Using Monte Carlo method to include polygenic effects in calculation of SNP-BLUP model reliability

H. Ben Zaabza,* E. A. Mäntysaari, and I. Strandén
Natural Resources Institute Finland (Luke), FI-31600 Jokioinen, Finland

ABSTRACT

An SNP-BLUP model is computationally scalable even for large numbers of genotyped animals. When genetic variation cannot be completely captured by SNP markers, a more accurate model is obtained by fitting a residual polygenic effect (RPG) as well. However, inclusion of the RPG effect increases the size of the SNP-BLUP mixed model equations (MME) by the number of genotyped animals. Consequently, the calculation of model reliabilities requiring elements of the inverted MME coefficient matrix becomes more computationally challenging with increasing numbers of genotyped animals. We present a Monte Carlo (MC)-based sampling method to estimate the reliability of the SNP-BLUP model including the RPG effect, where the MME size depends on the number of markers and MC samples. We compared reliabilities calculated using different RPG proportions and different MC sample sizes in analyzing 2 data sets. Data set 1 (data set 2) contained 19,757 (222,619) genotyped animals, with 11,729 (50,240) SNP markers, and 231,186 (13.35 million) pedigree animals. Correlations between the correct and the MC-calculated reliabilities were above 98% even with 5,000 MC samples and an 80% RPG proportion in both data sets. However, more MC samples were needed to achieve a small maximum absolute difference and mean squared error, particularly when the RPG proportion exceeded 20%. The computing time for MC SNP-BLUP was shorter than for GBLUP. In conclusion, the MC-based approach can be an effective strategy for calculating SNP-BLUP model reliability with an RPG effect included.

Key words: breeding value, genomic evaluation, model reliability

INTRODUCTION

Genomic information has become increasingly important in predicting breeding values in dairy cattle breeding (Schaeffer, 2006). Two equivalent models have been proposed for this purpose, namely genomic relationship matrix-based BLUP (**GBLUP**) and SNP-BLUP (Strandén and Garrick, 2009). In a genetic evaluation, the model reliabilities of EBV can be used to compare the accuracy of EBV across generations (Misztal et al., 2013). The main computational work in calculating EBV model reliabilities is often related to the inversion of mixed model equations (**MME**). The size of the MME increases with the number of genotyped animals in GBLUP but is bounded by the number of markers used in SNP-BLUP. In GBLUP, the need to invert the genomic relationship matrix **G** increases the computing time further. When the number of genotyped animals (n) increases beyond the number of markers (m), the computing time needed for MME inversion is roughly proportional to $(n/m)^3$ for GBLUP over SNP-BLUP. Thus, SNP-BLUP can be expected to be computationally scalable even for large numbers of genotyped animals.

Typically, genetic variation cannot be completely captured by SNP markers because of the incomplete linkage disequilibrium between the QTL and the genome-wide markers used to predict genetic values (Liu et al., 2011). However, GBLUP and SNP-BLUP models can be improved by augmenting them with residual polygenic (**RPG**) effects. For example, using an RPG proportion of 20% in multi-step genomic BLUP is known to improve the reliability of the prediction and reduce bias (Guarini et al., 2018). Liu et al. (2016) showed that GBLUP and SNP-BLUP genomic models that include an RPG effect are equivalent, although inclusion of the RPG effect in the models increases the MME size by the number of genotyped animals. Hence, the reliability calculation requiring elements of the inverse of the MME coefficient matrix will become more computationally challenging in both models the more genotyped animals are included.

Received July 12, 2019.

Accepted February 4, 2020.

*Corresponding author: hafedh.benzaabza@luke.fi

Approximation is commonly used when the computation of exact model reliability by direct inversion of the left-hand side of the MME becomes unfeasible. Several approximation methods for non-genomic evaluations have been suggested (Misztal and Wiggans, 1988; Gengler and Misztal, 1996; Harris and Johnson, 1998; Jamrozik et al., 2000; Tier and Meyer, 2004), and some have even been proposed for single-step GBLUP (Misztal et al., 2013; Edel et al., 2019). Sargolzaei et al. (2014) presented a model reliability approximation approach for SNP-BLUP and GBLUP. They introduced 2 methods, using (1) a subset of training individuals or (2) a subset of SNP markers. The authors reported that there were some biases in the mean and standard deviation of the approximated reliability associated with these methods. To our knowledge, no approximate approaches have been published for GBLUP or SNP-BLUP with RPG effects included.

In this study, we present a Monte Carlo (MC)-based approximation method to include the RPG effect into the SNP-BLUP model. This MC approach increases the MME size by the number of MC samples instead of the number of RPG effects. The objectives of our study were (1) to calculate the individual animal reliability in SNP-BLUP with an RPG effect using the presented method, (2) to compare the obtained model reliabilities for different MC sample sizes to the correct values, and (3) to test different proportions of the RPG effect in reliability calculations for 2 genotyped cattle data sets.

MATERIALS AND METHODS

Reliability of Genomic Models with a Polygenic Effect

Consider a GBLUP model,

$$\mathbf{y} = \mathbf{1}_n\mu + \mathbf{u} + \mathbf{e}, \quad [1]$$

where $\mathbf{y}$ is a vector of observations, $\mathbf{1}_n$ is a vector of ones, μ is the unknown overall mean, $\mathbf{u}$ is the vector of unknown breeding values, $\mathbf{e}$ is the residual vector, and all vectors are of size $n \times 1$. It is assumed that the random effects $\mathbf{u}$ and $\mathbf{e}$ are independent, and that $\mathbf{e} \sim N(\mathbf{0}, \mathbf{R})$, where $\mathbf{R}$ is a diagonal residual variance or covariance matrix containing weights (VanRaden, 2008). Further, we assume that $\mathbf{u} \sim N(\mathbf{0}, \mathbf{G}\sigma_u^2)$, where $\mathbf{G}$ is the genomic relationship matrix and σ_u^2 is the genetic variance. Let us denote the genomic relationship matrix as $\mathbf{G} = \mathbf{Z}_c\mathbf{Z}_c'$, where $\mathbf{Z}_c$ is the centered and scaled marker matrix of size $n \times m$. In a typical centering and scaling of $\mathbf{Z}_c$ (VanRaden, 2008), the genotypes of marker k are coded as $\frac{1}{s_k}(0 - 2p_k)$, $\frac{1}{s_k}(1 - 2p_k)$, and $\frac{1}{s_k}(2 - 2p_k)$ for

genotypes BB, AB, and AA, respectively, where p_k is the (base population) allele frequency of marker k . The common scaling factors are $s_k = \sqrt{\sum_{l=1}^m 2p_l(1-p_l)}$ and $s_k = \sqrt{m[2p_k(1-p_k)]}$, often called VanRaden methods 1 and 2, respectively.

A polygenic effect can be accounted for in GBLUP either by an additional RPG effect or by combining the genomic and polygenic breeding value effects. We used the latter approach because it allows a simple equivalent SNP-BLUP model. Let the genomic relationship matrix combining the genomic and polygenic covariances be

$$\mathbf{G}_w = (1-w)\mathbf{Z}_c\mathbf{Z}_c' + w\mathbf{A}_{22}, \quad [2]$$

where w is the proportion of genetic variance not captured by markers and $\mathbf{A}_{22}$ is the pedigree relationship matrix of size $n \times n$ for genotyped animals.

The MME for the GBLUP model [1] using $\mathbf{G}_w$ is

$$\begin{bmatrix} \mathbf{1}_n'\mathbf{R}^{-1}\mathbf{1}_n & \mathbf{1}_n'\mathbf{R}^{-1} \\ \mathbf{R}^{-1}\mathbf{1}_n & \mathbf{R}^{-1} + \mathbf{G}_w^{-1}\sigma_u^{-2} \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{1}_n'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{R}^{-1}\mathbf{y} \end{bmatrix},$$

where we assume that $\mathbf{G}_w$ is non-singular. Denote the coefficient matrix of this MME by $\mathbf{C}$ and its inverse matrix elements as $\mathbf{C}^{-1} = \begin{bmatrix} \mathbf{C}^{\mu\mu} & \mathbf{C}^{\mu u} \\ \mathbf{C}^{u\mu} & \mathbf{C}^{uu} \end{bmatrix}$. The reliability of

the EBV of animal j can be calculated as $r_j^2 = 1 - \frac{\mathbf{PEV}_j}{\mathbf{G}_{jj}\sigma_u^2}$,

where $\mathbf{PEV}_j$ is diagonal element j in the prediction error variance (PEV) matrix $\mathbf{C}^{uu}$, and $\mathbf{G}_{jj}$ is diagonal element j in the genomic relationship matrix $\mathbf{G}_w$. Note that the computation of the PEV matrix $\mathbf{C}^{uu}$ requires 2 inversions: (1) inversion of the genomic relationship matrix of size n and (2) inversion of the MME coefficient matrix of size $n+1$.

