C.M., Schaeffer, L.R., Liu, Z., Burnside, E.B., and Kolstad, B. 1998. J. Dairy Sci. 81:550-562.

Legarra, A., and Reverter, A. 2018. *Genet. Sel. Evol.* 50:53.

Liu, Z., Reinhardt, F., Bünger, A., and Reents, R. 2004. *J. Dairy Sci.* 87:1896-1907.

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

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

Liu, Z., and Masuda, Y. 2021. *Interbull Bulletin* 56:41-51.

Mäntysaari, E. A., Liu, Z., and VanRaden., P. M. 2010. *Interbull Bulletin* 41:10–14.

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.