Approximating Genomic Reliabilities for National Genomic Evaluation

Z. Liu[1], P. M. VanRaden[2], M.H. Lidauer[3], M. P. Calus[4], H. Benhajali[5], H. Jorjani[5] and V. Ducrocq[6]

*[1]IT solutions for Animal Production (**vit**), Heinrich-Schröder-Weg 1, D-27283 Verden, Germany; [2]AnimalGenopmics and Improvement Lab, USDA, Beltsville, MD, USA; [3]Natural Resources Institute Finland (Luke), Finland; [4] Wageningen University & Research, The Netherlands; [5]Interbull Centre, Swedish University of Life Sciences, Uppsala, Sweden; [6]INRA, France*

Abstract

With the introduction of standard methods for approximating effective daughter/data contribution by Interbull in 2001, conventional EDC or reliabilities contributed by daughter phenotypes are directly comparable across countries and used in routine conventional evaluations. In order to make published genomic reliabilities comparable across countries and consistent with conventional reliabilities, a working group for genomic reliability calculation developed a new method that is feasible for any number of genotyped animals and also adjusts theoretical model genomic reliabilities based on genomic validation results. The first step of the proposed reliability method calculates reliabilities contributed by SNP genotypes via an efficient software *snp_blup_rel*. This new genomic reliability method accounts for the residual polygenic effect in genomic evaluation and is applicable to both single-step and multi-step genomic models. The adjustment procedure makes the changes in genomic reliabilities reflecting the changes in GEBV and ensures candidates genomic reliabilities from an early evaluation being consistent with later genomic reliabilities when the animals have received phenotype data. The proposed reliability method was applied to a large German Holstein population. Adjustment factors for the theoretical model genomic reliabilities were derived based on a genomic validation study via Interbull *GEBV Test*. Results from the test implementation for German Holsteins demonstrated high efficiency and feasibility of the proposed genomic reliability method. Several aspects have been discussed for future optimisations. All involved countries were requested to test the software *snp_blup_rel* and proposed genomic reliability method. Depending on the country feedbacks, the software and the proposed genomic reliability method will be fine-tuned towards an official implementation by all the involved countries.

Key words: genomic reliability, genomic evaluation, genomic validation, single-step genomic model

Introduction

For approximating effective daughter contribution (EDC) of bulls, Interbull introduced standardized procedures in 2001 for conventional evaluation (Interbull, 2001), although the calculation of total reliability of conventional evaluation that includes also parental contribution of bulls has not been fully harmonized. In contrast to the conventional reliabilities, published values of genomic reliability are less comparable across countries because of a lack of harmonized calculation procedures and also due to differences in national methods for approximating genomic reliability values. When countries still use a two-step genomic model for genomic evaluation, genomic reliabilities must be consistent with conventional reliabilities between the two separate evaluations. Additionally, we must make sure that genomic reliability values of animals in different life stages must be consistent across evaluations as well: from being a selection candidate, to getting own phenotype data and to becoming a reference animal.

To address the problems and challenges related to the genomic reliabilities, Interbull decided in 2014 to set up a working group aiming to develop standardized statistical procedures for approximating genomic reliabilities. Two reports of the Genomic Reliability Working Group were presented (Harris 2015a; 2015b), focusing on theoretical

investigation about genomic validation R^2 value and genomic reliability via simulation. Main conclusions from the two studies were that the genomic validation R^2 value and genomic reliability were two different measures of accuracy of genomic prediction in populations under selection, and that the difference between them reduced as the validation R^2 value increased. Genomic reliabilities are a measure of the standard error of genomic breeding values, whereas genomic validation R^2 values are a measure of the reliability of the selection in which those breeding values are used (Bijma, 2012). From the perspective of Interbull, genomic reliabilities are the measure of interest.

An accurate method for calculating genomic reliabilities must have the following desired features: 1) accounting for a residual polygenic effect; 2) feasible for any number of genotyped animals; 3) applicable to single-step genomic models; 4) efficient for frequent genomic evaluation; and 5) approximated genomic reliabilities being consistent with the genomic validation prediction error variance. The aim of this study was to develop standard procedures for approximating genomic reliabilities for national genomic evaluation.

Materials and Methods

Countries have implemented their own methods for approximating genomic reliabilities, e.g. Harris and Johnson (2010), and VanRaden *et al.* (2011) for multi-step genomic models; Misztal *et al.* (2013) and Taskinen *et al.* (2013) for single-step genomic models. These genomic reliability methods differ both in reliability definition: theoretical model reliability vs realized reliability, as well as in accuracy and efficiency. On the basis of these currently available methods we have developed a new genomic reliability method.

