

Effects of ignoring inbreeding in model-based accuracy for BLUP and SSGBLUP

Ignacio Aguilar¹  | Eduardo N. Fernandez² | Agustin Blasco³  | Olga Ravagnolo¹ | Andres Legarra⁴ 

¹Instituto Nacional de Investigación Agropecuaria (INIA), Montevideo, Uruguay

²Cátedra de Mejora y Conservación de Recursos Genéticos e Instituto de Investigación sobre Producción Agropecuaria, Ambiente y Salud, Facultad de Ciencias Agrarias, UNLZ, Buenos Aires, Argentina

³Institute for Animal Science and Technology, Universitat Politècnica de València, València, Spain

⁴UMR GenPhySE, INRA Toulouse, Castanet Tolosan, France

Correspondence

Ignacio Aguilar, Instituto Nacional de Investigación Agropecuaria (INIA), Montevideo, Uruguay.
Email: iaguilar@inia.org.uy

Funding information

FEDER; INRA; Universidad Nacional de Lomas de Zamora; European Unions' Horizon 2020 Research & Innovation Programme, Grant/Award Number: N°72787

Abstract

Model-based accuracy, defined as the theoretical correlation between true and estimated breeding value, can be obtained for each individual as a function of its prediction error variance (PEV) and inbreeding coefficient F , in BLUP, GBLUP and SSGBLUP genetic evaluations. However, for computational convenience, inbreeding is often ignored in two places. First, in the computation of reliability = $1 - \text{PEV} / (1 + F)$. Second, in the set-up, using Henderson's rules, of the inverse of the pedigree-based relationship matrix $\mathbf{A}$. Both approximations have an effect in the computation of model-based accuracy and result in wrong values. In this work, first we present a reminder of the theory and extend it to SSGBLUP. Second, we quantify the error of ignoring inbreeding with real data in three scenarios: BLUP evaluation and SSGBLUP in Uruguayan dairy cattle, and BLUP evaluations in a line of rabbit closed for >40 generations with steady increase of inbreeding up to an average of 0.30. We show that ignoring inbreeding in the set-up of the $\mathbf{A}$ -inverse is equivalent to assume that non-inbred animals are actually inbred. This results in an increase of apparent PEV that is negligible for dairy cattle but considerable for rabbit. Ignoring inbreeding in reliability = $1 - \text{PEV} / (1 + F)$ leads to underestimation of reliability for BLUP evaluations, and this underestimation is very large for rabbit. For SSGBLUP in dairy cattle, it leads to both underestimation and overestimation of reliability, both for genotyped and non-genotyped animals. We strongly recommend to include inbreeding both in the set-up of $\mathbf{A}$ -inverse and in the computation of reliability from PEVs.

1 | INTRODUCTION

The purpose of genetic evaluations is to predict, with some uncertainty, the breeding value of animals. Model-based accuracy (or its square, reliability) is used as a measure of risk in choosing parents of the next generation, and it condenses in a single number the uncertainty related to its breeding value. The measure of accuracy from BLUP theory is regularly used and reported in breeding evaluations (Miształ & Wiggans, 1988). The advent of genomic selection needs

methods to ascertain individual accuracies (Edel, Pimentel, Erbe, Emmerling, & Götz, 2019), and the increasing selection for complex traits (e.g. feed efficiency or methane emissions (Pryce et al., 2015)) needs measures of individuals accuracies in small to medium size data sets. For historical reasons of simplicity, inbreeding is often ignored in computations of accuracy from prediction error variance (PEV). Furthermore, often, pedigree inbreeding is also ignored in the computation of the inverse relationship matrix ($\mathbf{A}$ -inverse) using Henderson's (1976) rules, which results in

an approximated BLUP and a further level of approximation for the computation of model-based individual accuracies. As a result, even for small data sets where exact computations of PEVs are feasible, reported accuracies are often approximate.

The objective of this paper is to present correct ways of considering inbreeding in the computation of accuracy and compare with results from ignoring inbreeding, in a single-step GBLUP and traditional pedigree-based BLUP context, using two data sets: a dairy cattle data set with genotyped individuals (using BLUP and SSGBLUP) and a closed rabbit selection line with a large number of generations and steady increase of inbreeding (using BLUP).

2 | MATERIAL AND METHODS

2.1 | Theory

Accuracy (*acc*) is a model-based, individual measure of precision of the estimated breeding value (EBV). It is typically defined (e.g. Van Vleck, 1993a, 1993b) as the correlation, on repeated conceptual sampling, of the true breeding value of one individual (*u*) with its estimate ($\hat{u}$). For BLUP models, and assuming that the model is true, Henderson (1975, 1982, 1984) showed that $Cov(u, \hat{u}) = Var(\hat{u})$. Also, $Cov(u, \hat{u}) = Var(\hat{u}) = Var(u) - PEV$ where $PEV = Var(\hat{u} - u)$ and can be numerically obtained from the corresponding element on the inverse of the mixed model equations (either by sparse inversion or Monte Carlo methods). We will work with reliability (*rel*), the square of accuracy. According to the definition of a correlation:

$$rel = acc^2 = \frac{Cov(u, \hat{u})^2}{Var(\hat{u}) Var(u)} = \frac{Var(u) - PEV}{Var(u)} = 1 - \frac{PEV}{Var(u)}$$

It is often, but wrongly, assumed, even in textbooks (e.g. Bijma, 2012; Misztal & Wiggans, 1988; Mrode & Thompson, 2005) that $Var(u) = \sigma_u^2$, the genetic variance, whereas $Var(u)$ depends on each individual. An example is inbred individuals who tend to be more extreme. More precisely, for individual *i*, $Var(u_i) = A_{ii}\sigma_u^2$ if pedigree is used (classical BLUP), $Var(u_i) = G_{ii}\sigma_u^2$ (GBLUP) and $Var(u_i) = H_{ii}\sigma_u^2$ (SSGBLUP), where A_{ii} , G_{ii} and H_{ii} represent measures of self-relationships, that is, they are equal to $(1 + F_i)$ where F_i is the inbreeding coefficient. In BLUP and SSGBLUP, pedigree- or genomic-based measures of inbreeding are used. Note that, in selected populations, there is a reduction in the genetic variance due to selection and Bulmer effect (Bijma, 2012) but this is never considered in reporting model-based reliabilities and we will not do so. With this proviso, the correct expression for the computation of model-based reliability, defined as the squared correlation between *u* and $\hat{u}$ is.

$$rel_i = 1 - \frac{PEV_i}{(1 + F_i) \sigma_u^2}$$

where F_i is a measure of inbreeding. This derivation is implicit in Henderson's results but, to our knowledge, it was first explicitly published by (Van Vleck, 1993a, 1993b) although it has certainly been used before (Tier, Schneberger, Hammond, & Fuchs, 1991). In conventional pedigree evaluations, ignoring the term F_i in the denominator results in an underestimation of reliability for inbred animals in BLUP evaluations as F_i can only be positive.

