Implementation of Genomic Evaluation in German Holsteins

***F. Reinhardt*[1], *Z. Liu*[1], *F. Seefried*[1] *and G. Thaller*[2]**
[1]*vit w.V., Heideweg 1, 27283 Verden, Germany*
[2]*Christian-Albrechts-University, Inst. Animal Breeding and Husbandry, 24908 Kiel, Germany*

Abstract

Accurate selection of breeding animals can be enhanced by using genomic information. For genomic evaluation of German Holsteins, we developed a genomic model that estimated effects of 45,181 SNP markers, obtained from Illumina bovine chip SNP50, and a residual polygenic effect jointly. Variance of the residual polygenic effect was determined with a single parameter which can vary between traits. All markers were assumed to have equal variance. Using genotypes of 4570 German Holstein bulls with daughters, SNP marker effects were estimated for a total of 44 traits belonging to seven trait groups. Our genomic evaluation system comprises, besides routine conventional cow evaluation, four evaluations: a pure genomic evaluation using a BLUP marker model, a subset conventional evaluation including genotyped animals only, and a combined conventional and genomic evaluation with a BLUP method, and a conventional evaluation for calculating pedigree index for young genotyped animals without phenotypes. Reliabilities of direct genomic values were obtained by inverting genomic relationship matrix, and extra genomic effective daughter contributions were calculated from the gain in reliability of using genomic relationship rather than average relationship among genotyped animals. A routine genomic evaluation in August 2009 included 5982 genotyped animals with 1091 young calves. A validation study was conducted by treating 655 genotyped bulls born in the last birth year 2004 as young calves. A relatively high gain in realised reliability was found for all traits in the validation, except some traits with very low heritability. Due to the use of male pedigree index, estimated genomic breeding values of young calves appeared to be not overestimated. Observed risk of a wrong culling did not seem to be high, but the realised power of making a right selection on an individual animal basis appeared to be not satisfactory. By comparing SNP effect estimates across genomic evaluations, we found that variance of the SNP effects increased significantly with more genotyped animals added to reference population. Even for the BLUP marker model with equal marker variance, SNP effect estimates differentiated from one another markedly as the training population enlarged. A significant improvement in reliability was seen between estimated genomic breeding values and conventional pedigree index.

1. Introduction

Ever since Meuwissen *et al.* (2001) introduced the concept of genomic evaluation, great efforts have been put on developing and implementing genomic evaluation, e.g. using the Illumina bovine chip SNP50k, in dairy cattle worldwide (VanRaden 2008). In Germany, a national genome project started in 2006 with the goal of implementing genomic selection for German Holstein breed. Early research projects (Liu *et al.*, 2009a) were focused on comparing statistical models using simulated genomic data. As more and more animals were genotyped, statistical methods and procedures were needed to analyse the real data. The objectives of this study were to develop a genomic model and a genomic evaluation system for German Holsteins, and to validate the genomic evaluation system based on the field data.

2. Materials and Methods

2.1. A genomic evaluation model

A statistical model was applied to both genotypic and phenotypic data of genotyped animals:

$$q_i = \mu + v_i + \sum_{j=1}^{p} z_{ij} u_j + e_i \qquad [1]$$

where q_i is a deregressed proof or daughter yield deviation of a bull *i*, μ is a general mean, v_i is residual polygenic effect of bull *i*, p is number of fitted bi-allelic SNP markers

($j = 1, \cdots, p$), z_{ij} is genotype value (-1 and 1 for two homozygotes and 0 for heterozygote) of marker *j* of bull *i*, u_j is random regression coefficient for marker *j*, and e_i is residual effect for the record of bull *i*. A small fraction of genetic variance, currently set to less than 1%, was assumed for modelling the residual polygenic effect. Fitting the residual polygenic effect accounted for the fact that markers may not explain all genetic variance and it can also avoid the problem that the markers captured the relationship among animals if the genomic model did not include the polygenic effect. The fitted polygenic effect of the genomic models was analysed in the same way as in conventional genetic evaluation, i.e. using full pedigree and identical grouping of phantom parent groups.

