

OPTIMIZING GENOMIC PREDICTION TOOLS: NOVEL MODELING AND
APPROXIMATION METHODS FOR IMPROVED GENETIC EVALUATIONS IN LARGE-
SCALE DATASETS

by

JOE-MENWER TABET

(Under the Direction of Daniela Lourenco)

ABSTRACT

The increasing number of phenotyped and genotyped animals in genetic evaluations presents significant challenges in delivering routine, accurate breeding values to breeders. Overcoming these challenges requires developing new strategies, methods, and approximations to ensure high-quality predictions while minimizing computational demands. This dissertation addresses key issues, such as developing modeling strategies to account for differences in divergent populations and exploring the use of single-step genomic best linear unbiased prediction (ssGBLUP) as an alternative to traditional multistep methods. With the large number of genotyped animals, particularly in the dairy and beef industries, genomic evaluations have become increasingly complex, necessitating alternative strategies. One such approach is the use of indirect predictions (IP) instead of genomic estimated breeding values (GEBVs), offering a faster method to deliver breeding values for breeders by excluding animals that do not contribute to the evaluation and calculating their IP based on single nucleotide polymorphisms (SNP) effects derived from previous evaluations. Along with IP, methods for efficiently calculating reliabilities—a metric representing the amount of information used to predict breeding values—

were also investigated. The computational cost of inverting matrices in the mixed model equations makes it impractical, even without genomic information. This challenge has driven the development of approximation methods to calculate reliabilities. However, complex models can overestimate these approximations, potentially compromising the selection process. Hence, developing correct approximation methods to improve the quality of reliabilities in genetic evaluations is imperative.

INDEX WORDS: Reliability, ssGBLUP, Indirect Predictions, Genomics, Algorithm for proven and young

OPTIMIZING GENOMIC PREDICTION TOOLS: NOVEL MODELING AND
APPROXIMATION METHODS FOR IMPROVED GENETIC EVALUATIONS IN LARGE-
SCALE DATASETS

by

JOE-MENWER TABET

BVSc., Lebanese University, Lebanon, 2016

M.S, Polytechnic University of Valencia, Spain, 2022

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial
Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2025

© 2025

Joe-Menwer Tabet

All Rights Reserved

OPTIMIZING GENOMIC PREDICTION TOOLS: NOVEL MODELING AND
APPROXIMATION METHODS FOR IMPROVED GENETIC EVALUATIONS IN LARGE-
SCALE DATASETS

by

JOE-MENWER TABET

Major Professor:
Committee:

Daniela Lourenco
Ignacy Misztal
Romdhane Rekaya
Jorge Hidalgo
Andres Legarra

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
May 2025

DEDICATION

To my late father, my amazing mother, and my three beautiful sisters: I am forever grateful for your encouragement and support, which have helped me realize my potential and achieve new milestones in life.

ACKNOWLEDGEMENTS

I would first like to thank my advisor, **Dr. Daniela Lourenco**, for believing in me and giving me the opportunity to achieve one of my dreams. Through her exceptional guidance and unwavering support, I am here today, completing my Ph.D. journey. Thanks to her, I was introduced to numerous opportunities that not only taught me valuable lessons but also provided me with the experience needed to excel in my research. To **Dr. Andres Legarra**, when I began my Ph.D. journey, you were someone I looked up to and always aspired to work with. I am deeply grateful for the knowledge and experience you have shared with me, and I am honored to be considered your student. I also want to extend my sincere thanks to **Dr. Ignacy Misztal** for his support, valuable insights, and guidance. His contributions have paved the way for us to excel and make meaningful contributions to the field of Animal Breeding and Genetics. I will always be grateful for his influence on my academic and professional growth. To **Dr. Romdhane Rekaya**, who not only served as an inspiring teacher but also promoted an enriching environment where we could freely exchange ideas. His teachings encouraged me to push my limits, deepening my understanding of theories and techniques. I always say that because of him, I fell in love with Bayesian statistics. To **Dr. Jorge Hidalgo**, I am grateful for your humility, one of your greatest virtues. You have always welcomed me for discussions and never made me feel like I didn't belong among the group at UGA. Thank you not only for being my teacher but also for being a friend who was there when I needed one.

I would like to acknowledge **Dr. Ignacio Aguilar** for his valuable time and insightful expertise. I truly enjoyed all our conversations, whether about research, industry, or personal life. I am grateful to have had the opportunity to work with you.

To **Dr. Vivian Breen**, with whom my doctoral journey began—thank you for the valuable time and industry insights you shared with me. Your guidance was instrumental in shaping my experience.

To **Dr. Paul VanRaden** and **Dr. Duane Norman**, I sincerely appreciate the time we spent discussing various topics, especially the fascinating history of U.S. dairy breeding. You always treated me as a colleague rather than a student, and for that, I am deeply grateful. It has been an honor to work with you and gain irreplaceable experience.

To **Dr. Zulma Vitezica**, thank you for your unwavering support. Because of you, my visits to Toulouse were not only possible but also incredibly successful. From day one, you treated me as one of your own students and readily helped me with any doubts I had. I truly appreciate your kindness and mentorship.

To **Dr. Matias Bermann**, my first class at UGA was a calculus class in Spanish with you, and to this day, it remains one of the best classes I have taken. I still apply the knowledge I gained in that class. I am proud to say that I have learned so much from you—not only in calculus but also in algebra and coding. Thank you.

To **Dr. Fernando Bussiman**, words cannot fully express my gratitude. I can confidently say that if it weren't for you, my Ph.D. journey would have been entirely different. I don't even know where to begin to thank you. You were the first person to welcome me with open arms, and you have supported me ever since. I can't count the countless hours you spent helping, guiding,

and teaching me something new. I am the student I am today in part because of you, and I am truly honored to be considered your student.

To my '**cubicle neighbors**', **Dr. Jennifer Richter** and **Dr. Mary Kate Hollifield**, as well as **Dr. Taylor McWhorter** and **Dr. Pedro Ramos**—you made my days at the department so much easier. I am incredibly grateful for the friendships I built with you. You were the best part of my day, and I will always cherish the moments—both good and bad—that we shared together.

A special thank you to the rest of our lab group. Each of you has played a role in shaping me both personally and professionally, and I am grateful for everything I have learned from you. Many of my colleagues have become my closest friends, and I will always cherish the memories we have made together.

Lastly, I would like to express my deepest gratitude to my amazing family and friends for their unwavering support throughout this journey. Your encouragement has meant the world to me.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	v
LIST OF TABLES	xi
LIST OF FIGURES	xiii
CHAPTER	
1 INTRODUCTION	1
2 LITERATURE REVIEW	4
History of Genetic Evaluations	4
Challenges of Large-Scale Genomic Evaluations.....	11
Validation of Genomic Evaluations	15
References.....	19
3 COMBINING LARGE BROILER POPULATIONS INTO A SINGLE GENOMIC EVALUATION: DEALING WITH GENETIC DIVERGENCE.....	26
Abstract	27
Introduction.....	28
Materials and Methods.....	30
Results and Discussion	37
Conclusions.....	45
Acknowledgments.....	46
References.....	47

4	ALL-BREED SINGLE-STEP GBLUP EVALUATIONS FOR FERTILITY TRAITS IN U.S. DAIRY CATTLE	65
	Abstract	66
	Introduction	66
	Materials and Methods	68
	Results	76
	Discussion	83
	Conclusions	87
	Acknowledgments	88
	References	89
5	APPROXIMATING RELIABILITIES OF INDIRECT PREDICTIONS USING SNP EFFECTS FROM LARGE SSGBLUP EVALUATIONS	104
	Abstract	105
	Introduction	106
	Materials and Methods	108
	Results and Discussion	111
	Conclusions	116
	Acknowledgments	116
	References	117
6	CORRECTING OVERESTIMATIONS OF APPROXIMATE TRADITIONAL RELIABILITIES WHEN YOUNG GENOMIC BULLS ARE USED IN FEW HERDS	129
	Abstract	130

Introduction.....	131
Materials and Methods.....	133
Results and Discussion	139
Conclusions.....	142
Acknowledgments.....	142
References.....	143
7 CONCLUSIONS.....	156

LIST OF TABLES

	Page
Table 3.1: Descriptive statistics of lines 1 and 2	52
Table 3.2: Heritability (diagonal), genetic correlation (above-diagonal) for each trait.....	53
Table 3.3: Number of animals belonging to each genetic group	54
Table 3.4: Number of focal animals for each line and trait	55
Table 4.1: Descriptive statistics of the studied traits	96
Table 4.2: Number of core animals per breed and sex and total number of genotypes across breed	
97	
Table 4.3: Estimated bias, in genetic standard deviations, of (G)EBV comparing whole and partial datasets for ssGBLUP with MF/UPG at 5% and 10% residual polygenic effect (ssGBLUP _{UPG5} /ssGBLUP _{MF5} and ssGBLUP _{UPG10} /ssGBLUP _{MF10}), and traditional pedigree BLUP with MF and UPG (PBLUP _{MF} and PBLUP _{UPG})	98
Table 4.4: Estimated regression coefficient of (G)EBV from whole on partial datasets for ssGBLUP with UPG/MF at 5% and 10% residual polygenic effect (ssGBLUP _{UPG5} /ssGBLUP _{MF5} and ssGBLUP _{UPG10} /ssGBLUP _{MF10}), and traditional pedigree BLUP with MF and UPG (PBLUP _{MF} and PBLUP _{UPG}) with Linear Regression (LR) Improved Genomic Validation (IGV).....	99
Table 4.5: Correlation of (G)EBV from whole and partial datasets for ssGBLUP with MF at 5% and 10% residual polygenic effect (ssGBLUP _{UPG5} /ssGBLUP _{MF5} and	

ssGBLUP _{UPG10} /ssGBLUP _{MF10}), ssGBLUP with UPG (ssGBLUP _{UPG}) and traditional pedigree BLUP with MF and UPG (PBLUP _{MF} and PBLUP _{UPG})	100
Table 4.6: LR (Linear Regression) and IGV (Improved Genomic Validation) statistics for multistep predictions with genomics (MS _{Genomic}) and without genomics (MS _{BLUP}).....	101
Table 5.1: Distribution of phenotypes and genotypes per breed	121
Table 5.2: Number of core animals per breed and sex and total number of genotypes across breed	122
Table 5.3: Correlation between genomic reliabilities (REL_{GS3}) and reliability of Indirect predictions (REL_{IP}) for Holstein (HO), Jersey (JE), and Brown-Swiss (BS) from multi-breed and single-breed evaluations	123
Table 5.4: Means of genomic reliability (REL_{GS3}) and reliability of Indirect predictions (REL_{IP}) for Jersey (JE) and Brown-Swiss (BS) from multi-breed and single-breed evaluations.	124
Table 6.1: Reliability means, correlation, and regression of REL_{EXACT} on approximated reliabilities of Odd and Normal sires from simulated dataset.....	146
Table 6.2: Reliability means correlation and regression of REL_{EXACT} on approximated reliabilities of Odd and Normal sires from DataReduced.....	147
Table 6.3: Correlation and regression of REL_{EXACT} with approximated reliabilities for cows with records in DataReduced	148
Table 6.4: Reliability means, correlation, and regression of REL_{MW} on approximated reliabilities of Odd and Normal sires from DataFull.	149

LIST OF FIGURES

	Page
Figure 3.1: Principal component analysis (PCA) of the genomic relationship matrix	56
Figure 3.2: Prediction accuracy (acc_{LR}) with standard errors on the y-axis between Combined (both lines evaluated together without crossbred animals), Origin 1 (evaluation of animals from line 1 only) and Origin 2 (evaluation of animals from line 2 only) with M1/NOUPG (M1: Model with no account for animal's line of origin, NOUPG: no use of UPG or MF in the model) of BW (Body Weight), CY (Carcass Yield), LH (Leg Health) and MT (Mortality) on the x-axis	57
Figure 3.3: Prediction accuracy (acc_{LR}) with standard errors on the y-axis of M1 (Model with no account for animal's line of origin), M2 (Model with a vector of fixed effect based on animal's line of origin included in the model) and M3 (Contemporary groups and sex effects defined based on animal's line of origin) with NOUPG (no use of UPG or MF in the model) of BW (Body Weight), CY (Carcass Yield), LH (Leg Health) and MT (Mortality) on the x-axis	58
Figure 3.4: Prediction accuracy (acc_{LR}) with standard errors on the y-axis of M1 (Model with no account for animal's line of origin), M2 (Model with a vector of fixed effect based on animal's line of origin included in the model) and M3 (Model with contemporary groups and sex effects defined based on animal's line of origin) with UPG1 (UPG based on animals' line of origin) of BW (Body Weight), CY (Carcass Yield), LH (Leg Health) and MT (Mortality) on the x-axis	59

