Validating German Holstein single-step evaluations for test-day traits using Interbull's new GEBVtest software

Z. Liu, H. Alkholder and R. Reents

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

Abstract

To validate national genomic evaluation systems, particularly those based on a single-step model, Interbull further developed current GEBV test method and extended functionality of the GEBV test python software. As response variable of the linear regression analysis in the GEBV validation test, GEBV as well as deregressed GEBV of validation animals were considered, besides the current deregressed conventional EBV of validation bulls. The aims of this study were to validate the single-step evaluation of four test-day traits for German Holsteins using the newly optimized GEBV test software, and to compare alternative forms of dependent variable and diverse groups of validation animals for genomic validation. Phenotypic, genotypic, and pedigree data were obtained from official April 2021 evaluation for German Holsteins for this study. The single-step evaluations of all the test-day traits were shown to pass the new GEBV test, using dependent variable GEBV or deregressed GEBV for either the validation bulls or cows. For all the tested scenarios, regression slope b_1, genomic model R^2 and R^2 increase from a conventional model 2 to genomic model 1 all seemed to meet expectations. Notable variation was observed in the validation results across the subgroups of the validation animals, e.g. the validation bulls born in different years. Dependent variable deregressed GEBV or conventional EBV resulted in clearly lower R^2 values than GEBV, and the b_1 values deviated slightly more from 1. For the low-reliability validation cows, dependent variables GEBV and deregressed GEBV led to markedly different R^2 values of the gnomic model 1, though similar R^2 values were found for the high-reliability validation bulls. The deregressed GEBV seemed to be a more appropriate form of dependent variable for the GEBV test than the dependent variable GEBV, especially for the low-reliability validation cows. The new GEBV test software was proven to work as expected.

Key words: GEBV test, genomic validation, single-step model, test-day trait, random regression

Introduction

At the beginning of genomic selection in dairy cattle, Interbull developed a genomic validation test, the so-called GEBV test (Mäntysaari et al. 2010), to certify national genomic evaluation systems. As more and more countries have been upgrading their national genomic evaluation from a multi-step to a single-step genomic model, the GEBV test needs to be modified and optimized for the new genomic model. Dependent variable of the current GEBV test has been deregressed conventional EBV (DRP) of validation bulls, which may be subject to the bias caused by genomic pre-selection, because the conventional evaluation ignored the genomic information. Therefore, new statistical methods are required to identify proper dependent variables for the GEBV test.

By default, Interbull member countries chose genotyped bulls with daughters as validation animals for the GEBV test. However, genotyped cows with phenotype records may be an alternative of the validation animals, since many more cows or heifers have been genotyped in an increasing number of countries and the selection intensity on the cows or heifers was typically much lower than the genotyped bulls.

The aims of this study were 1) to validate the single-step evaluation of four test-day traits for German Holstein using the new GEBV test software (Sullivan 2022), and 2) to compare alternative forms of the dependent variable and validation animal groups for the genomic validation.

Materials and Methods

For German Holsteins, a single-step SNP BLUP model (ssSNPBLUP) was applied to national test-day and bull MACE data, obtained from April 2021 official evaluation for German Holsteins (Alkhoder et al. 2022). The same phenotype, genotype and pedigree data sets were used here for testing the new GEBV test software (Sullivan 2022). In total, 242,121,126 test-day records of 12,432,940 cows were jointly analysed with 138,770 Holstein bulls from the corresponding MACE evaluation. There were 949,636 genotyped Holstein animals considered in the single-step evaluation, some of them being young candidates or culled animals. The pedigree file of the single-step model contained 20,461,400 animals. Four test-day traits were chosen for this study: milk, fat, protein yields and somatic cell scores (SCS). Same as in the study by Alkhoder et al. (2022), 30% residual polygenic variance was assumed for each of the test-day traits. In the study by Alkhoder et al. (2022) a genomic validation was conducted by deleting test-day records in the last four years to simulate a forward prediction. Because no truncated MACE evaluation was available for the validation study, daughter-proven bulls born in the four youngest years were truncated from the full MACE bull data set (Alkhoder et al. 2022).