However, for genomic evaluations by GBLUP, there is a higher variability of inbreeding, and even "negative" inbreeding is possible ($G_{ii} < 1$), which indicate higher heterozygosity than population average (Legarra, Lourenco, & Vitezica, 2018). As for SSGBLUP, genotyped animals have self-relationships G_{ii} , whereas non-genotyped animals get an "improved" estimation of inbreeding in matrix **H** of the form $H_{ii} - 1$, where $H_{ii} = A_{ii} + a_{i,2}A_{22}^{-1}(G - A_{22})A_{22}^{-1}a_{2,i}$ (Legarra, Aguilar, & Misztal, 2009) where $a_{i,2} = a_{2,i}$ is the part of the *i*-th column of **A** that corresponds to the genotyped individuals. Xiang, Christensen, and Legarra (2017) used $rel = 1 - \frac{PEV}{(1 + F_i)\sigma_u^2}$ with $F_i = H_{ii} - 1$ obtained from **H** constructed

for crossbreds using metafounders, and they argued that ignoring this correction in this particular model with metafounders resulted in spurious results.

Although there are several reports for small data sets of model-based reliabilities using SSGBLUP, none uses the diagonal of **H**. The extent of the effect of ignoring the diagonal of **H** in the computation of model-based reliabilities is unknown, and it can (wrongly) either increase or decrease the reliabilities.

A second problem in computation of reliabilities is that frequently, Henderson's (1976) simplified rules are used for computation of **A**⁻¹ (Golden, Brinks, & Bourdon, 1991; Mehrabani-Yeganeh, Gibson, & Schaeffer, 2000; Van Vleck, 1993b), resulting in the use of an approximate matrix, that we call **A**^{*-1}, instead. However, inversion of **A**^{*-1} to obtain a relationship matrix of **A**^{*} does not yield the expected results. Quoting Golden et al. (1991) "Using an approximate **A**-inverse is not the same as assuming an individual has covariance to each of his parents of 0.5, and a diagonal element in **A** of 1". The use of either **A**^{*-1} implicitly assumes an incorrect matrix of "true" relationships, that is $(\mathbf{A}^{*-1})^{-1} = \mathbf{A}^*$, when we know that $\mathbf{A}^* \neq \mathbf{A}$. This results in incoherencies in the genetic model, different MME and therefore different PEV.

Based on our experience and some studies (Mehrabani-Yeganeh et al., 2000), in BLUP, using **A**^{*-1} instead of **A**⁻¹ seems to change very little the results of genetic evaluations (EBVs) and is the default in the BLUPF90 family of programs (Misztal et al., 2002), in MixBLUP (Ten Napel et al., 2017) and in PEST (Groeneveld, Kovac, & Wang,

1990). Software that correctly computes $\mathbf{A}^{-1}$ by default includes Wombat (Meyer, 2007) and ASReml (Gilmour, Gogel, Cullis, Thompson, & Butler, 2009).

Accordingly, the default in SSGBLUP computations was to compute.

$$\mathbf{H}^{*-1} = \mathbf{A}^{*-1} + \begin{pmatrix} 0 & 0 \\ 0 & \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} \end{pmatrix}$$

where $\mathbf{A}^{*-1}$ was computed *without* inbreeding in Henderson's rules but $\mathbf{A}_{22}^{-1}$ was computed inverting $\mathbf{A}_{22}$, that was computed *with* inbreeding in all cases. Again, the use of $\mathbf{H}^{*-1}$ (formed as

$$\mathbf{H}^{*-1} = \mathbf{A}^{*-1} + \begin{pmatrix} 0 & 0 \\ 0 & \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} \end{pmatrix})$$
 implicitly assumes an in-

correct matrix of "true" relationships, that is $(\mathbf{H}^{*-1})^{-1} = \mathbf{H}^* \neq \mathbf{H}$. This resulted in strong convergence problems for large data sets which included inbred animals (Matilainen, Strandén, Aamand, & Mäntysaari, 2018; Strandén, Matilainen, Aamand, & Mäntysaari, 2017). It is unclear how ignoring inbreeding when computing $\mathbf{A}^{*-1}$ or $\mathbf{H}^{*-1}$, affects the calculation of PEV (computed by inversion) and therefore accuracies.

Putting all together, ignoring inbreeding affects computation of individual accuracies in two different places: first, if the wrong denominator is used (in $rel_i = 1 - \frac{PEV_i}{(1+F_i)\sigma_u^2}$) and sec-

ond, if PEV is wrongly computed using either $\mathbf{H}^{*-1}$ or $\mathbf{A}^{*-1}$. Van Vleck (1993b) already pointed out both problems; however, there are no published examples in livestock data sets of the consequences of ignoring F_i in the reliability, and there is no description of this problem specifically in GBLUP or SSGBLUP evaluations.

2.2 | Data sets

The dairy cattle data set involves milk yield from the Uruguayan Holstein national genetic evaluation. Dairy cattle records were provided by the Uruguayan National Dairy Herd Improvement (MU—Mejoramiento y Control Lechero Uruguayo), and genealogical information was provided by the National Herdbook (ARU—Asociación Rural del Uruguay). Data consisted of 305 DIM milk yields from 925,821 records of 377,612 cows from lactation 1 to 5 since 1990 to 2018. Pedigree file was created using 3 generations of ancestors backwards from either phenotyped or genotyped animals and consisted of 511,576 animals. Figure 1 shows the average level of inbreeding for cows with phenotypes.

The model used was a single-trait model with repeated records for milk yield which included fixed effects of herd-year-season, lactation-age and random effects of permanent environment and animal additive genetic effect. This is not

the official evaluation which is a random regression model. Genetic groups were not used in this particular genetic evaluation, given that the purpose was to compute reliabilities (which are not well defined under a fixed genetic group model). Heritability of the trait was 0.20.