Reliabilities of direct genomic values

Calculating genomic reliability involves inverse of the genomic relationship matrix **G** which may contain a large number of genotyped animals. Levels of genomic reliabilities for candidates depend on, among other factors, candidates' genomic relationships with reference animals and reliability values of the reference animals. If the sire of a candidate is included in a reference population, the candidate's reliability of direct genomic value (DGV) is expected to be higher than candidates without sire in the reference population. Liu *et al.* (2010) derived an approximation method for DGV reliabilities of candidates based on average and maximum value of their genomic relationship coefficients with reference animals. Regression formulae were developed for the approximation without inverting the genomic relationship matrix with all genotyped animals included. A similar approximation procedure has been used for routine genomic evaluation in the US (Wiggans *et al.*, 2010).

Using a recursion algorithm, Misztal *et al.* (2014) developed an approximated inverse of the genomic relationship matrix for proven and young animals (APY). With proper selection of the so-called core animals, the inverse matrix can be fairly accurately set up for virtually any number of genotyped animals.

The above mentioned methods for approximating genomic reliabilities are based on the genomic relationship matrix which grows in size rapidly with time. Inverting the potentially very large **G** matrix represents a major computational bottleneck for calculating genomic reliabilities. Our new reliability method applies a SNP based genomic model that is equivalent to a GBLUP model, but it does not rely on the genomic relationship matrix. We assume, at this stage of genomic reliability calculation, a SNP BLUP model without a residual polygenic effect (RPG) for calculating DGV reliabilities:

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{Zg} + \mathbf{e} \qquad [1]$$

where **y** is a vector of deregressed EBV of reference bulls or cows; **1** is a vector of 1s; μ is the general mean; **g** is a vector of effects of *m* fitted SNP markers; **Z** is a design matrix for **g**; and **e** is a vector of residuals with:

$$[\mathrm{var}(\mathbf{e})]^{-1} = \mathbf{R}^{-1} = diag\{n_e \sigma_e^{-2}\} \qquad [2]$$

where n_e represents effective daughter contribution (EDC) of bulls or effective record contributions (ERC) of cows (Taskinen *et al.*, 2014) in case also cows are included in the reference population, expressed on animal basis not on the usual progeny basis, of the reference animals, and σ_e^2 is residual variance. Note that Interbull's EDC of bulls (Interbull, 2001) are expressed on a progeny basis derived from a sire model, whereas the EDC or ERC in eq. [2] are expressed on a animal basis from an animal model. Although the two sets of EDC/ERC are equivalent, i.e. leading to equal reliability values, Interbull's EDC of bulls need to be converted to the animal model basis for the reliability calculation. Mixed model equations for estimating the effects of model [1] are:

$$\begin{bmatrix} \mathbf{1'R^{-1}1} & \mathbf{1'R^{-1}Z} \\ \mathbf{Z'R^{-1}1} & \mathbf{Z'R^{-1}Z} + \sigma_u^{-2}\mathbf{B}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{\mathbf{g}} \end{bmatrix} = \begin{bmatrix} \mathbf{1'R^{-1}y} \\ \mathbf{Z'R^{-1}y} \end{bmatrix} \quad [3]$$

where σ_u^2 represents additive genetic variance and matrix **B** is defined similar to VanRaden (2008):

$$\mathbf{B} = \left(\sum_{j=1}^{m} 2p_j(1-p_j) \right)^{-1} \mathbf{I} \quad [4]$$

with p_j being allele frequency of SNP marker *j*.

The SNP BLUP model [3] comprises *m*+1 equations. Currently, the number of the fitted SNP markers, *m*, is usually around 50,000 in most countries. With the advance of parallel computing, fairly large matrices, like the coefficient matrix in eq. [3], can be efficiently inverted using the multi-threaded BLAS subroutines on multi-core computers. The inverse of the coefficient matrix of eq. [3] can be written as:

$$\mathbf{C}^{-1} = \begin{bmatrix} \mathbf{C}^{\mu\mu} & \mathbf{C}^{\mu g} \\ \mathbf{C}^{g\mu} & \mathbf{C}^{gg} \end{bmatrix}. \quad [5]$$

Reliability for DGV of a genotyped animal *i* is:

$$\Re_i^{SNP} = 1 - \mathbf{z}_i \mathbf{C}^{gg} \mathbf{z}_i' \sigma_u^{-2} \quad [6]$$

where $\mathbf{z}_i$ is a row in design matrix **Z** corresponding to the *i*-th genotyped animal. The genotyped animal *i* may be a reference animal or a candidate. As mentioned above, so far we assumed that the SNP genotypes explain all genetic variation. The DGV or SNP reliabilities need to be discounted when an RPG is considered.