The GEBV test (Mäntysaari et al. 2010) compared a genomic model (Model 1) to a conventional model (Model 2) for validation animals. Therefore, special conventional evaluations with the full and truncated data sets were conducted by using the same phenotype and pedigree data as for the corresponding single-step evaluations, except no genotype data being considered in the conventional evaluations. Per animal, there were a total of four estimated breeding values (EBV) available: from the full and truncated evaluations for the ssSNPBLUP model and the conventional random regression test-day model.

Reliability values of EBV of the conventional evaluations were approximated following a multi-trait reliability calculation method. Effective daughter contribution (EDC) of bulls or effective record contribution of cows with phenotype data were calculated for the single-step model with MACE data integrated. Genomic reliability values were approximated for GEBV of all the animals from the ssSNPBLUP model, following Interbull genomic reliability method (Liu et al. 2017). The approximated genomic reliabilities, adjusted based on the genomic validation results, were used in this study.

Scenarios for testing the new GEBV test

As dependent variable of the linear regression model in the GEBV test method, we tested these scenarios: GEBV of the full evaluation as described by Legarra and Reverter (2018), deregressed GEBV (VanRaden 2021), and deregressed EBV of conventional national evaluations for cows and conventional MACE evaluations for bulls (Liu and Masuda 2021).

As validation animals, we investigated two groups of animals: validation bulls with daughters and cows with own phenotype data. According to the GEBV test rule (Mäntysaari et al. 2010), validation bulls must have daughters in own country with EDC $\geq$ 20. The validation cows were genotyped cows with test-day records in the domestic population. The new GEBV test python software was modified to allow the validation cows with EDC < 20 to be considered.

National validation bulls and cows

Genotyped Holstein bulls, born in 2013 through 2016, were selected as validation animals, and they must have daughters in at least 10 herds in Germany in the full data set with at least 20 EDC and no daughters in the truncated data set according to the rule by Interbull GEBV test (Mäntysaari et al. 2010). Genotyped Holstein cows must have own test-day records in the full evaluation and were young candidates in the truncated evaluation. Table 1 shows the number of the Holstein validation bulls or cows by birth year.

Table 1. Numbers of Holstein validation bulls and cows by year of birth

Birth year of bulls	2013	2014	2015	2016
Number of bulls	557	489	429	180
Birth year of cows	2015	2016	2017	2018
Number of cows	12,083	49,029	63,245	56,032

Four evaluations were conducted for testing the new software: single-step and conventional evaluations using the full and truncated data sets. EBV or GEBV of the four evaluations were adjusted for the same base population that included cows born in 2015 through 2017. Because the four evaluations were expressed on the same base cow population, the option of no base adjustment was chosen when running the new GEBV test python software.

Results & Discussion

There were, in total, 1,655 Holstein national validation bulls selected (Table 1). The number of Holstein validation cows amounted to 180,389. Tables 2 and 3 show correlations of milk yield GEBV or EBV of the validation bulls or cows between any pair of the four genomic and conventional evaluations.

Table 2. – Correlations of milk yield GEBV or EBV between any pair of the four evaluations for the Holstein validation bulls

	GEBV trunc	EBV full	EBV trunc	DRP_ MACE
GEBV full evaluation	0.90	0.98	0.43	0.93
GEBV truncated evaluation		0.85	0.57	0.80
EBV full evaluation			0.45	0.95
EBV truncated evaluation				0.41
Deregressed MACE full evaluation (**DRP_MACE**)				

Table 3. – Correlations of milk yield GEBV or EBV between any pair of the four evaluations for the Holstein validation cows