Genotypes of 5,072 animals were available for analyses, including 2,246 and 2,826 genotypes for bulls and cows, respectively. Genotypes were obtained from the International Dairy and Beef SNP chip IDBv3 (Mullen et al., 2013) by Weatherby's Scientific Ltd. (Johnstown, Naas, Co.) and from the Illumina Bovine SNP50k Beadchip (Illumina Inc.). The 39,288 SNP markers in common from both panels were available for analyses. SNP with minor allele frequency greater of 0.05 from autosome chromosomes and samples with a call rate >0.90 were used. Missing SNP was imputed using FImpute (Sargolzaei, Chesnais, & Schenkel, 2014) with the UMD 3.1 assembly. Distribution of animals with phenotypes and genotypes is presented in Table 1. Genetic evaluation was carried out using BLUPF90 (Misztal et al., 2002), and for each scenario, prediction error variances (PEV) were obtained from the inverse of MME using FSPAK-YAMS (Masuda, Aguilar, Tsuruta, & Misztal, 2015). Therefore, different PEVs were obtained, according to each assumed relationship matrix.

The second data set is from the meat rabbit line A from the Universitat Politècnica de València, Spain. This is a closed selected line with complete records and pedigree since its foundation. The breeding objective of this maternal line is litter size ($h^2 = 0.10$); further description can be found in Fernandez, Sanchez, Martinez, Legarra, and Baselga (2017). The data set for this work included pedigree (40 generations totalling 5,668 animals) and 15,671 records (litter size) of animals born from 1980 to 2009. Inbreeding increases steadily at a rate of 0.0084 per generation, so that animals in 2009 have an average inbreeding coefficient of roughly 0.30. This data set is included to verify the effects of high accumulated inbreeding on the computations of accuracy. The linear model for genetic evaluation includes the fixed effects of inbreeding depression (with F as covariate), parity-lactation status of the female and random year-season, additive genetic and permanent effects; see reference above for details. The high cumulated inbreeding provokes that not fitting F as covariate in this data set yields biased estimates of genetic trend (Fernandez et al., 2017).

2.3 | Models

For the dairy cattle data, EBVs and accuracies were calculated using BLUP (only pedigree information) and SSGBLUP (pedigree and marker information). Two scenarios for computation of the inverse of the relationship matrix for the MME were tested:

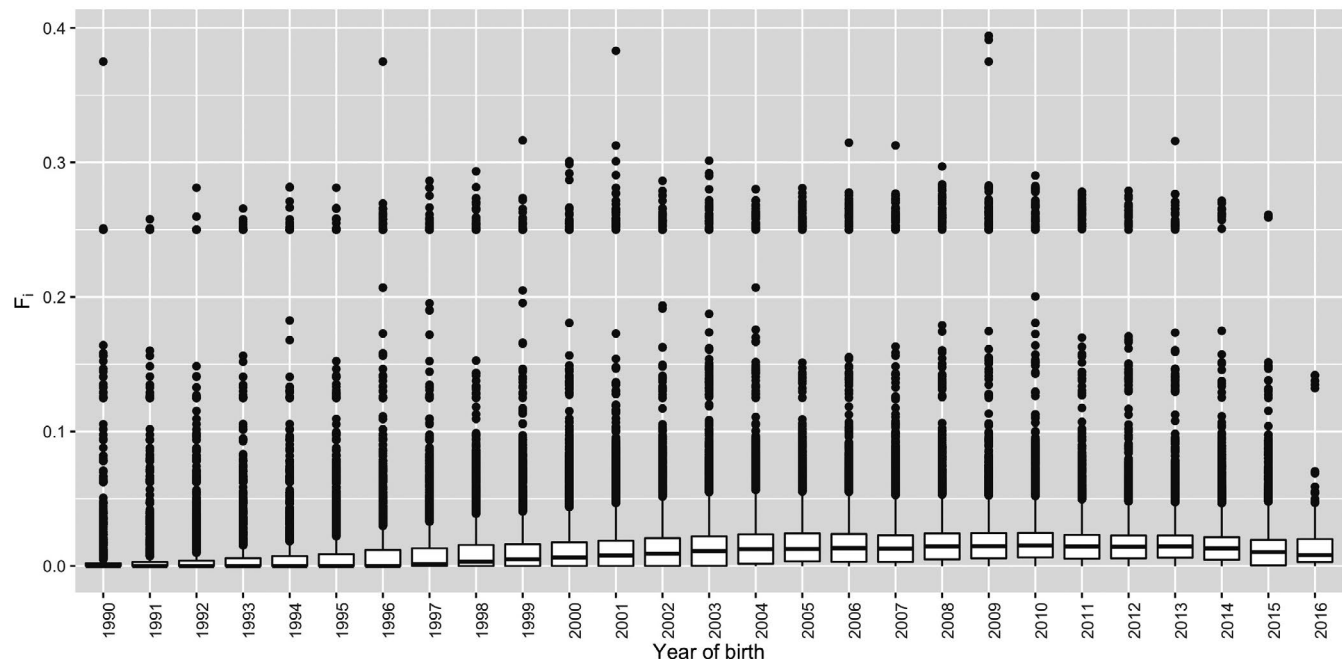


FIGURE 1 Pedigree inbreeding for Uruguayan dairy cattle

TABLE 1 Distribution of animals across categories in the Uruguayan Holstein data set

Category	Total	Genotyped	Ungenotyped
Cows	377,612	2,464	375,148
Sires	8,223	606	7,617
Other ancestors	124,089	350	123,739
Candidates to selection	1,652	1,652	0

- Correct A-inverse $\mathbf{A}^{-1}$ (alternatively, H-inverse: $\mathbf{H}^{-1}$) using Henderson's rules with inbreeding ("add_an_upginb" type of random effect in BLUPF90). This results in correct *PEV*.
- Incorrect A-inverse $\mathbf{A}^{*-1}$ (alternatively, H-inverse: $\mathbf{H}^{*-1}$) using Henderson's rules ignoring inbreeding ("add_animal" type of random effect in BLUPF90). This results in incorrect *PEV* that we will call *PEV**.

Matrix $\mathbf{G}$ involved in $\mathbf{H}$ was computed as described for $\mathbf{G}_w$ in Christensen, Madsen, Nielsen, Ostensen, and Su (2012) (Christensen et al., 2012), that is using VanRaden's (2008) method 1, adjusting for inbreeding and average relationship and blending with 0.05 of matrix $\mathbf{A}_{22}$.

In addition, BLUPF90 was modified to calculate reliabilities. Two computations of reliabilities were done:

- Correct, accounting for inbreeding in the denominator $rel = 1 - \frac{PEV}{(1+F_i)\sigma_u^2}$, where F_i may be from $\mathbf{A}$ or from $\mathbf{H}$.
- Incorrect, *not* accounting for inbreeding in the denominator, $rel^* = 1 - \frac{PEV^*}{\sigma_u^2}$

For rabbit, since there was no marker information, only $\mathbf{A}^{-1}$ or $\mathbf{A}^{*-1}$ was computed for BLUP, with resulting *PEV* or *PEV**, and accuracies were computed using as denominator $(1+F_i)\sigma_u^2$ or σ_u^2 . In particular, we computed the two reliabilities above (*rel* and *rel**) and also $rel^\# = 1 - \frac{PEV^*}{(1+F_i)\sigma_u^2}$ which uses

incorrect *PEV* and correct denominator.

