

## **Approximation of reliability of single-step genomic breeding values for dairy cattle in the Czech Republic\***

**Jiří Bauer<sup>1\*\*</sup>, Luboš Vostrý<sup>1,2</sup>, Josef Příbyl<sup>1</sup>,  
Alena Svitáková<sup>1,2</sup>, Ludmila Zavadilová<sup>1</sup>**

<sup>1</sup> Institute of Animal Science, Department of Genetics and Breeding of Farm Animals,  
Přátelství 815, 10401 Praha - Uhřetěves, Czech Republic

<sup>2</sup> Czech University of Life Sciences, Prague, Department of Genetics and Breeding,  
Kamýcká 129, 165 21, Praha 6 - Suchbátka, Czech Republic

*(Accepted September 13, 2014)*

The method for approximating reliability of genomic breeding values in the single-step genomic BLUP evaluation was tested on milk production data from dairy cattle in the Czech Republic. The reliability of young dairy bulls without progeny increased by an average of 0.02 for heritability of 0.25 after genomic information was included in the evaluation. The increase was minimal for proven bulls with many progeny because much information was already known about these individuals. The overall increase in the reliability for the entire population of dairy cattle was small, but noticeable; the small increment of the increase resulted from the unfavourable ratio of genotyped bulls and the size of the population (1:2500). A relatively small increase in reliability by genotyping was due to the small number of reference bulls with known genomic information.

**KEYWORDS:** dairy cattle / genomic selection / reliability / single-step GBLUP

Prediction of genomic breeding values is a promising procedure that can be used for a more accurate genetic evaluation of cattle. The greatest benefit of this method can be observed in young animals that do not have their own performance data. Two approaches are used to predict genomic breeding values: the multistep method described by VanRaden [2008] and the single-step method [Miształ *et al.* 2009, Christensen and Lund 2010].

---

\*This research was supported by the Ministry of Agriculture of the Czech Republic, Project No. QI111A167

\*\*Corresponding author: bauer.jiri@vuzv.cz

The single-step method has already been applied to cattle populations [e.g. Aguilar *et al.* 2010, Tsuruta *et al.* 2011], and Přibyl *et al.* [2012, 2013] used this method in the Czech Republic.

The reliability of predicted breeding values is an important issue and although it can be calculated by inverting the left hand side of the BLUP system of equations [VanRaden 2008]; it is not always feasible due to a large population size, which leads to enormous computational requirements. Misztal *et al.* [2011, 2013] proposed the single-step genomic breeding value reliability approximation procedure, which requires inversion of the genomic relationship matrix and section of the pedigree-based relationship matrix containing the genotyped animals.

As at present the single-step method for predicting genomic breeding values in the Czech Republic is being successfully developed [e.g. Přibyl *et al.* 2012, 2013] the method for calculating reliability of predicted genomic breeding values should be established as well. The aim of this study was to approximate reliability of non-genomic and single-step genomic breeding values for the dairy cattle population in the Czech Republic.

## Material and methods

Data regarding the first lactation of the Holstein cattle population in the Czech Republic was used in this study. In total, 729,341 lactations were recorded during the calving years 1991-2004 and 4-generation pedigree comprised 1,917,416 animals. The BovineSNP50 Beadchip V2 by Illumina was used in genotyping of 838 Holstein bulls. The data are similar to the data used in their study by Přibyl *et al.* [2012]. The trait evaluated in the dairy cattle was milk yield in the first lactation.

The computation procedure is based on the work by Misztal *et al.* [2013], who proposed a feasible method to approximate reliability of single-step genomic breeding values.

The reliability of genomic breeding values can be generally approximated based on contributions from the effective number of records from observations, pedigree and genomic information [Misztal *et al.* 2013], while the entire calculation was conducted as a sequence of these consecutive steps:

1. Approximation of reliability of non-genomic breeding values of the entire population using the iterative approach described in Misztal *et al.* [1993], considering the effective number of records originating from the pedigree and performance testing.
2. Updating reliability of genotyped animals by  $\mathbf{Q}^{-1}$ , the inversion of the sum of reliability from the first step in the form of the effective number of records, and the relationship matrix calculated by the formula [Misztal *et al.* 2013]:

$$\mathbf{Q}^{-1} = [\mathbf{D} + (\mathbf{I} + \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1}) \alpha]^{-1}$$

where  $\mathbf{D}$  is the contribution of records and pedigrees to reliability,  $\mathbf{I}$  is the identity

matrix;  $\mathbf{G}^{-1}$  is the inversion of the genomic relationship matrix;  $\mathbf{A}_{22}^{-1}$  is the inversion of the section of the relationship matrix that contains pedigree-based relationship information about genotyped animals,  $\alpha$  the ratio of error variance to additive genetic variance.