An SNP-BLUP model equivalent to the GBLUP model [1] using the genomic relationship matrix $\mathbf{G}_w$ [2] is

$$\mathbf{y} = \mathbf{1}_n\mu + \mathbf{Z}\mathbf{g} + \mathbf{e}, \quad [3]$$

where $\mathbf{Z} = [\sqrt{(1-w)}\mathbf{Z}_c \quad \sqrt{w}\mathbf{L}]$, the square matrix $\mathbf{L}$ is a Cholesky decomposition of $\mathbf{A}_{22}$ (i.e., $\mathbf{A}_{22} = \mathbf{L}\mathbf{L}'$), and $\mathbf{g}$ is $m_n \times 1$ vector of unknown marker or pseudo-marker effects, $m_n = m + n$. For example, when 1,000,000 animals are genotyped for 50,000 markers, the corresponding $\mathbf{Z}$ matrix is of size 1,000,000 $\times$ 1,050,000, and there are 1,050,000 (pseudo) marker ef-

fects. We assume that $\mathbf{g} \sim N(\mathbf{0}, \mathbf{I}_{m_n} \sigma_u^2)$. Note that $\text{Var}(\mathbf{Zg}) = \mathbf{ZZ}'\sigma_u^2 = \mathbf{G}_w\sigma_u^2 = \text{Var}(\mathbf{u})$, and thus σ_u^2 is the same additive genetic variance as in GBLUP. With no RPG effects ($w = 0$), the $\mathbf{Z}$ matrix contains columns of zeros for the effects related to RPG effects and, consequently, unnecessary unknowns. In contrast, in the GBLUP model, the RPG effect is included in the covariance structure of breeding values, and no additional unknowns are generated.

The equivalence of GBLUP and SNP-BLUP models with RPG effects is demonstrated in Liu et al. (2016), where the RPG is a separate effect in the SNP-BLUP model. According to Henderson (1984), 2 models are linearly equivalent when their phenotypic variance matrices $\text{Var}(\mathbf{y})$ are equal and their fixed effect predictions are also equal. For the GBLUP model [1], $\text{Var}(\mathbf{y}) = \sigma_u^2 \mathbf{G}_w + \mathbf{R} = \mathbf{V}$, and for the SNP-BLUP model [3], $\text{Var}(\mathbf{y}) = \sigma_u^2 \mathbf{ZZ}' + \mathbf{R} = \sigma_u^2 [(1-w)\mathbf{Z}_c\mathbf{Z}_c' + w\mathbf{LL}'] + \mathbf{R} = \mathbf{V}$. Because the fixed-effects model ($\mathbf{1}_n$) and $\text{Var}(\mathbf{y})$ matrices are the same in both models [1] and [3], their fixed-effect predictions $\mathbf{1}_n\hat{\mu}$ are equal and the models are linearly equivalent. The EBV from the GBLUP model are $\hat{\mathbf{u}}$, and from the SNP-BLUP model $\mathbf{Z}\hat{\mathbf{g}}$, which are equal due to the equality of their ($\mathbf{u}$ and $\mathbf{Zg}$) variances and the linear equivalence of the models. The Appendix shows the transformations needed from model [3] to a model with a separate RPG effect—that is, how to obtain the same model effects and PEV for these models.

The MME for the SNP-BLUP model [3] is

$$\begin{bmatrix} \mathbf{1}_n' \mathbf{R}^{-1} \mathbf{1}_n & \mathbf{1}_n' \mathbf{R}^{-1} \mathbf{Z} \\ \mathbf{Z}' \mathbf{R}^{-1} \mathbf{1}_n & \mathbf{Z}' \mathbf{R}^{-1} \mathbf{Z} + \mathbf{I}_u \sigma_u^{-2} \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{\mathbf{g}} \end{bmatrix} = \begin{bmatrix} \mathbf{1}_n' \mathbf{R}^{-1} \mathbf{y} \\ \mathbf{Z}' \mathbf{R}^{-1} \mathbf{y} \end{bmatrix},$$

where EBV are $\hat{\mathbf{u}} = \mathbf{Z}\hat{\mathbf{g}}$. As before, we denote the coefficient matrix of this MME by $\mathbf{C}_g$ and its inverse matrix elements as $\mathbf{C}_g^{-1} = \begin{bmatrix} \mathbf{C}_g^{\mu\mu} & \mathbf{C}_g^{\mu g} \\ \mathbf{C}_g^{g\mu} & \mathbf{C}_g^{gg} \end{bmatrix}$. The reliability of the EBV for animal j is $r_j^2 = 1 - \frac{\text{PEV}_j}{\mathbf{G}_{jj}\sigma_u^2}$, where PEV_j is diagonal element j in the PEV matrix $\mathbf{ZC}_g^{gg}\mathbf{Z}'$ or $\mathbf{z}_j\mathbf{C}_g^{gg}\mathbf{z}_j'$ with $\mathbf{z}_j$ equal to row j in $\mathbf{Z}$. Note that in this SNP-BLUP model, the lower computational cost of calculating the Cholesky decomposition $\mathbf{L}$ instead of inverting the genomic relationship matrix $\mathbf{G}_w$ in GBLUP is often lost due to the need to invert $\mathbf{C}_g$ of size m_n , which is larger than n . As an example, consider the 1,000,000 genotyped $\mathbf{Z}$ having a size of $1,000,000 \times 1,050,000$. The corresponding $\mathbf{C}_g$ is of size $1,050,001 \times 1,050,001$.

This is larger than the MME of the original GBLUP approach, whereas the MME of the original SNP-BLUP had size $50,001 \times 50,001$. We overcome this difficulty of large MME size by using a Monte Carlo approach, which is explained in the next section.

Monte Carlo Sampling to Estimate Numerator Relationship Matrix $\mathbf{A}$

The $\mathbf{A}_{22}$ matrix needed in model [3] to account for the RPG effect can be approximated by MC sampling. Let $\mathbf{a}_i$ be the breeding value vector of q related animals from MC sample i : $\mathbf{a}_i \sim N(\mathbf{0}, \mathbf{A})$, $i = 1, \dots, n_{\text{MC}}$, where n_{MC} is the number of MC samples and $\mathbf{A}$ is the pedigree-based numerator relationship matrix. The additive genetic variance can be assumed to be equal to 1. After ordering the animals such that parents precede their progeny, the breeding value for animal j can be generated by

$$\mathbf{a}_{i,j} = \frac{1}{2}(\mathbf{a}_{i,s} + \mathbf{a}_{i,d}) + \phi_{i,j},$$

where subscript s is the sire and d is the dam of animal j , and $\phi_{i,j}$ is the Mendelian segregation effect. The segregation effect is normally distributed with mean 0 and variance $\frac{1}{2}(1 - \bar{F})$, where $\bar{F}$ is the mean of parent inbreeding coefficients, $\bar{F} = \frac{1}{2}(F_s + F_d)$. The breeding value of an unknown parent is simulated from the standard normal density $N(0,1)$.

The variance of $\mathbf{a}$ is $\mathbf{A} = \text{Var}(\mathbf{a}) = E(\mathbf{aa}') - E(\mathbf{a})E(\mathbf{a})' = E(\mathbf{aa}')$, because $E(\mathbf{a}) = \mathbf{0}$. A simple MC-based estimator of $\mathbf{A}$ is $\hat{\mathbf{A}} = \frac{1}{n_{\text{MC}}} \sum_{i=1}^{n_{\text{MC}}} \mathbf{a}_i \mathbf{a}_i'$. It follows that using the independent samples of $\mathbf{a}_i$, the MC estimator equals

$$\hat{\mathbf{A}} = \frac{1}{n_{\text{MC}}} \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_{n_{\text{MC}}} \end{bmatrix} \begin{bmatrix} \mathbf{a}_1' \\ \vdots \\ \mathbf{a}_{n_{\text{MC}}}' \end{bmatrix} = \mathbf{U}\mathbf{U}', \quad [4]$$

where $\mathbf{U} = \frac{1}{\sqrt{n_{\text{MC}}}} \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_{n_{\text{MC}}} \end{bmatrix}$ is a $q \times n_{\text{MC}}$ matrix. A submatrix of $\mathbf{A}$, such as $\mathbf{A}_{22}$, can be approximated by using only the elements pertaining to the submatrix animals in the $\mathbf{a}_i$ vectors. Thus, MC sampling can be limited to the submatrix of genotyped animals and their ancestors, and thereafter, only the MC samples of genotyped animals will be used. Note that the rank of $\hat{\mathbf{A}}$ is $\min(n_{\text{MC}}, q)$, and the rank of $\hat{\mathbf{A}}_{22}$ is $\min(n_{\text{MC}}, n)$. Thus, $\hat{\mathbf{A}}_{22}$ is singular when $n_{\text{MC}} < n$.