Figure 3.5: Dispersion (b_1) with standard errors on the y-axis levels between Combined (both lines evaluated together without crossbred animals), Origin 1 (evaluation of animals from line 1 only) and Origin 2 (evaluation of animals from line 2 only) with M1/NOUPG (M1: Model with no account for animal's line of origin, NOUPG: no use of UPG or MF in the model) of BW (Body Weight), CY (Carcass Yield), LH (Leg Health) and MT (Mortality) on the x-axis 60

Figure 3.6: Dispersion levels (b_1) with standard errors on the y-axis of M1 (Model with no account for animal's line of origin), M2 (Model with a vector of fixed effect based on animal's line of origin included in the model) and M3 (Model with contemporary groups and sex effects were defined based on animal's line of origin) with MF2 (UPG and MF based on animal's line of origin and separated by sex) of BW (Body Weight), CY (Carcass Yield), LH (Leg Health) and MT (Mortality) on the x-axis 61

Figure 3.7: Bias with standard errors between Combined (both lines evaluated together without crossbred animals) results from partial dataset vs. Origin 1 (evaluation of animals from line 1 only) and Origin 2 (evaluation of animals from line 2 only) results from whole dataset of BW (Body Weight), CY (Carcass Yield), LH (Leg Health) and MT (Mortality) on the x-axis 62

Figure 3.8: Distribution of Body weight (BW) and Mortality (MT) GEBV (Genomic estimated breeding values) for each model and scenario (NOUPG: No UPG or MF in the model, UPG1/MF1: UPG and MF based on animals' line of origin, UPG2/MF2: UPG and MF based on animal's line of origin and separated by sex; M1: no account for any line of origin, M2: Vector of fixed effect based on animal's line of origin included in the model, M3: Contemporary groups and sex effects were defined based on animal's line of origin 63

Figure 3.9: Spearman Correlation plot of all models and scenarios for Body weight (NOUPG: No UPG or MF in the model, UPG1/MF1: UPG and MF based on animals' line of origin, UPG2/MF2: UPG and MF based on animal's line of origin and separated by sex; M1: no account for any line of origin, M2: Vector of fixed effect based on animal's line of origin included in the model, M3: Contemporary groups and sex effects were defined based on animal's line of origin).....	64
Figure 4.1: Genetic trend of CCR and DPR in Holstein and Jersey using Metafounders in ssGBLUP with 10% residual polygenic effect and PBLUP	102
Figure 4.2: Estimates of Genetic groups for unknown sires of foreign dams in Holstein and Jersey from the whole dataset for CCR with ssGBLUP _{MF10} and ssGBLUP _{UPG10} (ssGBLUP with MF/UPG and residual polygenic effect = 10%)	103
Figure 5.1: Multi-breed reliability for Holstein focal animals with 5% residual polygenic effect. Abbreviations: REL_{GS3} = reliability of genomic predicted breeding value; REL_{IP} = reliability of indirect prediction.	125
Figure 5.2: Multi-breed reliability for Jersey (JE) and Brown Swiss (BS) focal animals with 5% residual polygenic effect. Abbreviations: REL_{GS3} = reliability of genomic predicted breeding value; REL_{IP} = reliability of indirect prediction.....	126
Figure 5.3: Single-breed reliability for Brown Swiss focal animals with 5% residual polygenic effect. Abbreviations: REL_{GS3} = reliability of genomic predicted breeding value; REL_{IP} = reliability of indirect prediction.	127
Figure 5.4: Single-breed reliability for Jersey focal animals with 5% residual polygenic effect. Abbreviations: REL_{GS3} = reliability of genomic predicted breeding value; REL_{IP} = reliability of indirect prediction.	128

Figure 6.1: Distribution of daughters per sire	150
Figure 6.2: Distribution of herds per sire	151
Figure 6.3: Reliabilities of Odd sires from the simulated dataset.....	152
Figure 6.4: Reliabilities of Odd and Normal sires from DataReduced	153
Figure 6.5: Reliabilities of cows with records from DataReduced	154
Figure 6.6: Reliabilities of Odd and Normal sires from DataFull	155

CHAPTER 1

INTRODUCTION

Animal breeding has been practiced for centuries, dating back to the 18th century. Robert Bakewell (1726-1795), often regarded as the founder of animal breeding, introduced theories emphasizing the importance of accurate record-keeping. He pioneered the concept of progeny testing and advocated for “breeding the best to the best.” These early advancements laid the groundwork for animal breeding as a systematic discipline.

In the 20th century, as a modern science, animal breeding shifted into a new period, driven by the contributions of renowned scientists such as Ronald A. Fisher, Sewall Wright, Jay L. Lush, and Charles R. Henderson. These pioneers significantly advanced the field, leading to transformative changes in animal production. For instance, in 1895, the average 305-day milk production of a dairy cow in the United States was 1,088 kg; today, it exceeds 12,000 kg per cow. Similarly, in 1957, the average weight of a broiler chicken at 56 days was 902 g, a figure that has since quadrupled to approximately 4 kg. This remarkable progress has been achieved by identifying and selecting superior animals as parents for the next generation, ensuring their genetic merit is passed on to their offspring.

To identify animals with superior genetic potential, robust models and methods have been developed to estimate breeding values—a measure of an animal's genetic merit for passing traits to the next generation. Alongside breeding value estimation, reliability is a critical metric farmers and breeders use to make informed breeding decisions. Reliability reflects the confidence in the estimated breeding value, representing the amount of information used in its calculation. High

reliability (>80%) indicates that the breeding value can be trusted as by adding more data the value will not change considerably.

In 1963, Henderson introduced the "Best Linear Unbiased Prediction" (BLUP) method, which estimates breeding values by minimizing prediction error variance. Later, Henderson (1984) demonstrated that prediction error variance could be obtained from the diagonal elements of the inverted mixed model equations, facilitating the calculation of breeding values' reliability.

The early 2000s marked the beginning of the genomic era in animal breeding, transforming the field with the integration of genomic information. Meuwissen et al. (2001) demonstrated the potential of using genomic data to estimate breeding values with greater accuracy. Over the years, extensive research has advanced genomic technologies and genetic evaluation methodologies, leveraging this information. However, these advancements have brought significant computational challenges due to the high demands of incorporating genomic data. Consequently, researchers have focused on methods to reduce computational costs while maintaining accuracy. For example, the Algorithm for Proven and Young (APY; Misztal et al. (2014a)) reduces the dimensionality of the genomic relationship matrix (**G**), enabling faster matrix inversion. Additionally, approximation methods for reliability computation (Misztal and Wiggans, 1988; Tier and Meyer, 2004; Bermann et al., 2022a) have been developed to address the computational intensity of mixed model equation inversion. Validation methods have also been established to assess these strategies by measuring population accuracy, bias, and the dispersion of breeding values (Legarra and Reverter, 2018; VanRaden, 2021).

The objective of this dissertation was to model strategies for improving breeding value estimation and to develop reliability approximation methods to reduce computational costs and provide better quality approximations. Chapter 2 provides a comprehensive literature review.

Chapter 3 explores strategies for combining large broiler populations into a unified genetic evaluation to enhance the quality of breeding values. Chapter 4 examines the feasibility of implementing a single-step methodology for an all-breed U.S. dairy cattle population. Chapter 5 introduces an alternative reliability computation method aimed at reducing the computational cost of genetic evaluations. Chapter 6 investigates the potential overestimation of reliability approximation methods in young genomic bulls utilized in a limited number of herds. The general conclusions of this dissertation are presented in Chapter 7.

CHAPTER 2

LITERATURE REVIEW

HISTORY OF GENETIC EVALUATIONS

Animal breeding and genetics help producers identify genetically superior animals for economically important traits by applying statistical methods to evaluate and rank animals. This process, known as genetic evaluation, estimates an animal's genetic potential using performance, pedigree, and genomic data, improving population performance and profitability.

Advances in statistical modeling, selection index theory, and computational techniques have revolutionized genetic evaluations. Hazel (1943) adapted the selection index theory from plant to animal breeding. The fundamental idea of selection indices is to predict an individual's additive genetic value using all available information. This involves applying multiple linear regression based on the data of the individual and their relatives to predict the additive genetic value as follows:

$$a = \mathbf{b}'\mathbf{y} + \mathbf{e} \quad [2.1]$$

where a represents the animals' additive value, $\mathbf{b}$ is a vector of regression coefficients, $\mathbf{y}$ is a vector of records from the animal and its relatives, and $\mathbf{e}$ represents the error.

Soon after, Henderson (1963) introduced the Best Linear Unbiased Prediction (BLUP) method, which remains the most widely used tool. BLUP uses mixed model equations (MME) to simultaneously estimate fixed effects and predict an animal's genetic merit or additive genetic value while minimizing the prediction error variance (PEV). This method computes unbiased

predictors that are linear functions of the observations, provided the analytical model correctly accounts for all relevant environmental and genetic biases (Henderson, 1973).

The general mixed model equation used in BLUP is as follows:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e} \quad [2.2]$$

where $\mathbf{y}$ is a vector of observations, $\boldsymbol{\beta}$ is a vector of fixed effects, $\mathbf{u}$ is a vector of additive genetic random effects, $\mathbf{e}$ is a vector of residuals, and $\mathbf{X}$ and $\mathbf{Z}$ are incidence matrices for the fixed and random effects, respectively. The solution to this equation yields the ‘best unbiased’ predictor of the additive genetic effect $\mathbf{u}$, referred to as the estimated breeding values (EBV). EBV rank animals and support breeders in making informed breeding decisions.

BLUP assumes that genetic relationships among animals provide significant information for predicting breeding values. Henderson (1976) demonstrated how to incorporate these relationships into the estimation process using the ‘Numerator relationship matrix.’ Initially, BLUP evaluations only considered sires as random effects, giving rise to the ‘Sire models.’ Subsequent developments in genetic evaluation programs led to the ‘Animal model,’ where all the animals in the population are evaluated. These advancements were driven by increased data availability, improved data storage, and enhanced statistical and computational methods.

When applying the animal model to BLUP, certain assumptions are made about the random variables: $\mathbf{e} \sim N(0, \mathbf{R})$ where $\mathbf{R} = \mathbf{I}\sigma_e^2$, $\mathbf{I}$ is an identity matrix, σ_e^2 is the residual variance, and the additive genetic random effect $\mathbf{u} \sim N(0, \mathbf{G})$ where $\mathbf{G} = \mathbf{A}\sigma_u^2$, $\mathbf{A}$ is the numerator relationship matrix, and σ_u^2 is the variance of $\mathbf{u}$. The MME for the animal model are:

$$\begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} + \mathbf{A}^{-1}\lambda \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{bmatrix} \quad [2.3]$$

Where $\mathbf{A}^{-1}$ is the inverse of the numerator relationship matrix, and λ is a variance ratio $\frac{\sigma_e^2}{\sigma_u^2}$.

Henderson (1976) and Quaas (1976) developed an efficient algorithm for inverting the relationship matrix, simplifying the implementation of the animal model and making genetic evaluations a routine process.

Henderson (1975) demonstrated through the MME theory that a metric called reliability could be derived from the inverse of the coefficient matrix. The author showed that the elements associated with the genetic effect in the inverted coefficient matrix correspond to the prediction error variance (PEV). For an animal l , the reliability is calculated as:

$$REL_l = 1 - \frac{PEV_l}{(1 + F_l)\sigma_u^2} \quad [2.4]$$

where PEV_l is the prediction error variance and F_l is the inbreeding coefficient of animal l . This metric typically accompanies EBVs and provides producers with additional tools to make better breeding decisions and enhance population performance. These methods form the foundation of modern genetic evaluations, which remain used in breeding programs today.

Genomic technologies, such as DNA mapping, sequencing, and analysis, have transformed genetic evaluations. Early efforts, including the use of polymorphic markers like microsatellites to identify regions influencing economically significant traits (quantitative trait loci, or QTL), led to the development of marker-assisted selection (MAS) (Lande and Thompson, 1990). However, MAS had a limited impact due to the polygenic nature of traits.

With the advances in molecular genetic techniques and the discovery of single nucleotide polymorphisms (SNPs) genomic selection became a possibility, as proposed by Meuwissen et al.

(2001). In a simulation, they incorporated genetic markers into a genomic evaluation, increasing prediction accuracy, reducing generational interval by enabling earlier selection of animals and accelerating genetic progress. However, high genotyping costs and the lack of dense SNP panels initially limited the adoption of genomic selection. With the advancements in genomic technologies, including the design of the 50K SNP chip and the decreasing cost of genotyping, animal breeders began incorporating genomic information into genetic evaluations. The first genomic evaluation was published for Holstein and Jersey cattle in 2009 (Wiggans and Carrillo, 2022), incorporating genotype information to compute genomic estimated breeding values (GEBV).

