Brown Swiss Genomic Evaluation

B. Zumbach, H. Jorjani and J. Dürr
Interbull Centre, Box 7023, SE-750 07 Uppsala, Sweden

Abstract

In the framework of the Intergenomics project initiated by the European Brown Swiss Federation, a common reference population for genomic evaluation has been created to enable participating countries to estimate SNP effects more accurately. Genomic data of 3392 Brown Swiss bulls from 7 countries were evaluated for protein yield, somatic cell score and non-return rate based on GBLUP. Accuracy of GEBVs of bulls with at least 10 daughters for protein and somatic cell score were larger by about 5% than the conventional proofs, for non-return rate the difference was about 13%. For young bulls genomic reliabilities were about 2.5 times larger than the respective parent averages. Correlations between conventional proofs, DGVs and GEBVs were close to 1.

Introduction

Genomic evaluation is becoming integrated into national evaluation in an increasing number of countries. As the gain in accuracy through genomics depends on the size of the reference population (e.g., Goddard and Hayes, 2008; Hayes *et al.,* 2009) small populations as the Brown Swiss (BSW) breed are in disadvantage. With the Intergenomic Project conducted at the Interbull Centre in Uppsala, Sweden, a common reference population has been created to enable participating countries to estimate SNP effects more accurately. The objective of this paper is to present first results on protein yield, somatic cell score and Non-Return Rate.

Materials and Methods

Genomic data of 3,392 BSW bulls from 7 countries born between 1936 and 2008 were considered in the analysis (Table 1). Those animals were genotyped using Illumina Bovine SNP50 Bead Chip (Illumina, San Diego, CA). Minor allele frequency was set to 0.02 and call rate to 0.75. Furthermore, markers in perfect LD were excluded resulting in a total of 42,437 selected SNPs (Jorjani *et al.,* 2010).

In total, three MACE (Multiple Across Country Evaluation, see Schaeffer, 1994) runs were performed for generating phenotypic data: 1 run for genomic evaluation (Run 3) and 2 runs for validation (Run 1 and Run 2).

Table 1. Contribution of genotypes used for the genomic evaluation by country.

Country	Number of genotypes
Austria	136
Germany	517
France	86
Italy	745
Slovenia	188
Switzerland	1,091
United States of America	629
Total	3,392

For genomic evaluation (Run 3) national data for January 2010 routine evaluation were used as input for a special MACE run. This special run differed from the routine MACE run by lowering the cutoff year of proofs to 1920 and the pedigree cutoff year to 1900 in order to use the maximum information for the genotyped bulls. With regard to number of herds and daughters the usual BSW specific editing rules with a minimum of 10 was applied. For bulls without own proofs parent averages were calculated. All data were standardized to a mean = 0 and standard deviation =1.

Table 2 summarizes the genomic and phenotypic data used for the genomic evaluation (Run 3). The proportion of animals with daughter information was 94% for protein yield, 87% for somatic cell score and 69% for non-return rate.

The validation MACE runs were as follows: Run 1 (full data) was as the run for the genomic evaluation (Run 3), however without any restriction with regard to number of herds and daughters (minimum=0).

As at the time of the analysis no historical national data for validation were available in the reduced dataset (run 2) the proofs of 20% of the youngest bulls with an EDC of ≥20 were excluded and parent averages were used instead. The same relaxed editing rules as in Run 1 were applied.

The validation set contained 467 (non-return rate) to 608 bulls (protein yield) born after 1999 (Table 3). While all the proofs of these bulls were based on full information in run 1, the validation bulls' proofs in run 2 were more based on parent averages. The following 3 models have been tested:

MODEL0: EBV(run 1) = a + b DGV (run 2)

MODEL1: EBV(run 1) = a + b GEBV(run 2)

MODEL2: EBV(run 1) = a + b EBV (run 2)

Parameters used in validation are correlation, intercept, regression coefficient and R^2. In the absence of selection the expectation of intercept and slope are 0 and 1, respectively (Boichard *et al.*, 1995).