Expression [4] can be formulated differently by making the substitution $\mathbf{a}_i = \mathbf{L}\mathbf{x}_i$, where $\mathbf{L}$ is the Cholesky decomposition of $\mathbf{A} = \mathbf{L}\mathbf{L}'$, and $\mathbf{x}_i \sim N(\mathbf{0}, \mathbf{I}_q)$. Then, $\hat{\mathbf{A}} = \mathbf{L}\hat{\mathbf{I}}\mathbf{L}'$, where $\hat{\mathbf{I}} = \frac{1}{n_{\text{MC}}} \sum_{i=1}^{n_{\text{MC}}} \mathbf{x}_i \mathbf{x}_i'$. Matrix $\hat{\mathbf{I}}$ approaches the identity matrix when n_{MC} increases, because $E(\mathbf{x}\mathbf{x}') = E(\mathbf{x}\mathbf{x}') - E(\mathbf{x})E(\mathbf{x})' = \text{Var}(\mathbf{x}) = \mathbf{I}_q$, and $E(\mathbf{x}) = \mathbf{0}$. But, like $\hat{\mathbf{A}}$, matrix $\hat{\mathbf{I}}$ is singular when $n_{\text{MC}} < q$. Similarly, we can express $\hat{\mathbf{A}}_{22} = \mathbf{L}_2 \hat{\mathbf{I}}_2 \mathbf{L}_2'$, where $\mathbf{L}_2$ is the Cholesky decomposition of $\mathbf{A}_{22}$, $\mathbf{A}_{22} = \mathbf{L}_2 \mathbf{L}_2'$, $\hat{\mathbf{I}}_2 = \frac{1}{n_{\text{MC}}} \sum_{i=1}^{n_{\text{MC}}} \mathbf{v}_i \mathbf{v}_i'$, and $\mathbf{v}_i \sim N(\mathbf{0}, \mathbf{I}_n)$. Matrix $\hat{\mathbf{I}}_2$ is singular when $n_{\text{MC}} < n$.

Monte Carlo in Residual Polygenic Effect Approximation

The approximation [4] can be used in place of the $\mathbf{A}_{22}$ matrix in the genomic relationship matrix $\mathbf{G}_w$:

$$\mathbf{G}_w^* = (1 - w)\mathbf{Z}_c \mathbf{Z}_c' + w\mathbf{U}_{22} \mathbf{U}_{22}', \quad [5]$$

where $\mathbf{U}_{22}$ is an $n \times n_{\text{MC}}$ matrix of the MC samples for genotyped individuals in the $\mathbf{U}$ matrix. The approximate genomic relationship matrix can be expressed as $\mathbf{G}_w^* = \mathbf{S}\mathbf{S}'$, where $\mathbf{S} = [\mathbf{S}_Z \quad \mathbf{S}_U]$ and $\mathbf{S}_U = \sqrt{w}\mathbf{U}_{22}$. This allows the approximation of the genomic relationship matrix having pedigree relationship information, without a need to form the $\mathbf{A}$ matrix explicitly. We call this approach MC-GBLUP. Note that to invert $\mathbf{G}_w^*$, it has to be non-singular. The $\mathbf{G}_w^*$ matrix is non-singular when $\text{rank}(\mathbf{G}_w^*) = n$, which is true when $n_{\text{MC}} + \text{rank}(\mathbf{Z}_c) \geq n$. Note that the number of MC samples may need to be much higher than $n - \text{rank}(\mathbf{Z}_c)$ for the MC approach to be accurate. In the special case of a fully polygenic model ($w = 1$), at least n MC samples are necessary ($n_{\text{MC}} \geq n$).

The genomic relationship matrix [5] can be used to create an equivalent MC-SNP-BLUP model:

$$\mathbf{y} = \mathbf{1}_n \mu + \mathbf{S}\mathbf{g}_s + \mathbf{e}, \quad [6]$$

where $\mathbf{g}_s$ is a vector of $(m + n_{\text{MC}})$ unknown (pseudo) marker effects, of which m are due to the SNP markers and n_{MC} are dummy effects for approximating the breeding values where an RPG effect is included. The MME for the MC-SNP-BLUP model [6] is

$$\begin{bmatrix} \mathbf{1}_n' \mathbf{R}^{-1} \mathbf{1}_n & \mathbf{1}_n' \mathbf{R}^{-1} \mathbf{S} \\ \mathbf{S}' \mathbf{R}^{-1} \mathbf{1}_n & \mathbf{S}' \mathbf{R}^{-1} \mathbf{S} + \mathbf{I}_{\sigma_u^2} \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{\mathbf{g}}_s \end{bmatrix} = \begin{bmatrix} \mathbf{1}_n' \mathbf{R}^{-1} \mathbf{y} \\ \mathbf{S}' \mathbf{R}^{-1} \mathbf{y} \end{bmatrix}.$$

Denote the coefficient matrix of this MME by $\mathbf{C}_s$ and its inverse matrix elements as $\mathbf{C}_s^{-1} = \begin{bmatrix} \mathbf{C}_s^{\mu\mu} & \mathbf{C}_s^{\mu g} \\ \mathbf{C}_s^{g\mu} & \mathbf{C}_s^{gg} \end{bmatrix}$. The approximate reliability for the EBV of animal j is

$$r_j^2 = 1 - \frac{\mathbf{PEV}_j}{\mathbf{G}_{jj}^* \sigma_u^2}, \quad [7]$$

where $\mathbf{PEV}_j$ is diagonal element j in the PEV matrix, and $\mathbf{G}_{jj}^*$ is diagonal element j in the genomic relationship matrix $\mathbf{G}_w^* = \mathbf{S}\mathbf{S}'$. The PEV for animal j is $\mathbf{s}_j \mathbf{C}_s^{gg} \mathbf{s}_j'$, where $\mathbf{s}_j$ is row j in $\mathbf{S}$.

The reliability calculations in GBLUP and SNP-BLUP require the diagonal elements of the genomic relationship matrix $\mathbf{G}_w$. In the presented MC approaches, the diagonal elements of $\mathbf{G}_w^*$ in the denominator of the reliability formula $r_j^2 = 1 - \frac{\mathbf{PEV}_j}{\mathbf{G}_{jj}^* \sigma_u^2}$ included the approximate diagonal elements of $\mathbf{A}_{22}$ by MC sampling. The approximate $\mathbf{G}_{jj}^*$ can be replaced by the exact value $\mathbf{G}_{jj}$, because the diagonal values of $\mathbf{A}_{22}$ can be easily calculated. The diagonal values of $\mathbf{A}_{22}$ equal $1 + F_j$, where F_j is the inbreeding coefficient of animal j . We call the MC approach using $\mathbf{G}_{jj}^*$ the full MC approach, and the approach using $\mathbf{G}_{jj}$ the partial MC approach.

Data

The EBV reliabilities were calculated by applying both GBLUP and SNP-BLUP models. We used 2 data sets to examine the quality of the approximate MC-SNP-BLUP model reliability method. The first data set (hereafter called Data 1) consisted of 19,757 genotyped animals and 231,186 pedigree animals, comprising Finnish Red dairy cattle. The genotyped animals in Data 1 had 11,729 SNP markers, which were a sample from the markers used in the joint Nordic genomic evaluations. The second data set (hereafter called Data 2) consisted of a sample of 222,619 genotyped animals and 13.35 million pedigree animals from the Irish beef cattle carcass conformation evaluation, comprising several breeds (Mäntysaari et al., 2017). The number of markers in the Irish evaluation was 50,240. In our study, all genotyped animals were assumed to have 1 observation. For Data 1, no weights were used, and the heritability of the trait was set as 0.5. For Data 2, we used the effective record number as a weight, and assumed the heritability of the trait to be 0.27.

Implementation and Computations

Individual animal model reliabilities were calculated using GBLUP and MC-SNP-BLUP. The GBLUP model gave the exact reliabilities to which we compared those approximated by MC-SNP-BLUP. We tested 4 RPG weights: 1%, 20%, 50%, and 80%. The $\mathbf{Z}_c$ matrix was centered and scaled as in method [1] by VanRaden (2008), using the base population allele frequencies. These were estimated by the GLS approach described in McPeck et al. (2004) and Strandén et al. (2017).

The MC-SNP-BLUP model reliability calculations were tested with different numbers of MC samples. For Data 1, we used MC sample sizes of 1,000, 5,000, 10,000, and 20,000, and for Data 2, sample sizes of 5,000, 10,000, 20,000, 40,000, and 60,000. We report the correlation, maximum absolute difference, and mean squared error (MSE) between the model reliabilities from GBLUP and MC-SNP-BLUP. In addition, for each setup, we fitted the linear regression of the reliability from GBLUP on that from MC-SNP-BLUP. Furthermore, statistics on the difference between the MC-estimated diagonal elements of the $\mathbf{A}_{22}$ matrix and its true value were calculated using the 2 data sets. We computed the MSE as well as the signed absolute maximum difference, where the sign indicates either positive or negative bias.