Because daughter yield deviations (DYD) resulting from a multi-trait model, e.g. random regression test-day model (Liu *et al.*, 2004), cannot be optimally analysed with the single trait genomic model, a single trait deregression procedure considering full animal pedigree was applied to derive deregressed proofs (DPRF) for genotyped bulls. Effective daughter contributions (EDC) were used as weighting factor for DPRF in the genomic evaluation. During the development of the genomic evaluation system, alternative genomic models (Liu *et al.*, 2009a) differing in prior variance functions of the fitted markers were compared (data not shown here) using real genomic data for German Holsteins. We found that the models fitting fewer markers, i.e. non-linear models assigning high variances to big markers, gave poorer goodness of fit than the genomic models fitting all available markers, e.g. BLUP model EQ assuming equal marker variance or nonlinear model E1. This finding and others suggested that the assumption of infinitesimal model for quantitative traits may be closer to the true genetic inheritance model. Due to the negligible difference in accuracy between the BLUP model EQ and non-linear model E1, we preferred the BLUP model, which assumed equal marker variance, for reasons such as consistency with conventional polygenic genetic evaluation model, accurate reliability calculation and robustness against diverse underlying true genetic models between traits.

2.2. The genomic evaluation system

Our genomic model fitted a residual polygenic effect, and its variance proportion was determined with a single parameter, which can be easily modified without changing anything else in the whole genomic evaluation system. For setting up the genomic relationship matrix, allele frequencies of base population were estimated using the gene content method (Gengler *et al.*, 2007). Reliabilities of direct genomic values (DGV) were calculated by inverting genomic relationship matrix (VanRaden 2008). Conventional estimated breeding values (EBV) and DGV were combined using a BLUP approach (Ducrocq and Liu, 2009) instead of a selection index method on an animal by animal basis.

Four genetic evaluations were required for a complete genomic evaluation. Besides a conventional evaluation, denoted as E1, including all animals, e.g. with a random regression test-day model (Liu *et al.*, 2004), a pure genomic evaluation (E2) using model 1, a subset conventional evaluation involving only genotyped animals (E3), and a conventional evaluation combining with genomic information (E4), were conducted. The subset conventional Evaluation E3, containing the same genotyped animals as E2, was needed in order to calculate the reliability gain by using genomic relationship rather than average relationship among genotyped animals (VanRaden, 2008). For genotyped young animals without phenotypes, Evaluation E4 provided genomic enhanced pedigree index (GPI). Animals providing phenotypic data were bulls with daughters plus genotyped animals with pseudo-records in Evaluation E4.

Reliability of EBV from Evaluation E2 was computed by indirect inversion of mixed model equation coefficient matrix. Extra genomic EDC was calculated for each genotyped animal as:

$$\psi = \alpha\left(\frac{R_{\hat{a}}^{2[E2]}}{1 - R_{\hat{a}}^{2[E2]}} - \frac{R_{\hat{a}}^{2[E3]}}{1 - R_{\hat{a}}^{2[E3]}}\right) \qquad [2]$$

where ψ is the extra genomic EDC, α is the ratio of residual to sire variance, $R_{\hat{a}}^{2[E2]}$ and $R_{\hat{a}}^{2[E3]}$ represent reliability of Evaluations E2 and E3, respectively.

For each genotyped animal, a pseudo-record was generated with:

$$q = \hat{\mu} + (\hat{a} - \hat{\mu}) / R^2_{\hat{a}} \qquad [3]$$

where q represents the generated record which is analogue to deregressed proof, $\hat{\mu}$ is estimated general mean of genotyped animals in reference population, $\hat{a}$ is estimated DGV, sum of all SNP effect estimates, for the animal, and $R^2_{\hat{a}}$ denotes reliability of $\hat{a}$ from Evaluation E2. For genotyped animals with phenotypes, the pseudo-record was a weighted average of deregressed proof from conventional evaluation and q from Formula 3:

$$q_{comb} = (\psi\, q + \psi_C\, q_C) / (\psi + \psi_C) \qquad [4]$$

where q_{comb} denotes combined pseudo-record, q_C is deregressed proof from conventional evaluation, and ψ_C is EDC from conventional evaluation. Weighting factor for q_{comb} is sum of both EDC: $\psi + \psi_C$.