This evaluation involves constructing the genomic relationship matrix (**G**) to integrate genomic information in the genetic evaluation. Proposed by VanRaden (2008), the **G** matrix is defined as:

$$\mathbf{G} = \frac{\mathbf{ZZ}'}{2 \sum p_i(1 - p_i)} \quad [2.5]$$

where **Z** is a matrix centered by twice the frequency of the allele i (p_i) based on the genotyped animals present in the population, base allele frequencies or allele frequencies equal 0.5. The **G** matrix is analogous to the pedigree relationship (**A**) from the MME, with the exception that relationships between animals are Identical by State (IBS) rather than Identical by Descent (IBD). The diagonal elements of **G** reflect homozygous loci counts, while off-diagonal elements represent shared alleles among individuals.

The first genomic selection method implemented was based on a multistep procedure (VanRaden, 2008; VanRaden et al., 2009). Due to limited number of genotyped animals, most of the animals in the population could not receive GEBV. The multistep process began with using a pedigree-based BLUP evaluation to obtain EBV. The EBV of genotyped animals were deregressed

to calculate daughter yield deviations (DYDs), which were used in a genomic BLUP (GBLUP) to estimate SNP effects and compute direct genomic values (DGV) for genotyped animals. This method was initially adopted because bulls were the primary animals genotyped, and these animals could not be recorded for all performance traits. Meanwhile, non-genotyped animals received traditional EBV. Finally, the DGV and EBV for genotyped animals were combined to produce GEBV. Despite its successful implementation and contribution to genetic progress, this method was limited to simple models (Miszta et al., 2009) and introduced bias and errors in the estimations (Petry and Ducrocq, 2009; Masuda et al., 2018; Jibrila et al., 2021).

To overcome the challenges of multistep evaluations, Miszta et al. (2009) proposed a single-step method that combines all available information into a single genetic evaluation. This method constructs a realized relationship matrix (**H**) by combining pedigree (**A**) and genomic (**G**) relationship matrices. Legarra et al. (2009) defined **H** as:

$$\mathbf{H} = \mathbf{A} + \begin{bmatrix} \mathbf{A}_{12}\mathbf{A}_{22}^{-1}(\mathbf{G} - \mathbf{A}_{22})\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{A}_{12}\mathbf{A}_{22}^{-1}(\mathbf{G} - \mathbf{A}_{22}) \\ (\mathbf{G} - \mathbf{A}_{22})\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & (\mathbf{G} - \mathbf{A}_{22}) \end{bmatrix} \quad [2.6]$$

where **A**₂₂ is the pedigree relationship matrix for genotyped animals, **A**₂₂⁻¹ its inverse, **A**₁₂ is the pedigree relationship between non-genotyped and genotyped animals, and other matrices are as defined above. Despite the complexity of **H**, its inverse has a simpler form, as presented by Aguilar et al. (2010):

$$\mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix} \quad [2.7]$$

where **G**⁻¹ is the inverse of the **G**.

Today, the single-step genomic BLUP (ssGBLUP) is the preferred method for routine livestock evaluations due to its efficiency and ability to handle both genotyped and non-genotyped animals.

Improving genetic evaluations involves not only developing estimation methods but also the application of the correct model that accounts for all the variability in the data to provide the best unbiased predictors. A case in point is the application of “Unknown Parent Group” (UPG) and “Metafounders” (MF) to address missing pedigree information.

In most livestock populations, it is common to have animals with no known pedigree (missing parents), often referred to as the base population. In closed populations where no crossbreeding or importation of animals occurs, a single base population is considered. However, most livestock populations do not meet these criteria, requiring the consideration of multiple base populations. For instance, the U.S. dairy cattle genetic evaluation includes six different breeds and their crossbreds. Treating all as one base population would lead to erroneous breeding values. Similarly, in poultry production, genetic lines are often imported and bred to different lines to increase variability.

To address all these differences, animals with missing parents are assigned to genetic groups based on criteria like sex, breed, and year. This approach is termed Unknown Parent Group or UPG. Each genetic level or group is treated as unrelated and of infinite size, so inbreeding is zero (Quaas, 1988). The UPGs were modeled as fixed effects, and the breeding value of an individual can be decomposed as:

$$\mathbf{u}^* = \mathbf{u} + \mathbf{Qg} \quad [2.8]$$

where $\mathbf{Q}$ is a matrix of expected gene contributions from a genetic group and $\mathbf{g}$ is a vector of fixed group means.

With the development of BLUP, Quaas and Pollak (1981) reformulated the MME to account for UPGs directly in the inverse of the numerator relationship matrix:

$$\mathbf{A}^* = \begin{bmatrix} \mathbf{Q}'\mathbf{A}^{-1}\mathbf{Q} & -\mathbf{Q}'\mathbf{A}^{-1} \\ -\mathbf{A}^{-1}\mathbf{Q} & \mathbf{A}^{-1} \end{bmatrix} \quad [2.9]$$

Misztal et al. (2013) extended the UPG model to ssGBLUP by substituting $\mathbf{A}^{-1}$ by $\mathbf{A}^*$ leading to:

$$\mathbf{H}^* = \mathbf{A}^* + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix} \quad [2.10]$$

This method reduced inflation in breeding values and improved the accuracy of prediction. However, careful definition of genetic groups remains crucial (Misztal et al., 2013; Tsuruta et al., 2019). Other methods were developed to consider UPGs in ssGBLUP evaluations, enhancing the breeding values (Masuda et al., 2022; Strandén et al., 2022).

With the introduction of genomic evaluation and selection, it was proven that the assumption of unrelated populations was not correct, and relationships within and across these groups can exist. A new concept of “Metafounders” or MF was introduced that serve as proxies for base populations, with relationships modeled using a matrix $\mathbf{\Gamma}$ that estimates covariances between the means of these groups, thereby accounting for their similarity. The $\mathbf{\Gamma}$ matrix defines relationships within and across MF, ensuring allele substitution effects are defined across groups (Christensen et al., 2015).

Methods to estimate the relationship across base populations were developed, such as the method of moments, maximum likelihood, or generalized least squares (Legarra et al., 2015; Garcia-Baccino et al., 2017). For two MF x and y , the elements of $\mathbf{\Gamma}$ are:

$$\gamma_{xy} = \frac{2}{n} (2\mathbf{p}_x - \mathbf{1})(2\mathbf{p}_y - \mathbf{1})' \quad [2.11]$$

where γ_{xy} is the relationship coefficient between x and y , $\mathbf{p}_x$ and $\mathbf{p}_y$ their allele frequency vectors and n the number of markers. Assuming the mean of $\mathbf{p}_x$ and $\mathbf{p}_y = 0.5$, $\gamma_{xy} = 8 \text{ cov}(p_{x,k}, p_{y,k})$ representing the covariance of allele frequencies at the k^{th} locus (Garcia-Baccino et al., 2017).

MF models improve compatibility between **A** and **G**, by addressing missing pedigree issues. They are particularly useful in multibreed or crossbred evaluations, enhancing both dispersion and bias in estimations (Bermann et al., 2021; Kluska et al., 2021; Masuda et al., 2022). Chapter 3 examines strategies for integrating large broiler populations into a single genetic evaluation framework to improve the accuracy and quality of breeding value by implementing these methods.

When modeling effects that influence data, several considerations must be followed. Data often accumulates over time (e.g., months or years) or under varying environmental conditions (e.g., high vs. low temperatures). To address these factors, specific effects have been developed for genetic models. One key component is Contemporary groups (CG), which usually represent subsets of time-by-location effects. For instance, animals raised on different farms, in different years, or managed differently by various producers form distinct CG. Within a CG, animals are generally of similar age and managed alike.

Another important consideration is the herd-sire effect (HS), a random effect with variance σ_{hs}^2 , introduced to account for preferential treatment of animals from the same sire within a herd (e.g., daughters of the same sire might have access to better quality feed than other animals in the herd). Neglecting this effect would introduce bias in genetic evaluations and overestimate reliabilities (Norman, 1974; Meyer, 1987).

CHALLENGES OF LARGE-SCALE GENOMIC EVALUATIONS

The adoption of genomic information in genetic evaluations has been widely praised for its significant impact on genetic progress and prediction accuracy. This has driven research into developing low-cost SNP panels to make the technology accessible to all producers. For instance, the U.S. dairy industry adopted genomic evaluation in 2009 and now has over ten million

genotyped animals (<https://uscdcb.com/>), while the American Angus Association has around 2 million genotypes of animals in its database (Pedro Ramos, Angus Genetics Inc., personal communication). These examples highlight the importance of genomic data in improving genetic evaluations.

However, the rapid growth in genomic information has introduced computational challenges. BLUP-based methods such as GBLUP and ssGBLUP require matrix inversion, which becomes computationally infeasible as the number of genotyped animals increases. The inversion of $\mathbf{G}$ has a cubic cost with the number of genotyped animals and is not feasible for over 150,000 animals (Fragomeni et al., 2015). Additionally, inverting the coefficient matrix of the MME can be computationally demanding, limiting the ability to calculate reliabilities.

To address these challenges, Misztal et al. (2014a) proposed the Algorithm for Proven and Young (APY), which considers the reduced dimensionality of $\mathbf{G}$ to construct a sparse representation of $\mathbf{G}^{-1}$ by recursions on a “core” subset of genotyped animals with the size of the rank of $\mathbf{G}$.

The $\mathbf{G}^{-1}$ with APY ($\mathbf{G}_{\text{APY}}^{-1}$) is constructed as:

$$\mathbf{G}_{\text{APY}}^{-1} = \begin{bmatrix} \mathbf{G}_{\text{cc}}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \begin{bmatrix} -\mathbf{G}_{\text{cc}}^{-1}\mathbf{G}_{\text{cn}} \\ \mathbf{I} \end{bmatrix} \mathbf{M}_{\text{nn}}^{-1} \begin{bmatrix} -\mathbf{G}_{\text{nc}}\mathbf{G}_{\text{cc}}^{-1} & \mathbf{I} \end{bmatrix} \quad [2.12]$$

where $\mathbf{G}_{\text{cc}}^{-1}$ is the inverse of the genomic relationship matrix for core animals, $\mathbf{G}_{\text{cn}}$ is the genomic relationship matrix between core and non-core animals and $\mathbf{M}_{\text{nn}}^{-1}$ is the inverse of the relationship matrix between non-core animals. The $\mathbf{M}_{\text{nn}}^{-1}$ is calculated as:

$$m_{\text{nn},l} = \text{diag} \left(g_{ll} - \mathbf{g}'_{lc} \mathbf{G}_{\text{cc}}^{-1} \mathbf{g}_{cl} \right) \quad [2.13]$$

where g_{ll} is the diagonal element of $\mathbf{G}_{\text{nn}}$ for the non-core animal l , and $\mathbf{g}_{lc}$ is the relationship between the non-core animal l and all the other animals in the core set. When APY is used, a direct inversion is only required for the matrix of relationships between core animals ($\mathbf{G}_{\text{cc}}$), which has a

small dimension. This algorithm reduces computational time from weeks to a few hours compared to the direct inversion of $\mathbf{G}$ (Masuda et al., 2016; Misztal, 2016). The APY has been successfully applied in dairy cattle (Tsuruta et al., 2021; Cesarani et al., 2022), beef cattle (Lourenco et al., 2015), pigs (Lourenco et al., 2016), and broiler chicken (Simeone et al., 2012; Bermann et al., 2021).

Another strategy is removing genotyped animals that no longer contribute to evaluations, such as inactive bulls or young animals without records. These animals' GEBVs are replaced with indirect predictions (IP) derived from SNP effects calculated in previous evaluations. This reduces computational costs while maintaining reliable predictions (Garcia et al., 2020; Bermann et al., 2022a). Calculating the IP of those animals is as follows:

$$\hat{\mathbf{a}} = \lambda \mathbf{DZ}'\mathbf{G}^{-1}\hat{\mathbf{u}} \quad [2.14]$$

Where $\hat{\mathbf{a}}$ is a vector of SNP effects, $\hat{\mathbf{u}}$ is a vector of GEBVs, $\lambda = \frac{\sigma_a^2}{\sigma_u^2}$ is the ratio of SNP variance σ_a^2 to additive genetic variance σ_u^2 , $\mathbf{D}$ is a diagonal matrix of SNP weights ($\mathbf{D} = \mathbf{I}$), and $\mathbf{Z}$ is a matrix of centered gene content. Once $\hat{\mathbf{a}}$ is available, IP can be computed as $\mathbf{IP} = \mathbf{Z}\hat{\mathbf{a}}$ for any genotyped animals, mainly those removed from the evaluation. Many studies investigated the use of IP from a ssGBLUP evaluation with (Lourenco et al., 2018; Garcia et al., 2020; Bermann et al., 2022a) or without (Lourenco et al., 2015) APY and found that using IP can be beneficial as they can be used as quick genomic predictions for certain animals without running a complete evaluation.