Data of all three runs were analyzed with the GBLUP package provided by Paul VanRaden based on the methodology described by VanRaden (1998).

Results and Discussion

Genomic evaluation results

Reliabilities for bulls with own proofs varied according to country scale and trait: Differences among country scales were lowest for protein yield and highest for cow conception 1. Genomic reliabilities increased most for country scales with lower conventional reliability, i.e. the highest increase in reliability through genomics was for USA for non-return rate (Table 4). There was already a considerable increase in reliability for proven bulls of about 6% (protein yield) to 18% (non-return rate). Bulls born after 2003 had highest increases in reliability (Figures 1 to 3).

The increase in reliability for bulls without daughter information is shown in Figure 4. Genomic reliabilities were about 2.5 times larger than the corresponding reliabilities based on parent averages only for all the traits. Genomic reliabilities were in the range of those reported for Holstein (e.g., de Roos *et al.,* 2009; Schenkel *et al.,* 2009; Reinhardt *et al.,* 2009). As Interbull has no performance records for dams the parent averages were relatively low.

Correlations between EBVs of bulls with data and GEBVs for protein yield were about 0.97 on all country scales. The same is true for somatic cell score with the exception of the USA scale where the correlation is 0.95. The correlations for non-return rate were slightly lower: 0.93 on CHE and DEU scale and 0.92 on the USA scale. Correlations between EBVs of bulls with data and DGVs were a bit lower: They ranged from 0.90 (non-return rate USA) over 0.95 (somatic cell score) to 0.97 (protein yield DEU). Estimates on the USA scale for somatic cell score and non-return rate were lower by about 0.02 compared to the other countries. These results are in accordance with those reported by Szyda *et al.* 2009. Correlations between DGVs and GEBVs (all data) were larger than 0.99 for all traits and on all country scales.

Validation results

The estimated values for intercept and slope were mainly different from the 0 and 1, respectively. This deviation might be mainly due to selective genotyping. Therefore, the expected values under consideration of selection have to be estimated which will be done in a later study.

The slope estimates from Model 0 and Model 1 were in the range of values that have been reported for Holstein and Fleckvieh populations (e.g., Berry *et al.*, 2009; Gredler *et al.*, 2009). The slope estimates for Model 2 in this study were generally high, especially for protein yield and somatic cell score which has to be seen with caution.

R^2 values for Models 0 and 1 were very similar and ranged from about 0.40 (protein yield) to 0.50 or higher (somatic cell score and non-return rate) which is in agreement with the results by Gredler *et al.* (2009) for Fleckvieh and those of Lund and Su (2009) for Holstein. R^2 values for Model 2 were clearly lower with the exception of protein yield. This shows an improvement of the genomic model over the conventional one. The unexpected values for protein could be partly due to the 'truncation' of national data sets in lack of historical national data, i.e. the parent averages include more information than they would have had four years before. A validation based on historical national data is planned.

References

Berry, D.P., Kearny, F. & Harris, B.L. 2009. Genomic selection in Ireland. *Interbull Bulletin 39,* 29-34.

de Roos, A.P.W., Schrooten, C., Mullaart, E., van der Beek, S., de Jong, G. & Vooskamp, W. 2009. Genomic selection at CRV. *Interbull Bulletin 39,* 47-50.

Goddard, M. 2009. Genomic selection: prediction of accuracy and maximization of long term response. *Genetica 136,* 245–257.

Gredler, B., Nirea, K.G., Solberg, T.R., Egger-Danner, C., Meuwissen, T.H.E. & Sölkner, J. 2009. Genomic selection in Fleckvieh/Simmental – First results. *Interbull Bulletin 40,* 209-213.

Hayes , B., Bowman, P.J., Chamberlain, A.J. & Goddard, M.E. 2009. Invited review: Genomic selection in dairy cattle: Progress and challenges. *J. Dairy Sci. 92,* 433-443.