	GEBV trunc	EBV full	EBV trunc	DRP_ NAT
GEBV full evaluation	0.95	0.85	0.55	0.71
GEBV truncated evaluation		0.73	0.61	0.54
EBV full evaluation			0.64	0.83
EBV truncated evaluation				0.32
Deregressed national full evaluation (**DRP_NAT**)				

It can be seen in Table 2 that GEBV correlation between the full and truncated evaluation for the validation bulls, 0.90, is much higher than the EBV correlation, 0.45, indicating the genomic information increased the evaluation stability for the validation bulls. A similar pattern in the correlations is also observed for the validation cows in Table 3. The EBV or GEBV correlations between the full and truncated evaluations are higher for the validation cows than for the validation bulls. This may be explained by the fact that the contribution by own test-day records to the cows' EBV or GEBV of the full evaluation is much smaller than the contribution by daughters' phenotype data to the validation bulls' EBV or GEBV of the full evaluation.

Validation results using GEBV as dependent variable

For the validation bulls or cows, their full evaluation GEBV, $GEBV_{full}$, were regressed on their GEBV from the truncated evaluation, $GEBV_{trunc}$, following the LR method (Legarra and Reverter 2018). Same as for the current GEBV test (Mäntysaari et al. 2010), the following two models were analysed:

Model 1: $GEBV_{full} = b_0 + b_1 * GEBV_{trunc}$
Model 2: $GEBV_{full} = b_0 + b_1 * EBV_{trunc}$

with weights on the dependent variable being cow's or bull's reliability converted from her own test-day records or his EDC by phenotype data of his daughters, respectively. Model 1 (M1) described the GEBV regression of the full on truncated evaluation, whereas Model 2 (M2) regressed the GEBV of the full evaluation on conventional EBV of the truncated evaluation, EBV_{trunc}.

Table 4 shows the validation results provided with the new GEBV test software. All the four traits are shown to pass the GEBV test for the validation bulls. The regression slope b_1 values are all close to 1, indicating no significant over- or under-prediction of the single-step GEBV. The model R^2 values range from 0.71 for protein yield to 0.80 for milk and fat yields. The model R^2 increase (ΔR^2) from the conventional parental average EBV_{trunc} (M2) to $GEBV_{trunc}$ (M1), indicating reliability gain contributed by the genomic information, varies from 0.47 for protein yield to 0.62 for milk yield.

Table 4. – Validation results using the dependent variable GEBV of the validation bulls for the four test-day traits

	Model 1		M1-M2	Test
Trait	b_1	R^2	ΔR^2	result
Milk yield	1.01	0.80	0.62	Pass
Fat yield	1.00	0.80	0.50	Pass
Protein yield	0.95	0.71	0.47	Pass
SCS	0.99	0.78	0.54	Pass

Validation results for the bulls by birth years

The validation bulls differed in distance to the genomic reference population. Therefore, the validation results may vary among the bulls born in different birth years. The GEBV test software was applied to the validation bulls separately for each of the birth years. Figure 1 shows the regression slope b_1 value of the validation bulls born in a year between 2013 and 2016. Due to the lower number of bulls born in 2016, the b_1 values are more variable among the traits than in the other years. It seems that the regression slope b_1 deviates a little more from 1, as the validation bulls are becoming younger, i.e., departing further away from the reference population except for the last birth year 2016. Among the four test-day traits, protein yield has the lowest b_1 values, suggesting larger over-prediction, though statistically insignificant, than the other traits.

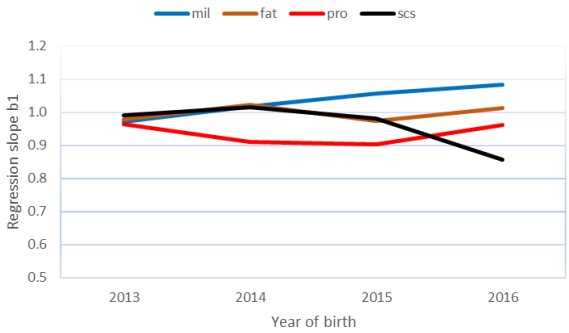

Figure 1. Regression slope of the genomic model 1 for the validation bulls born in the same year.