3. Derivation of reliabilities of genomic breeding values for genotyped animals by:

$$\text{reliability}_i = 1 - \alpha q_{ii}$$

where  $q_{ii}$  are diagonal elements of the  $\mathbf{Q}^{-1}$  matrix.

4. Addition of the contribution of genotyping to the reliability of non-genotyped animals by the procedure similar to the first step procedure. Reliabilities of genotyped bulls were kept constant to avoid the contribution from the relationships among these bulls being counted twice and to avoid changes in the already fully conveyed values.

The relaxation parameter was applied in the first and last steps to obtain less oscillation between the iterations and better convergence. The relaxation parameter  $\omega$  was set to 0.5 and the effective number of records for animal  $i$  at the beginning of iteration  $n+2$  was:  $d_i^{n+2} = \omega * (d_i^n + d_i^{n+1})$ . Entry values for the  $\mathbf{G}$  and  $\mathbf{A}_{22}$  matrices were acquired from the calculations carried out by Pribyl *et al.* [2012]. The  $\mathbf{G}$  matrix had already been scaled to the mean of the diagonal elements equal to 1 by the procedure described by Forni *et al.* [2011] and shifted to  $\mathbf{A}_{22}$  matrix values as proposed by Vitezica *et al.* [2011]. Heritability of the first lactation yield was 0.25.

Results of all the steps were recorded and evaluated based on the values for descriptive statistics, the distribution of reliabilities and their span. Differences in reliability approximations between non-genomic and genomic values were also noted. To ensure a better understanding of the situation, two subcategories of genotyped dairy bulls were monitored in the analysis of the results: young bulls without progeny and 'proven' bulls with performing daughters.

## Results and discussion

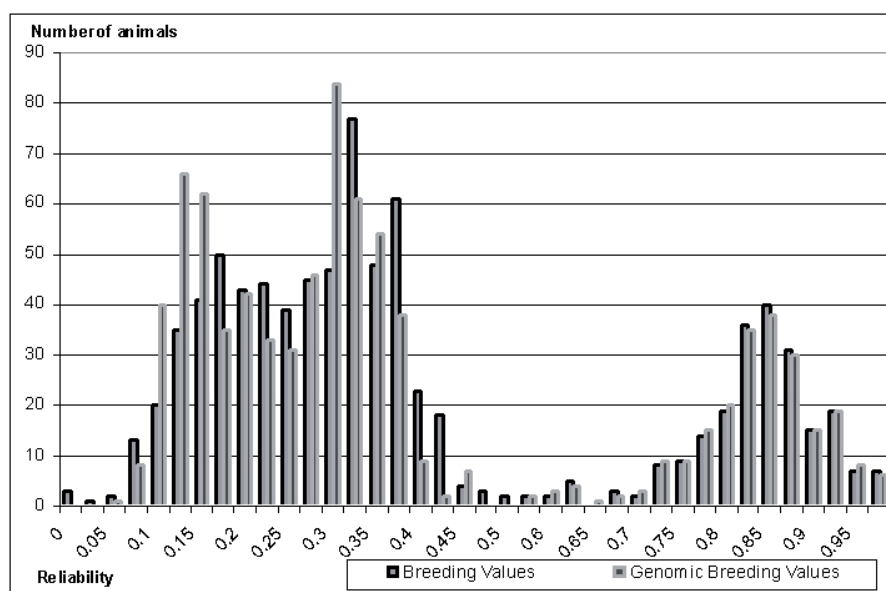
The subset of genotyped proven bulls had an average contribution of 85 daughters to the reliability. The standard deviation was 49.31. From these 217 proven bulls, 199 had more than 20 daughters and 167 had more than 50 daughters.

Descriptive statistics of approximated reliabilities for non-genomic and genomic breeding values of dairy cattle are shown in Table 1. All the approximated reliabilities fell within a defined interval of the reliability function, but three genotyped bulls exhibited reliabilities of non-genomic and genomic breeding values equal to zero. None of these bulls had daughters.