Jorjani, H., Zumbach, B., Jakobsen, J., Hjerpe, E. & Dürr, J. 2010. BSW genomic data preparation. Interbull Meeting, Riga, Latvia, May 31 to June 4, 2010.

Lund, M.S. & Su, G. 2009. Genomic selection in the Nordic countries. *Interbull Bulletin 39,* 39-42.

Reinhardt, F., Liu, Z., Seefried, F. & Thaller, G. 2009. Implementation of genomic evaluation in German Holsteins. *Interbull Bulletin 40,* 219-225.

Schaeffer, L.R. 1994. Multiple-Country Comparison of Dairy Sires. *J. Dairy Sci. 77,* 2671-2678.

Schenkel, F.S., Sargolazei, M., Kistemaker, G., Jansen, G.B., Sullivan, P., Van Doormaal, B.J., VanRaden, P.M. & Wiggans, G.R. 2009. Reliability of genomic evaluation of Holstein Cattle in Canada. *Interbull Bulletin 39,* 51-58.

Szyda, J., Zarnecki, A. & Kaminski, S. 2009. The Polish genomic breeding value estimation project. *Interbull Bulletin 39,* 43-46.

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

Table 2. Data for genomic evaluation (Run 3).

Genotyped animals	Birth year	No. animals	With own proofs[1)]		
			protein yield	Somatic cell score	Non-return rate
	< 1960	13	13	13	13
	1960-1979	167	164	139	113
	1980-1989	427	425	362	260
	1990-1999	1,743	1,709	1,612	1,286
Bulls with	2000	288	282	271	211
daughter	2001	251	249	239	186
information	2002	223	217	206	185
	2003	88	87	75	68
	2004	38	36	26	17
	2005	55	12	11	5
PA only	2006-2008	99	0	0	0
All		3,392	3,194	2,954	2,344

[1)] EBVs including daughter performances.

Table 3. Genomic data for validation (Run 1 and Run 2).

Genotyped animals	Birth year	No. animals		
		protein yield	Somatic cell score	Non-return rate
	< 1960	13	13	13
	1960-1979	162	138	114
Reference	1980-1989	425	364	263
population	1990-1999	1,692	1,616	1,339
	2000	214	271	216
Validation set	1999-2005	608	564	467
Bulls without proof	1969-2008	278	426	980

Table 4. Mean conventional (rel) and genomic reliabilities (grel) for bulls with daughter information for different traits and different country scales.

Country	**Protein yield**		**Somatic cell score**		**Non-return rate**	
scale	*rel*	*grel*	*rel*	*grel*	*rel*	*grel*
CHE	81.7	86.9	82.3	87.3	58.4	71.4
DEU	80.3	85.8	81.7	86.8	58.2	71.3
FRA	74.5	81.9	79.3	85.0		
ITA	78.0	84.1	79.0	84.8		
SVN	73.7	81.5				
USA	74.8	82.2	69.9	79	44.4	61.5