3 | RESULTS AND DISCUSSION

In both data sets, ignoring inbreeding in the set-up of $\mathbf{A}^{-1}$ or $\mathbf{H}^{-1}$ resulted in negligible changes in EBVs (correlations higher than 0.99).

3.1 | Dairy cattle: Effect of ignoring inbreeding in the relationship matrices

When ignoring inbreeding in the set-up of $\mathbf{A}^{-1}$, the (incorrectly built) inverses of the numerator relationships are $\mathbf{A}^{*-1}$ and $\mathbf{H}^{*-1}$. Use of these approximate matrices implies that the (wrongly) assumed relationship matrices are $\mathbf{A}^* = (\mathbf{A}^{*-1})^{-1}$ and $\mathbf{H}^* = (\mathbf{H}^{*-1})^{-1}$. To quantify the extent of the error of using $\mathbf{A}^{*-1}$ and $\mathbf{H}^{*-1}$ instead of $\mathbf{A}^{-1}$ and $\mathbf{H}^{-1}$, we plot the diagonals of $\mathbf{A}$ and $\mathbf{H}$ (which contain correct estimates of inbreeding) versus the diagonals of $\mathbf{A}^*$ and $\mathbf{H}^*$ (which contain incorrect estimates) (Figure 2). It can be seen that ignoring inbreeding in Henderson's rules results in higher self-relationships, thus implicitly assuming animals to be inbred when they in truth

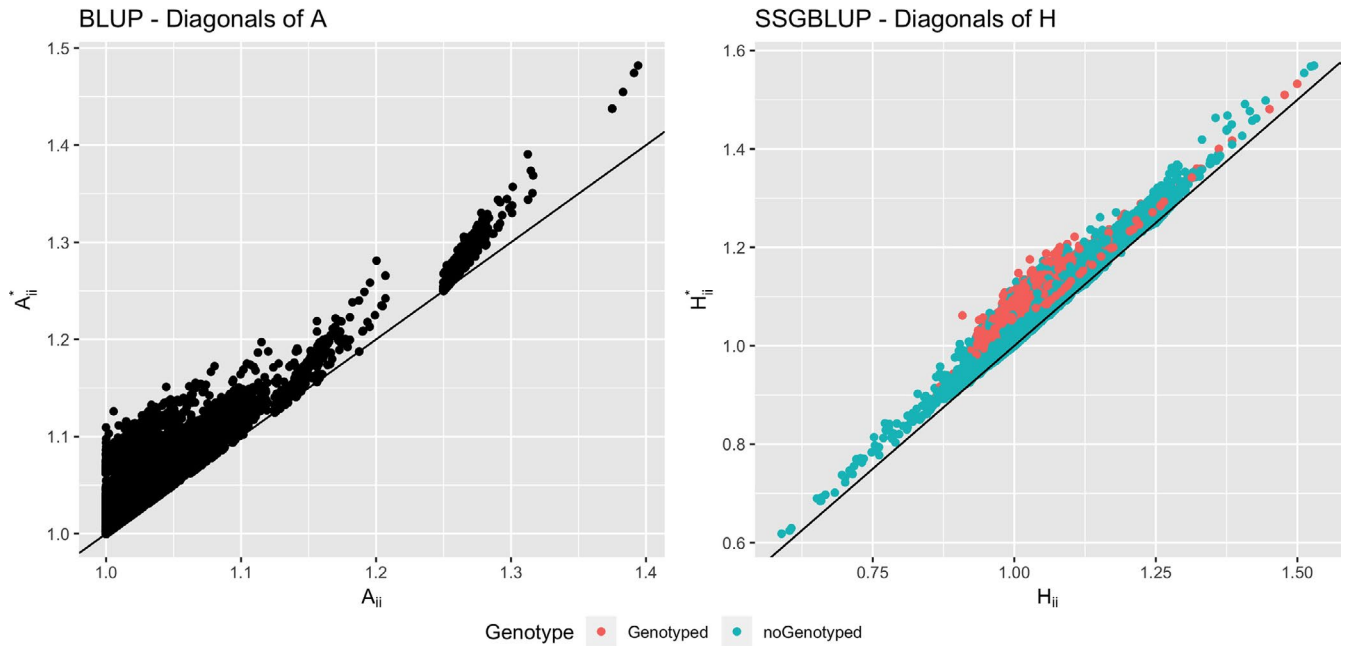


FIGURE 2 True self-relationships versus. implied self-relationships (using Henderson's rules and ignoring inbreeding) for Uruguayan dairy cattle