For genotyped young animals, pedigree index was required for combination with DGV, if no conventional EBV were available. Pseudo-record q and associated EDC ψ were set to a tiny positive number and Evaluation E4 was run with the modified genomic input data. Resulting EBV from the modified evaluation E4, denoted as E4pa, were equal to parental average for the young genotyped animals without own phenotype. Due to concerns about overestimated conventional EBV of bull dams, animals providing data in Evaluations E4 and E4pa were bulls with daughters, no cows with records. Phenotypic data of bull dams were not directly considered in these evaluations, but their records were considered in DYD or DPRF of their sires. Therefore, the pedigree index derived from Evaluation E4pa is actually male pedigree index based on all available pedigree information, it is not a full parental average of sire and dam EBV.

Reliabilities of DGV from Evaluation E2 and reliabilities of EBV from the subset conventional Evaluation E3 were calculated by inverting mixed model equations, thanks to the relatively low number of genotyped animals. In contrast, reliabilities of EBV of Evaluations E4 and E4pa were approximated using a similar reliability calculation method as Liu *et al.* (2004), since the much higher number of animals in both evaluations made direct matrix inversion infeasible.

2.3. Marker and phenotypic data processing

Genomic data of German Holsteins comprised about 2500 bulls born from 1998 to 2002 and 500 cows from the German national genome project (GenoTrack) and about 2500 bulls born before 1998 and in 2003 and 2004 from routine genotyping, those animals were genotyped using Illumina chip Bovine SNP50 BeadChip. Minor allele frequency was set to 0.01 and call rate to 0.90. Total number of selected SNP markers was 45,181 for genomic evaluations. Special handling was done for 533 markers on X chromosome with regard to estimating marker allele effect and base-population frequency, because male animals have only one allele. Call rate for animals was set a minimum of 0.95.

Deregressed proofs and EDC were obtained from most recent conventional evaluation for all bulls with at least 10 EDC. A total of 44 traits from seven trait groups were considered in genomic evaluation: milk production (3 traits), udder health (1 trait), function longevity (1 trait), calving (4 traits), female fertility (6 traits), workability (4 traits) and conformation (25 traits). Total merit index and sub-indices for each trait group were calculated together with the approximated relibilities. Either individual traits or indices had three forms of breeding values available: direct genomic, conventional and combined of both components.

2.4. Data materials for genomic evaluations

Table 1 summarises genomic and phenotypic data used for a routine genomic evaluation in August 2009. A total of 4572 genotyped bulls had daughters with milk yield proofs. This genomic evaluation included also 1091 genotyped animals born between 2005 and 2009.

For validating the genomic evaluation system all genotyped bulls born in 2004 were chosen as validation animals, and 3684 genotyped bulls born before 2004 were selected to estimate SNP effects, simulating genomic evaluation four years ago, i.e. August 2005. Table 2 shows the number of genotyped bulls in training and validation sets.

Table 1. Genomic and phenotypic data for a genomic evaluation in August 2009.

Genotyped animals	Birth year	No. animals	with EBV milk yield
Bulls with daughters	≤1997	612	611
	1998	379	379
	1999	447	447
	2000	489	489
	2001	489	488
	2002	495	489
	2003	966	955
	2004	1014	714
Without phenotypes	2005-2009	1091	0
All		5982	4572

Table 2. Genomic data for a validation evaluation.