Statistics on the concordance between the correct elements of $\mathbf{A}_{22}$ and those approximated by MC sampling, separately for diagonal and off-diagonal elements, were investigated with Data 1. The $\mathbf{A}_{22}$ matrix approximated by different MC sample sizes was subtracted from the true $\mathbf{A}_{22}$ matrix. From this difference, we computed the following descriptive statistics: (1) average of diagonals (dBias), (2) square root of average of squared diagonals (dsMSE), (3) average of off-diagonal elements (off-bias), and (4) square root of average of squared values of off-diagonal elements (offsMSE) in 3 reliability groups (low, medium, and high). The 3 reliability groups each had an equal number of genotyped animals, where animals were assigned to the groups by their GBLUP-based reliabilities when the RPG weight was 80%. Thus, all reliabilities in the low group were less than those in the medium and high groups, and the reliabilities in the medium group were all less than those in the high group.

We report the wall-clock computing times for the different steps in MC-SNP-BLUP, as well as the total computing time. The latter includes the reported steps and final computing times for the reliabilities, of which a significant proportion was due to computation of $\text{PEV}(\mathbf{s}_j \mathbf{C}_s^{gg} \mathbf{s}_j')$ for each animal. The computing times are independent of the size of the RPG weight but depen-

dent on the MME size. To diminish the effect of fluctuations in computing times due to variability in total computer usage by other users, we report the average computing times over the RPG weights used.

To make and invert the required MME coefficient matrices, we used LAPACK (Anderson et al., 1999) subroutines from the Intel Math Kernel Library (Intel, 2014). We applied single precision for the MC samples and for making the MME coefficient matrix, but double precision for inversion of the MME matrix and for the reliability calculations. We used a multi-core computer with 2 Intel Xeon E5-2680 v. 2 processors (2.8 GHz; Intel Corp., Santa Clara, CA) for the computations and limited them to 10 CPU cores.

RESULTS

MC-SNP-BLUP and GBLUP Model Reliability Estimations

The correlations of model reliabilities between exact GBLUP and approximate MC-SNP-BLUP-based models for Data 1 are given in Table 1. In GBLUP, the RPG information was included into the genomic relationship matrix, whereas in MC-SNP-BLUP, the marker effect coefficients were augmented by MC-generated breeding values. Increasing the number of MC samples increased the correlation between model reliabilities, which became clearer the larger the RPG proportion. For example, the correlation for MC-SNP-BLUP with an RPG proportion of 1% [MC-SNP-BLUP(1)] was 1.000 for all MC sample sizes. For MC-SNP-BLUP(80), the correlations were 0.974, 0.997, 0.999, and 0.999 for MC sample sizes 1,000, 5,000, 10,000, and 20,000, respectively. For 1,000 MC samples, the correlations between MC-SNP-BLUP and GBLUP reliabilities increased as the proportion of the RPG effect decreased: 0.974, 0.992, 0.999, and 1.000 for RPG proportions of 80%, 50%, 20%, and 1%, respectively. These correlations became stronger when the MC sample size was increased to 5,000: 0.997, 0.999, 1.000, and 1.000 for RPG proportions of 80%, 50%, 20%, and 1%, respectively. Further increasing the number of MC samples to 10,000 increased the correlations between MC-SNP-BLUP and GBLUP model reliabilities only slightly for all RPG proportions.

The regression coefficients (slopes) of most of the scenarios did not differ significantly from 1, and the intercepts were close to 0, with the exception of scenarios with a high RPG weight and small MC sample size. With more MC samples, the slopes and intercepts approached 1 and 0, respectively. For example, with an RPG of 80%, the slope was 1.98 (intercept -0.89), 1.23 (-0.19), 1.12 (-0.1), and 1.06 (0.05) for MC sample

sizes of 1,000, 5,000, 10,000, and 20,000, respectively. The smaller the RPG, the closer the intercept was to 0 and the slope to 1. For an MC sample size of 1,000, the slope was 1.98 (intercept -0.89), 1.36 (-0.33), 1.06 (-0.06), and 1.00 (0.00) for RPG weights of 80%, 50%, 20%, and 1%, respectively.

The maximum absolute difference and MSE between GBLUP and MC-SNP-BLUP model reliabilities for Data 1 are shown in Table 1. The MSE associated with MC-SNP-BLUP(1) was practically 0 in all MC sample sizes. For the remaining RPG proportions, the MSE decreased as the number of MC samples increased, which is consistent with the correlation results. However, we found large relative differences in the MSE values depending on the MC sample sizes, ranging from $116 \cdot 10^{-5}$ with 1,000 MC samples to $2 \cdot 10^{-5}$ with 10,000 MC samples for MC-SNP-BLUP(20). The relative change in the MSE was much smaller, from $2 \cdot 10^{-5}$ to $1 \cdot 10^{-5}$, when the sample size increased from 10,000 to 20,000. Similarly, the larger the MC sample size, the lower the maximum error. For MC-SNP-BLUP(20), increasing the sample size from 1,000 to 10,000 reduced the maximum error by up to 6 times. The reductions in the maximum error and the MSE were relatively modest when the number of MC samples increased from 10,000 to 20,000.

Figures 1 and 2 illustrate the effect of MC sample size on model reliability for RPG proportions of 20% and 80% in MC-SNP-BLUP when analyzing Data 1. In the scenario with a large RPG proportion and small MC sample size, reliabilities from MC-SNP-BLUP were markedly inflated when the reliability values were less than 0.6. The inflation diminished consistently for reliabilities higher than 0.6. With 5,000 and 10,000 MC samples, the MC-SNP-BLUP model reliability estimations showed approximately 30% and 50% less inflation compared with estimations using 1,000 MC samples (Figure 2). For a 20% RPG proportion (Figure 1), the MC-SNP-BLUP model reliabilities deviated considerably less from GBLUP than for an 80% RPG proportion. As shown in Figure 2, using MC-SNP-BLUP with a small MC sample size gave reliabilities that were more inflated, compared with the exact GBLUP values, than those derived from SNP-BLUP without an RPG effect. Compared with GBLUP model reliabilities, the maximum difference (0.33 vs. 0.47) and MSE (0.040 vs. 0.093) were lower for SNP-BLUP than for MC-SNP-BLUP using 1,000 MC samples, although the correlations were approximately the same (0.97).

Table 2 contains statistics on GBLUP versus MC-SNP-BLUP for Data 2. The correlation estimates between MC-SNP-BLUP and GBLUP model reliabilities were generally higher than 0.990, with the exception of the correlation (0.983) between GBLUP and MC-SNP-BLUP(80) using 5,000 MC samples. With 5,000 MC

Table 1. Correlation (r), maximum difference (max), and mean squared error (MSE) between correct reliability from genomic relationship matrix-based BLUP (GBLUP) and approximation by Monte Carlo (MC)-SNP-BLUP under different numbers of MC samples (N_{MC}) and different residual polygenic proportions (w) in analysis of data set 1 (Data 1)¹

N_{MC}	w	r	Max	MSE ($\times 10^{-5}$)	b_0	b_1
1,000	0.01	1.000	0.00	0	0.00	1.00
	0.20	0.999	0.06	116	-0.06	1.06
	0.50	0.992	0.24	2,259	-0.33	1.36
	0.80	0.974	0.47	9,297	-0.89	1.98
5,000	0.01	1.000	0.00	0	-0.00	1.00
	0.20	1.000	0.02	7	0.01	1.02
	0.50	0.999	0.07	207	-0.08	1.09
	0.80	0.997	0.17	1,072	-0.19	1.23
10,000	0.01	1.000	0.00	0	0.00	1.00
	0.20	1.000	0.01	2	-0.01	1.01
	0.50	0.999	0.04	58	0.04	1.05
	0.80	0.999	0.09	322	-0.10	1.12
20,000	0.01	1.000	0.00	0	0.00	1.00
	0.20	1.000	0.01	1	0.00	1.00
	0.50	1.000	0.02	16	0.02	1.02
	0.80	0.999	0.05	85	0.05	1.06

¹Data 1 comprised 19,757 genotyped animals by 11,729 SNP markers. Intercept (b_0) and slope (b_1) of regression of correct reliability from GBLUP on approximation by MC-SNP-BLUP.

samples, the MSE of the approximation by MC-SNP-BLUP decreased along with a lower RPG proportion: 0.16945, 0.04323, and 0.00242 for RPG proportions of 80%, 50%, and 20%, respectively. The MSE was markedly reduced when the MC sample size increased to

Figure 1. Exact model reliability by genomic relationship matrix-based BLUP (GBLUP) versus approximate reliability by Monte Carlo (MC)-SNP-BLUP or by SNP-BLUP for residual polygenic effect (RPG) proportion of 20%.