Similarly, if IPs are available, it is valuable to have a reliability measure comparable to that of GEBVs. Liu (2017) demonstrated how to derive SNP effects' prediction error covariance (PEC) in a SNP BLUP model. Since SNP-BLUP and GBLUP are equivalent models, the same approach

can be applied to SNP effects obtained using ssGBLUP. Derivations for SNP PEC under ssGBLUP were described by Gualdrón Duarte et al. (2014) and Aguilar et al. (2019):

$$\text{Var}(\hat{\mathbf{a}}) = \frac{1}{2 \sum p_i(1 - p_i)} \mathbf{Z}' \mathbf{G}^{-1} (\mathbf{G} \sigma_u^2 - \mathbf{C}^{\mathbf{u}_2 \mathbf{u}_2}) \mathbf{G}^{-1} \mathbf{Z} \frac{1}{2 \sum p_i(1 - p_i)} \quad [2.15]$$

With $\mathbf{C}^{\mathbf{u}_2 \mathbf{u}_2}$ the GEBVs prediction error variance matrix for genotyped animals. Given the computational challenges of obtaining $\mathbf{G}^{-1}$, the sparse approximation $\mathbf{G}_{\text{APY}}^{-1}$ which requires less computational demand can be used (Garcia et al., 2022). Bermann et al. (2022a) showed that, with APY, estimating $\text{Var}(\hat{\mathbf{a}})$ can be reduced to components associated with the core animals:

$$\text{Var}(\hat{\mathbf{a}}) = \frac{1}{2 \sum p_i(1 - p_i)} \mathbf{Z}'_c \mathbf{G}_{\text{cc}}^{-1} (\mathbf{G}_{\text{cc}} \sigma_u^2 - \mathbf{C}^{\mathbf{u}_{2c} \mathbf{u}_{2c}}) \mathbf{G}_{\text{cc}}^{-1} \mathbf{Z}_c \frac{1}{2 \sum p_i(1 - p_i)} \quad [2.16]$$

where $\mathbf{Z}_c$ is the centered gene content matrix for core animals, and $\mathbf{C}^{\mathbf{u}_{2c} \mathbf{u}_{2c}}$ is the GEBVs' prediction error variance for core animals. However, computing $\mathbf{C}^{\mathbf{u}_{2c} \mathbf{u}_{2c}}$, still requires inverting a high-dimensional coefficient matrix. Bermann et al. (2022b) presented an approximation to obtain the PEV:

$$\mathbf{C}_{\text{approx}}^{\mathbf{u}_{2c} \mathbf{u}_{2c}} = (\mathbf{D}_c + \lambda \mathbf{G}^{\text{cc}} - \lambda \mathbf{G}^{\text{cn}} (\lambda \mathbf{M}_{\text{nn}}^{-1} + \mathbf{D}_n)^{-1} \lambda \mathbf{G}^{\text{nc}})^{-1} \quad [2.17]$$

$\mathbf{D}_c$ and $\mathbf{D}_n$ are the core and non-core blocks of the diagonal matrix based on effective record contributions and $\lambda = \frac{\sigma_e^2}{\sigma_u^2}$.

Substituting this approximation, $\text{Var}(\hat{\mathbf{a}})$ becomes:

$$\text{Var}(\hat{\mathbf{a}}) = \frac{1}{2 \sum p_i(1 - p_i)} \mathbf{Z}'_c \mathbf{G}_{\text{cc}}^{-1} (\mathbf{G}_{\text{cc}} \sigma_u^2 - \mathbf{C}_{\text{approx}}^{\mathbf{u}_{2c} \mathbf{u}_{2c}}) \mathbf{G}_{\text{cc}}^{-1} \mathbf{Z}_c \frac{1}{2 \sum p_i(1 - p_i)} \quad [2.18]$$

Finally, the reliability of IP for animal l is calculated as:

$$REL_{IP_l} = 1 - \frac{\mathbf{z}_l \text{Var}(\hat{\mathbf{a}}) \mathbf{z}_l'}{\text{Var}(\mathbf{u})} \quad [2.19]$$

where $\mathbf{z}_l$ is the gene content vector for animal l , and $\text{Var}(\mathbf{u})$ is the additive ge. Chapter 5 evaluates the performance of this approximation by comparing it to a genomic reliability approximation method presented by Bermann et al. (2022b).

Computational limitations also affect MME inversions, even in pre-genomic evaluations. Iterative methods like Gauss-Seidel or Preconditioned Conjugate Gradient can solve MME without direct inversion but cannot calculate prediction error variances (PEV) or reliabilities. Misztal and Wiggans (1988) presented an iterative algorithm to approximate PEV in an animal model suitable for large non-genomic data sets. The approximation calculates each animal's effective number of records and pedigree contributions by ignoring the off-diagonals and absorbing the non-genetic effect model (i.e., subclasses, permanent environmental effect). Another approximation method proposed by Tier and Meyer (2004) approximates the PEV sequentially. First, diagonals for animals with records are adjusted for a limited subclass size (i.e., CG), then the diagonal values of parents are updated based on limited information from their offspring, and finally, the offspring's diagonal values are adjusted according to the adjusted values of their parents ensuring no double-counting of the same animal. However, these approximations may lead to overestimation, particularly in models with several effects. Chapter 6 explores solutions to mitigate such issues, ensuring better and more efficient approximations in genomic evaluations.

VALIDATION OF GENOMIC EVALUATIONS

Measuring the quality of breeding values is crucial for animal breeding. Frequent genetic evaluation changes require tools to assess how these changes impact predictions. Hence, developing methods to rank, understand, and quantify prediction model behavior is essential to breeders.

Cross-validation typically involves creating two datasets: a whole dataset (w), considered to provide “true” values by using all available information, and another partial dataset (p), which excludes information (e.g., phenotypic or genomic) to test how predictors from the partial dataset align with those from the whole dataset. Well-established methods exist for validating pedigree-based BLUP or multistep evaluations. For instance, the standard Interbull GEBV test regresses DYD obtained from pedigree-based BLUP using the whole dataset on GEBV from the partial dataset (Mäntysaari et al., 2010). However, the results of this method might be influenced by biased EBVs caused by preselection when selection is based on GEBVs (Patry and Ducrocq, 2009; Masuda et al., 2018; Jibrila et al., 2021). Also, deriving DYD for traits with low heritability or maternal effects is difficult (Legarra and Reverter, 2018).

Legarra and Reverter (2018) introduced the Linear Regression (LR) method to address these issues. This approach compares GEBVs from whole and partial datasets for focal animals (e.g., youngest genotyped animals or sires with a specific number of daughters with records). The main advantage of the LR method is its applicability to all traits, including low heritable and maternal effect traits, and its flexibility to analyze any type of animal (parents, youngest animals). This method calculates four main statistics:

- 1- Accuracy: refers to population accuracy, a function of the correlation between the true breeding value (TBV) and GEBV, indicating the magnitude of change in GEBV with additional information.
- 2- Bias: is the mean difference of GEBVs of the focal animals between the two datasets, with an expected value of zero indicating unbiased predictions.
- 3- Dispersion: Assessed via the regression coefficient (b_1) of GEBVs from the whole dataset on the partial dataset; an expected value of 1 suggests no over- or under-dispersion.

- 4- Correlation: represents the correlation between the GEBVs of focal animals from both datasets, reflecting the relationship between the two and indicating the expected changes over time.

These statistics are calculated using the following formulas:

$$accuracy = \sqrt{\frac{cov(\hat{\mathbf{u}}_w, \hat{\mathbf{u}}_p)}{(1 - \bar{F})\sigma_u^2}} \quad [2.20]$$

$$bias = \frac{\bar{\hat{\mathbf{u}}}_p - \bar{\hat{\mathbf{u}}}_w}{\sigma_u} \quad [2.21]$$

$$b_1 = \frac{cov(\hat{\mathbf{u}}_w, \hat{\mathbf{u}}_p)}{var(\hat{\mathbf{u}}_p)} \quad [2.22]$$

$$correlation = r(\hat{\mathbf{u}}_w, \hat{\mathbf{u}}_p) \approx \frac{acc_p}{acc_w} \quad [2.23]$$

where, $\hat{\mathbf{u}}_w$ and $\hat{\mathbf{u}}_p$ are the vectors of GEBVs for focal animals from whole and partial datasets, respectively, $\bar{F}$ is the average inbreeding coefficient of focal animals, σ_u^2 (σ_u) is the additive genetic variance (standard deviation); and $\bar{\hat{\mathbf{u}}}_w$ and $\bar{\hat{\mathbf{u}}}_p$ are the GEBV averages from whole and partial datasets of focal animals, respectively. A limitation of this method is the dependency between GEBVs from whole and partial datasets, which can influence results. To address this, VanRaden (2021) proposed the “improved genomic validation” method. It uses deregressed GEBV ($dGEBV$) in the regression, extracting new information gained in reliability between the partial and whole datasets. The $dGEBV$ is calculated as:

$$dGEBV = GEBV_p + \frac{GEBV_w - GEBV_p}{rel_{dif}} \quad [2.24]$$

where rel_{dif} is the difference in GEBV reliability between the whole and partial dataset calculated from the effective record contribution, $GEBV_f$ and $GEBV_p$ are the GEBVs obtained from the whole and partial dataset, respectively. The $dGEBV$ is then regressed on $GEBV_p$:

$$dGEBV = b_0 + b_1 GEBV_p + e \quad [2.25]$$

And b_1 measures GEBV dispersion.

Both methods have been compared and yielded similar results without change in the direction of the results (Himmelbauer et al., 2023).

REFERENCES

- Aguilar, I., A. Legarra, F. Cardoso, Y. Masuda, D. Lourenco, and I. Misztal. 2019. Frequentist p-values for large-scale-single step genome-wide association, with an application to birth weight in American Angus cattle. *Genetics Selection Evolution* 51(1):28. doi: 10.1186/s12711-019-0469-3
- Aguilar, I., I. Misztal, D. L. Johnson, A. Legarra, S. Tsuruta, and T. J. Lawlor. 2010. Hot topic: a unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *J Dairy Sci* 93(2):743-752. doi: 10.3168/jds.2009-2730
- Bermann, M., D. Lourenco, V. Breen, R. Hawken, F. Brito Lopes, and I. Misztal. 2021. Modeling genetic differences of combined broiler chicken populations in single-step GBLUP. *J Anim Sci* 99(4)doi: 10.1093/jas/skab056
- Bermann, M., D. Lourenco, N. S. Forneris, A. Legarra, and I. Misztal. 2022a. On the equivalence between marker effect models and breeding value models and direct genomic values with the Algorithm for Proven and Young. *Genetics Selection Evolution* 54(1):52. doi: 10.1186/s12711-022-00741-7
- Bermann, M., D. Lourenco, and I. Misztal. 2022b. Efficient approximation of reliabilities for single-step genomic best linear unbiased predictor models with the Algorithm for Proven and Young. *Journal of Animal Science* 100(1)doi: 10.1093/jas/skab353
- Cesarani, A., D. Lourenco, S. Tsuruta, A. Legarra, E. L. Nicolazzi, P. M. VanRaden, and I. Misztal. 2022. Multibreed genomic evaluation for production traits of dairy cattle in the United States using single-step genomic best linear unbiased predictor. *J Dairy Sci* 105(6):5141-5152. doi: 10.3168/jds.2021-21505

- Christensen, O. F., A. Legarra, M. S. Lund, and G. Su. 2015. Genetic evaluation for three-way crossbreeding. *Genetics Selection Evolution* 47(1):98. doi: 10.1186/s12711-015-0177-6
- Fragomeni, B. O., D. A. L. Lourenco, S. Tsuruta, Y. Masuda, I. Aguilar, A. Legarra, T. J. Lawlor, and I. Misztal. 2015. Hot topic: Use of genomic recursions in single-step genomic best linear unbiased predictor (BLUP) with a large number of genotypes. *Journal of Dairy Science* 98(6):4090-4094. doi: 10.3168/jds.2014-9125
- Garcia, A. L. S., Y. Masuda, S. Tsuruta, S. Miller, I. Misztal, and D. Lourenco. 2020. Indirect predictions with a large number of genotyped animals using the algorithm for proven and young. *J Anim Sci* 98(6)doi: 10.1093/jas/skaa154
- Garcia-Baccino, C. A., A. Legarra, O. F. Christensen, I. Misztal, I. Pocrnic, Z. G. Vitezica, and R. J. Cantet. 2017. Metafounders are related to F (st) fixation indices and reduce bias in single-step genomic evaluations. *Genet Sel Evol* 49(1):34. doi: 10.1186/s12711-017-0309-2
- Gualdrón Duarte, J. L., R. J. C. Cantet, R. O. Bates, C. W. Ernst, N. E. Raney, and J. P. Steibel. 2014. Rapid screening for phenotype-genotype associations by linear transformations of genomic evaluations. *BMC Bioinformatics* 15(1):246. doi: 10.1186/1471-2105-15-246
- Hazel, L. N. 1943. THE GENETIC BASIS FOR CONSTRUCTING SELECTION INDEXES. *Genetics* 28(6):476-490. doi: 10.1093/genetics/28.6.476
- Henderson, C. R. 1963. Selection Index and Expected Genetic Advance, *Statistical Genetics and Plant Breeding*. National Academy of Sciences No. 982. National Research Council Publication. p. 141-163.
- Henderson, C. R. 1973. In *Proceedings of Animal Breeding and Genetics Symposium in Honor of Dr Jay L. Lush*. In: *Animal Breeding and Genetics Symposium*. p 10-41.