Figure 2 displays the model R^2 value of the genomic model 1 for the validation bulls born in each of the four birth years. The model R^2 values decrease slightly from the oldest to youngest birth year, suggesting that GEBV reliability is reduced when the validation bulls are younger and departing further from the genomic reference population. As the regression slope b_1, the model R^2 has lowest value for protein yield among the four traits.

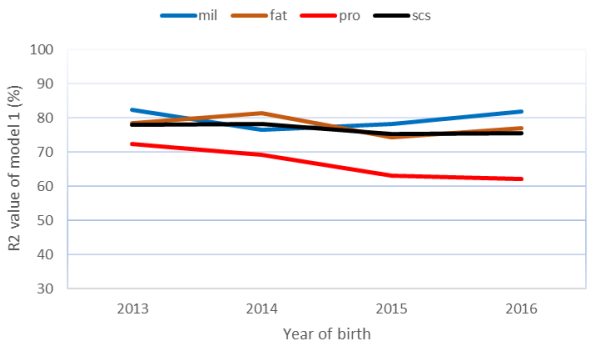

Figure 2. Model R^2 value of the genomic model 1 for the validation bulls born in the same year.

Figure 3 shows the model R^2 value increase (ΔR^2) from the conventional model 2 to the genomic model 1 for the validation bulls born in different years. The younger the validation bulls, the higher the increase of the model R^2 value. The trend in the model R^2 value increase suggests that the youngest validation bulls benefit the most from the genomic data for genomic prediction.

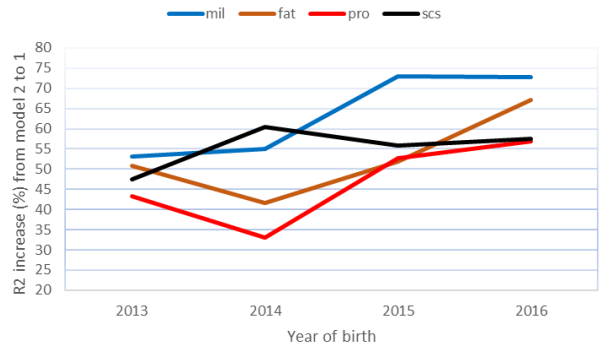

Figure 3. Model R^2 value increase from the conventional model 2 to genomic model 1 for the validation bulls born in the same year.

Validation results for the validation cows

Validation results for the cows are given in Table 5. Based on the criteria of the GEBV test, all the four traits are shown to pass the validation test, because the regression slope b_1 values are nearly 1 and the model R^2 increase (ΔR^2) are clearly greater than 0. The R^2 values of model 1 are surprisingly high, ranging from 0.88 to 0.91, for the validation cows, which may indicate a high autocorrelation of GEBV of the cows between the truncated and full evaluations. The much higher R^2 values of the validation cows than the validation bulls can be explained by the fact that own test-day records of the validation cows contribute typically less to their GEBV than many daughters of the validation bulls to GEBV of the bulls. Therefore, the extremely high R^2 values of the genomic model 1 must not be interpreted as high reliabilities of GEBV of the validation cows.

Table 5. – Validation results using the dependent variable GEBV of the validation cows for the four test-day traits

	Model 1		M1-M2	Test
Trait	b_1	R^2	ΔR^2	result
Milk yield	1.03	0.89	0.59	Pass
Fat yield	1.03	0.91	0.55	Pass
Protein yield	1.01	0.88	0.48	Pass
SCS	1.02	0.91	0.66	Pass

Within each birth year of the validation cows, the regression slope b_1 and R^2 values of model 1 show a much flatter trend over the four birth years than the validation bulls. As far as the R^2 increase from the model 2 to model 1 concerned, the youngest validation cows have the largest increase in the R^2 value.