Genotyped bulls	Birth year	No. animals	Sum
Training set (reference population)	≤1997	588	3684
	1998	360	
	1999	424	
	2000	440	
	2001	445	
	2002	477	
	2003	950	
Validation set	2004	655	4339

3. Results

3.1. Validation results

A genomic evaluation was conducted by removing the 655 genotyped bulls born in 2004. SNP marker effects were re-estimated using the sub-set of data, simulating a genomic evaluation in August 2005. For the validation bulls, conventional pedigree index (PI) as well as genomic enhanced GPI were calculated in Evaluation E4pa and E4, respectively. Those PI and GPI were compared to deregressed proofs from conventional August 2009 evaluation for those 2004 bulls, in order to validate the genomic evaluation system. Relationship coefficients of the validation bulls were calculated with the training animals and they were correlated with their reliabilities of DGV. We found significant correlations of the reliability of DGV with average value of relationship (0.46) and maximum value of relationship (0.36) of validation animal to all animals in the reference training population. Validation bulls with genotyped sires in the training set had higher reliability of DGV than those without genotyped sires. It is obvious from this study that validation result depends on the distance between training and reference populations.

Tables 3 and 4 give results of the validation study. Reliability of PI did not vary much across traits. For three production traits, gain in realised reliability is about 30%, with highest gain for fat yield, possibly due to the DGAT gene influence. Low heritability traits, such fertility traits and stillbirth, have lowest gain in realised reliability, this can be partially explained by the fact that reference bulls had much lower reliability of conventional EBV than for the other traits. The remaining traits had a realised gain in reliability between 14% to 24%, with an exception of udder depth of 31%. In general, the gain in realised reliability was high, which was contributed by the big size of reference population for SNP effect estimation.

Table 3. Reliabilities for 2004 validation bulls for all but type traits .

Trait	Pedigree index	Realised reliability GPI	Gain
Milk yield	36	68	32
Fat yield	36	72	35
Protein yield	36	64	28
Somatic cell scores	36	54	19
Longevity	34	51	17
NR56 heifer	33	40	7
Days open	34	42	8
Stillbirth maternal	33	42	9
Milking speed	32	58	25

Based on the validation data, we calculated relative risk of making a wrong culling (Table 5) or a wrong selection (Table 6) using milk yield estimates. The validation bulls were selected or culled based on their GPI. And their GPI were compared to conventional EBV four years later, obtained from August 2009 evaluation. A total of 556 validation bulls were culled at GPI < 1 standard deviation (Table 5), 56 of the culled bulls actually had conventional EBV in August 2009 greater than 1.0 genetic standard deviation above average, corresponding to a risk of 10.1%, which was equal to Type I error of rejecting a null hypothesis that was true.

Table 4. Reliabilities for 2004 validation bulls for type traits.

Trait	Pedigree index	Realised reliability GPI	Gain
Stature	33	53	20
Angularity	33	54	21
Rump angle	35	48	14
Udder depth	32	63	31
Front teat placement	32	52	20
Body depth	32	53	20
Chest width	32	60	27
Fore udder attachm.	32	59	27
Overall body	33	48	16
Rear leg set	32	48	17
Rear udder height	35	51	16
Rump width	33	52	20
Udder support	34	56	21
Milk type	33	51	19
Rear teat placement	35	53	18
Locomotion	30	40	10
	32	52	20

Table 5. Risk of a wrong culling calculated with 655 validation bulls born in 2004.

Min. GPI for culling	No. culled animals	Frequency and percentage of bulls with 2009 conventional proofs being: >1.0s	>1.5s	>2.0s	>2.5s
-0.5s[§]	193	2	1	1	
	29%	1.0%	0.5%	0.5%	
0.0s	306	6	1	1	
	47%	2.0%	0.3%	0.3%	
0.5s	441	22	4	3	1
	67%	5.0%	0.9%	0.7%	0.2%
1.0s	556	56	16	6	1
	85%	10.1%	2.9%	1.1%	0.2%

[§] s stands for standard deviation.