In contrast to the previous methods, like Liu *et al.* (2010), Wiggans *et al.* (2010) and the APY algorithm, the computational costs for approximating reliabilities of DGV for any number of genotyped animals depend mainly on the number of fitted SNP markers, much less on the number of reference animals or on the number of genotyped animals.

Strandén and Mäntysaari (2015) developed a software package, called *snp_blup_rel*, based on the SNP BLUP model [1]. The software package (Mäntysaari and Strandén, 2016) inverts the matrix in eq. [5] with the multi-threaded BLAS subroutines for calculating DGV reliabilities via formula [6]. The authors have kindly provided the software to national genetic evaluation centers conducting official genomic evaluation.

Information sources for genomic reliabilities

Three sources of information contribute to the total conventional reliability of an animal: own phenotype data of the animal, parental and progeny contributions via pedigree. Contribution of genotypes can be viewed as the extra source contributed by the genomic relationship matrix after subtracting the numerator relationship matrix of the genotyped animals (Aguilar *et al.*, 2010):

$$\begin{aligned} \mathbf{H}^{-1} &= \begin{bmatrix} \mathbf{A}^{11} & \mathbf{A}^{12} \\ \mathbf{A}^{21} & \mathbf{A}^{22} + \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix} \\ &= \mathbf{A}^{-1} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix} \end{aligned} \quad [7]$$

where subscript 1 represents non-genotyped animals and subscript 2 genotyped animals. It can be seen from eq. [7] that the main difference in reliabilities between genomic and conventional evaluations is caused by the extra

genomic contribution: $\mathbf{G}^{-1} - \mathbf{A}_{22}^{-1}$. Therefore, the emphasis of genomic reliability calculation is put on the calculation of the genomic contribution under the assumption of available conventional reliabilities.

Calculation of the genomic contribution

Reliabilities of DGV/SNP of any genotyped animals can be calculated using formula [6] by inverting the coefficient matrix of eq. [3] for reference animals only, with the multi-threaded BLAS subroutines. This is equivalent to inverting the following matrix under a GBLUP model for all genotyped animals:

$$\left[\begin{bmatrix} \mathbf{R}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \sigma_u^{-2}\mathbf{G}^{-1}\right]. \qquad [8]$$

In parallel to eq. [8], we can approximate conventional reliabilities for all the genotyped animals based on:

$$\left[\begin{bmatrix} \mathbf{R}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \sigma_u^{-2}\mathbf{A}_{22}^{-1}\right]. \qquad [9]$$

The pedigree relationship matrix $\mathbf{A}_{22}$ does not need to be explicitly inverted. Instead, by adding non-genotyped relatives to the pedigree of the genotyped animals we can apply regular conventional reliability method to approximate conventional reliabilities for all the genotyped animals. For a genotyped animal *i*, genomic contribution is defined as:

$$\lambda \left(\frac{\mathfrak{R}_i^{DGV}}{1-\mathfrak{R}_i^{DGV}} - \frac{\mathfrak{R}_i^{A22}}{1-\mathfrak{R}_i^{A22}}\right) \qquad [10]$$

where $\lambda = (1-h^2)/h^2$, and $\mathfrak{R}_i^{A22}$ is the conventional reliability for the *i*-th genotyped animal based on matrix [9]. The genomic reliability $\mathfrak{R}_i^{DGV}$ represents the contribution by SNP genotypes which is calculated via eq. [6]. $\mathfrak{R}_i^{DGV}$ may also be affected by other factors, such as the proportion of residual polygenic variance in the total additive genetic variance, accuracy of genotypes imputed from lower-density chips.

Adjustment for the theoretical model genomic reliabilities

The theoretical genomic reliabilities calculated above depend on model assumptions of the genomic and conventional models, they tend to be higher than those realized reliabilities, calculated from validation R^2 values derived from genomic validation with truncated data (Harris *et al.*, 2015a and 2015b). Therefore, those theoretical model genomic reliabilities must be adjusted using genomic validation following Interbull's *GEBV Test* (Mäntysaari *et al.*, 2010).