Validation results using the deregressed GEBV as dependent variable

As another alternative of dependent variable of the GEBV test, deregressed GEBV (DGEBV, VanRaden 2021) of the validation bulls or cows were investigated using the two models:

Model 1: $DGEBV_{full} = b_0 + b_1 * GEBV_{trunc}$
Model 2: $DGEBV_{full} = b_0 + b_1 * EBV_{trunc}$.

Table 6 shows validation results of the four traits for the Holstein validation bulls. In comparison to Table 4, regression slope b_1 values using DGEBV as dependent variable deviate slightly more from 1 than using GEBV. The R^2 values of the genomic model 1 are lower than those of the dependent variable GEBV. The R^2 value increase (ΔR^2) from model 2 to model 1 is only marginally lower than those of the dependent variable GEBV. Same as the case of the dependent variable GEBV, all the four traits of German Holsteins are shown to pass the GEBV test using the dependent variable DGEBV.

Table 6. – Validation results using deregressed GEBV of the validation bulls as dependent variable

Trait	Model 1 b_1	R^2	M1-M2 ΔR^2	Test result
Milk yield	1.01	0.77	0.60	Pass
Fat yield	1.00	0.75	0.48	Pass
Protein yield	0.94	0.64	0.43	Pass
SCS	0.98	0.71	0.50	Pass

Validation results for the validation cows

The validation results using the DGEBV as dependent variable are given in Table 7 for the Holstein validation cows. In contrast to the scenario of GEBV as dependent variable (Table 5), we see that the R^2 values of model 1 are much lower, ranging from 0.54 for protein yield to 0.69 for SCS. The much lower R^2 values of model 1 with the dependent variable DGEBV may indicate that DGEBV is a more appropriate form of dependent variable for the GEBV test than the dependent variable GEBV, especially for the low-reliability validation cows. In addition, we see that the R^2 values of Model 1 for the validation cows are markedly lower than those for the validation bulls in Table 6, which indeed meets our expectation. The R^2 value increase (ΔR^2) from model 2 to model 1 is lower than that for the validation bulls (Table 6), except the trait SCS. The regression slope b_1 values deviate more from 1 than the scenario of GEBV as dependent variable (Table 5). All the four traits are shown to pass the GEBV test with the dependent variable DGEBV for the validation cows.

Table 7. – Validation results using deregressed GEBV of the validation cows as dependent variable

Trait	Model 1 b_1	R^2	M1-M2 ΔR^2	Test result
Milk yield	1.07	0.66	0.46	Pass
Fat yield	1.06	0.68	0.45	Pass
Protein yield	1.02	0.54	0.30	Pass
SCS	1.05	0.69	0.52	Pass

Validation results for the cows by birth years

The GEBV test was applied to the validation cows by birth year, instead of across all the birth years jointly, using the DGEBV as response variable. Figure 4 shows trends in b_1 values across birth years of the validation cows. In comparison to Figure 1 for the validation bulls, the regression lines for the validation cows are much smoother, due to the higher number of validation cows. The regression slope b_1 seems to have a slightly decreasing trend over the birth years.

Figure 5 shows the trend in R^2 value of model 1 for the validation cows within birth year, with DGEBV as response variable. As the distance to the genomic reference population getting larger, the R^2 values of model 1 tend to be a little bit smaller, i.e., younger validation cows have lower genomic reliabilities than the older ones.