Table 2. Correlation (r), maximum difference (max), and mean squared error (MSE) between correct reliability from genomic relationship matrix-based BLUP (GBLUP) and approximation by Monte Carlo (MC)-SNP-BLUP under different numbers of MC samples (N_{MC}) and different residual polygenic proportions (w) in analysis of data set 2 (Data 2)¹

N_{MC}	w	r	Max	MSE ($\times 10^{-5}$)	b_0	b_1
5,000	0.20	0.999	0.07	242	0.09	1.10
	0.50	0.994	0.30	4,323	-0.43	1.46
	0.80	0.983	0.54	16,945	-1.23	2.35
10,000	0.20	1.000	0.05	81	0.05	1.05
	0.50	0.998	0.19	1,822	-0.24	1.26
	0.80	0.993	0.39	8,280	-0.63	1.70
20,000	0.20	1.000	0.03	24	0.03	1.03
	0.50	0.999	0.11	622	-0.13	1.14
	0.80	0.998	0.24	3,142	-0.31	1.36
40,000	0.20	1.000	0.01	6	0.01	1.01
	0.50	1.000	0.06	184	-0.06	1.07
	0.80	0.999	0.13	995	-0.15	1.18
60,000	0.20	1.000	0.01	3	-0.00	1.01
	0.50	1.000	0.04	81	-0.04	1.05
	0.80	1.000	0.09	481	-0.10	1.18

¹Data 2 comprised 222,619 genotyped animals by 50,240 SNP markers. Intercept (b_0) and slope (b_1) of regression of correct reliability from GBLUP on approximation by MC-SNP-BLUP.

40,000: 0.00995, 0.00184, and 0.00006 for RPG proportions of 80%, 50%, and 20%, respectively. Increasing the number of MC samples from 40,000 to 60,000 decreased the MSE only slightly.

The results obtained with Data 2 were similar to those obtained with Data 1, with the major difference

that more MC samples were needed with Data 2 to attain as high correlations and as small MSE as those obtained using Data 1. For example, a correlation of 0.999 between MC-SNP-BLUP(80) and GBLUP was achieved with 10,000 MC samples when using Data 1, but obtaining the same correlation with Data 2 required 40,000 MC samples. Similarly, when Data 2 was used, more MC samples were needed to attain an equally low MSE to that achieved with Data 1. For example, an MC sample size of 5,000 was sufficient to give an MSE of 0.01702 using Data 1, whereas with Data 2, 40,000 MC samples were needed to obtain an MSE of 0.00995.

Difference Between the True and Estimated A_{22} Matrices

The denominator values in the model reliability formula are from the diagonal of the A_{22} matrix. We used MC approximation, even though the exact value was available. Table 3 shows the concordance between true and MC-approximated A_{22} diagonal elements under different MC sample sizes and RPG proportions in the analyses of Data 1 and Data 2. These are averages over different proportions of RPG. Similar results were obtained in both data sets. Increasing the number of MC samples decreased both the maximum difference and the MSE. Given that the errors associated with 10,000 MC samples were already negligible, increasing the MC sample size to 20,000 did not greatly improve the MC estimate.

Table 4 gives statistics on the differences in the diagonal and off-diagonal elements between the true and approximated A_{22} matrices under different MC sample sizes for Data 1. The larger the number of MC samples, the smaller the values of dsMSE and offsMSE were. For example, the offsMSE estimates were at least 0.0323,

Figure 2. Exact model reliability by genomic relationship matrix-based BLUP (GBLUP) versus approximate reliability by Monte Carlo (MC)-SNP-BLUP or by SNP-BLUP for residual polygenic effect (RPG) proportion of 80%.

Table 3. Signed absolute maximum difference (max) and mean squared error (MSE) between true A_{22} diagonal elements and Monte Carlo (MC)-approximated A_{22} diagonal elements under different numbers of MC samples (N_{MC}) in analysis of data sets 1 and 2 (Data 1, Data 2)¹

N_{MC}	Data 1		Data 2	
	Max	MSE ($\times 10^{-5}$)	Max	MSE ($\times 10^{-5}$)
1,000	-0.21	216	—	—
5,000	-0.10	43	-0.10	41
10,000	-0.06	21	-0.07	20
20,000	0.04	11	0.05	10
40,000	—	—	0.04	5
60,000	—	—	-0.03	3

¹Data 1 comprised 19,757 genotyped animals and 11,729 SNP markers. Data 2 comprised 222,619 genotyped animals and 50,240 SNP markers. Reported numbers are maximum values from different residual polygenic proportions.

Table 4. Average difference between the true and approximated $\mathbf{A}_{22}$ matrices with 1,000, 5,000, 10,000, and 20,000 Monte Carlo (MC) samples by low (average $r^2 = 0.19$), medium (average $r^2 = 0.32$), and high (average $r^2 = 0.38$) genomic relationship matrix-based BLUP model reliability groups¹

r^2 group	dBias $\times 10^3$	dsMSE $\times 10^2$	off-bias $\times 10^3$	offsMSE $\times 10^2$
1,000 MC samples				
Low	-0.47	4.47	-1.10	3.23
Medium	0.62	4.51	0.06	3.24
High	2.48	4.49	0.06	3.24
5,000 MC samples				
Low	-0.59	2.06	-0.55	1.45
Medium	-0.68	2.04	-0.36	1.45
High	-0.15	2.08	-0.68	1.46
10,000 MC samples				
Low	-0.49	1.47	-0.56	1.02
Medium	-1.00	1.45	-0.64	1.02
High	-1.77	1.44	-0.97	1.03
20,000 MC samples				
Low	0.82	1.02	0.59	0.72
Medium	0.68	1.03	0.57	0.72
High	0.90	1.04	0.79	0.73

¹dBias = average of diagonals; dsMSE = square root of average of squared diagonals; off-bias = average of off-diagonal elements; offsMSE = square root of average of squared off-diagonal elements.

0.0145, 0.0102, and 0.0072 for MC sample sizes of 1,000, 5,000, 10,000, and 20,000, respectively. The average dBias and off-bias values of the exact minus estimated $\mathbf{A}_{22}$ matrices did not show equally consistent patterns, although an increase in the number MC samples often reduced the average bias. The animals were divided into 3 model reliability groups, according to the results from the GBLUP model. As already mentioned, the lower the GBLUP model reliability, the more it was overestimated by the MC-SNP-BLUP model.

Computing Times

Tables 5 and 6 show the average computing times required for the various computational tasks in MC-SNP-BLUP by different RPG proportions and MC sample sizes for Data 1 and 2. Making the MC samples was the fastest task, accounting for 5% (4 s) to 20% (71 s) of the total computing time for Data 1 and from 8% (40 min) to 17% (220 min) for Data 2. The time required to make the MME increased almost quadratically as the MME size increased, increasing as the number of MC samples increased. For example, the ratios of computing times for 20,000 MC samples over 5,000 MC samples were 3.5 and 1.7 for Data 1 and Data 2, respectively. However, the ratios of the squared MME size corresponding to these MC samples were 3.6 and 1.6 for Data 1 and Data 2, respectively. Inversion of the MME coefficient matrix was the most expensive task,

Table 5. Computing times (wall clock, in seconds) for calculating model reliability in Monte Carlo (MC)-SNP-BLUP under different MC sample sizes during various computing steps¹

Computing step	1,000	5,000	10,000	20,000
MC sampling	4	18	33	71
Making the MME	18	28	45	97
Inverting the MME	15	33	66	239
Total	78	131	220	536

¹Reported times are averages over 4 residual polygenic proportion cases. Number of genotyped animals = 19,757 and number of SNP = 11,729. MME = mixed model equations.

taking up from 25% (33 s) to 45% (239 s) of the total computing time for Data 1 with 5,000 and 20,000 MC samples, respectively—close to a cubic proportional increase in MME size. For Data 2, however, inverting the coefficient matrix of the MME was often computationally less expensive than making the MME. The time requirement for MME matrix inversion increased from 16 min (5,000 MC samples) to 206 min (60,000 MC samples), whereas making the MME took from 40 min (5,000 MC samples) to 220 min (60,000 MC samples). The total computing time increased by a factor of 4.1 for Data 1 and by 1.8 for Data 2 as the number of MC samples increased from 5,000 to 20,000.

Table 7 gives the GBLUP model computing times for the 2 data sets. For Data 1, the total computing time required for the reliability calculations was less in the MC-SNP-BLUP model (Table 5) than in GBLUP, with the exception of MC-SNP-BLUP using 20,000 MC samples. For Data 2, GBLUP was always slower than MC-SNP-BLUP (Table 6), and even the largest MC sample size took up only 27% of the total computing time required by the GBLUP model. The computing time needed to make the $\mathbf{G}$ matrix was less than 1 min for Data 1 but more than 4 h for Data 2. The proportional time required to compute the reliabilities (i.e., the total time minus the time for the other computing steps) decreased from 57% for Data 1 to 13% for Data 2.

Table 6. Computing times (wall clock, in minutes) for calculating model reliability in Monte Carlo (MC)-SNP-BLUP under different MC sample sizes during various computing steps¹

Computing step	5,000	10,000	20,000	40,000	60,000
MC sampling	11	19	38	78	141
Making the MME	40	47	69	106	220
Inverting the MME	16	21	33	74	206
Total	140	170	252	432	861

¹Reported times are averages over 3 residual polygenic proportion cases. Number of genotyped animals = 222,619 and number of SNP = 50,240. MME = mixed model equations.