VanRaden and O'Connell (2018) developed a procedure for adjusting the theoretical model genomic reliabilities using validated GEBV according to the Interbull's *GEBV Test*. For validation bulls, two sets of GEBV are available: u_L for a later, complete genomic evaluation with daughters' phenotypes included, u_E for an early, truncated genomic evaluation with no daughters available yet. We can derive an expected change in genomic reliabilities based on the two sets of GEBV of the validation bulls:

$$E(\Delta\mathfrak{R}) = \mathrm{var}(\hat{u}_L - \hat{u}_E)/\sigma_u^2. \qquad [11]$$

This population parameter $E(\Delta\mathfrak{R})$ should be highly correlated with the validation R^2 value from the *GEBV Test*. Let us define average genomic reliability of the validation bulls from the later, complete evaluation as $\overline{\mathfrak{R}}_L$, then genomic reliability of the early evaluation for the validation population is expected, on average, to be:

$$E(\mathfrak{R}_E) = \overline{\mathfrak{R}}_L - E(\Delta\mathfrak{R}). \qquad [12]$$

Denote $\mathfrak{R}_{E\_i}$ as a model genomic reliability of the early, truncated evaluation for a validation bull *i*, we convert the early genomic reliability to EDC for all the validation bulls to obtain an average of the EDC:

$$\overline{\varphi}_E = \tfrac{1}{n}\lambda \sum_{i=1}^{n} \mathfrak{R}_{E\_i}/(1-\mathfrak{R}_{E\_i}) \qquad [13]$$

where *n* is the number of validation bulls. The expected genomic reliability from the early evaluation $E(\Re_E)$ is converted to EDC:

$$E(\varphi_E) = \lambda\, E(\Re_E)/(1-E(\Re_E)) \,. \qquad [14]$$

Using the two EDC values we can derive an adjustment factor for converting the theoretical model to realized genomic EDC:

$$f = E(\varphi_E)/\overline{\varphi}_E \,. \qquad [15]$$

The genomic EDC adjustment factor $f<1$ or $f>1$ indicates an overestimation or underestimation of the early genomic EDC/reliabilities, respectively. This multiplicative adjustment procedure affects not only average but also variance of the final, realized genomic reliability values.

In fact, this adjustment procedure is applicable to any two evaluations, as long as the GEBV are validated e.g. via Interbull's *GEBV Test* (Mäntysaari *et al.*, 2010).

A new genomic reliability method

Based on the individual components of genomic reliabilities given before, we have developed an unified, standardized method for approximating genomic reliabilities for national genomic evaluations. The new standard genomic reliability method consists of six steps, with Step 5 being optional.

Step 1: Reliabilities of SNP genotypes
For a genotyped animal *i*, reliability of its SNP genotypes, $\Re_i^{SNP}$, is calculated with eq. [6]. The SNP reliability is calculated under the assumption that the SNP markers explain all genetic variation. The software *snp_blup_rel* (Mäntysaari & Strandén, 2016) is recommended for calculating the SNP reliabilities.

Step 2: Reliabilities of DGV
When a residual polygenic effect is assumed in national genomic evaluation and genotypes of low-density chips are imputed to a standard chip, reliability of DGV for animal *i* can be approximated as:

$$\Re_i^{DGV} = (1-k) r_{IMP}^2 \Re_i^{SNP} \,. \qquad [16]$$

where k represents the proportion of residual polygenic variance in total additive genetic variance (σ_u^2), and r_{IMP} is the accuracy of genotype imputation. The formula [16] applies to genotyped animals not belonging to the reference population. For a reference animal, its DGV reliability is computed as:

$$\Re_i^{DGV} = r_{IMP}^2 \Re_i^{SNP} \,. \qquad [17]$$

The DGV reliability of the *i*-th genotyped animal can be converted to EDC:

$$\varphi_i^{DGV} = \lambda\, \Re_i^{DGV}/(1-\Re_i^{DGV}) \,. \qquad [18]$$

Step 3: Adjusting the theoretical reliabilities
The multiplicative adjustment procedure (eq. [11] to [15]) can be used to adjust the theoretical EDC of DGV, φ_i^{DGV}, to a realized one with:

$$\varphi_{real\_i}^{DGV} = f * \varphi_i^{DGV} \,. \qquad [19]$$

This multiplicative EDC adjustment procedure changes average as well as variance of the DGV EDC. For the very first time of implementing this adjustment procedure, the multiplicative factor f is usually unknown and can be set to 1 for the initialisation step. After the multiplicative factor f is determined, using the adjustment procedure, *Step 3* can be repeated to obtain the realised DGV EDC ($\varphi_{real\_i}^{DGV}$).