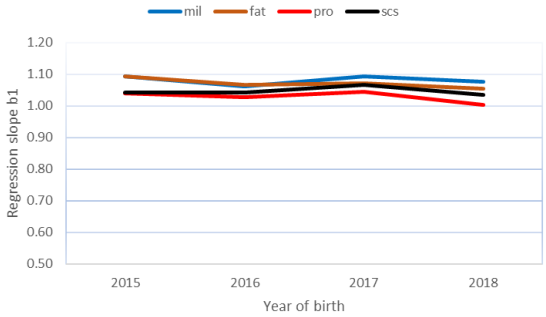

Figure 4. Regression slopes of the model 1 for the validation cows with DGEBV as response variable

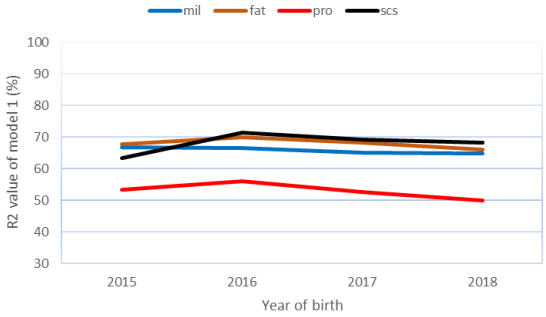

Figure 5. R^2 values of the model 1 for the validation cows with DGEBV as response variable

In Figure 6 we can see the R^2 increase (ΔR^2) from the conventional model 2 to the genomic model 1 for the validation cows born in different years, using their DGEBV as dependent variable of the GEBV test. A clear increasing trend can be seen in ΔR^2 over the

birth years of the validation cows. The youngest validation cows born in 2018 have the highest R^2 increase for all the test-day traits.

GEBV versus deregressed GEBV

As response variable of the GEBV test, both GEBV and DGEBV of the validation animals have been tested in our study. For the genomic model 1, the validation bulls have lower R^2 values (Table 4) than the validation cows (Table 5) with the response variable GEBV. However, when DGEBV is used as the dependent variable, the R^2 values of the validation bulls (Table 6) are clearly higher than those of the validation cows (Table 7). Though the model R^2 values cannot be directly compared between the two different dependent variables GEBV and deregressed GEBV, the R^2 values of the two forms of dependent variable can be explained that own test-day records of the validation cows contribute much less to their GEBV than many daughters with phenotype data to GEBV of the validation bulls. This indeed raises the question if GEBV can be regarded as an optimal dependent variable for the low-reliability validation cows.

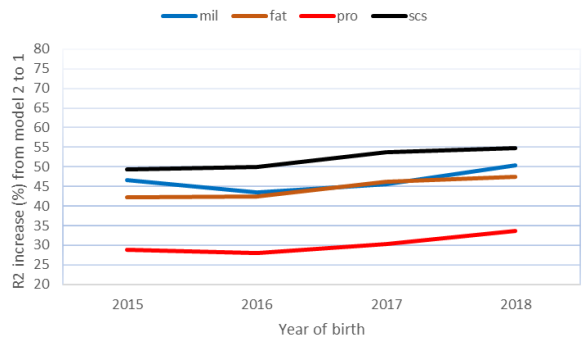

Figure 6. Model R^2 value increase from the conventional model 2 to genomic model 1 for the validation cows with DGEBV as response variable

Based on our results, DGEBV seems to be a more appropriate form of dependent variable than GEBV itself for the purpose of GEBV test, especially for the low-reliability validation animals. However, the DGEBV calculation (VanRaden 2021) relies on accurate and comparable genomic reliabilities across all Interbull member countries which can be approximated using the Interbull genomic reliability method (Liu et al. 2017). The current way of calculating DGEBV (VanRaden 2021) did not use pedigree or genomic relationships among all the genotyped or non-genotyped animals. Rather the deregression process was done on an animal-by-animal basis. An alternative way of computing the deregressed GEBV, particularly from the single-step SNP BLUP model, were developed by Liu and Masuda (2021), which considered both the pedigree and genomic relationships among all the animals and solved DGEBV with an iterative procedure. More importantly, the GEBV deregression method (Liu and Masuda 2021) estimated DGEBV using only data from the full evaluation, independently from the truncated genomic evaluation.

Validation results using the deregressed conventional EBV as dependent variable

As a third form of response variable for the GEBV test, conventional DRP of the validation animals were investigated: deregressed MACE EBV of the validation bulls (DRP_MACE) and deregressed national EBV of the validation cows (DRP_NAT).

