

Experience with single-step GEBV validation in CRV Holstein-Friesian breeding program

I. Jibrila and H. Eding

AEU CRV u.a., P.O. Box 454, 6800 AL Arnhem, The Netherlands

E-mail: herwin.eding@crv4all.com

Abstract

Validation of genomic evaluations is an important step in the development cycle. The usual way to validate is to regress deregressed conventional EBV of young bulls from a full evaluation on predicted GEBV from a reduced evaluation, where some recent years of data have been omitted. Some alternative approaches to validation of (single step) GEBV were tested. The types of validation performed were regression of GEBV or deregressed GEBV on predicted GEBV of either validation bulls or validation cows. Additionally we compared validations on DEBV or DGEVB for some traits using either in house validation methods or the newly implemented Interbull validation software. The results indicate that validations on DGEVB produce satisfactory results. Validation using cow GEBV can be used when the number of validation bulls is too limited (< 300 bulls). Validation on DEBV seems no longer to be a good approach, since input GEBV are often more reliable than the DEBV they are tested against.

Introduction

Validation of genomic evaluation methods is an important step in the development of such methods. Additionally it is an important part of the procedure to accept (G)EBV from evaluation methods by Interbull. The standard Inter-bull method of genomic validation is a regression of conventional realized deregressed EBV (DEBV) from a full dataset on genomic predictions (GEBV of validation bulls without progeny in a reduced dataset). However, since the introduction of this method several generations of genotyped dairy cattle have entered evaluations. The question is whether the current method of validation is still valid. We tested several alternative methods of validation as well as the validation procedures implemented in the latest Interbull software (Sullivan, 2023).

Methods

General validation method

For routine in house validations we use two sets of GEBV, from two single step evaluations:

1) A FULL evaluation using all available data. GEBV from this evaluation, or their deregressed proof (DGEVB; VanRaden *et al.*, 2009) are used as the dependent variable Y in validation regressions.

2) A reduced (RED) evaluation in which the four most recent years of data have been omitted. Animals with only genotypes (and pedigree) in RED are selected to serve as validation animals. For the validation two regressions are performed:

$$Y = a_0 + a_1 PA_{RED}$$

$$Y = b_0 + b_1 GEBV_{RED}$$

Since validation animals have no own observation or progeny in the RED data, the $GEBV_{RED}$ is a function of the parent average PA_{RED} and the genotype. The difference in R^2 from these regressions is a measure of the information added by the genotype.

Validation on (D)GEBV using bulls or cows

We compared validation on (D)GEBV using validation bull or validation cows. Validation animals were selected on the condition that no offspring/own observation are present in RED

and own observations (cows) or at least 20 daughters (bulls) are available in the FULL data. The main statistics of interest was the regression factor b_i and is presented here.

The traits presented here are traits for which standard validations do not always produce satisfactory results. They are:

- Udder health
 - Clinical mastitis (CM)
 - Subclinical mastitis (SCM)
- Livability
 - Maternal stillbirth in cows (MSB)
 - Direct stillbirth in cows (DSB)
- Milk fever/ketosis
 - Milk fever (MFE)
 - Clinical ketosis (KET)
- Reproductive disorders
 - Retained placenta (RET)
 - Endometritis (EMT)
 - Metritis (MT)
 - Cystic ovaries (COV)
 - Anoestrus (ANO)

Validation on DGEBV versus DEBV

To test validations on either deregressed GEBV or deregressed conventional EBV, and compare results from in house validations to results obtained with the new Interbull validation software (Sullivan, 2023), we did a number of additional validations, where the FULL evaluation was either a genomic single step evaluation (GEBV) or a conventional evaluation (EBV) which were deregressed using VanRaden *et al.* (2009).

A RED dataset with genomic predictions was made by performing a single step genomic evaluation on data in which the 4 most recent years of observations were omitted. Selection of bulls for validation was done according to standard practice (no offspring in RED, a min. of 20 dau in FULL). These validations were done on the following traits:

- Milking speed (MS)
- Temperament (TEM)
- Maternal stillbirth of cows (MSB)
- Direct stillbirth from cows (DSB)
- Clinical mastitis (CM)

Table 1. Regression factors b_0 from bull validations on DGEBV and GEBV

Trait	h ²	N _{bulls}	regression	
			Bulls	
			DGEBV	GEBV
SCM	0,056	763	1,03	0,99
CM	0,062	775	0,89	0,86
MSB	0,005	485	1,12	1,03
DSB	0,006	571	0,96	0,88
MFE	0,035	51	0,59	0,61
KET	0,096	115	0,62	0,60
RET	0,064	272	0,69	0,65
EMT	0,060	256	0,67	0,65
MT	0,049	224	0,61	0,60
COV	0,029	210	0,85	0,79
ANO	0,034	293	0,70	0,67