- 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. doi: 10.2307/2529339
- Henderson, C. R. 1984. Applications of linear models in animal breeding
- Himmelbauer, J., H. Schwarzenbacher, C. Fuerst, and B. Fuerst-Waltl. 2023. Comparison of different validation methods for single-step genomic evaluations based on a simulated cattle population. *Journal of Dairy Science* 106(12):9026-9043. doi: 10.3168/jds.2023-23575
- Jibrila, I., J. Vandenplas, J. Ten Napel, R. F. Veerkamp, and M. P. L. Calus. 2021. Avoiding preselection bias in subsequent single-step genomic BLUP evaluations of genomically preselected animals. *J Anim Breed Genet* 138(4):432-441. doi: 10.1111/jbg.12533
- Kluska, S., Y. Masuda, J. B. S. Ferraz, S. Tsuruta, J. P. Eler, F. Baldi, and D. Lourenco. 2021. Metafounders May Reduce Bias in Composite Cattle Genomic Predictions. *Front Genet* 12:678587. doi: 10.3389/fgene.2021.678587
- Lande, R., and R. Thompson. 1990. Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics* 124(3):743-756. doi: 10.1093/genetics/124.3.743
- Legarra, A., I. Aguilar, and I. Misztal. 2009. A relationship matrix including full pedigree and genomic information. *J Dairy Sci* 92(9):4656-4663. doi: 10.3168/jds.2009-2061
- Legarra, A., O. F. Christensen, Z. G. Vitezica, I. Aguilar, and I. Misztal. 2015. Ancestral Relationships Using Metafounders: Finite Ancestral Populations and Across Population Relationships. *Genetics* 200(2):455-468. doi: 10.1534/genetics.115.177014
- Legarra, A., D. A. Lourenco, and Z. G. Vitezica. 2018. Bases for genomic prediction. *Short course* 1(1):1-141.

- Legarra, A., and A. Reverter. 2018. Semi-parametric estimates of population accuracy and bias of predictions of breeding values and future phenotypes using the LR method. *Genet Sel Evol* 50(1):53. doi: 10.1186/s12711-018-0426-6
- Liu, Z. 2017. Approximating genomic reliabilities for national genomic evaluation. *Interbull Bulletin* (51)
- Lourenco, D. A., S. Tsuruta, B. O. Fragomeni, Y. Masuda, I. Aguilar, A. Legarra, J. K. Bertrand, T. S. Amen, L. Wang, D. W. Moser, and I. Misztal. 2015. Genetic evaluation using single-step genomic best linear unbiased predictor in American Angus. *J Anim Sci* 93(6):2653-2662. doi: 10.2527/jas.2014-8836
- Lourenco, D. A. L., S. Tsuruta, B. O. Fragomeni, C. Y. Chen, W. O. Herring, and I. Misztal. 2016. Crossbreed evaluations in single-step genomic best linear unbiased predictor using adjusted realized relationship matrices¹. *Journal of Animal Science* 94(3):909-919. doi: 10.2527/jas.2015-9748
- Mäntysaari, E., Z. Liu, and P. M. Van Raden. 2010. Interbull validation test for genomic evaluations. In: *Interbull Annual Meeting*. p 5.
- Masuda, Y., I. Misztal, S. Tsuruta, A. Legarra, I. Aguilar, D. A. L. Lourenco, B. O. Fragomeni, and T. J. Lawlor. 2016. Implementation of genomic recursions in single-step genomic best linear unbiased predictor for US Holsteins with a large number of genotyped animals. *Journal of Dairy Science* 99(3):1968-1974. doi: <https://doi.org/10.3168/jds.2015-10540>
- Masuda, Y., P. M. VanRaden, I. Misztal, and T. J. Lawlor. 2018. Differing genetic trend estimates from traditional and genomic evaluations of genotyped animals as evidence of preselection bias in US Holsteins. *Journal of Dairy Science* 101(6):5194-5206. doi: <https://doi.org/10.3168/jds.2017-13310>

- Masuda, Y., P. M. VanRaden, S. Tsuruta, D. A. L. Lourenco, and I. Misztal. 2022. Invited review: Unknown-parent groups and metafounders in single-step genomic BLUP. *Journal of Dairy Science* 105(2):923-939. doi: 10.3168/jds.2021-20293
- Meuwissen, T. H., B. J. Hayes, and M. E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157(4):1819-1829. doi: 10.1093/genetics/157.4.1819
- Meyer, K. 1987. Estimates of variances due to sire $\times$ herd interactions and environmental covariances between paternal half-sibs for first lactation dairy production. *Livestock Production Science* 17:95-115.
- Misztal, I. 2016. Inexpensive Computation of the Inverse of the Genomic Relationship Matrix in Populations with Small Effective Population Size. *Genetics* 202(2):401-409. doi: 10.1534/genetics.115.182089
- Misztal, I., A. Legarra, and I. Aguilar. 2009. Computing procedures for genetic evaluation including phenotypic, full pedigree, and genomic information. *J Dairy Sci* 92(9):4648-4655. doi: 10.3168/jds.2009-2064
- Misztal, I., A. Legarra, and I. Aguilar. 2014. Using recursion to compute the inverse of the genomic relationship matrix. *J Dairy Sci* 97(6):3943-3952. doi: 10.3168/jds.2013-7752
- Misztal, I., Z. G. Vitezica, A. Legarra, I. Aguilar, and A. A. Swan. 2013. Unknown-parent groups in single-step genomic evaluation. *J Anim Breed Genet* 130(4):252-258. doi: 10.1111/jbg.12025
- Misztal, I., and G. R. Wiggans. 1988. Approximation of Prediction Error Variance in Large-Scale Animal Models. *Journal of Dairy Science* 71:27-32. doi: [https://doi.org/10.1016/S0022-0302\(88\)79976-2](https://doi.org/10.1016/S0022-0302(88)79976-2)

- Norman, H. D. 1974. Factors That Should Be Considered in a National Sire Summary Model1. Journal of Dairy Science 57(8):955-962. doi: [https://doi.org/10.3168/jds.S0022-0302\(74\)84992-1](https://doi.org/10.3168/jds.S0022-0302(74)84992-1)
- Patry, C., and V. Ducrocq. 2009. Bias due to genomic selection. In: Interbull workshop on genomic selection, NA, France. p 77-82.
- Quaas, R. L. 1976. Computing the Diagonal Elements and Inverse of a Large Numerator Relationship Matrix. Biometrics 32(4):949-953. doi: 10.2307/2529279
- Quaas, R. L. 1988. Additive Genetic Model with Groups and Relationships. Journal of Dairy Science 71(5):1338-1345. doi: [https://doi.org/10.3168/jds.S0022-0302\(88\)79691-5](https://doi.org/10.3168/jds.S0022-0302(88)79691-5)
- Quaas, R. L., and E. J. Pollak. 1981. Modified Equations for Sire Models with Groups. Journal of Dairy Science 64(9):1868-1872. doi: [https://doi.org/10.3168/jds.S0022-0302\(81\)82778-6](https://doi.org/10.3168/jds.S0022-0302(81)82778-6)
- Simeone, R., I. Misztal, I. Aguilar, and Z. G. Vitezica. 2012. Evaluation of a multi-line broiler chicken population using a single-step genomic evaluation procedure. J Anim Breed Genet 129(1):3-10. doi: 10.1111/j.1439-0388.2011.00939.x
- Strandén, I., G. P. Aamand, and E. A. Mäntysaari. 2022. Single-step genomic BLUP with genetic groups and automatic adjustment for allele coding. Genetics Selection Evolution 54(1):38. doi: 10.1186/s12711-022-00721-x
- Tier, B., and K. Meyer. 2004. Approximating prediction error covariances among additive genetic effects within animals in multiple-trait and random regression models. Journal of Animal Breeding and Genetics 121(1):77-89. doi: 10.1111/j.1439-0388.2003.00444.x
- Tsuruta, S., D. A. L. Lourenco, Y. Masuda, T. J. Lawlor, and I. Misztal. 2021. Reducing computational cost of large-scale genomic evaluation by using indirect genomic prediction. JDS Communications 2(6):356-360. doi: <https://doi.org/10.3168/jdsc.2021-0097>

- Tsuruta, S., D. A. L. Lourenco, Y. Masuda, I. Misztal, and T. J. Lawlor. 2019. Controlling bias in genomic breeding values for young genotyped bulls. *J Dairy Sci* 102(11):9956-9970. doi: 10.3168/jds.2019-16789
- VanRaden, P. M. 2008. Efficient methods to compute genomic predictions. *J Dairy Sci* 91(11):4414-4423. doi: 10.3168/jds.2007-0980
- VanRaden, P. M. 2021. Improved genomic validation including extra regressions. *Interbull Bulletin* 56:65-69.
- VanRaden, P. M., C. P. Van Tassell, G. R. Wiggans, T. S. Sonstegard, R. D. Schnabel, J. F. Taylor, and F. S. Schenkel. 2009. Invited review: reliability of genomic predictions for North American Holstein bulls. *J Dairy Sci* 92(1):16-24. doi: 10.3168/jds.2008-1514
- Wiggans, G. R., and J. A. Carrillo. 2022. Genomic selection in United States dairy cattle. *Front Genet* 13:994466. doi: 10.3389/fgene.2022.994466

CHAPTER 3

COMBINING LARGE BROILER POPULATIONS INTO A SINGLE GENOMIC

EVALUATION: DEALING WITH GENETIC DIVERGENCE¹

¹Joe-Menwer Tabet, Fernando Bussiman, Vivian Breen, Ignacy Misztal, Daniela Lourenco.
Submitted to Journal of Animal Science.

ABSTRACT

Combining breeding populations that have diverged at some point is a conventional practice, particularly in the poultry industry, where generation intervals are short and genetic evaluations should be frequently available. This study aimed to assess the feasibility of combining large, distantly genetically connected broiler populations into a single genomic evaluation within the single-step GBLUP framework. The pedigree data for broiler lines 1 and 2 consisted of 428,790 and 477,488 animals, being 156,088 and 186,387 genotyped, respectively. Phenotypic data for Body weight (kg), Carcass Yield (%), Mortality (1-2), and Feet Health (1-7) were collected for 397,974 animals in line 1 and 458,881 in line 2. A four-trait model was employed for the analyses, and genetic differences between the populations were addressed through different approaches: introducing an additional fixed effect accounting for the line of origin (M2) or making each fixed effect origin-specific (M3). Those models were compared against a conventional model (M1) that did not account for animal origin in the evaluation. Unknown parent groups (UPG) and Metafounders (MF) were fit to account for the genetic differences in M1, M2, and M3; they were set based on the animal's line of origin and sex. Accuracy, bias, and dispersion were used to assess the performances of the models using the Linear Regression method. Validations were performed separately within individual lines and collectively after combining the two lines to better assess the advantages of combining the two populations. Overall, the accuracy increased when the two populations were combined compared to the accuracies obtained from evaluating each line individually. Notably, there were no apparent differences among the models regarding accuracy and dispersion. Regarding bias, using models M2 or M3 with UPG yielding the least biased estimates in the combined evaluation. Thus, when combining different populations into a single

genomic evaluation, accounting for the genetic and non-genetic differences among the lines ensures accurate and less biased predictions.

INTRODUCTION

Commercial breeding programs often combine populations with diverse genetic or environmental backgrounds to simplify the genetic evaluation process, obtain breeding values in a multi-population base, increase diversity, and enhance genetic progress. Combining all available information has the potential to yield more accurate and unbiased predictions (VanRaden et al., 2007). Selecting a proper model becomes essential when incorporating different populations into the same evaluation. Neglecting these differences could result in biased Genomic Estimated Breeding Values (GEBV) (Macedo et al., 2020), compromising the selection of parents of future generations. Among the strategies that exist to address these genetic and non-genetic differences in the evaluations is, introducing a fixed effect accounting for the animal's origin (Bermann et al., 2021). Correctly accounting for genetic and non-genetic differences between populations is important when combining different populations into a joint evaluation; a study by Cesarani et al. (2022), where levels of fixed effects were redefined to be defined within-breed. They showed that failure to do so resulted in lower validation reliabilities and increased inflation of GEBV. Within-population fixed effects consider that different populations, breeds, or lines respond individually to identifiable environmental factors.