Table 8 shows validation results using DRP_MACE as response variable with:

Model 1: $DRP\_MACE_{full} = b_0 + b_1 * GEBV_{trunc}$
Model 2: $DRP\_MACE_{full} = b_0 + b_1 * EBV_{trunc}$.

Because there were no deregressed GEBV from the single-step model available yet (Liu and Masuda 2021), we could test only the deregressed conventional MACE EBV. In comparison to the DGEBV results in Table 6, we see that the regression slope b_1 values deviate more from 1 in Table 8. The R^2 values of model 1 with DRP_MACE as response variable are lower than those with DGEBV as dependent variable (Table 6). The R^2 increase (ΔR^2) from the conventional model 2 to the

genomic model 1 is also smaller with dependent variable DRP_MACE than DGEBV. Nevertheless, all the traits are shown to pass the GEBV test with DRP_MACE as dependent variable.

Table 9 shows validation results for the cows with dependent variable DRP_NAT:

Model 1: $DRP\_NAT_{full} = b_0 + b_1 * GEBV_{trunc}$
Model 2: $DRP\_NAT_{full} = b_0 + b_1 * EBV_{trunc}$.

The regression slope b_1 values of the genomic model 1 in Table 9 are significantly less than 1 and lower than those with dependent variable DGEBV (Table 7) or those with dependent variable GEBV (Table 5). The markedly smaller regression slope values suggest that the $GEBV_{trunc}$ or EBV_{trunc} from the truncated evaluations may have too high variance for the response variable DRP_NAT. One can also argue that the dependent variable DRP_NAT of the validation cows may have too low variance for the truncated early evaluations $GEBV_{trunc}$ or EBV_{trunc}. As a matter of fact, no foreign daughter data of bulls in the MACE evaluation were integrated in the conventional national random regression test-day model evaluation, from which the DRP_NAT of the cows were derived. In contrast, MACE data of the bulls were evaluated jointly with national cow test-day data in the single-step evaluation as well as in the special conventional evaluation. Therefore, the regression slope $b_1<1$ is caused more by the relatively low variance of the response variable DRP_NAT of the validation cows than by the relatively high variance of the $GEBV_{trunc}$ or EBV_{trunc}. The R^2 values of the genomic model 1 are substantially less than those with dependent variable DGEBV (Table 7) or GEBV (Table 5). The R^2 increase (ΔR^2) from the conventional model 2 to the genomic model 1 is, by far, the smallest among all the forms of dependent variable. For the validation cows, all the traits are shown to fail the GEBV test when the national deregressed EBV DRP_NAT are used as dependent variable in the GEBV test.

Table 8. – Validation results using deregressed MACE EBV of the validation bulls as dependent variable

	Model 1		M1-M2	Test
Trait	b_1	R^2	ΔR^2	result
Milk yield	1.03	0.65	0.49	Pass
Fat yield	1.05	0.69	0.42	Pass
Protein yield	0.98	0.54	0.36	Pass
SCS	0.98	0.62	0.42	Pass

Many countries started large-scale female genotyping on a routine basis only some years ago, thus the history of female animal genotyping is rather short in the countries. This may be a limiting factor for conducting genomic validation by truncating data in last four years, as recommended by the current rules of the Interbull GEBV test or trend validation test. Therefore, deleting phenotype data of fewer than four years would make a forward prediction more realistic for the genomic validation.