Similar to Table 5, risk of a wrong selection was calculated for the validation bulls. A total of 99 bulls were selected, based on their GPI value > 1.0 standard deviation above average. Because 47 of them had conventional EBV in August 2009 less than 1.0 standard deviation above the mean, the risk of a wrong selection was 47% (type II error). In contrast to the relatively low type I error in Table 5, the power of making a correct selection is still unsatisfactory on an individual animal level. However, the risk of making wrong selection can be reduced or the power of selecting right animals increased significantly, if genotyped animals are considered jointly, e.g. a group of genotyped calves.

Table 6. Risk of a wrong selection calculated with 655 validation bulls born in 2004.

Min. GPI for culling	No. selected animals	Frequency and percentage of bulls with 2009 conventional proofs being: < 0.0s	< 0.5s	< 1.0s	<1.3s
1.3s[§]	57	5	14	25	29
	8.7%	8.8%	25%	44%	51%
1.0s	99	10	31	47	55
	15%	10%	31%	47%	56%
0.5s	214	36	87	128	153
	33%	17%	41%	60%	72%

[§] s stands for standard deviation.

3.2. Genomic evaluation results

In a routine genomic evaluation of August 2009, 23572 bulls with EDC ≥10 and 7077 genotyped animals were jointly evaluated.

Among the genotyped animals, there were 4637 bulls with daughters in milk and the remaining were young calves or genotyped females. As explained before, GEBV or GPI was calculated by combining conventional with direct genomic information using a BLUP approach. Figure 1 shows genetic trends in DGV, EBV (or PI in case of young animals), and GEBV (or GPI) in milk yield of German Holsteins. Note that the DGV curve was based on a much lower number of genotyped animals than the EBV or GEBV curves. There were much lower number of animals in birth years 2005 and 2006 and the years until 1998 for DGV. It can be seen that DGV curve overlapped completely with EBV or GEBV for the years from 1998 to 2004, because most of training bulls were born in this period. Older genotyped bulls have higher genetic trend of DGV than EBV or GEBV, this is due to the selection of sires of genotyped bulls born after 1997 and they had higher genetic merit than average of all bulls in those years. Genetic trend of DGV is lower than PI or GPI for youngest bulls or calves born from 2006 onwards, though the difference becomes smaller for the last three years 2007-2009. For young genotyped animals without phenotypes, no overestimation of PI and GPI is observed overall, this can be explained by the use of male pedigree index which circumvents the overestimated bull dam proof problem. A reasonable percentage of the young calves born since 2007 had foreign sires which had no daughter information in German national evaluation, this may cause the lower level of PI or GPI of the young calves.

3.3. SNP effect estimates

During the development of the German genomic evaluation system a number of test runs were conducted, which enabled to compare SNP effect estimate across evaluations. Table 7 shows comparison of SNP effect estimates of milk yield from six genomic test runs differing in number of genotyped bulls in training set. As the number of reference bulls increased from 735 to 4339, variance of SNP effect estimated increased more than four fold. Estimate of the SNP marker, which had the largest effect, increased from 1.34 to 5.22 continuously, expressed as standard deviation of average marker, as the size of reference population enlarged. The correlation of SNP marker estimates decreased between two runs, when the number of genotyped bulls differed more, as expected. Note that the correlation of SNP effect estimates is much lower than correlation of DGVs. Even under the BLUP model assuming equal marker variance, markers can have very different estimates, and these marker estimates differentiate more as more new genotyped animals are added to reference population.

Table 7. Comparison of SNP effect estimates from different test evaluations.

No. bulls (milk, kg)	SNP var[§]	SNP effect[$]	Correlation of all SNP estimates between evaluations				
			B	C	D	E	F
735 (A)	100	1.34	.81	.56	.50	.46	.43
1088 (B)	149	1.96		.69	.61	.55	.53
1939 (C)	261	3.29			.83	.72	.69
3081 (D)	371	4.15				.86	.84
3684 (E)	438	4.86					.95
4339 (F)	478	5.22					

[§]Relative variance of SNP effect estimates
[$]Largest (same) SNP effect estimate in standard deviation of average marker

4. Discussion

A BLUP model assuming equal marker variances was chosen for routine genomic evaluation for German Holsteins. This choice was based on several reasonings and findings, e.g. robustness of BLUP model against underlying distribution of true QTL or number of segregating QTL for quantitative traits. The infinitesimal model of quantitative traits make the BLUP marker model attractive, since small markers are also considered jointly with big ones. It can be seen clearly in Table 7 that differences in marker effect estimates became ever bigger, i.e. large markers differentiating more from small ones, with more genotyped bulls added to the reference population. Even under the assumption of equal marker variance, markers do allow to have quite

different estimates and their difference increases with more training animals.