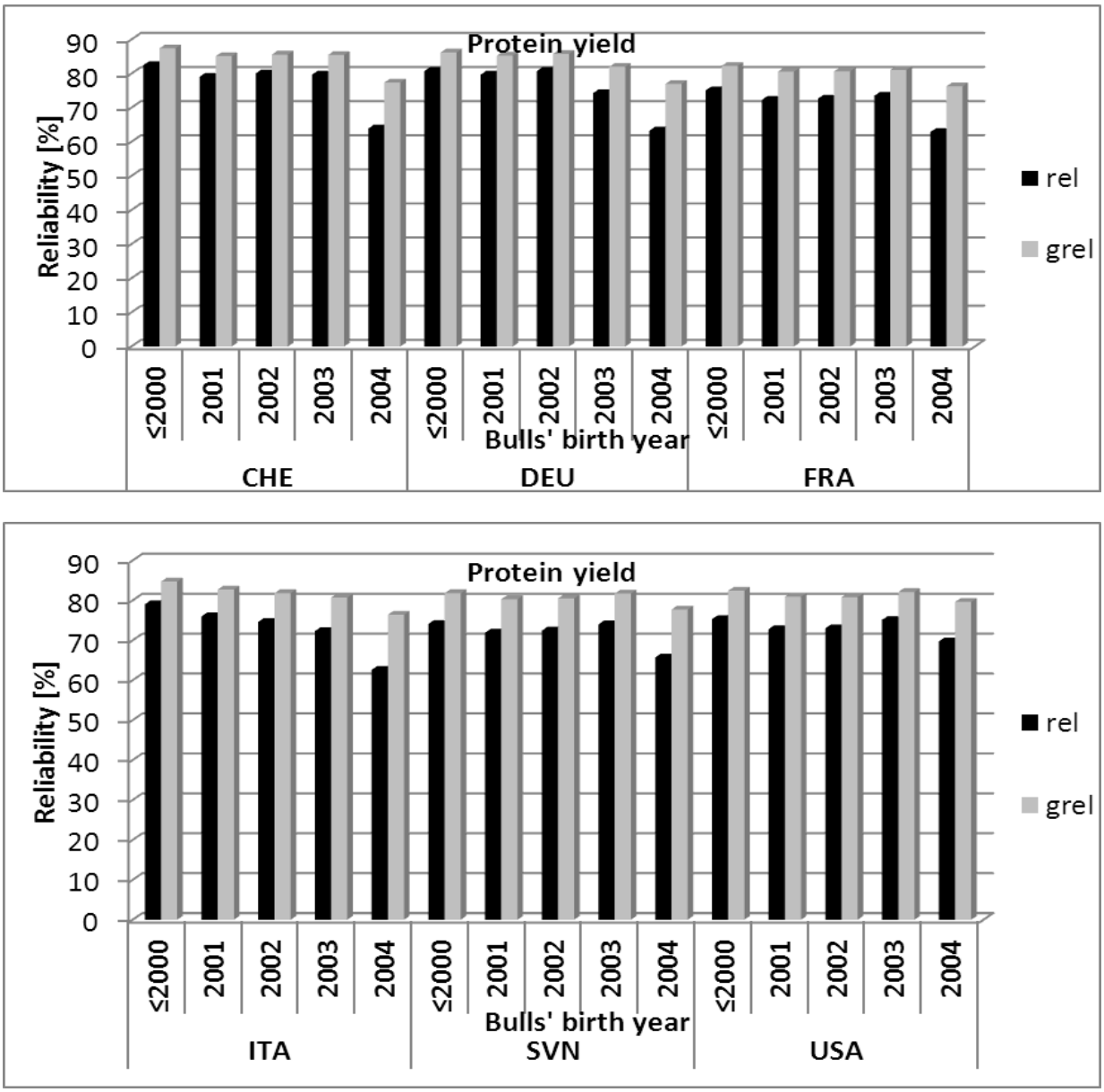

Figure 1. Conventional (rel) and genomic reliabilities (grel) for bulls with daughter information by year of birth on different country scales for protein yield.
Birth year 2004 also includes also records from birth year 2005