Table 9. – Validation results using deregressed national EBV of the validation cows as dependent variable

	Model 1		M1-M2	Test
Trait	b_1	R^2	ΔR^2	result
Milk yield	0.89	0.29	0.19	Fail
Fat yield	0.89	0.27	0.16	Fail
Protein yield	0.82	0.21	0.11	Fail
SCS	0.86	0.16	0.10	Fail

The second criterion of the current Interbull GEBV test requires that the R^2 increase from the conventional model 1 to genomic model 2 must be greater than 0, $\Delta R^2 > 0$. A truncated (and/or a full) conventional evaluation is thus needed for the GEBV test. The original idea of this test criterion was to quantify if the use of genomic information led to an increase in accuracy of genomic prediction. Because the special conventional evaluation will not be done routinely in the single-step evaluation by the countries, conducting this special convention evaluation, with the full or truncated data or both, does not seem to be justified, fifteen years after the introduction of the genomic selection in dairy cattle.

Alternative ways of testing the increase in prediction accuracy need to be sought.

As the validation animals in our study, we chose both validation bulls and validation cows. These two groups of validation animals behaved differently in the GEBV test. We could also select foreign bulls without domestic daughters as a new group of validation animals. According to the current rule of the GEBV test, the validation animals must have own phenotype data, validation bulls with daughters or validation cows with own records, in the full evaluation. However, we could extend the definition of the validation animals to young animals that do not have own phenotypic data in the full evaluation yet. For example, we could compare GEBV of second-generation candidates in the truncated evaluation to their GEBV as first-generation candidates in the full evaluation for the GEBV test.

The new GEBV test software (Sullivan 2022) already provided several useful features. Extra regression terms (VanRaden 2021) could be added to the current simple linear regression model for enabling the countries to detect potential bias in their genomic evaluation and to improve their national genomic models. In this study, we grouped the validation bulls or cows by birth year to investigate the impact of distance of the validation animals to the reference population. A fixed effect of birth year may be added to the simple regression model, the genomic model 1, to see if the distance of the validation animals to the genomic reference population may have an impact on the bias and accuracy of their GEBV. Other factors, like country of origin or generation number of the validation animals, may be considered in an extended GEBV test, too.

Conclusions

The newly developed GEBV test software by Interbull allowed GEBV or deregressed GEBV to be used as dependent variable of the linear regression model for genomic validation, besides the original deregressed conventional EBV. We applied the new GEBV test python software to the single-step evaluation of German Holsteins for four test-day traits. We compared three forms of dependent variable: GEBV, deregressed GEBV and deregressed conventional EBV, for both validation bulls and validation cows as the targeted group. The new GEBV software provided validation results as expected for all the tested scenarios, in terms of regression slope b_1, R^2 value of the genomic model 1 and R^2 value increase from the conventional model 2 to the genomic model 1. We observed notable variation in the validation results among the subgroups of the validation animals, e.g., born in different years. In general, the dependent variable GEBV resulted in higher model R^2 value than the DGEBV or deregressed EBV, especially for the low-reliability validation cows. The deregressed GEBV appeared to be a more appropriate form of response variable for the GEBV test than the GEBV itself.

Acknowledgments

Our colleagues, Dr. Stephan Rensing and Dr. Johannes Heise, are thanked for discussion.

References

Alkhoder, H., Liu, Z., Segelke, D. and Reents, R. 2022. Application of a single-step SNP BLUP random regression model to test-day yields and somatic cell scores in German Holsteins. Interbull Bulletin 57:74-83.

Legarra, A. and Reverter A. 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.

Liu, Z. and Masuda, Y. 2021. A deregression method for single-step genomic model using all genotype data. Interbull Bulletin 56:41-51.

Liu, Z., VanRaden, P.M., Lidauer, M.H., Calus, M.P., Benhajali, H., Jorjani, H. and Ducrocq, V. 2017. Approximating genomic reliabilities for national genomic evaluation. Interbull Bulletin 51:75-85.

Mäntysaari, E. A., Liu, Z. and VanRaden, P. M. 2010. Interbull validation test for genomic evaluations. Interbull Bulletin 41:10-14.

Sullivan, P. 2022. The new python software for GEBV test.

VanRaden, P.M. 2021. Improved genomic validation including extra regressions. Interbull Bulletin 56:65-69.