Generally, the difference between reliabilities of young and proven bulls was very noticeable. As it is shown in Figure 1, reliability approximations are separated into two groups, with young bulls most frequently exhibiting a reliability approximation around 0.25 and proven bulls characterised by a much higher value (near 0.87).

**Table 1.** Reliability approximations generated for breeding values and genomic enhanced breeding values in dairy cattle

| Data                   | Approximation of reliability of | No.       | Mean  | SD    | Median | Min.  | Max.  |
|------------------------|---------------------------------|-----------|-------|-------|--------|-------|-------|
| Whole population       | BV                              | 1,917,416 | 0.322 | 0.130 | 0.330  | 0     | 0.998 |
| Genotyped bulls        | BV                              | 838       | 0.397 | 0.275 | 0.305  | 0.060 | 0.987 |
| Genotyped young bulls  | BV                              | 621       | 0.244 | 0.095 | 0.257  | 0.060 | 0.578 |
| Genotyped proven bulls | BV                              | 217       | 0.835 | 0.075 | 0.843  | 0.594 | 0.987 |
| Genotyped bulls        | GEBV                            | 838       | 0.411 | 0.268 | 0.325  | 0     | 0.986 |
| Genotyped young bulls  | GEBV                            | 621       | 0.263 | 0.099 | 0.270  | 0     | 0.584 |
| Genotyped proven bulls | GEBV                            | 217       | 0.837 | 0.074 | 0.844  | 0.602 | 0.987 |
| Whole population       | GEBV                            | 1,917,416 | 0.322 | 0.131 | 0.330  | 0     | 0.998 |

**Fig. 1.** Approximation of reliability of genomic and non-genomic breeding values for genotyped bulls.

Differences in approximations of non-genomic and genomic breeding values for all genotyped bulls are shown in Figure 2. Differences between approximations of reliabilities for non-genomic and genomic breeding values were more noticeable in young bulls than in proven bulls: 559 of 621 young bulls exhibited diverse reliabilities. These changes were both positive and negative, although the number of young bulls with increased reliability was drastically higher. The decrease in reliability of breeding values observed in some young bulls as a result of inclusion of genomic information was probably a consequence of the more realistic relationship coefficients obtained

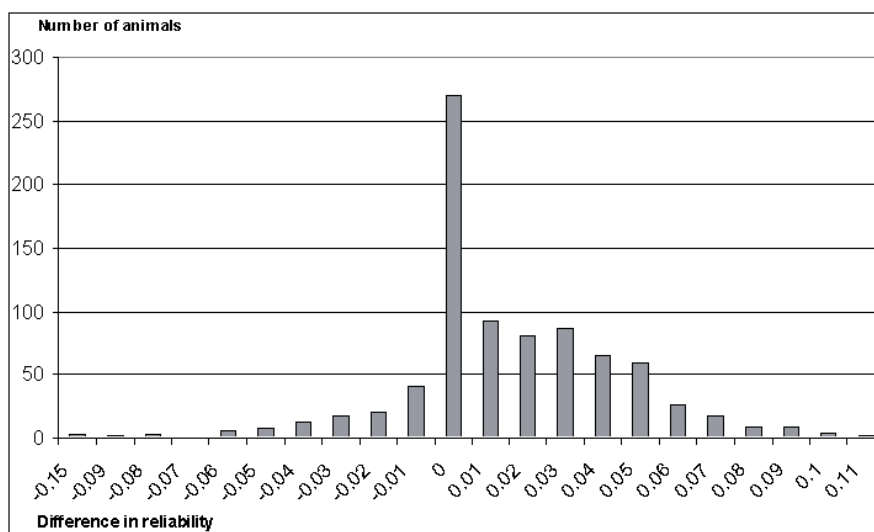


Fig. 2. Differences in approximation of reliability between genomic and non-genomic breeding values for genotyped bulls.

when expected relationship is replaced by more accurate genomic information, including some information about Mendelian sampling.

The increase in reliability of young genotyped bulls corresponds well with the correlation published by Přibyl *et al.* [2012] on the same data, with the correlation between predicted breeding values before and after testing increasing from 0.495 to 0.544 after inclusion of genomic data in single-step genomic BLUP.