Table 2. Regression factors b_0 from cow validations on DGEBV and GEBV

Trait	h ²	N _{cows}	regression	
			Cows	
			DGEBV	GEBV
SCM	0,056	89 776	1,12	1,02
CM	0,062	14 698	1,08	0,96
MSB	0,005	71 515	1,33	1,06
DSB	0,006	66 131	1,22	0,94
MFE	0,035	3 565	1,09	0,94
KET	0,096	8 619	0,87	0,75
RET	0,064	17 672	0,90	0,78
EMT	0,060	17 519	0,93	0,79
MT	0,049	14 543	1,01	0,81
COV	0,029	12 186	1,00	0,85
ANO	0,034	18 325	0,89	0,77

Results

In Table 1 the values of the regression coefficient b_0 are given for the bull validations using DGEBV vs GEBV of the traits mentioned above. In Table 2 the values of b_0 are given for cow validations. It can be seen from Table 1 that validations consistently failed for traits where the number of validation bulls was smaller than 300. There was no great difference between validation on DGEBV or GEBV, although

regression on DGEBV tended to give better results overall in bulls.

From Table 2 it can be seen that milk fever and reproductive disorder traits successfully validated when validation cow GEBV were regressed on DGEBV. For stillbirth traits however a substantial underestimation was observed.

Validation on conventional deregressed EBV

In Table 3 the results on regression factors from the validations are presented. Validations were either the in house AEU validation or the validation as implemented in the latest Interbull GEBV test (Sullivan, 2023). What can be seen from this table is that validation results were comparable when regression was on DGEBV. In fact, all traits passed validation successfully in both the AEU and Interbull method of validation. Regression of DEBV on predicted GEBV, however, showed varying results with no traits successfully validating in both AEU and ITB systems of validation.

Table 3. Regression factors of validations of GEBV predictions on deregressed GEBV or EBV using methods from AEU and Interbull (ITB)

Trait	regression			
	on GEBV		on EBV	
	AEU	ITB	AEU	ITB
MS	1,07	1,10	0,74	0,89
TEM	0,92	0,80	0,61	1,16
DSB	0,93	0,70	0,55	0,43
MSB	0,99	0,76	0,64	0,71
CM	0,85	0,79	0,56	1,48

Discussion

DGEBV versus GEBV validation

The regression of DGEBV on validation animal GEBV showed more consistent results than did the regressions of (non-deregressed) GEBV, especially in cow validations. The results from Table 1 indicate that for a successful validation on bulls a sufficient

number of validation bulls is essential. The results indicate a number larger than 450 validation bull is a prerequisite for successful validation.

Bull versus cow validation

For traits with sufficient numbers of validation bulls (SCM, CM, MSB and DSB) the differences in regression factors between bull and cow validations were not substantial, although bull validations did show less deviations from 1 compared to cow validations.

The regression factors for cow validation on DGEBV did show larger deviations from 1. This might be an indication that the deregression procedure for these traits is problematic for deregressed GEBV of low reliability (note also the very low h^2 of MSB and DSB). For other traits, however, especially those traits where the number of validation bulls was limited, the results from cow validations seemed to produce useful results.

Validation on DEBV versus DGEBV

Testing validation of predicted GEBV on realized conventional DEBV showed uneven results in both AEU and Interbull methodology. This may be a result of the fact that nowadays we have multiple generations of genotyped animals. This means that for validation animals not only the prediction of the mendelian sampling term changes by adding genotype information. The parent average of an animal also changes, particularly if its parents are relatively young and genotyped. Hence the validation regresses relatively reliable predicted GEBV on deregressed proofs from less reliable EBV, whereas a validation on DGEBV ensures the values tested against are more reliable than the input for the regression. Hence it would seem to be preferable to validate against deregressed proofs from realized GEBV.

References

- Sullivan P., 2023, 'New GEBVtest Program',
Interbull Technical Workshop, Maccarese,
Rome, Italy
- VanRaden, P. M., C. P. VanTassell, G. R.
Wiggans, T. S. Sonstegard, R. D. Schnabel,
J. F. Taylor, and F. S. Schenkel, 2009.
Invited review: Reliability of genomic
predictions for North American Holstein
bulls. *J. Dairy Sci.* 92:16-24.