Step 4: Calculating the genomic EDC gain
To calculate the genomic EDC gain defined in eq. [10], firstly we need to approximate the conventional reliabilities, $\Re_i^{A22}$, for every genotyped animal using phenotype data of the reference animals and pedigree for all genotyped animals. Conventional reliability approximation is usually composed of three steps (Liu *et al.*, 2004): calculation of animal's own data contribution, accumulation of progeny contribution by processing pedigree from youngest to oldest animals, and calculation of parental contribution by processing pedigree from the oldest to

youngest animals. Conventional EDC of the genotyped animal *i* is converted from the reliability with:

$$\varphi_i^{A22} = \lambda \frac{\Re_i^{A22}}{1-\Re_i^{A22}}, \quad [20]$$

which resembles the ERC of animal *i* on an animal model basis. A gain in EDC contributed by genotype of *i*-th animal compared to its conventional EDC is then computed as shown in [10]:

$$\varphi_i^{gain} = \varphi_{real\_i}^{DGV} - \varphi_i^{A22}. \quad [21]$$

If $\varphi_i^{gain} < 0$ for any reason, set $\varphi_i^{gain} = 0$.

Step 5: Propagation of genomic information to non-genotyped relatives

This step is optional, because in some cases a publication of reliabilities of non-genotyped animals enhanced with relatives' genomic information may not be required. The propagation of genomic information to non-genotyped relatives may involve tens of millions of animals. In order to avoid double counting of genomic information, the genomic EDC gain only of the reference animals, φ_i^{gain}, is treated as data for approximating reliability of the non-genotyped relatives. Similar as for the conventional reliability calculation (Liu *et al.*, 2004), the first step of calculation of data contribution involves here only the reference animals. Then the progeny contribution and parental contribution are approximated by processing the pedigree file. At end of the reliability calculation, each non-genotyped animal receives a reliability value, $\Re_i^{propg}$, which is converted to EDC:

$$\varphi_i^{propg} = \lambda \frac{\Re_i^{propg}}{1-\Re_i^{propg}}. \quad [22]$$

Because the propagation process does not account for the linkage disequilibrium break-down between parents and progeny, a maximum value is imposed on to the propagated genomic EDC of the non-genotyped animals:

$$\varphi_i^{propg} \le \max(\varphi_i^{gain}) \quad [23]$$

where $\max(\varphi_i^{gain})$ represents maximum genomic EDC gain value among all candidates.

Step 6: Final reliabilities enhanced with genomic information

Again we assume that all countries use a proper method to approximate conventional reliability using pedigree and phenotype data. For every animal included in the conventional evaluation, its conventional reliability, $\Re_i^{conv}$, is converted to EDC with:

$$\varphi_i^{conv} = \lambda \frac{\Re_i^{conv}}{1-\Re_i^{conv}}. \quad [24]$$

If a single-step genomic model (Aguilar *et al.*, 2010) is used in national evaluation, the conventional reliabilities $\Re_i^{conv}$ can be obtained by ignoring the genomic information.

For a genotyped animal, final EDC contributed by conventional and genomic information is:

$$\varphi_i^{final} = \varphi_i^{conv} + \varphi_i^{gain}, \quad [25]$$

and for a non-genotyped animal:

$$\varphi_i^{final} = \varphi_i^{conv} + \varphi_i^{propg}. \quad [26]$$

Final reliability enhanced with genomic information is then:

$$\Re_i^{final} = \frac{\varphi_i^{final}}{\varphi_i^{final} + \lambda}. \quad [27]$$

Results & Discussion

An application to German Holsteins

To test its accuracy and efficiency, the proposed genomic reliability method was applied to genotype and phenotype data of German Holsteins from May 2017 evaluation. The number of fitted SNP markers was 45,613 (Alkhoder *et al.,* 2014). The EuroGenomics bull reference population consisted of 35,533 Holstein bulls. A total of 314,608 genotyped Holstein animals were analyzed. All MACE traits were evaluated. In addition, six national traits, with 11,792 Holstein bulls included in reference population, were selected as well for the test implementation. For testing the propagation to non-genotyped animals, c.a. 20 millions of cows with test-day records and 7 millions of ancestors from routine test-day model evaluation were evaluated for three milk production traits.