DISCUSSION

Our study investigated an MC sampling method for approximating the reliabilities of an SNP-BLUP model with an RPG effect. We obtained correlations above 0.99 between the exact values from GBLUP and the corresponding approximations from the MC-SNP-BLUP model for all studied MC sample sizes, except for the lowest number of MC samples and highest RPG weight. For Data 1, a correlation of 0.97 was achieved using 1,000 MC samples (Table 1), but for Data 2, 5,000 MC samples were required to give a correlation of 0.98 (Table 2). The correlations agreed reasonably well with the maximum absolute difference and MSE, and showed high concordance even at small MC sample sizes. This was because correct ordering of the animals with respect to model reliability was achieved already with a small number of MC samples. Ideally, the intercept and slope of regression should be close to 0 and 1, respectively. In our study, these parameters deviated less from expectations with higher numbers of MC samples and lower RPG weights. The slope was always above 1, indicating inflation (overestimation of differences) in the approximation of MC-SNP-BLUP model reliability.

The model reliabilities from the SNP-BLUP model without an RPG effect were less inflated than those from the MC-SNP-BLUP model when only a small number of MC samples were used. This suggests that a very small MC sample size in MC-SNP-BLUP may give poorer reliabilities than SNP-BLUP without an RPG effect. Because the reliabilities from SNP-BLUP including the RPG effect should be lower than without the RPG effect, it is easy to test whether the MC sample size is too small.

Our results are based on full MC approaches, where MC sampling is used in calculating PEV and the diagonal of the genomic relationship matrix in the model reliability formula. We also tested the partial MC approach, where the diagonal elements of the genomic

relationship matrix are calculated exactly—that is, without MC samples. Despite using the correct denominator values $\mathbf{G}_w$ instead of $\mathbf{G}_w^*$ in [7] for MC-SNP-BLUP, we failed to observe lower deviations from the correct model reliability values by GBLUP. Quite the contrary, the correlations between true and approximate reliabilities from the full MC approach were slightly higher than those from the partial MC approach. Correspondingly, the maximum differences obtained from the full MC approach were slightly lower than those from the partial approach (results not provided). The MSE, however, did not show a consistent pattern. Nevertheless, these results suggest that the inflation can be explained by a too-low PEV when using a small MC sample size.

The model reliabilities for MC-SNP-BLUP with an RPG effect were always higher than those obtained from GBLUP, particularly with a small MC size. This upward bias seems to be a feature of our MC-SNP-BLUP approach when the number of MC samples is less than the number of genotyped animals. Note, however, that the upward bias should not be confused with an MC sampling error. A small number of MC samples may, by chance, give a model reliability that is less than the true model reliability for an animal. The MC-SNP-BLUP model derivation showed how MC sampling can be used to approximate the $\mathbf{A}_{22}$ matrix. The difference between the true and estimated $\mathbf{A}_{22}$ matrices decreased when the number of MC samples increased (Table 4), and the 3 reliability groups showed similar average values in the $\mathbf{A}_{22}$ difference matrix. As already mentioned, the upward bias in the reliabilities seems to be due to too low-values of PEV. Thus, based on the results, the bias in reliabilities was not related to the expected values of reliabilities but, instead, to sampling variation in the estimation of $\mathbf{A}_{22}$.

We presented a MC-GBLUP model approach equivalent to the MC-SNP-BLUP model where the genomic relationship matrix $\mathbf{G}_w = (1 - w)\mathbf{Z}_c\mathbf{Z}_c' + w\hat{\mathbf{A}}_{22}$, with $\hat{\mathbf{A}}_{22} = \mathbf{L}_2\hat{\mathbf{I}}_2\mathbf{L}_2'$, as described in the Materials and Methods section. The Monte Carlo estimate $\hat{\mathbf{I}}_2$ approached the identity matrix when n_{MC} increased. Assume that all animals have 1 observation and that $h^2 = 0.5$ and $\mathbf{R} = \mathbf{I}$, as in the Data 1 analysis. Then, ignoring the general mean effect, the average PEV is the mean of the diagonal of MME coefficient matrix $(\mathbf{I} + \mathbf{G}_w^{-1})^{-1}$. The MC sample size has the largest effect when the RPG proportion $w = 1$; that is, PEV is $(\mathbf{I} + \mathbf{G}_w^{-1})^{-1} = (\mathbf{I} + \mathbf{A}_{22}^{-1})^{-1}$. Because the trace of a matrix is equal to the sum of its eigenvalues, the sum of PEV is equal to the sum of the eigenvalues of $(\mathbf{I} + \mathbf{A}_{22}^{-1})^{-1}$. This sum is equal to $f = \sum \lambda_i / (\lambda_i + 1)$, where λ_i is the i th

Table 7. Computing times (wall clock) for calculating model reliability in genomic relationship matrix-based BLUP¹

Computing step	Data 1 (s)	Data 2 (min)
Making the $\mathbf{G}$ matrix	53	276
Inverting the $\mathbf{G}$ matrix	50	1,116
Making the MME	2	216
Inverting the MME	52	1,092
Total	365	3,090

¹Reported times are averages over 4 residual polygenic proportion cases for data set 1 (Data 1) and 3 residual polygenic proportion cases for data set 2 (Data 2). MME = mixed model equations. $\mathbf{G}$ is the genomic relationship matrix.

eigenvalue of the $\mathbf{A}_{22}$ matrix. Note that the expected value of λ_i is 1, if the animals are not inbred. Using the Taylor series expansion, we can note that the expected value of f depends not only on the expected values of λ_i but also on their variance. The smaller the sum f is, the larger the variance of the eigenvalues will be. Note that $\hat{\mathbf{A}}_{22} = \mathbf{L}_2 \hat{\mathbf{I}}_2 \mathbf{L}_2'$ is a sum of the squares of random vectors with degrees of freedom n_{MC} .

In their paper on the bending of covariance matrices, Hayes and Hill (1981) applied the multivariate theory of the roots of the sample matrices. In variance-covariance matrices, the larger roots are biased upward but the smaller roots downward, and pairs of equal roots are spread apart. The magnitude of the bias increases as the sample sizes (i.e., degrees of freedom) decrease, the numbers of variables increase, and the population roots become closer in value to each other. Thus, when the MC sample size increases, the variance of the eigenvalues of $\hat{\mathbf{A}}_{22} = \mathbf{L}_2 \hat{\mathbf{I}}_2 \mathbf{L}_2'$ decreases, until it reaches the variance of the eigenvalues of $\mathbf{A}_{22}$. Therefore, the trace of $(\mathbf{I} + \hat{\mathbf{A}}_{22}^{-1})^{-1}$ and, thereafter, PEV for the animals also decrease with increasing n_{MC} . When the $\hat{\mathbf{A}}_{22}$ matrix is singular (i.e., $n_{MC} < n$), the same derivation based on the eigenvalue decomposition can be performed using the selection index approach: $PEV = \hat{\mathbf{A}}_{22} - \hat{\mathbf{A}}_{22}(\mathbf{I} + \hat{\mathbf{A}}_{22})^{-1}\hat{\mathbf{A}}_{22}$.

Another way to explain the inflation is that the MC simulation increased genetic relatedness when the number of MC samples was small. When $n_{MC} < n$, the rank of $\hat{\mathbf{I}}_2$ is equal to n_{MC} , and $\hat{\mathbf{I}}_2$ has linear dependencies between columns and rows. In other words, $\hat{\mathbf{A}}_{22} = \mathbf{L}_2 \hat{\mathbf{I}}_2 \mathbf{L}_2'$ generates relationships between animals, even for animals unrelated by pedigree information. Because the formulas to calculate reliability assume the MC-SNP-BLUP model to be the true model and the record (observations) information remains the same as in the GBLUP model, MC-SNP-BLUP may consider the extra linear dependencies between animals incorrectly as increased genetic relatedness. Thus, although the average differences between the matrices $\mathbf{A}_{22}$ and $\hat{\mathbf{A}}_{22}$ were small (Table 4), $\hat{\mathbf{A}}_{22}$ may have introduced an excess of genetic relationships between genotyped animals through linear dependencies when $n_{MC} < n$. Consequently, when considering individual animal reliability, the MC-SNP-BLUP model assumes more information to be available from the other genotyped animals than in GBLUP, and the model reliability will thus be inflated.

According to the above explanations, the observed inflation will be larger with lower exact reliability, because the excess information is, in relative terms, more important for low- than for high-reliability EBV. Figure 2 illustrates how a small MC sample size can

give higher model reliabilities from the MC-SNP-BLUP model than from SNP-BLUP, which may be attributable to stronger genetic relatedness due to the MC pedigree relationships ($\hat{\mathbf{A}}_{22}$) rather than to the markers.