Proven bulls exhibited a very small increase in reliability. Because a large amount of information was known about performance of their daughters, the genotyping information for these bulls did not contribute additional valuable information capable of improving reliability.

The small increase in reliability of breeding values generally true for all genotyped bulls is probably a result of the small number of reference bulls in the study, as the increase in reliability of breeding values due to genotyping is approximately linearly dependent upon the number of reference bulls [VanRaden *et al.* 2009].

The differences of reliability approximations for non-genomic and genomic breeding values for the entire population were very small, because of the small number of genotyped bulls relative to the entire population.

The approximation of reliability of breeding values increase can be credited to inclusion of genomic information according to the method of Misztal *et al.* [2013] for ssGBLUP, especially in the case of young bulls without a sufficient number of progeny. Reliability of non-genotyped relatives is also improved. An increase in reliability should be more evident with the rising number of genotyped bulls in the reference population.

## REFERENCES

1. AGUILAR I., MISZTAL I., JOHNSON D.L., LEGARRA A., TSURUTA S., LAWLOR T.J., 2010 – Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *Journal of Dairy Science* 93, 743-752.
2. CHRISTENSEN O.F., LUND M.S., 2010 – Genomic prediction when some animals are not genotyped. *Genetics Selection Evolution* 42 (2), doi: 10.1186/1297-9686-42-2.
3. FORNI S., AGUILAR I., MISZTAL I., 2011 – Different genomic relationship matrices for single-step analysis using phenotypic, pedigree and genomic information. *Genetics Selection Evolution* 43, (1), doi:10.1186/1297-9686-43-1.
4. MISZTAL I., LEGARRA A., SHORT T.H., 1993 – Implementation of Single- and Multiple-Trait Animal Models for Genetic Evaluation of Holstein Type Traits. *Journal of Dairy Science* 76 (5), 1421-1432.
5. MISZTAL I., LEGARRA A., AGUILAR I., 2009 – Computing procedures for genetic evaluation including phenotypic, full pedigree, and genomic information. *Journal of Dairy Science* 92, 4648-4655.
6. MISZTAL I., TSURUTA S., AGUILAR I., LEGARRA A., LAWLOR T.J., 2011 – Approximation of genomic accuracies in single-step genomic evaluation, Interbull Bulletin 44. Stavanger. 103-109.
7. MISZTAL I., TSURUTA S., AGUILAR I., LEGARRA A., VAN RADEN P.M., LAWLOR T.J., 2013 – Methods to approximate reliabilities in single-step genomic evaluation. *Journal of Dairy Science* 96 (1), 647-654.
8. PŘIBYL J., HAMAN J., KOTT T., PŘIBYLOVÁ J., ŠIMEČKOVÁ M., VOSTRÝ L., ZAVADILOVÁ L., ČERMÁK V., RŮŽIČKA Z., ŠPLÍCHAL J., VERNER M., MOTYČKA J., VONDRÁŠEK L., 2012 – Single-step prediction of genomic breeding value in a small dairy cattle population with strong import of foreign genes. *Czech Journal of Animal Science* 57 (4), 151-159.
9. Přibyl J., Madsen P., Bauer J., Přibyl J., Šimečková M., Vostrý L., Zavadilová L., 2013 – Contribution of domestic production records, Interbull estimated breeding values, and single nucleotide polymorphism genetic markers to the single-step genomic evaluation of milk production. *Journal of Dairy Science* 96 (2), 1865-1873.
10. TSURUTA S., I. MISZTAL I. AGUILAR LAWLOR T.J., 2011 – Multiple-trait genomic evaluation of linear type traits using genomic and phenotypic data in US Holsteins. *Journal of Dairy Science* 94, 4198-4204.
11. VANRADEN P.M., 2008 – Efficient methods to compute genomic predictions. *Journal of Dairy Science* 91, 4414-4423.
12. VANRADEN P. M., VAN TASSELL C.P., WIGGANS G.R., SONSTEGARD T.S., SCHNABEL R.D., TAYLOR J.F., SCHENKEL F.S., 2009 – Invited review: Reliability of genomic predictions for North American Holstein bulls. *Journal of Dairy Science* 92, 16-24.
13. VITEZICA Z.G., AGUILAR I., MISZTAL I., LEGARRA A., 2011 – Bias in genomic predictions for populations under selection. *Genetic Research* (Camb.), 93, 357-366.