Another method is the use of Unknown Parent Groups (UPG) (Thompson, 1979; Quaas, 1988) which has been commonly used in animal breeding schemes to overcome issues related to missing pedigree since not accounting for that leads to assigning the same genetic base for animals without parents, which is not correct (Westell et al., 1988). These UPG are usually assigned based

on criteria like breed, line, sex, birth year, and selection pathway, effectively representing the average genetic value of animals within the same genetic group. Beyond addressing missing pedigree information, UPG can also accommodate breed differences (Legarra et al., 2007).

An alternative approach proposed by Legarra et al. (2015) involves Metafounders (MF), a method that uses genotypes to assess population differences. While UPG assumes an unrelated founder population, MF take into account the existence of relationships within and across these founder populations. The relatedness within and between MF is quantified by a matrix gamma (Γ), reflecting the relationships within and between populations. Under a single-trait model, MF have exhibited better results when compared to other methods to model missing pedigree in simulated data (Bradford et al., 2019). Garcia-Baccino et al. (2017) showcased that by including MF in the genetic model, the bias of genomic prediction could be reduced without any loss of accuracy.

In the MF approach, the genomic relationship matrix $\mathbf{G}$ is centered by 0.5 allele frequency, and Γ is used to create connections among populations through the pedigree relationship matrix for all ($\mathbf{A}$) and genotyped ($\mathbf{A}_{22}$) animals. In the UPG approach, $\mathbf{G}$ is often centered by current allele frequencies. When genotypes from different populations are combined, an overall current allele frequency is used; however, this can create some incompatibilities between $\mathbf{G}$ and $\mathbf{A}_{22}$. Makgahlela et al. (2014) and Lourenco et al. (2016) used population-specific allele frequencies to construct $\mathbf{G}$ when combining different populations. Although it improved the compatibility of $\mathbf{G}$ and $\mathbf{A}_{22}$, there was no increase in accuracy or reduction in bias. Bermann et al. (2021) reported that MF and UPG were enough to account for population differences. However, in their study, a very small foreign population (1,484 individuals) was combined with the main breeding population (240,929 individuals), which may not well represent the current practice.

The primary objective of this study was to develop a strategy for incorporating two different broiler populations of comparable size into one single genomic evaluation using cross-classified effects accounting for animal origin, UPG, or MF. Then, we assessed the accuracy of prediction, bias, and dispersion across these models, aiming for faster, unbiased predictions.

MATERIALS AND METHODS

Animal care and use committee approvals were unnecessary because data were obtained from preexisting databases.

Data

Cobb-Vantress Inc. (Siloam Springs, AR) provided the data used in this study. The pedigree includes 906,278 animals from 2 different lines (Line 1= 428,790 and Line 2= 477,488 animals). In the broiler industry, it is common to introduce external genetics into a population to increase variability and benefit from heterosis. The two populations involved consist of 47 mating groups (MG) each, with mating groups representing different generations. The two genetic lines were crossed from MG 21 to MG 29, and their offspring were classified as part of line 2 (i.e., crossbred progeny was treated as purebred from line 2). After MG 29, the two populations were separated again and bred independently while remaining in the same environment. The animals were genotyped with a 60k single nucleotide polymorphisms (SNP) panel with a total number of genotyped animals equal to 342,475 (line 1 = 156,088 and line 2 = 186,387). The traits under evaluation were body weight (BW) in kg (recorded once at 35 days of age), carcass yield (CY) in percentage, mortality (MT) as either 1 (alive) or 2 (dead), and leg health (LH) with seven categories ranging between 1 (normal) and 7 (severely abnormal). Table 1 presents a summary statistic of the lines 1 and 2 of each trait used in this study.

Quality control on genotypes was done through PREGSF90 software (Misztal et al., 2014b), where SNP or animals with call rate smaller than 0.90, minor allele frequency less than 0.05, departure from Hardy-Weinberg equilibrium (difference between observed and expected heterozygous frequency) higher than 0.15 or monomorphic SNP were removed. Mendelian conflicts (discrepant homozygous SNP) were assessed in parent-progeny pairs. SNPs were excluded if the conflict rate exceeded 10% across all pairs, and progenies were removed if the rate exceeded 1% of the total SNP. Also, animals with duplicated genotypes were excluded, resulting in 49,088 markers and 341,430 animals (line 1 = 155,043 and line 2 = 186,387).

Variance components estimation

Variance components were estimated using the Expectation-Maximization Restricted Maximum Likelihood algorithm implemented in the BLUPF90+ software (Misztal et al., 2014b). To estimate variance components, a four-trait pedigree-based animal model was used, including all phenotypic and pedigree information:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_1\mathbf{p}_m + \mathbf{Z}_2\mathbf{a} + \mathbf{e} \quad [3.1]$$

where: $\mathbf{y}$ is the vector of phenotypes (BW, CY, MT, and LH); $\boldsymbol{\beta}$ is the vector of fixed effects containing contemporary group and sex; $\mathbf{p}_m$ is the vector of maternal permanent environmental random effects (only for BW); $\mathbf{a}$ is a vector additive genetic random effects; $\mathbf{e}$ is a vector of random residuals; $\mathbf{X}$, $\mathbf{Z}_1$, and $\mathbf{Z}_2$ are incidence matrices relating the vector $\mathbf{y}$ to $\boldsymbol{\beta}$, $\mathbf{p}_m$, and $\mathbf{a}$, respectively. Random effects were assumed to follow a multivariate normal distribution with a mean zero and the following (co)variance structure:

$$\text{Var} \begin{bmatrix} \mathbf{p}_m \\ \mathbf{a} \\ \mathbf{e} \end{bmatrix} = \begin{bmatrix} \mathbf{I} \otimes \boldsymbol{\Sigma}_{p_m} & \mathbf{0} & \mathbf{0} \\ \text{sym.} & \mathbf{A} \otimes \boldsymbol{\Sigma}_a & \mathbf{0} \\ & & \mathbf{I} \otimes \boldsymbol{\Sigma}_e \end{bmatrix} \quad [3.2]$$

where: $\mathbf{A}$ is the pedigree relationship matrix, $\mathbf{I}$ is an identity matrix of proper order, $\otimes$ is the Kronecker product, and Σ_{pm} , Σ_a , and Σ_e are the (co)variance matrices of maternal permanent environmental effects, additive genetic effects, and residuals, respectively, with the following structure:

$$\Sigma_{pm} = \begin{bmatrix} \sigma_{pmBW}^2 & 0 & 0 & 0 \\ & 0 & 0 & 0 \\ & & 0 & 0 \\ \text{sym.} & & & 0 \end{bmatrix} \quad [3.3]$$

$$\Sigma_a = \begin{bmatrix} \sigma_{aBW}^2 & \sigma_{aBWaCY} & \sigma_{aBWaMT} & \sigma_{aBWaLH} \\ & \sigma_{aCY}^2 & \sigma_{aCYaMT} & \sigma_{aCYaLH} \\ & & \sigma_{aMT}^2 & \sigma_{aMTaLH} \\ \text{sym.} & & & \sigma_{aLH}^2 \end{bmatrix} \quad [3.4]$$

$$\Sigma_e = \begin{bmatrix} \sigma_{eBW}^2 & \sigma_{eBWecY} & \sigma_{eBWemT} & \sigma_{eBWelH} \\ & \sigma_{ecY}^2 & \sigma_{ecYemT} & \sigma_{ecYelH} \\ & & \sigma_{emT}^2 & \sigma_{emTelH} \\ \text{sym.} & & & \sigma_{elH}^2 \end{bmatrix} \quad [3.5]$$

Results of the variance component estimation are presented in Table 2.

Genomic estimated breeding values

Breeding values were estimated using the BLUP90IOD2OMP1 software (Misztal et al., 2014b) that implements single-step genomic BLUP (ssGBLUP) (Legarra et al., 2009; Aguilar et al., 2010; Christensen and Lund, 2010), combining pedigree and genomic relationships in a matrix $\mathbf{H}$, whose inverse ($\mathbf{H}^{-1}$) replaces the inverse of the pedigree relationship matrix ($\mathbf{A}^{-1}$) in the mixed model equations as in Aguilar et al. (2010). The genomic relationship matrix ($\mathbf{G}$) was constructed as the type I from VanRaden (2008). To obtain the same genetic base, mean diagonals and off-diagonals of $\mathbf{G}$ were adjusted to those from the pedigree relationship matrix for genotyped animals

($\mathbf{A}_{22}$) (Chen et al., 2011). To avoid singularity problems, the $\mathbf{G}$ was blended with 5% of $\mathbf{A}_{22}$ (McWhorter et al., 2023).

The algorithm for proven and young (APY) (Misztal et al., 2014a) was used to invert $\mathbf{G}$ because the number of genotyped animals surpassed 50k. The number of core animals in this study was defined based on the dimensionality of the genomic information available as the number of largest eigenvalues explaining 99% of the variance in $\mathbf{G}$. When all the genotypes were used, meaning genotyped animals from both lines, the core dataset comprised 12,165 randomly selected genotyped animals. Regarding separate evaluations for lines 1 and 2 (Origin 1 and 2), the core dataset comprised 10,000 and 11,000 randomly selected animals, respectively.

Models

In this study, we proposed three four-trait models to combine the two lines into a single genomic evaluation.

The first model denoted as M1, did not account for any line or origin of animals and was the same as the model used for the variance components estimation (1) but included genomic information:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_1\mathbf{p}_m + \mathbf{Z}_2\mathbf{a} + \mathbf{e} \quad [3.6]$$

where all components were previously described.

In model 2 (M2), a vector of fixed effects was included in the model to account for the animal's line of origin as follows:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{W}\boldsymbol{\alpha} + \mathbf{Z}_1\mathbf{p}_m + \mathbf{Z}_2\mathbf{a} + \mathbf{e} \quad [3.7]$$

with $\boldsymbol{\alpha}$ representing a vector accounting for each animal's line of origin; the levels were determined based on the animal's parents (i.e., parental origin). In the end, five levels were defined to accommodate every combination of parental origin for each animal. Animals with both parents from the same line (i.e., from line 1 or 2) were included in their respective levels (level 1 = 397,975,

level 2 = 411,861). Animals with parents from different lines (crossbred progeny) were placed in separate levels; for instance, level 3 (N = 13,235) included animals that had their parents from different lines; in this case, the male parent was from line 1, and the female parent was from line 2, as for level 4 (N = 2,849) is the opposite case (male parents from 2 and female parents from line 1). Finally, animals with parents of unknown origin were grouped into one level (level 5 = 30,865).

W is an incidence matrix relating **y** to the vector **α** .

In model 3 (M3), the fixed effects (contemporary group and sex) were made line-specific by concatenating **α** to the effects in **β** . After running the above models to include both lines in the same evaluations, ssGBLUP evaluations were also run for each line separately according to Model 1. These models will be referred to as Origin 1 for line 1 evaluation and Origin 2 for line 2 evaluation. Origin 1 and Origin 2 were compared to a third model, referred to as Combined, which includes animals from both lines 1 and 2, excluding crossbreds between the two.

Unknown Parent Groups and Metafounders

UPG and MF can be used in mixed model equations to account for genetic differences among defined groups and missing pedigree information. UPG are assumed to be unrelated and non-inbred. MF rely on the idea that the founder population, in part, share gametes; hence, these populations share a common ancestor, and relationships among them exist, which are, in turn, ignored by UPG (Legarra et al., 2015).

Initially, UPG for ssGBLUP models were defined as Misztal et al. (2013), where the Quaas Pollack (QP) transformation (Quaas and Pollak, 1981) is applied to **A**, **G**, and **A₂₂**. In the context of MF, the relatedness of animals in the base population can be represented by γ in the case of a single population and **Γ** when more than one population is considered. Positive relationships between MF signify overlapping founder populations, whereas negative values indicate population

divergence (resulting from increased heterozygosity). A value of zero implies that the founder populations are unrelated. For multiple populations, $\mathbf{\Gamma}$ comprises within- and across-founder population relationships estimated using SNP markers through the generalized least squares approach, as described in Garcia-Baccino et al. (2017) through an in-house software called GAMMAF90. Legarra et al. (2024) showed that MF are a useful concept to characterize relationships within and across populations and can be interpreted as a measure of genetic similarities between populations. Garcia-Baccino et al. (2017) presented the relationship between γ and the fixation index F_{st} that is a measure of genetic differentiation due to the population structure, where $F_{st} = \frac{\frac{\gamma_b + \gamma_{b'} - \gamma_{bb'}}{8} - \frac{\gamma_{bb'}}{4}}{\frac{1}{2} - \frac{\gamma_{bb'}}{4}}$ with b and b' representing two different populations. This relationship holds when F_{st} is considered a measure of homozygosity.