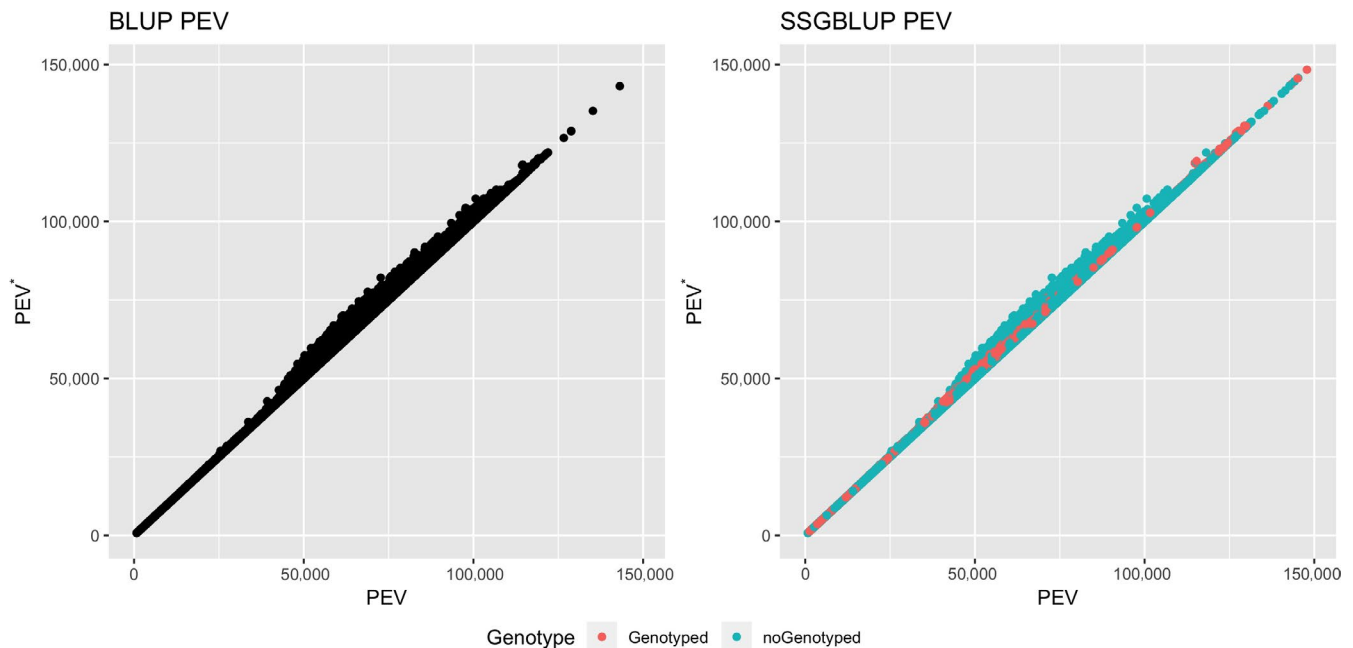


FIGURE 3 Correct PEV versus incorrect PEV* (using Henderson's rules and ignoring inbreeding) for Uruguayan dairy cattle

are non-inbred. This is in agreement with Golden et al. (1991). Furthermore, inbred animals are considered more inbred than they truly are. Correlations $r(\text{diag}(A), \text{diag}(A^*)) = 0.94$ and $r(\text{diag}(H), \text{diag}(H^*)) = 0.87$ indicate quite a strong modification in the assumptions of genetic composition.

For SSGBLUP, ignoring inbreeding in the set-up of A^{-1} , not only generates wrong diagonal elements for A^* , but also modifies diagonal elements for genotyped individuals in H^* . In addition, the block in $(H^{*-1})^{-1}$ corresponding to genotyped

individuals is no longer G (the genomic relationship matrix), but something different.

If an animal is (implicitly) assumed to be inbred by using simplified A^{*-1} or H^{*-1} , its a priori uncertainty is higher, and this in turn results in an increase of the PEV computed from the inverse of the MME (that we called PEV^*). Even though this is the case, the effect of ignoring inbreeding on the numerical values of PEV^* was found to be small, that is $PEV^* - PEV$ tended to be negligible. For each extra 10%

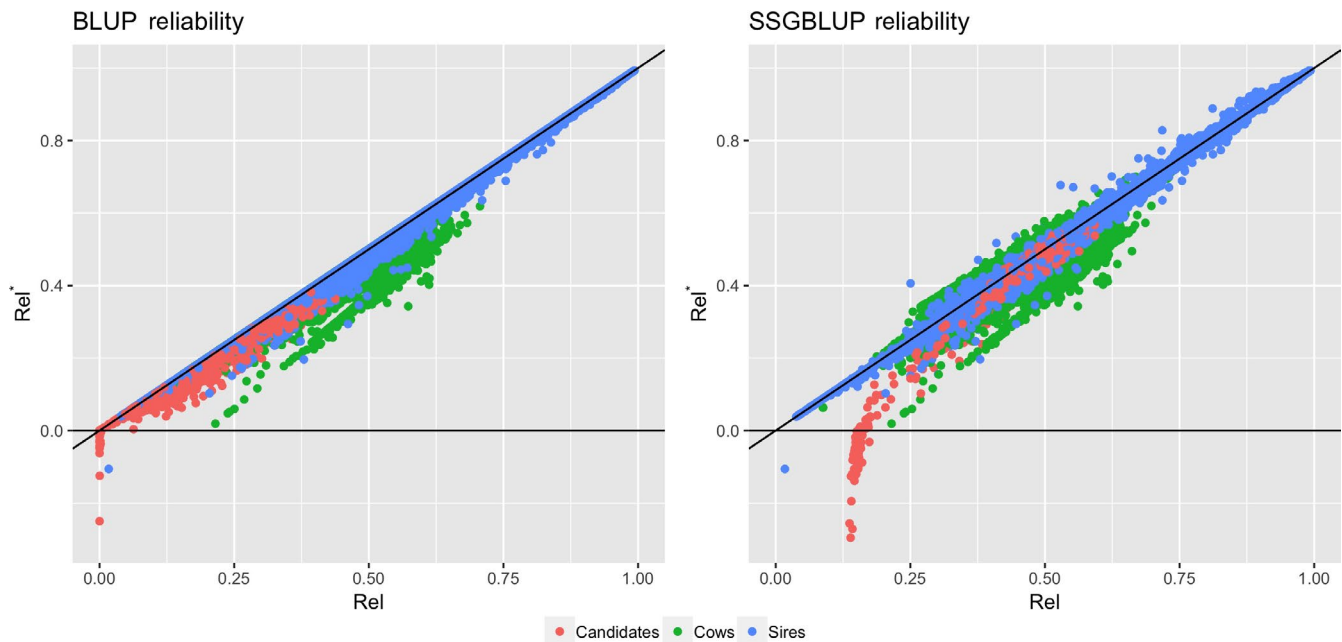


FIGURE 4 Correct reliability dividing PEV by inbreeding (X-axis) or not (Y-axis) for Uruguayan dairy cattle. Colours indicate candidates to selection, cows and bulls

“implied wrong” inbreeding (as in Figure 2 for some animals with zero inbreeding), there is an extra increase of 0.5% in PEV. Figure 3 shows (correct) PEV versus. (incorrect) PEV^* . There is actually quite good agreement and little dispersion. Correlation for PEV and PEV^* is greater than 0.99, and there is no observable bias (i.e. PEV^* is on average neither larger nor smaller than PEV) and the regression of PEV on PEV^* is very close to 1.