The reliability calculation was run on a Linux server with Intel Xeon CPU E5-2690 v2 @ 3.00GHz, with 2 CPU families having 10 cores each. Among all the steps of the proposed genomic reliability method, Step 1 took the most computing time and RAM / disk usage, even with the fast software *snp_blup_rel* utilizing the highly efficient BLAS subroutines. The Step 5 of propagation to about 27 millions of non-genotyped animals required much less computing resources or time, by using the conventional reliability method (Liu *et al.,* 2004). The Step 1 for calculating $\mathfrak{R}_i^{SNP}$ was further divided into two parts in order to process all genotyped animals efficiently. In the first part of Step 1, the coefficient matrix of eq. [3] was set up and inverted using only the reference bulls, followed by calculating $\mathfrak{R}_i^{SNP}$ for all the 314,608 genotyped animals in the second part. The software *snp_blup_rel* required for the first part a total clock time of c.a. 60 minutes running on 10 cores and the peak RAM usage was 38 GB. For the second part of processing all the genotyped animals, the total clock time was about 82 minutes by using 10 cores. The peak RAM usage was 121 GB due to the selection of memory intensive option. The above statistics of computing requirements were for one of the MACE traits. The six national traits needed fewer resources for the first part due to their smaller reference populations.

A genomic validation study

The complete bull reference population with 35,533 Holstein bulls was truncated by birth years of the bulls for a genomic validation study. The truncated reference population contained 31,428 reference bulls born before 2010. A total of 894 German domestic bulls born in 2010 to 2012 were treated as validation animals. Interbull *GEBV Test* was conducted for all the traits. For the validation bulls, GEBV from the truncated as well as the complete genomic evaluations were validated via the *GEBV Test* for all the traits. Genomic reliabilities were calculated for the two evaluations.

Preliminary results of approximated genomic reliabilities obtained from the validation study and the full genomic evaluation seemed to be logical, though the proposed genomic reliability is subject to further fine-tunings. Average value of expected genomic reliabilities from the early, truncated evaluation, $E(\mathfrak{R}_E)$, for the validation bulls as candidates was clearly higher for the MACE traits with a larger reference population than for the six national traits with much fewer reference bulls. Expected genomic reliability changes for the national validation bulls from being candidates to daughter-proven, $E(\Delta\mathfrak{R})$, were lower, on average, for the MACE traits than for the six national traits with a much smaller reference population.

Allele frequencies of SNP markers

Allele frequencies have an impact on the reliabilities of SNP/DGV of genotyped animals (eq. [6]). When a frequency $p_j = 0.5$ was used for all the SNP markers, some reference bulls received too low $\mathfrak{R}_i^{SNP}$ that was less than their pure daughter contribution. Some of those reference bulls happened to have extreme values of diagonals of the genomic relationship matrix **G** (unpublished data). Therefore, we cannot recommend at present using the equal

allele frequencies for the genomic reliability calculation. Ideally, allele frequencies of base population should be used in the reliability calculation. However, estimation of the base population of allele frequencies (Gengler *et al.*, 2007) is computationally very demanding, because for each SNP marker a set of mixed model equations must be solved which size increases with the number of genotyped animals. A simple estimate from the current genotyped population with all genotyped animals or from reference population only may suffice for the reliability approximation. Allele frequencies estimated using more genotyped animals tend to be more stable over evaluations and will cause less change in the genomic reliabilities. It is recommended that countries use the same allele frequencies for the genomic reliability approximation as for routine genomic evaluation or SNP effect estimation.

Conventional reliability calculation

The proposed reliability method put an emphasis on the approximation of genomic reliability under the assumption that countries use accurate methods to calculate conventional reliabilities. This assumption was made for the conventional reliability $\mathfrak{R}_i^{conv}$ for all animals with phenotypic data as well as $\mathfrak{R}_i^{A22}$ for genotyped animals.

Frequency of genomic reliability calculation

The most time-consuming step of the proposed genomic reliability calculation method, i.e. the calculation of reliabilities of SNP/DGV, requires inverting a matrix with a dimension equal to the number of fitted SNP markers. Although the highly efficient multi-threaded BLAS subroutines were used for the matrix inversion on a multi-core computer, this step took the most computing time among all the six steps based on the application to German Holstein data. However, the current version of the applied software originally was written for research purposes and yet is not optimized for large applications.

Usually, countries conduct routine genomic evaluation much more frequently, e.g. once a week, daily or just-in-time (Alkhoder *et al.*, 2014), than they update reference population for SNP effect estimation, which depends on frequencies of national conventional or MACE evaluations. The first part of the most time-consuming Step 1 of the proposed method needs to be done only when the reference population has been updated. The second part of Step 1 is required only for newly genotyped animals or those genotyped animals present in previous evaluation with significant change in (imputed) genotypes, and computations of DGV for the individual animals are fast. Through these measures we can significantly reduce the computational costs.