In the pedigree, there were 4,383 animals with both (or only one) parents unknown, 2,147 from line 1 and 2,236 from line 2. These were the animals to which the UPG and MF were applied. Each method had two definitions, requiring four scenarios to be tested, including UPG and MF. The scenarios were noted as follows:

- I. UPG1 or MF1: UPG and MF were based on the animal's line of origin.
- II. UPG2 or MF2: UPG and MF were based on the animal's line of origin and sex, where the UPG or MF of the same line were separated by sex.

The number of animals in each defined group is shown in Table 3. In addition to the four scenarios described earlier, there was a scenario where genetic differences were not modeled, denoted as NOUPG.

Validation

After conducting the evaluations, all models and scenarios were compared by the Linear Regression method (LR) proposed by Legarra and Reverter (2018). This method involves first

creating two datasets: a whole dataset that contains all the available information and a partial dataset where we removed the phenotypes of focal animals, their siblings, and contemporaries. Focal animals were the youngest genotyped animals in the population. Next, three metrics were estimated: the prediction accuracy (acc_{LR}), the estimated bias, and the dispersion (b_1) for the focal animals. acc_{LR} is estimated as a function of the covariance between GEBV of focal animals from whole and partial datasets). This statistic reflects population accuracy rather than individual accuracies, and it holds significance when comparing the predictive ability of the models and scenarios.

$$acc_{LR} = \sqrt{\frac{\text{cov}(\hat{\mathbf{u}}_p, \hat{\mathbf{u}}_w)}{(1 - \bar{F})\hat{\sigma}_u^2}} \quad [3.8]$$

Bias is defined as the difference of means of the GEBV of the focal animals from both the partial and whole datasets, with an expected value of 0, indicating that the estimations are unbiased. In our study, bias is corrected by adjusting the genetic base to the mean of the genomic estimated breeding values for animals belonging to MG 15.

$$\text{bias} = \overline{\hat{\mathbf{u}}_p} - \overline{\hat{\mathbf{u}}_w} \quad [3.9]$$

Bias lower than 0 indicates underestimation, and higher than 0 means overestimation of partial GEBV.

The b_1 is the regression coefficient of the GEBV from the whole dataset on the GEBV of the partial dataset for the focal candidates, with an expected value of 1 indicating no over nor under-dispersion in the predictions.

$$b_1 = \frac{\text{cov}(\hat{\mathbf{u}}_w, \hat{\mathbf{u}}_p)}{\text{var}(\hat{\mathbf{u}}_p)} \quad [3.10]$$

Any deviation from the expected values would compromise the selection process and affect the genetic progress of a population. Overestimation of breeding values leads to the conception that young animals are superior selection candidates compared to older animals, leading to an erroneous selection of younger animals. The underestimation of breeding values results in the selection of a low number of young animals. Standard errors of the metrics were reported as described by Bermann et al. (2024).

All these statistics were computed for the focal animals, the youngest phenotyped and genotyped animals within the population. That led to selecting animals from the last seven MG from both lines. Table 4 presents the count of focal animals chosen for each trait and line.

Spearman's correlation coefficients were calculated for the focal animals to assess how their rankings, based on estimated breeding values, changed across different models and scenarios. Values close to one indicate that the rankings remained unchanged.

RESULTS AND DISCUSSION

Gamma estimation

The results of $\mathbf{\Gamma}$ based on the animal's line of origin (Line 1 or 2), were as follows:

$\mathbf{\Gamma} = \begin{bmatrix} 0.53 & 0.39 \\ 0.39 & 0.53 \end{bmatrix}$. The within-relationships (or level of homozygosity) for lines 1 and 2 were equal to 0.53. The across-line relationship between lines 1 and 2 was 0.39, indicating that these two populations are connected at the base level; their founders overlap and have not completely diverged. This result aligns with the history of the two populations, where they were connected at one point before separating again. The $\mathbf{\Gamma}$ estimations indicated a moderate correlation between the two. If the given matrix had been comprised of zeros, the model would be equivalent to the UPG model when using only the $\mathbf{A}$ matrix, where these correlations are ignored in the setup. Models

used to estimate GEBV typically assume that missing parents in the population are sampled from an unrelated, infinite base population. However, the estimated relationships suggest that the base populations of these two lines are related while still showing genetic diversity between them.

Another indicator of population genetic differentiation is the percentage of variation explained by the first principal component. To determine the number of core animals for each line, we performed a singular value decomposition (SVD) of **Z**, the matrix of centered gene content, which is equal to the square root of the eigenvalues of the genomic relationship matrix (**G**). We calculated the eigenvalues and obtained the first principal components for the following groups: lines 1 and 2 (with all crossbreds removed), crossbred animals alone, and all the genotyped animals together. The first principal component (PC1) explained for line 1: 1.16%, line 2: 4.46%, crossbred animals: 3.31%, and all animals together: 7.79% of the variation (Figure 1). The values between lines 1 and 2 show genetic differences between them. Line 2 showed a higher percentage that reflects the divergence of the two populations. The percentage of variation of PC1 for the crossbreds falls between values of line 1, as expected. The high percentage when all animals are considered together indicates genetic differentiation that adds more information about the differences between lines 1 and 2 even though they share some genetic overlap, as shown with the MF estimation. Due to computational limitations with large datasets (> 300k genotypes), we approximated the PCA plot (Figure 1) using FastPCA available on plink2 (Purcell and Chang; Chang et al., 2015).

Validation

In Figure 2, the Combined results refer to the analysis when both lines were evaluated together, ignoring crossbred animals. Origin 1 and Origin 2 represented evaluations when the complete dataset was divided into two sets based on the animal's line of origin, creating Origin 1

and Origin 2, respectively, which only contain animals that belong to each line. We present the results of the focal animals from M1 NOUPG to assess their performance when evaluated together or separately. When both lines were combined, there was an increase in prediction accuracy in all traits except for LH. This increase was most pronounced in BW, where the accuracy increased to 0.83, compared to Origin 1 (0.51) and Origin 2 (0.64) individually. In the case of LH, the combined evaluation yielded a prediction accuracy of 0.67 as opposed to Origin 1, with 0.39, and Origin 2, with 0.78. Combining the lines was advantageous for animals originating from line 1, but for line 2, it resulted in an accuracy decrease of 0.11. This increase is attributed to the increased information available for the analysis. A study by Kluska et al. (2021) found that the increase in prediction accuracy within the Montana composite population primarily resulted from increased genomic information rather than accounting for different breeds in the evaluation. These findings are consistent with the outcomes of the study, where combining both lines led to increased phenotypic and genotypic information.

Similar outcomes were observed for M2 and M3 with scenario NOUPG (Figure 3), where these models addressed non-genetic differences between the two lines by introducing a vector of fixed effects (M2) or concatenating the vector to the models' fixed effects (M3) accounting for the animals' line of origin. These adjustments showed equivalent levels of prediction accuracy. Consequently, there are no differences between the models regarding prediction accuracy when accounting for non-genetic differences between the two lines. This lack of differences could be attributed to both populations being under the same environment. Figure 4 compares the prediction accuracy of the three models when UPG1 was utilized. There was no difference between UPG and MF; both methods produced similar results. Therefore, only results for UPG1 will be presented

and considered a general representation of the prediction accuracy when the evaluation used genetic groups to account for the genetic differences between the two populations.

There was a marginal improvement in prediction accuracy upon incorporating UPG or MF in the evaluations. The highest increase was in BW, which rose from 0.83 (when no genetic differences were considered in M1/NOUPG; Figure 2) to 0.85 (when these differences were accounted for in M1/UPG1; Figure 4). When both non-genetic and genetic differences (UPG or MF) were included in M2 and M3, the prediction accuracy for BW increased from 0.83 to 0.87. Differences between M2 and M3 regarding the use of genetic groups, changes were observed across all traits except for MT, where there was a decrease of 0.02 with MF1 compared to UPG1.

Hence, considering genetic differences between the two lines within the models did not yield a significant increase in the prediction accuracy. Bradford et al. (2019) presented similar findings in their study, demonstrating that the increase in accuracy was primarily influenced by the trait's heritability rather than including MF or UPG in a ssGBLUP evaluation. Their results align with the outcomes of this study, as they observed minor improvements in accuracies attributed to the inclusion of MF in their models. Kluska et al. (2021) reported that implementing MF did not improve accuracies within a composite beef population. However, it did result in reduced dispersion levels for certain traits. Additionally, the model exhibited the least biased genomic prediction. Also, when examining the pattern of missing parents in the pedigree, we found that nearly all phenotyped animals had known parents. However, missing pedigree information was more common in earlier generations, which could explain the minimal impact of UPG and MF on prediction accuracy.

Figure 5 presents b_1 for focal animals when both lines were combined in the same evaluation in M1 with scenario NOUPG (Combined) alongside the b_1 of focal animals when each

population was evaluated separately (Origin 1 and Origin 2). For Combined evaluation, the regression coefficient increased for BW, CY, and LH, with the highest increase observed in LH, going from 0.93 in Origin 1 and 2 to 0.99. As for MT, the regression coefficient b_1 in Combined was equal to Origin 2 (0.81); for Origin 1, the coefficient was equal to 0.72. This highlights that line 1 benefited from including both lines as their GEBV became less inflated.

When evaluations were conducted separately, there was no over- or under-dispersion of GEBV for BW, CY, and LH. However, when the two populations were combined, the coefficient approached one. This trend was consistent across all traits, highlighting the feasibility of an evaluation that includes both populations.

Figure 6 presents the estimated dispersion levels of GEBV for focal animals across the three models under the MF2 scenario, as dispersion was consistent across all scenarios. No dispersion (b_1 approaching 1) was observed for BW, CY, and LH. All three models consistently exhibited negligible dispersion of GEBV, with values hovering around 1. This suggests that considering genetic differences (UPG or MF) led to dispersion levels approaching 1, indicating no inflation (nor deflation) of the GEBV.

b_1 remained consistent across all scenarios, whether or not genetic differences were considered in the models. In the NOUPG scenario, dispersion values were close to 1, showing minimal GEBV inflation across models (M1-M2-M3/NOUPG). Across all scenarios, M1 consistently showed higher values than M2 and M3, with differences ranging from 0.01 to 0.04. This consistency suggests that not accounting for non-genetic differences produced similar results to when these effects were considered. However, the MF1 and MF2 scenarios showed slightly higher values than NOUPG, indicating that including genetic differences helped reduce GEBV inflation.

In scenarios accounting for genetic differences, higher levels of b_1 were consistently attained with MF than with UPG. The estimated coefficients were 0.01 to 0.02 higher with MF1/2 than UPG1/2 in M2 and M3 for BW and CY. This observation implies that MF generally led to values slightly closer to one (Supplemental Table S1).

Across all traits, the regression coefficient maintained a stable value around one, indicating no over- or under-dispersion; however, for MT, focal animals were consistently over-dispersed GEBV across all models and scenarios. For MT, the highest regression coefficient (0.81) was observed in NOUPG and MF2 scenarios in M1. The intricate nature of the trait, the ambiguity surrounding its definition, and the fact that only alive animals had genotypes may have contributed to the excessive overdispersion. One concern regarding the inclusion of genetically divergent populations into the same evaluation is the GEBV dispersion, which can occur due to the populations being subject to different selection purposes (Cesarani et al., 2022). However, this study showed that including both lines together led to no over- or under-dispersion of GEBV.

Figure 7 presents the bias obtained from separate evaluations of lines 1 and 2 using the whole dataset (i.e., Combined partial dataset vs. Origin 1 whole dataset and Combined partial dataset vs. Origin 2 whole dataset). Note that corresponding focal animals from Combined and Origin datasets were compared to assess if evaluating both populations alters the scale of GEBVs. Deviations from zero, especially in BW and CY, were observed. For BW, the bias between the partial and whole datasets from Origin 1 was 0.04, whereas the bias using the Combined partial dataset and Origin 1 whole dataset was -0.61. This suggests that the GEBV generated by each evaluation operates on different scales when combined or separated. This is a significant finding as the GEBV scale is crucial when selecting schemes because it would allow a fair comparison of GEBV across the population and informed selection decisions (Piccoli et al., 2018). Consequently,

to account for this difference in scales, an across-line adjustment is necessary for the breeding values before proceeding with the selection based on the outcomes of the combined evaluation.