Genomic evaluation using the GBLUP and SNP-BLUP models has been found to give the same EBV irrespective of the approach used to center the marker genotype matrix when the model has a fixed general mean (Strandén and Christensen, 2011). However, the centering or allele coding approach applied to the marker matrix can have affect model reliabilities. When the genotypes were centered in Strandén and Christensen (2011) to have column averages of 0, the reliabilities were higher than by any other studied allele coding approach. When the RPG effect was included in the model using MC-SNP-BLUP [6], no centering was applied to the simulated values in the $\mathbf{U}_{22}$ matrix. However, the base population assumption of 0 is likely to have led to close to 0 column averages over the MC samples. Thus, the MC-simulated BV may have affected the MME like another allele coding approach, such that the mean of the model changes and the PEV matrix gives smaller-than-true error variance, which leads to an inflation of model reliability.

As shown in Tables 1 and 2, 10,000 and 60,000 MC samples gave practically the same maximum difference in model reliability (0.09) for Data 1 and Data 2, respectively, in the MC-SNP-BLUP(80) analysis. The results indicate that the goodness of the MC-SNP-BLUP model reliability approximation depends on the number of MC samples used, which, in turn, depends on the number of genotyped animals. However, the population structure may have a large effect as well. Data 1 contained genotyped animals from 1 breed, whereas Data 2 had a multiple-breed structure. Thus, the population structure in Data 1 is the one commonly used in genetic evaluations, whereas Data 2 represents a more challenging population for the presented MC-based approach.

The results suggest that the MC-SNP-BLUP model can be useful in approximating the reliability of EBV of individual animals when the RPG effect is included in the SNP-BLUP model. With a small number of MC samples, the MC-SNP-BLUP model reliabilities showed an upward bias. It would appear that the minimum MC sample size depends on the number of genotyped animals, the RPG proportion, and the population structure. Thus, further investigations are required to give an unequivocal explanation for the upward bias and to determine the sufficient number of MC samples.

A typical problem in the calculation of model reliabilities by approximation is inflation—that is, overly high reliabilities. Our MC-based approach was no exception. For Data 1, the approximate values by MC-SNP-BLUP(50) with 5,000 MC samples agreed well

with the correct values from the GBLUP model. When 5,000 MC samples were used in MC-SNP-BLUP(80), the approximated reliabilities agreed well with the correct values for reliabilities higher than 0.6. Increasing the MC sample size to 20,000 improved the reliability approximation considerably, especially for the lowest reliability values. This would suggest that the correlation between GBLUP- and MC-SNP-BLUP-approximated reliabilities is high even with a small number of MC samples, although the approximate reliabilities can still be inflated. Increasing the number of MC samples lowered this inflation, which was higher with lower reliability. In addition, the larger the RPG weight, the more MC samples were needed to reduce inflation. Thus, the MC approach may require an impractically large number of samples when using a completely polygenic model (i.e., a pedigree-based animal model). With an RPG proportion of 20% or less, about 10,000 MC samples can be sufficient to give reliabilities close enough to the true ones for practical purposes when analyzing data from a single breed population.

Our results are consistent with those reported in a simulation-based study by Hickey et al. (2009), who found that the accuracy of PEV approximation is affected by the level of the true PEV and the MC sample size. The authors compared 10 sampling-based formulations for PEV approximation. They found that the different formulations converged to the exact PEV at different rates but that the difference between the approaches tended to become negligible with a higher number of MC samples. Gengler and Misztal (1996) suggested a correction method to lessen inflation in the approximation of model reliability. They applied multiple diagonalization to a multiple-trait model with missing data, and reported that underestimation only occurred for lower-reliability sires, indicating that the approximation did not affect sires with a large number of progeny.

Tables 5 and 6 illustrate the increase in total computing time with increasing MC sampling size. Guo and Thompson (1994) pointed out the need to make a reasonable compromise in an MC approach between 2 conflicting objectives: between high approximation accuracy and short computing time. In the analysis of Data 1 (Table 1), increasing the MC sample size from 10,000 to 20,000 had no practical importance in terms of the MSE, especially with RPG weights of 20% or less. In case of a high RPG weight, the inflation of model reliability could be remedied, for example, by increasing the number of MC samples.

Even though parallel computing is easily applied to any MC approach, we did not consider it in our sampling procedure, because the computing time required to generate the MC samples was not significant. The

short computing time was due to the simple computing steps in the breeding value simulation and our decision to limit the pedigree used in the simulation to the ancestors of genotyped animals. For example, the pedigree used in Data 2 was pruned from the total 13.35 million to 850,407 before the MC sampling step. For the other computing steps, such as making and inverting the MME matrix, parallel computing was efficiently employed using the library subroutines.

Data 1 included fewer than 20,000 genotyped animals, and the number of MC samples was at most 20,000. In practice, we could have calculated SNP-BLUP-based reliabilities without MC sampling by using the Cholesky-decomposed matrix of the relationship matrix $\mathbf{A}_{22}$, as explained in the Materials and Methods section, but this would have taken up more computing time than GBLUP. For Data 2, however, the Cholesky approach might be computationally faster than the GBLUP approach. Calculation of the $\mathbf{A}_{22}$ matrix took approximately 12 h using Data 2, whereas MC-SNP-BLUP took from 2 h (5,000 MC samples) to 8 h (40,000 MC samples).

Inverting the coefficient matrix of the MME was often computationally less expensive than making the MME in the analysis of Data 2 but not of Data 1. There may be several reasons for this. First, it might be due to the large number of genotyped animals ($n > m + n_{MC}$) in Data 2. The computing time needed for MME matrix inversion increases in cubic of the size of the matrix $m + n_{MC}$, and making the MME requires the product of 2 rectangular matrices, demanding about $n \times (m + n_{MC})^2$ floating point operations. Second, we applied the single-precision subroutine SGEMM to make the MME, and double-precision subroutines DPOTRF and DPOTRI were used for the inversion. Double-precision subroutines may be computationally better optimized than single-precision computations in a 64-bit operating system computer. Third, certain differences in these 2 steps may have limited the parallelizable part of the program in the algorithms used in the BLAS subroutines when the MME size increased. Finally, the matrix sizes in the analysis of Data 2 were quite large, such that memory access to the matrices may have slowed down the computations. Even though the MME coefficient matrix was stored in double precision, it still required less memory than the marker matrix stored in single precision in all MC sample sizes except the largest (60,000). Hence, the memory requirements increased at different paces. With 5,000 (60,000) samples, about 24 (92) GB of memory was needed to store the MME coefficient matrix, but 46 (91) GB was needed to store the marker matrix.

Only a few results have been published demonstrating a comparison between approximated and theoretical

exact genomic reliabilities. Edel et al. (2019) proposed an approximation of model reliability for single-step GBLUP based on an equivalent reformulation of single-step GBLUP as a marker-based model as well as calculation of the PEV of the marker effects. To reduce the complexity of the matrix operations, the authors limited their examination only to those non-genotyped ancestors of genotyped animals that contributed relevant information for PEV calculation of the marker effects. In our study, MC sampling was limited to genotyped animals and their ancestors, even though the full pedigree was available.

Sargolzaei et al. (2014) developed 2 methods for SNP-BLUP model reliability estimation. They used GBLUP and ridge-regression methods (subset of SNP-BLUP) under different numbers of animals and SNP markers. The authors reported correlations from 0.858 to 0.896 between approximated and theoretical reliability using their first method based on a subset of training individuals. For their second method, which was based on a subset of equidistant SNP, the corresponding correlations were from 0.839 to 0.999. According to Karaman et al. (2016), none of the proposed methods for accuracy estimation is universally best, since the performance of the approximations is affected by genetic architecture, heritability of the trait, and size of the reference population. All this has relevance in the MC-SNP-BLUP approach as well.

CONCLUSIONS

The MC sampling approach allowed reduction of the size of the MME, which largely determines the computational cost in reliability calculations in data analysis involving a large number of genotyped animals. The computing time required to invert the MME coefficient matrix was roughly in proportion $\left(\frac{n}{m + n_{MC}}\right)^3$ for the GBLUP over the MC-SNP-BLUP model. Our results indicate that MC-SNP-BLUP approximation gives a high correlation with the exact GBLUP model reliability with as few as 5,000 MC samples. However, more MC samples were needed to give a small maximum absolute difference or MSE when the RPG effect had a high weight. The proposed method proved to be particularly useful when the number of genotyped animals is larger than the number of SNP markers plus MC samples. In addition, it is much faster than the GBLUP approach. These promising results suggest that the MC approach may prove useful in the calculation of genomic reliability from the SNP-BLUP model with an RPG effect included.