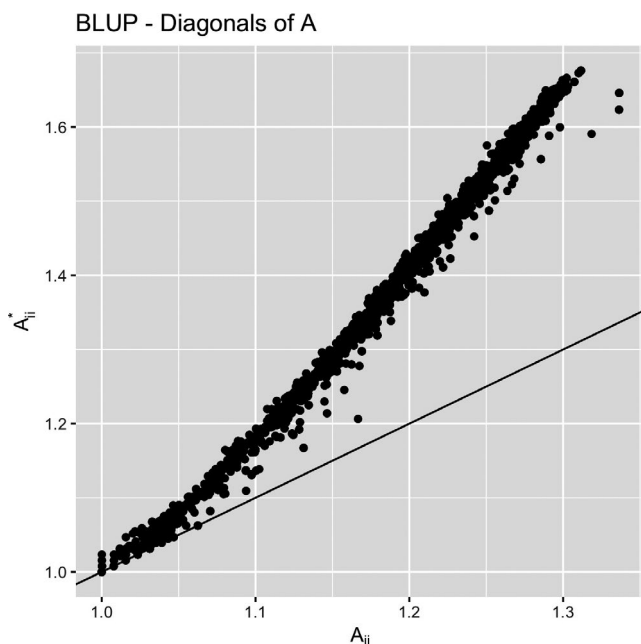


FIGURE 5 True self-relationships (A_{ii}) versus. implied self-relationships (A_{ii}^*) (using Henderson's rules and ignoring inbreeding) in the “A” rabbit line

3.2 | Dairy cattle: Effect of ignoring inbreeding in the computation of the reliability

In Figure 4, we present the results comparing the (worse possible) expression of repeatability, $rel^* = 1 - \frac{PEV^*}{\sigma_u^2}$ with the (correct) $rel = 1 - \frac{PEV}{(1+F_i)\sigma_u^2}$. We choose not to present intermediate

cases, for example with incorrect numerator but correct denominator as this makes presentation cumbersome. In the first case, inbreeding is ignored both in the construction of $\mathbf{A}^{-1}$ and $\mathbf{H}^{-1}$ (and therefore in the MME) and in the computation of reliability from PEV, which is the default option of many software such as BLUPF90 or PEST. The second case is the theoretically sound option. It can be seen that ignoring inbreeding systematically underestimates reliability in BLUP while in SSGBLUP there is over and underestimation of reliability.

Another issue with ignoring F_i is that some animals will obtain a PEV^* greater than σ_u^2 having thus a negative reliability. This will also mean that for highly inbred animals with little information (and therefore large PEV^*), accuracy will not be computed since the square root of $rel^* = 1 - \frac{PEV^*}{\sigma_u^2}$ is not a real number for $rel^* < 0$.

3.3 | Rabbit

The effect of ignoring inbreeding in the relationship matrices is dramatic and shown in Figure 5. In fact, the assumed

inbreeding is much higher, which in turn leads to underestimation of PEV^* as shown below. The overall correlation between PEV and PEV^* is 0.92. For this data set, given strong selection, we present average computed reliabilities per generation. Figure 6 (top) shows stabilization of reliability ($rel = 1 - \frac{PEV}{(1+F_i)\sigma_u^2}$) after a few generations. As the line be-

comes inbred, there is an increase in both $PEV = Var(u - \hat{u})$ and an increase in $Var(u) = (1+F)\sigma_u^2$. However, if PEV^* is computed (ignoring inbreeding coefficients in the set-up of A^{-1}), then PEV^* increases too much, leading to apparently

decreasing reliabilities ($rel^\# = 1 - \frac{PEV^*}{(1+F_i)\sigma_u^2}$) with time (middle). This trend is even more marked if reliabilities are computed using the wrong denominator, (i.e. as $rel^* = 1 - \frac{PEV^*}{\sigma_u^2}$),

in which case there are negative values of reliability, which are not consistent with the theory (bottom). All in all, an analyst of this data would be puzzled because the amount of information does not change with generations, yet reliability seems to decrease if incorrect expressions are used.

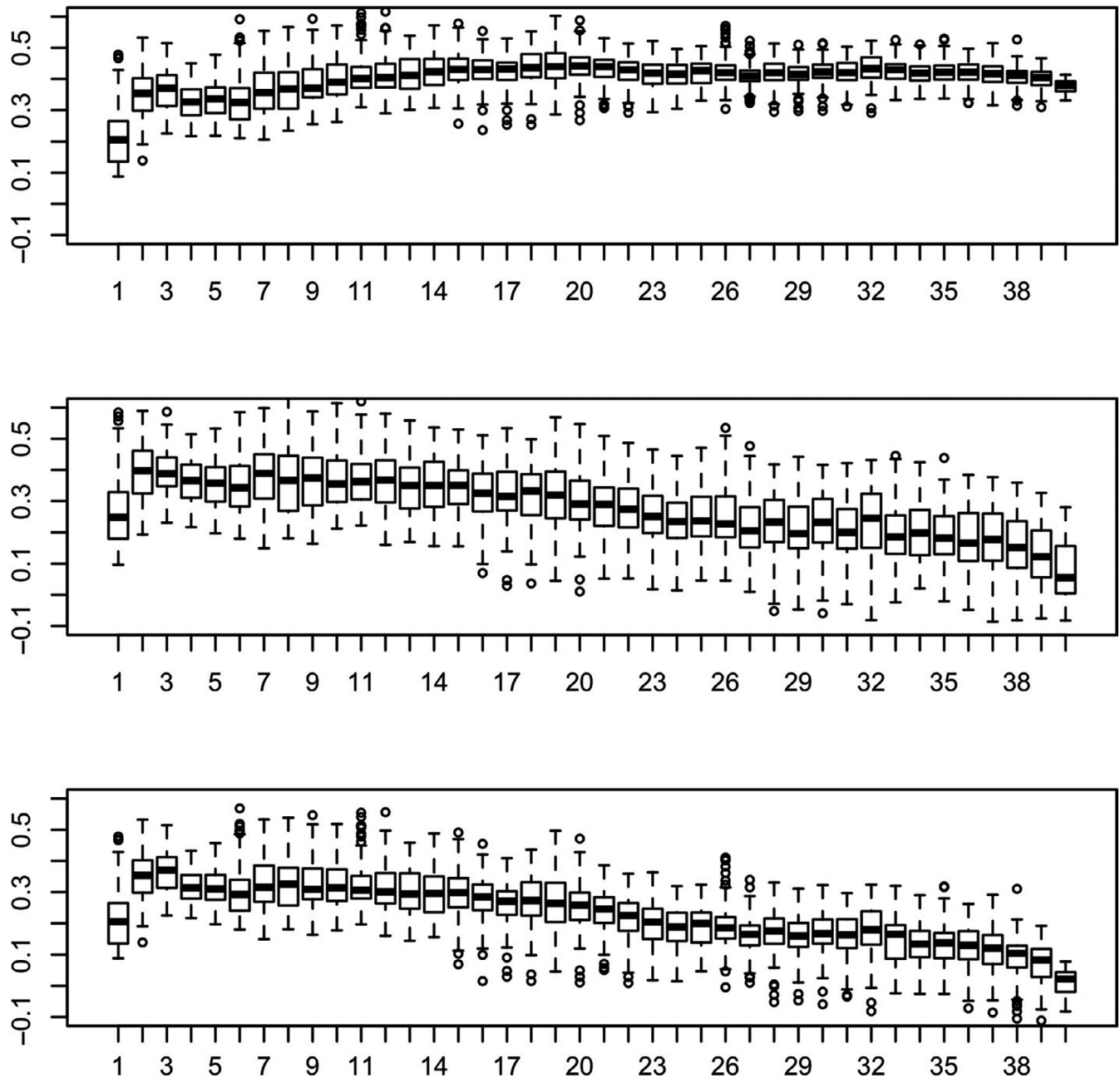


FIGURE 6 Exact (top), approximate using PEV^* and dividing by $1+F$ (middle), and approximate using PEV^* and not dividing by $1+F$ (bottom), reliabilities as a function of generation number in the “A” rabbit line