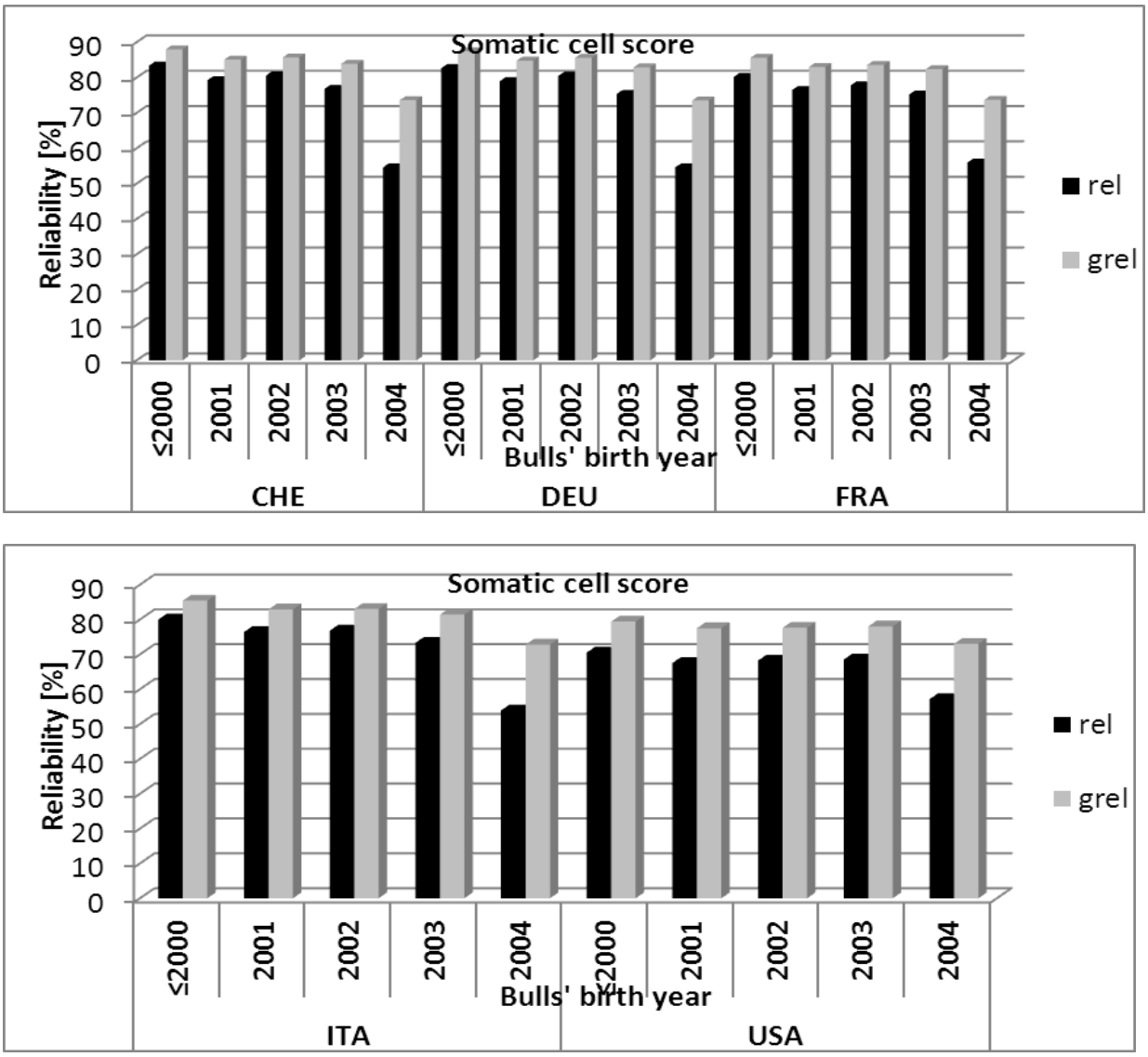

Figure 2. Conventional (rel) and genomic reliabilities (grel) for bulls with daughter information by year of birth on different country scales for somatic cell score.
Birth year 2004 also includes also records from birth year 2005