ACKNOWLEDGMENTS

The authors acknowledge Viking Genetics (Randers, Denmark) and Nordic Cattle Genetic Evaluation (Aarhus, Denmark) for providing the Finnish dairy cattle genotype data and the Irish Cattle Breeding Federation (ICBF; Bandon, Ireland) for providing the beef cattle data. The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

REFERENCES

- Anderson, E., Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen. 1999. LAPACK users' guide, 3rd ed. SIAM, Philadelphia, PA.
- Edel, C., E. C. G. Pimentel, M. Erbe, R. Emmerling, and K.-U. Götz. 2019. Short communication: Calculating analytical reliabilities for single-step predictions. *J. Dairy Sci.* 102:3259–3265. <https://doi.org/10.3168/jds.2018-15707>.
- Gengler, N., and I. Misztal. 1996. Approximation of reliability for multiple-trait animal models with missing data by canonical transformation. *J. Dairy Sci.* 79:317–328. [https://doi.org/10.3168/jds.S0022-0302\(96\)76366-X](https://doi.org/10.3168/jds.S0022-0302(96)76366-X).
- Guarini, A. R., D. A. L. Lourenco, L. F. Brito, M. Sargolzaei, C. F. Baes, F. Miglior, I. Misztal, and F. S. Schenkel. 2018. Comparison of genomic predictions for lowly heritable traits using multi-step and single-step genomic best linear unbiased predictor in Holstein cattle. *J. Dairy Sci.* 101:8076–8086. <https://doi.org/10.3168/jds.2017-14193>.
- Guo, S. W., and E. A. Thompson. 1994. Monte Carlo estimation for large complex pedigrees. *Biometrics* 50:417–432. <https://doi.org/10.2307/2533385>.
- Harris, B., and D. Johnson. 1998. Approximate reliability of genetic evaluations under an animal model. *J. Dairy Sci.* 81:2723–2728. [https://doi.org/10.3168/jds.S0022-0302\(98\)75829-1](https://doi.org/10.3168/jds.S0022-0302(98)75829-1).
- Hayes, J. F., and W. G. Hill. 1981. Modification of estimates of parameters in the construction of genetic selection indices ('Bending'). *Biometrics* 37:483–493. <https://doi.org/10.2307/2530561>.
- Henderson, C. R. 1984. Applications of linear models in animal breeding. University of Guelph Press, Guelph, Canada.
- Hickey, J. M., R. F. Veerkamp, M. P. L. Calus, H. A. Mulder, and R. Thompson. 2009. Estimation of prediction error variances via Monte Carlo sampling methods using different formulations of the prediction error variance. *Genet. Sel. Evol.* 41:23. <https://doi.org/10.1186/1297-9686-41-23>.
- Intel. 2014. Intel Math Kernel Library reference manual. Accessed Jul. 11, 2017. <https://software.intel.com/en-us/mkl-reference-manual-for-fortran>.
- Jamrozik, J., L. R. Schaeffer, and G. B. Jansen. 2000. Approximate accuracies of prediction from random regression models. *Livest. Prod. Sci.* 66:85–92. [https://doi.org/10.1016/S0301-6226\(00\)00158-5](https://doi.org/10.1016/S0301-6226(00)00158-5).
- Karaman, E., H. Cheng, M. Z. Firat, D. J. Garrick, and R. L. Fernando. 2016. An upper bound for accuracy of prediction using GBLUP. *PLoS One* 11:e0161054. <https://doi.org/10.1371/journal.pone.0161054>.
- Liu, Z., M. E. Goddard, B. J. Hayes, F. Reinhardt, and R. Reents. 2016. Technical note: Equivalent genomic models with a residual polygenic effect. *J. Dairy Sci.* 99:2016–2025. <https://doi.org/10.3168/jds.2015-10394>.
- Liu, Z., F. R. Seefried, F. Reinhardt, S. Rensing, G. Thaller, and R. Reents. 2011. Impacts of both reference population size and inclusion of a residual polygenic effect on the accuracy of genomic

- prediction. *Genet. Sel. Evol.* 43:19. <https://doi.org/10.1186/1297-9686-43-19>.
- Mäntysaari, E. A., R. D. Evans, and I. Strandén. 2017. Single-step genomic evaluation for many more genotyped animals. *J. Anim. Sci.* 95:4728–4737. <https://doi.org/10.2527/jas2017.1912>.
- McPeck, M. S., X. Wu, and C. Ober. 2004. Best linear unbiased allele frequency estimation in complex pedigrees. *Biometrics* 60:359–367. <https://doi.org/10.1111/j.0006-341X.2004.00180.x>.
- Misztal, I., S. Tsuruta, I. Aguilar, A. Legarra, P. M. VanRaden, and T. J. Lawlor. 2013. Methods to approximate reliabilities in single-step genomic evaluation. *J. Dairy Sci.* 96:647–654. <https://doi.org/10.3168/jds.2012-5656>.
- Misztal, I., and G. R. Wiggins. 1988. Approximation of prediction error variance in large-scale animal models. *J. Dairy Sci.* 71:27–32. [https://doi.org/10.1016/S0022-0302\(88\)79976-2](https://doi.org/10.1016/S0022-0302(88)79976-2).
- Sargolzaei, M., L. R. Schaeffer, J. P. Chesnais, G. Kistemaker, G. R. Wiggins, and F. S. Schenkel. 2014. Approximation of reliability of direct genomic breeding values. Page 3 in *Proc. 10th World Cong. Genet. Appl. Livest. Prod.*, Vancouver, Canada.
- Schaeffer, L. R. 2006. Strategy for applying genome-wide selection in dairy cattle. *J. Anim. Breed. Genet.* 123:218–223. <https://doi.org/10.1111/j.1439-0388.2006.00595.x>.
- Strandén, I., and O. F. Christensen. 2011. Allele coding in genomic evaluation. *Genet. Sel. Evol.* 43:25. <https://doi.org/10.1186/1297-9686-43-25>.
- Strandén, I., and D. J. Garrick. 2009. Technical note: Derivation of equivalent computing algorithms for genomic predictions and reliabilities of animal merit. *J. Dairy Sci.* 92: 2971–2975.
- Strandén, I., K. Matilainen, G. P. Aamand, and E. A. Mäntysaari. 2017. Solving efficiently large single-step genomic best linear unbiased prediction models. *J. Anim. Breed. Genet.* 134:264–274.
- Tier, B., and K. Meyer. 2004. Approximating prediction error covariances among additive genetic effects within animals in multiple-trait and random regression models. *J. Anim. Breed. Genet.* 121:77–89. <https://doi.org/10.1111/j.1439-0388.2003.00444.x>.
- VanRaden, P. M. 2008. Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91:4414–4423. <https://doi.org/10.3168/jds.2007-0980>.

ORCIDS

- H. Ben Zaabza <https://orcid.org/0000-0003-4648-3813>
 E. A. Mäntysaari <https://orcid.org/0000-0003-0044-8473>
 I. Strandén <https://orcid.org/0000-0003-0161-2618>

APPENDIX

The residual polygenic (RPG) effect is often considered as a separate effect in SNP-BLUP (Liu et al., 2016):

$$\mathbf{y} = \mathbf{1}_n\mu + \mathbf{Z}_c\mathbf{g}_0 + \mathbf{a} + \mathbf{e}, \quad [\text{A1}]$$

where $\mathbf{g}_0 \sim N(\mathbf{0}, \mathbf{I}\sigma_u^2(1-w))$, $\mathbf{a} \sim N(\mathbf{0}, \mathbf{A}_{22}\sigma_u^2w)$, and w is the proportion of residual polygenic variance. Model [3] can be similarly expressed using separate effects:

$$\mathbf{y} = \mathbf{1}_n\mu + \mathbf{Z}_g\mathbf{g}_z + \mathbf{Z}_a\mathbf{g}_L + \mathbf{e}, \quad [\text{A2}]$$

where $\mathbf{Z}_g = \sqrt{1-w}\mathbf{Z}_c$, $\mathbf{Z}_a = \sqrt{w}\mathbf{L}$, $\mathbf{g}_z \sim N(\mathbf{0}, \mathbf{I}\sigma_u^2)$, and $\mathbf{g}_L \sim N(\mathbf{0}, \mathbf{I}\sigma_u^2)$.

Models [A1] and [A2] are linearly equivalent, and the estimated breeding values from these models are equal; that is, $\hat{\mathbf{u}} = \mathbf{Z}_c\hat{\mathbf{g}}_0 + \hat{\mathbf{a}} = \mathbf{Z}_g\hat{\mathbf{g}}_z + \mathbf{Z}_a\hat{\mathbf{g}}_L$. However, the models have different model effects, and thus, the MME coefficient matrices are different as well. A linear transformation of estimates is needed to calculate estimates of the other model: $\hat{\mathbf{g}}_0 = \sqrt{1-w}\hat{\mathbf{g}}_z$ and $\hat{\mathbf{a}} = \sqrt{w}\mathbf{L}\hat{\mathbf{g}}_L$. When calculating the PEV, the inverse matrix elements need to be pre- and post-multiplied by appropriate transformation matrices to give the same PEV. In the case of model [A1], the multiplier matrix is $[\mathbf{Z}_c \quad \mathbf{I}]$, but in model [A2] it is $\mathbf{Z} = [\mathbf{Z}_g \quad \mathbf{Z}_a]$.