4 | DISCUSSION

In the dairy cattle data set, with low levels of inbreeding, there is little effect of inclusion (or not) of inbreeding in the set-up of $\mathbf{A}^{-1}$ and $\mathbf{H}^{-1}$ on the actual values of PEV . In the rabbit data set, with high levels of inbreeding, ignoring inbreeding in the set-up of $\mathbf{A}^{-1}$ leads to highly inflated values of PEV^* and therefore to too low figures for reliability. We therefore strongly recommend to *always* include inbreeding in the set-up of $\mathbf{A}^{-1}$ and $\mathbf{H}^{-1}$, because not doing it may result in larger underestimation of accuracy and its cost is negligible. The cost of computing pedigree-based inbreeding (which is all that is needed for correct $\mathbf{A}^{-1}$ and $\mathbf{H}^{-1}$) is negligible compared to the overall cost of the genetic evaluation.

In both data sets, there are errors if the correct denominator $(1 + F_i)$ is not used. These errors underestimate reliability in BLUP and under and over-estimate reliability in SSGBLUP. The overestimation is because some animals are more heterozygote than the average of the population ($H_{ii} < 1$), and therefore less variable a priori. The fact that some animals have low apparent reliability may inhibit breeders from using them; for instance, animals with low rel^* , say 0.40 (accuracy = 0.63) may have actual rel around 0.55 (accuracy = 0.75). The correct denominator must therefore be used. Computing F_i with pedigree has negligible cost. However, computation of inbreeding in single-step analysis from H_{ii} mixing pedigree and genomic relationships is not obvious. For genotyped animals, H_{ii} is simply G_{ii} computed from genotypes, which is straightforward as a cross-product on the animal's genotype (VanRaden, 2008). For non-genotyped animals, $H_{ii} = A_{ii} + a_{i2}A_{22}^{-1}(G - A_{22})A_{22}^{-1}a_{2i}$ (Legarra et al., 2009), this is a cumbersome expression to compute. Other options include sparse inversion of $\mathbf{H}^{-1}$ (as in this work and in Xiang et al. (2017)) or indirect methods such as Colleau, Palhière, Rodríguez-Ramilo, and Legarra (2017) and Legarra, Aguilar, and Colleau (2020).

Neither of the data sets has high heritability. High heritability would result in smaller PEVs and lower error in the different approximations that ignore inbreeding—for instance, negative reliabilities would possibly not be observed.

All in all, the practical benefits of using correctly inbreeding in genetic evaluations and in computation of reliabilities are the following. It compares animals fairly while an inbred animal is penalized if inbreeding is ignored. It allows for finer selection decision (i.e. genomic “heterogeneity” is correctly accounted for). It avoids negative reliabilities, which are puzzling to report. Last, using correctly inbreeding protects from unexpected problems, such as lack of convergence due to not consideration of inbreeding in the $\mathbf{H}^{-1}$ matrix in SSGBLUP (Matilainen et al., 2018; Strandén et al., 2017).

ACKNOWLEDGMENTS

Work supported by FEDER (Poctefa funds, project ARDI), INRA SelGen metaprogram (project GDivSelGen) and project Lomas CYT FCA 22, Universidad Nacional de Lomas de Zamora (Argentina). This project has received funding from the European Unions' Horizon 2020 Research & Innovation Programme under grant agreement N°772787 -SMARTER. We thank reviewers for insightful comments.

DATA AVAILABILITY STATEMENT

Data in this work are used with the permission of (dairy cattle) the Uruguayan National Dairy Herd Improvement (MU—Mejoramiento y Control Lechero Uruguayo, Montevideo Uruguay), and genealogical information was provided by the National Herdbook (ARU—Asociacion Rural del Uruguay, Montevideo, Uruguay), and (rabbit) from the Universitat Politècnica de València, and is not publicly available.

ORCID

Ignacio Aguilar  <https://orcid.org/0000-0002-1038-4752>

Agustin Blasco  <https://orcid.org/0000-0003-3558-0810>

Andres Legarra  <https://orcid.org/0000-0001-8893-7620>

REFERENCES

- 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. *Journal of Animal Breeding and Genetics*, 129(5), 345–358. <https://doi.org/10.1111/j.1439-0388.2012.00991.x>
- Christensen, O., Madsen, P., Nielsen, B., Ostensen, T., & Su, G. (2012). Single-step methods for genomic evaluation in pigs. *Animal*, 6, 1565–1571. <https://doi.org/10.1017/S1751731112000742>
- Colleau, J.-J., Palhière, I., Rodríguez-Ramilo, S. T., & Legarra, A. (2017). A fast indirect method to compute functions of genomic relationships concerning genotyped and ungenotyped individuals, for diversity management. *Genetics Selection Evolution*, 49, 87. <https://doi.org/10.1186/s12711-017-0363-9>
- Edel, C., Pimentel, E. C. G., Erbe, M., Emmerling, R., & Götz, K.-U. (2019). Short communication: Calculating analytical reliabilities for single-step predictions. *Journal of Dairy Science*, 102, 3259–3265. <https://doi.org/10.3168/jds.2018-15707>
- Fernandez, E. N., Sanchez, J. P., Martinez, R., Legarra, A., & Baselga, M. (2017). Role of inbreeding depression, non-inbred dominance deviations and random year-season effect in genetic trends for prolificacy in rabbit closed lines. *Journal of Animal Breeding and Genetics*, 6, 441–452.
- Gilmour, A. R., Gogel, B. J., Cullis, B. R., Thompson, R., & Butler, D. (2009). *ASReml user guide release 3.0*. Hemel Hempstead, UK: VSN International Ltd.
- Golden, B. L., Brinks, J. S., & Bourdon, R. M. (1991). A performance programmed method for computing inbreeding coefficients from large data sets for use in mixed-model analyses. *Journal of Animal Science*, 69(9), 3564–3573. <https://doi.org/10.2527/1991.6993564x>
- Groeneveld, E., Kovac, M., & Wang, T. (1990). PEST, a general purpose BLUP package for multivariate prediction and estimation.