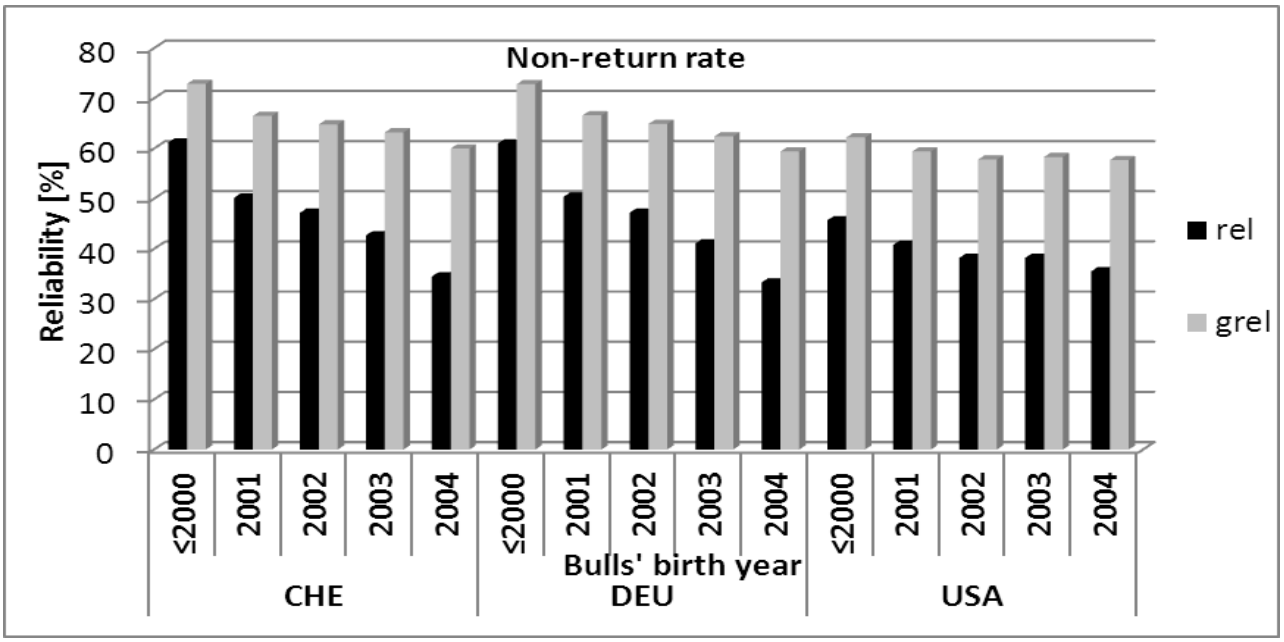

Figure 3. Conventional (rel) and genomic reliabilities (grel) for bulls with daughter information by year of birth on different country scales for non-return rate.
Birth year 2004 also includes also records from birth year 2005