We developed a BLUP approach to optimally combine information from conventional and direct genomic evaluations. This BLUP approach allows automatic propagation of genomic information to non-genotyped ancestors or progeny. It can also be used for correcting the bias caused by genomic pre-selection in conventional evaluation (Liu *et al.*, 2009b). It considers all animals, genotyped and non-genotyped, simultaneously. A more accurate way of generating pseudo-record than Formula 3 was proposed by Ducrocq and Liu (2009). By using degressed proofs of bulls and DGV of genotyped animals, the BLUP approach of Evaluation E4 can avoid the problem of overestimated bull dam proofs in genomic evaluation, which can be seen in Figure 1.

The one-step approach of genomic evaluation (Misztal *et al.*, 2009) appeared to be more appealing than our genomic evaluation system with four runs. In fact, these two approaches are well related to each other. Evaluation E2 or subset conventional evaluation E3 correspond to inverse matrix $\mathbf{G}^{-1}$ or $\mathbf{A}_{22}^{-1}$ in the single-step approach. Advantages of the one-step approach may possibly be made unattractive by the difficulty of inverting the ever increasing size of the genomic relationship matrix $\mathbf{G}$.

Realised gain in reliabilities of GPI from the validation study was high with exception of fertility traits with very low heritability, which was contributed by the high number of genotyped bulls in reference population. However, the realised reliabilities GPI were lower than theoretical ones, this may be explained by the way the genomic relationship was set up, in which an assumption was made that all markers were in complete linkage disequilibrium with QTL. Methods should be developed to solve the problem of overestimated reliabilities of DGV.

5. References

Ducrocq, V. & Liu, Z. 2009. Combining genomic and classical information in national BLUP evaluations. Interbull meeting, Barcelona, Spain. *Interbull Bulletin 40*, 172-177.

Gengler, N., Mayeres, P. & Szydalowski, M. 2007. A simple method to approximate gene content in large pedigree populations: application to the myostatin gene in dual-purpose Belgian Blue cattle. *Animal 1:01*, 21-28. doi: 10.1017/S1751731107392628.

Liu, Z., Reinhardt, F., Bünger, A. & Reents, R. 2004. Derivation and Calculation of Approximate Reliabilities and Daughter Yield-Deviations of a Random Regression Test-Day Model for Genetic Evaluation of Dairy Cattle. *J. Dairy Sci. 87*, 1896-1907.

Liu, Z., Seefried, F., Reinhardt, F. & Reents, R. 2009a. Dairy cattle genetic evaluation using genomic information. *Interbull Bulletin 39*, 23-28.

Liu, Z., Seefried, F., Reinhardt, F. & Reents, R. 2009b. A simple method for correcting the bias caused by genomic pre-selection in conventional genetic evaluation. Interbull meeting, Barcelona, Spain. *Interbull Bulletin 40*, 184-188.

Meuwissen, T.H., Hayes, B.J. & Goddard, M.E. 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics 157*, 1819-1829.

Misztal, I., Aguilar, I., Johnson, D., Legarra, A., Tsuruta, S. & Lawlor, T. 2009. A unified approach to utilize phenotypic, full pedigree and genomic information for a genetic evaluation of Holstein final score. Interbull meeting, Barcelona, Spain. *Interbull Bulletin 40*, 240-243.

VanRaden, P.M. 2008. Efficient methods to compute genomic predictions. *J. Dairy Sci. 91*, 4414-4423.

Figure 1. Genetic trends of DGV, EBV and GEBV in milk yield