- Proceedings of the 4th World Congress on Genetics Applied to Livestock Production, Edinburgh, 13, 488–491.
- Henderson, C. R. (1975). Best Linear Unbiased Estimation and Prediction under a Selection Model. *Biometrics*, 31, 423–447.
- Henderson, C. R. (1976). A simple method for computing the inverse of a numerator relationship matrix used in prediction of breeding values. *Biometrics*, 32(1), 69–83. <https://doi.org/10.2307/2529339>
- Henderson, C. R. (1982). Best linear unbiased prediction in populations that have undergone selection. In R. A. Barton, & W. C. Smith (Eds.), *Proceedings of the world congress on sheep and beef cattle breeding* (pp. 191–201). Palmerston North: Dunmore Press.
- Henderson, C. R. (1984). *Applications of linear models in animal breeding*. Guelph, Canada: University of Guelph.
- Legarra, A., Aguilar, I., & Colleau, J. J. (2020). Short communication: Methods to compute genomic inbreeding for ungenotyped individuals. *Journal of Dairy Science*, <https://doi.org/10.3168/jds.2019-17750>.
- Legarra, A., Aguilar, I., & Misztal, I. (2009). A relationship matrix including full pedigree and genomic information. *Journal of Dairy Science*, 92(9), 4656–4663. <https://doi.org/10.3168/jds.2009-2061>
- Legarra, A., Lourenco, D. A. L., & Vitezica, Z. G. (2018). Bases for genomic prediction. Retrieved from <http://genoweb.toulouse.inra.fr/~alegarra/>
- Masuda, Y., Aguilar, I., Tsuruta, S., & Misztal, I. (2015). Technical note: Acceleration of sparse operations for average-information REML analyses with supernodal methods and sparse-storage refinements. *Journal of Animal Science*, 93(10), 4670–4674. <https://doi.org/10.2527/jas.2015-9395>
- Matilainen, K., Strandén, I., Aamand, G. P., & Mäntysaari, E. A. (2018). Single step genomic evaluation for female fertility in Nordic Red dairy cattle. *Journal of Animal Breeding and Genetics*, 135(5), 337–348. <https://doi.org/10.1111/jbg.12353>
- Mehrabani-Yeganeh, H., Gibson, J. P., Schaeffer, L. R., (2000). Including coefficients of inbreeding in BLUP evaluation and its effect on response to selection. *Journal of Animal Breeding and Genetics*, 117(3), 145–151. <https://doi.org/10.1046/j.1439-0388.2000.00241.x>
- Meyer, K. (2007). WOMBAT—A tool for mixed model analyses in quantitative genetics by restricted maximum likelihood (REML). *Journal of Zhejiang University. Science. B*, 8(11), 815–821. <https://doi.org/10.1631/jzus.2007.B0815>
- Misztal, I., Tsuruta, S., Strabel, T., Auvray, B., Druet, T., & Lee, D. H. (2002). *BLUPF90 and related programs (BGF90)*. Montpellier, France: 7th World Congress on Genetics Applied to Livestock Production, CD-ROM Communication N° 28-07.
- Misztal, I., & Wiggans, G. R. (1988). Approximation of prediction error variance in large-scale animal models. *Journal of Dairy Science*, 71, 27–32. [https://doi.org/10.1016/S0022-0302\(88\)79976-2](https://doi.org/10.1016/S0022-0302(88)79976-2)
- Mrode, R., & Thompson, R. (2005). *Linear models for the prediction of animal breeding values*, Ed. 2. Wallingford, UK: CAB International.
- Mullen, M. P., McClure, M. C., Kearney, J. F., Waters, S. M., Weld, R., Flynn, P., ... Berry, D. P. (2013). Development of a custom SNP chip for dairy and beef cattle breeding, parentage and research. *Interbull Bulletin*, 47, 58–66.
- Pryce, J. E., Gonzalez-Recio, O., Nieuwhof, G., Wales, W. J., Coffey, M. P., Hayes, B. J., & Goddard, M. E. (2015). Hot topic: Definition and implementation of a breeding value for feed efficiency in dairy cows. *Journal of Dairy Science*, 98(10), 7340–7350. <https://doi.org/10.3168/jds.2015-9621>
- Sargolzaei, M., Chesnais, J., & Schenkel, F. S. (2014). A new approach for efficient genotype imputation using information from relatives. *BMC Genomics*, 15(1), 478. <https://doi.org/10.1186/1471-2164-15-478>
- Strandén, I., Matilainen, K., Aamand, G. P., & Mäntysaari, E. A. (2017). Solving efficiently large single-step genomic best linear unbiased prediction models. *Journal of Animal Breeding and Genetics*, 134(3), 264–274. <https://doi.org/10.1111/jbg.12257>
- Ten Napel, J., Vandenplas, J., Lidauer, M., Strandén, I., Taskinen, M., Mäntysaari, E., Veerkamp, R. F. (2017). MiXBLUP, user-friendly software for large genetic evaluation systems—Manual V2. Retrieved from: https://www.mixblup.eu/documents/Manual%20MiXBLUP%202.1_June%202017_V2.pdf
- Tier, B., Schneberger, M., Hammond, K., & Fuchs, W. C. (1991). Determining the accuracy of estimated breeding values in multiple trait animal models. Proceedings of the 9th AAABG Conference, 239–242
- Van Vleck, L. D. (1993a). *Selection index and introduction to mixed model methods*. Boca Raton, Florida: CRC Press.
- Van Vleck, L. D. (1993b). Variance of prediction error with mixed model equations when relationships are ignored. *Theoretical and Applied Genetics*, 85(5), 545–549. <https://doi.org/10.1007/BF00220912>
- VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *Journal of Dairy Science*, 91(11), 4414–4423. <https://doi.org/10.3168/jds.2007-0980>
- Xiang, T., Christensen, O. F., & Legarra, A. (2017). Technical note: Genomic evaluation for crossbred performance in a single-step approach with metafounders. *Journal of Animal Science*, 95(4), 1472–1480. <https://doi.org/10.2527/jas2016.1155>

How to cite this article: Aguilar I, Fernandez EN, Blasco A, Ravagnolo O, Legarra A. Effects of ignoring inbreeding in model-based accuracy for BLUP and SSGBLUP. *J Anim Breed Genet*. 2020;137:356–364. <https://doi.org/10.1111/jbg.12470>