We showed how to adjust the theoretical model genomic reliabilities to realized ones by using validated GEBV from Interbull's *GEBV Test*. Ideally, we should update the genomic reliability adjustment factors, whenever a *GEBV Test* is required.

Aspects for further development

Genomic selection has fundamentally changed the structure of breeding population in dairy cattle. The generation interval from sire to son has reduced from 4 or 5 years to 2 years. When new born candidates receive their first genomic evaluation, their sire or even grandsires may not have daughters in milk. This leads to a wider distance between the selection candidates and the bull reference population. Those candidates are called the second- or third-generation candidates. Because those candidates farther away from the bull reference population dominate the ranking of genomic evaluation, genomic reliabilities must account properly for the wider distance. Although SNP genotypes of those second- or third-generation candidates show lower genomic relationship to the reference population, the genomic reliability adjustment via the *GEBV Test* considers only the first-generation candidates, which needs to be extended to the second- or third-generation candidates (Liu *et al.*, 2016). In theory, when the genomic reliabilities of first-generation candidates have been adjusted

using the proposed procedure, we may use those validated GEBV and adjusted genomic reliabilities of the first-generation candidates as if they were from a later, complete genomic evaluation with daughter phenotypes. GEBV and genomic reliabilities of the second-generation candidates are treated as from an early, truncated genomic evaluation. By doing so, we may be able to adjust genomic reliabilities for the second- or even third-generation candidates.

By now, most countries still apply a multi-step genomic model for routine genomic evaluation. In order to pass Interbull's *GEBV Test*, DGV or GEBV of candidates are scaled by some countries in a way that the *GEBV Test* criterion, b_1=1, can be met. The scaling factor may be added to the DGV reliability formula [16] to reflect the changed variance of DGV or GEBV.

Genomic evaluation is mostly done on a single-trait model basis with a trait definition usually taken from MACE evaluation in dairy breeds. With the introduction of the single-step genomic model, genomic evaluation will gradually move back to a multi-trait model which is usually used for the current conventional evaluation. Our genomic reliability method must be extended to the multi-trait model, similar to the conventional reliability method applied to the multi-trait conventional models (Liu *et al.*, 2004).

The proposed genomic reliability method needs to be validated and verified thoroughly. Using matrix inversion to obtain exact theoretical genomic reliabilities may be a way to investigate the accuracy of the proposed method. Via simulation true reliabilities may be calculated by correlating GEBV with simulated true breeding values. Those true reliabilities may be compared to the approximated genomic reliabilities using the proposed method. We cordially invite all international partners to conduct their own validation and comparison of the proposed genomic reliability method.

The software snp_blup_rel

National genetic evaluation centers that conduct official genomic evaluation are requested to use this software for genomic reliability calculation. Using the same software ensures that all involved countries calculate DGV reliabilities in the same way. The first author developed own source programs and obtained equal SNP reliabilities, $\mathfrak{R}_i^{SNP}$, as the software *snp_blup_rel*. This indirectly verified that the software *snp_blup_rel* gives correct results. Additionally, this software takes advantage of multiple cores of computers to speed up the calculation of the genomic reliabilities. The current file format for input genotype data may be modified to be more compact. The output inverse of the coefficient matrix, eq. [5], may be stored in a binary file format instead of the current text format for faster I/O processes. A smaller file size can be achieved by changing from the co-ordinate ij-value format to a dense lower triangular format. With increasingly more cows added to national genomic reference populations in Holstein breeds, the incidence matrix $\mathbf{Z}$ in eq. [3] will be too large to be stored in computer memory for setting up the matrix products like $\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}$ or $\mathbf{Z}'\mathbf{R}^{-1}\mathbf{1}$. This problem can be solved by grouping the reference animals and set up the matrix products group by group, which can then be easily accumulated then across the groups. By doing so, we can process any number of reference animals for calculating the SNP/DGV reliabilities.

The involved countries need to test the provided software *snp_blup_rel* and share their experience with the others. Countries must develop own source programs to perform the Steps 3 to 6 for testing the proposed genomic reliability method. Based on country feedbacks, the software *snp_blup_rel* and the new genomic reliability method will be fine-tuned and optimized towards an official implementation by all the involved countries.