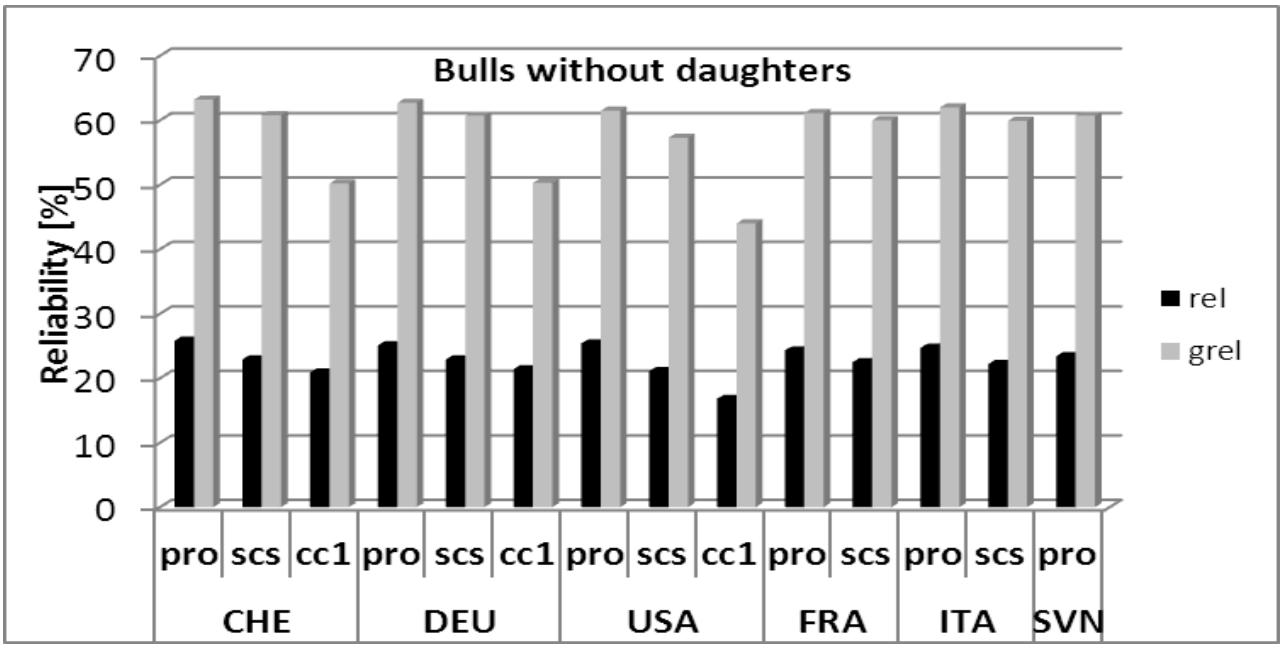

Figure 4. Conventional (rel) and genomic reliabilities (grel) for bulls without daughter information on different country scales for protein yield (pro), somatic cell score (scs) and non-return rate (cc1).

Table 5. Regression coefficients and R^2 of validation models for protein yield, somatic cell score and non-return rate on different country scales

Trait	Model	Country scale	Intercept	Slope	R2
Protein yield	M0	CHE	0.36±0.02	0.78±0.04	0.41
		DEU	0.42±0.03	0.75±0.04	0.42
		FRA	0.33±0.03	0.71±0.03	0.40
		ITA	0.37±0.03	0.71±0.04	0.40
		SVN	0.30±0.03	0.71±0.03	0.41
		USA	0.36±0.03	0.69±0.04	0.38
	M1	CHE	0.26±0.03	0.71±0.03	0.42
		DEU	0.29±0.03	0.69±0.03	0.43
		FRA	0.26±0.03	0.67±0.03	0.41
		ITA	0.29±0.03	0.67±0.03	0.40
		SVN	0.27±0.03	0.65±0.03	0.41
		USA	0.28±0.03	0.65±0.03	0.38
	M2	CHE	0.64±0.02	1.49±0.06	0.53
		DEU	0.20±0.03	1.50±0.06	0.52
		FRA	-0.02±0.03	1.56±0.06	0.53
		ITA	0.10±0.03	1.46±0.06	0.50
		SVN	1.13±0.03	1.54±0.06	0.49
		USA	0.02±0.03	1.52±0.06	0.49
Somatic cell score	M0	CHE	0.07±0.03	0.85±0.04	0.49
		DEU	0.06±0.03	0.85±0.04	0.48
		FRA	0.05±0.03	0.85±0.04	0.47
		ITA	0.10±0.03	0.82±0.04	0.44
		USA	0.00±0.03	0.81±0.04	0.45
	M1	CHE	0.09±0.03	0.80±0.03	0.49
		DEU	0.08±0.03	0.81±0.03	0.49
		FRA	0.08±0.03	0.81±0.04	0.48
		ITA	0.11±0.03	0.78±0.04	0.45
		USA	0.01±0.04	0.82±0.04	0.46
	M2	CHE	0.01±0.03	1.42±0.08	0.35
		DEU	0.00±0.03	1.45±0.08	0.35
		FRA	0.21±0.03	1.47±0.08	0.36
		ITA	0.34±0.03	1.46±0.08	0.37
		USA	-24.92±1.37	1.47±0.08	0.37
Non-return rate	M0	CHE	-0.04±0.03	0.74±0.03	0.50
		DEU	-0.04±0.03	0.76±0.03	0.51
		USA	-0.10±0.03	0.62±0.03	0.57
	M1	CHE	-0.03±0.03	0.71±0.03	0.53
		DEU	-0.02±0.03	0.73±0.03	0.53
		USA	-0.09±0.03	0.60±0.02	0.59
	M2	CHE	-0.13±0.04	1.14±0.08	0.28
		DEU	-0.31±0.04	1.16±0.09	0.26
		USA	-0.04±0.04	1.10±0.08	0.27