Conclusions

Genomic reliability calculation needs to be standardised for a better comparison across countries. A new genomic reliability approximation method has been developed that further ensures national genomic reliabilities being comparable between countries. An adjustment procedure made the genomic reliabilities of candidates more consistent with later reliabilities when they have received own phenotypes. Additionally, this adjustment procedure guaranteed the changes in genomic reliabilities correctly reflecting the changes in GEBV. A main component of the genomic reliability calculation involves the calculation of DGV reliabilities, which is accomplished by using the software *snp_blup_rel*. This common software provides a unified and efficient way for calculating the DGV reliabilities applicable for all countries participating in across country genomic evaluations. A test implementation to the large German Holstein population demonstrated that the new proposed genomic reliability method is efficient and feasible for any number of genotyped animals. The approximated genomic reliability values and the expected changes in genomic reliabilities in the adjustment step seemed to be logical. Independent, proper validation and verification of the reliability method need to be done for an official implementation by all the involved countries.

Acknowledgements

Bevin Harris who served as former convener of the Genomic Reliability Working Group is thanked for his leadership and work. Esa Mäntysaari and Ismo Strandén from Luke, Finland, are kindly acknowledged for providing the software *snp_blup_rel* to all the involved national genetic evaluation centres.

References

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. *J. Dairy Sci. 93,* 743–752.

Alkhoder, H., Liu, Z., Bach, Th., Pasman, E. & Reinhardt, F. 2014. A continuous genomic evaluation system for German Holsteins. *Interbull Bulletin 48,* 100–104.

Bijma, P. 2012. Accuracies of estimated breeding values from ordinary genetic evaluations do not reflect the correlation between true and estimated breeding values in selected populations. *J. Anim. Breed. Genet. 129,* 345-358.

Gengler, N., Mayeres, P. & Szydlowski, 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,* 21–28.

Misztal, I., Tsuruta, S., Aguilar, I., Legarra, A., VanRaden, P.M. & Lawlor, T.J. 2013. Methods to approximate reliabilities in single-step genomic evaluation. *J. Dairy Sci. 96,* 647-654.

Misztal, I., Legarra, A. & Aguilar, A. 2014. Using recursion to compute the inverse of the genomic relationship matrix. *J. Dairy Sci. 97,* 3943-3952.

Harris, B.L. & Johnson, D.L. 2010. Genomic predictions for New Zealand dairy bulls and integration with national genetic evaluation. *J. Dairy Sci. 93,* 1243-1252.

Harris, B., Calus, M., VanRaden, P.M., Lidauer, M., Ducrocq, V. & Liu, Z. 2015a First report of the genomic reliability working group. *Interbull Technical Workshop,* 24-25 Feburary, 2015, Verden, Germany.

Harris, B., Calus, M., VanRaden, P.M., Lidauer, M., Ducrocq, V. & Liu, Z. 2015b Second report of the genomic reliability working group. *Interbull Annual Meeting*, 9-12 July, 2015, Orlando, Florida, U.S.

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. 2010. Approximating reliabilities of estimated direct genomic values. *Interbull Bulletin 41*, 40-47.

Liu, Z., Alkhoder, H., Reinhardt, F. & Reents, R. 2016. Accuracy and bias of genomic prediction for second-generation candidates. *Interbull Bulletin 50,* 17-23.

Mäntysaari, E.A., Liu, Z. & VanRaden, P.M. 2010. Interbull validation tests for genomic validations. *Interbull Bulletin 41,* 17-21.

Mäntysaari, E.A. & Strandén, I. 2016. The software snp_blup_rel: Inside out. October 2016. *Interbull document*.

Strandén, I. & Mäntysaari, E.A. 2015. Effect of SNP chip density on model reliability of genomic evaluation. *EAAP Annual Meeting*, Warsaw, Poland, September 2015.

Taskinen, M., Mäntysaari, E.A., Aamand, G.P. & Strandén, I. 2014. Comparison of breeding values from single-step and bivariate blending methods. In: *Proceedings, 10th world congress of genetics applied to livestock production.* 3 p.

Taskinen, M., Mäntysaari, E.A., Lidauer, M.H., Knürr, T., Pösö, T., Su, G., Aamand, G.P. & Strandén, I. 2013. Comparison of model reliabilities from single-step and bivariate blending methods. *Interbull Bulletin 47,* 246-251.

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

VanRaden, P.M. & O'Connell, J. 2018. Validation of genomic reliabilities and gains from phenotypic updates. *Interbull Annual Meeting*, Auckland, New Zealand.

Wiggans, G.R. & P.M. VanRaden. 2010. Improved reliability approximation for genomic evaluations in the United States. *J. Dairy Sci. 93 (E-Suppl. 1):* 533. (Abstr.)