Figure 8 displays the distributions of GEBV for BW and MT from the evaluation of both lines together. CY and LH exhibited similar results, so only the distributions of BW and MT were presented. The distributions show that α in M2 and M3 led to means approaching zero, indicating less biased estimates than M1. These findings align with Bermann et al. (2021), who fitted a vector of fixed effects in their model to account for differences between broiler populations, resulting in unbiased and less dispersed genomic estimated breeding values. These visual representations in Figure 8 underscore the effectiveness of incorporating such adjustments in mitigating bias and enhancing the prediction accuracy.

When examining scenarios within each model, UPG and MF consistently led to distributions centered around zero. This trend was observed in M2 for all scenarios accounting for genetic differences and in M3 for scenarios UPG1, UPG2, and MF1. It is widely recognized that incorporating UPG in genetic evaluations reduces bias by addressing genetic differences between distinct populations, particularly in cases of strongly selected populations. However, caution is warranted in defining UPG, as an incorrect assignment of groups may lead to suboptimal estimations and subsequent bias. Consequently, accurate estimation of UPG with sufficient information is crucial to prevent potential issues (Tsuruta et al., 2014; Tsuruta et al., 2019). Kluska et al. (2021) showed that MF can reduce bias in genomic evaluations, as seen in the same study where MF led to the least biased GEBV.

However, in MF2 of M3, the distributions deviate from zero for BW (Figure 8) and CY (results not shown), with a slight deviation for MT, indicating bias in the estimates for these traits. This deviation could be due to the low number of genotyped animals with phenotypes linked to

MF within each line. This led to inaccurate estimates of MF effects, mainly when MF were defined based on the line of origin and sex. As for UPG, the estimates rely solely on phenotyped animals. Bermann et al. (2021) reported similar findings, where genomic predictions displayed slight bias and under-dispersion with MF, attributed to the uneven distribution of observations among MF. Kluska et al. (2021) showed that MF benefitted the prediction by alleviating the inflation of GEBV in some traits. Furthermore, they emphasized the significance of accurately estimating the effects of MF in evaluations. This accuracy is contingent upon the matrix ($\mathbf{\Gamma}$), which, in turn, relies on the number of genotyped animals with phenotypes associated with MF.

This study introduces three distinct models to address the joint evaluation of two different lines. M1 does not account for population differences between the two lines. In M2, a vector of fixed effects (α) is introduced to accommodate these differences. In M3, the fixed effects are concatenated with the vector (α), transforming them into line-specific fixed effects. The most effective strategy was to include a vector of fixed effects accounting for the line of origin. No differences between M2 and M3, as both account for the non-genetic discrepancies, resulting in more accurate and less biased GEBV. These findings align with Cesarani et al. (2022), who obtained less biased and more precise predictions by incorporating breed-specific effects within their models in a multibreed evaluation of U.S. dairy cattle.

Incorporating UPG or MF, addresses the genetic differences between these lines. Missing pedigree information and uncertainties regarding the structure of the base of the population can introduce biases in estimation, potentially compromising the selection process. UPG or MF offers a constructive solution by accounting for the genetic differences between these populations, with MF accounting for relationships within and between these genetic groups, which UPG ignores (Legarra et al., 2015). Given the history of the populations, it is essential to consider the differences

in the base population to avoid biases in the genetic trends that can accumulate and increase among selection candidates (Macedo et al., 2022).

Another approach for comparing models and scenarios is the Spearman correlation among the focal animals. This correlation assesses the similarity between the ranks of the focal animals across the different models and scenarios. A low correlation would suggest changes in ranking, disrupting the existing selection process when updating the evaluation models. Figure 9 presents a correlation plot across all models and scenarios only for BW since all the traits yielded similar results, with a scale ranging from 0.98 to 1.00. The lowest correlation was 0.98 between M3/MF1 and M1/NOUPG. This shows no changes in the rank when both lines are included in the same evaluation.

CONCLUSIONS

The primary aim of this study was to propose an efficient strategy for including two genetically divergent broiler populations in the same ssGBLUP evaluation. This approach aims to accelerate the evaluation process, enabling quicker breeding and management decisions while obtaining better accuracy and reducing bias. Accounting for non-genetic differences (M2 and M3) was proven to be an effective strategy, where comparable levels of prediction accuracy were achieved. Incorporating UPG or MF based on line of origin and sex highlighted the importance of accounting for genetic differences between the lines. While MF exhibited reduced bias and dispersion, proper estimation of their effects and accurate definition of groups are important to achieve optimal results. For instance, in MF2, biased estimations were observed, suggesting that UPG is a suitable alternative. Attention should be paid to the different scales of the GEBV between

the two lines, to ensure fair and reasonable candidate selection, a within-line adjustment before the selection process is recommended.

ACKNOWLEDGMENTS

This study was supported by Cobb-Vantress, Inc (Siloam Springs, AR). We appreciate discussions from Fernando Brito Lopes during this study as well as other Cobb-Vantress Inc. team for collecting the data.

REFERENCES

- Aguilar, I., I. Misztal, D. L. Johnson, A. Legarra, S. Tsuruta, and T. J. Lawlor. 2010. Hot topic: a unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *J Dairy Sci* 93(2):743-752. doi: 10.3168/jds.2009-2730
- Bermann, M., A. Legarra, A. A. Munera, I. Misztal, and D. Lourenco. 2024. Confidence intervals for validation statistics with data truncation in genomic prediction. *Genet Sel Evol* 56(1):18. doi: 10.1186/s12711-024-00883-w
- Bermann, M., D. Lourenco, V. Breen, R. Hawken, F. Brito Lopes, and I. Misztal. 2021. Modeling genetic differences of combined broiler chicken populations in single-step GBLUP. *J Anim Sci* 99(4)doi: 10.1093/jas/skab056
- Bradford, H. L., Y. Masuda, P. M. VanRaden, A. Legarra, and I. Misztal. 2019. Modeling missing pedigree in single-step genomic BLUP. *J Dairy Sci* 102(3):2336-2346. doi: 10.3168/jds.2018-15434
- Cesarani, A., D. Lourenco, S. Tsuruta, A. Legarra, E. L. Nicolazzi, P. M. VanRaden, and I. Misztal. 2022. Multibreed genomic evaluation for production traits of dairy cattle in the United States using single-step genomic best linear unbiased predictor. *J Dairy Sci* 105(6):5141-5152. doi: 10.3168/jds.2021-21505
- Chang, C. C., C. C. Chow, L. C. Tellier, S. Vattikuti, S. M. Purcell, and J. J. Lee. 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* 4(1)doi: 10.1186/s13742-015-0047-8
- Chen, C. Y., I. Misztal, I. Aguilar, A. Legarra, and W. M. Muir. 2011. Effect of different genomic relationship matrices on accuracy and scale. *J Anim Sci* 89(9):2673-2679. doi: 10.2527/jas.2010-3555

- Christensen, O. F., and M. S. Lund. 2010. Genomic prediction when some animals are not genotyped. *Genet Sel Evol* 42(1):2. doi: 10.1186/1297-9686-42-2
- Garcia-Baccino, C. A., A. Legarra, O. F. Christensen, I. Misztal, I. Pocrnic, Z. G. Vitezica, and R. J. Cantet. 2017. Metafounders are related to F (st) fixation indices and reduce bias in single-step genomic evaluations. *Genet Sel Evol* 49(1):34. doi: 10.1186/s12711-017-0309-2
- Kluska, S., Y. Masuda, J. B. S. Ferraz, S. Tsuruta, J. P. Eler, F. Baldi, and D. Lourenco. 2021. Metafounders May Reduce Bias in Composite Cattle Genomic Predictions. *Front Genet* 12:678587. doi: 10.3389/fgene.2021.678587
- Legarra, A., I. Aguilar, and I. Misztal. 2009. A relationship matrix including full pedigree and genomic information. *J Dairy Sci* 92(9):4656-4663. doi: 10.3168/jds.2009-2061
- Legarra, A., M. Bermann, Q. Mei, and O. F. Christensen. 2024. Redefining and interpreting genomic relationships of metafounders. *Genet Sel Evol* 56(1):34. doi: 10.1186/s12711-024-00891-w
- Legarra, A., J. K. Bertrand, T. Strabel, R. L. Sapp, J. P. Sánchez, and I. Misztal. 2007. Multi-breed genetic evaluation in a Gelbvieh population. *J Anim Breed Genet* 124(5):286-295. doi: 10.1111/j.1439-0388.2007.00671.x
- Legarra, A., O. F. Christensen, Z. G. Vitezica, I. Aguilar, and I. Misztal. 2015. Ancestral Relationships Using Metafounders: Finite Ancestral Populations and Across Population Relationships. *Genetics* 200(2):455-468. doi: 10.1534/genetics.115.177014
- Legarra, A., and A. Reverter. 2018. Semi-parametric estimates of population accuracy and bias of predictions of breeding values and future phenotypes using the LR method. *Genet Sel Evol* 50(1):53. doi: 10.1186/s12711-018-0426-6

- Lourenco, D. A. L., S. Tsuruta, B. O. Fragomeni, C. Y. Chen, W. O. Herring, and I. Misztal. 2016. Crossbreed evaluations in single-step genomic best linear unbiased predictor using adjusted realized relationship matrices¹. *Journal of Animal Science* 94(3):909-919. doi: 10.2527/jas.2015-9748
- Macedo, F. L., J. M. Astruc, T. H. E. Meuwissen, and A. Legarra. 2022. Removing data and using metafounders alleviates biases for all traits in Lacaune dairy sheep predictions. *Journal of Dairy Science* 105(3):2439-2452. doi: <https://doi.org/10.3168/jds.2021-20860>
- Macedo, F. L., O. F. Christensen, J. M. Astruc, I. Aguilar, Y. Masuda, and A. Legarra. 2020. Bias and accuracy of dairy sheep evaluations using BLUP and SSGBLUP with metafounders and unknown parent groups. *Genet Sel Evol* 52(1):47. doi: 10.1186/s12711-020-00567-1
- Makgahlela, M. L., I. Strandén, U. S. Nielsen, M. J. Sillanpää, and E. A. Mäntysaari. 2014. Using the unified relationship matrix adjusted by breed-wise allele frequencies in genomic evaluation of a multibreed population. *J Dairy Sci* 97(2):1117-1127. doi: 10.3168/jds.2013-7167
- McWhorter, T. M., M. Bermann, A. L. S. Garcia, A. Legarra, I. Aguilar, I. Misztal, and D. Lourenco. 2023. Implication of the order of blending and tuning when computing the genomic relationship matrix in single-step GBLUP. *J Anim Breed Genet* 140(1):60-78. doi: 10.1111/jbg.12734
- Misztal, I., A. Legarra, and I. Aguilar. 2014a. Using recursion to compute the inverse of the genomic relationship matrix. *J Dairy Sci* 97(6):3943-3952. doi: 10.3168/jds.2013-7752
- Misztal, I., S. Tsuruta, D. Lourenco, Y. Masuda, I. Aguilar, A. Legarra, and Z. G. Vitezica. 2014b. Manual for BLUPf90 family of programs.

https://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90_all.pdf

(Accessed

04/04/2024.

- Misztal, I., Z. G. Vitezica, A. Legarra, I. Aguilar, and A. A. Swan. 2013. Unknown-parent groups in single-step genomic evaluation. *J Anim Breed Genet* 130(4):252-258. doi: 10.1111/jbg.12025
- Piccoli, M. L., L. F. Brito, J. Braccini, F. V. Brito, F. F. Cardoso, J. A. Cobuci, M. Sargolzaei, and F. S. Schenkel. 2018. A comprehensive comparison between single- and two-step GBLUP methods in a simulated beef cattle population. *Canadian Journal of Animal Science* 98(3):565-575. doi: 10.1139/cjas-2017-0176
- Purcell, S. M., and C. C. Chang. PLINK 2.0.
- Quaas, R. L. 1988. Additive Genetic Model with Groups and Relationships. *Journal of Dairy Science* 71(5):1338-1345. doi: [https://doi.org/10.3168/jds.S0022-0302\(88\)79691-5](https://doi.org/10.3168/jds.S0022-0302(88)79691-5)
- Quaas, R. L., and E. J. Pollak. 1981. Modified Equations for Sire Models with Groups. *Journal of Dairy Science* 64(9):1868-1872. doi: [https://doi.org/10.3168/jds.S0022-0302\(81\)82778-6](https://doi.org/10.3168/jds.S0022-0302(81)82778-6)
- Thompson, R. 1979. Sire Evaluation. *Biometrics* 35(1):339-353. doi: 10.2307/2529955
- Tsuruta, S., D. A. L. Lourenco, Y. Masuda, I. Misztal, and T. J. Lawlor. 2019. Controlling bias in genomic breeding values for young genotyped bulls. *J Dairy Sci* 102(11):9956-9970. doi: 10.3168/jds.2019-16789
- Tsuruta, S., I. Misztal, D. A. Lourenco, and T. J. Lawlor. 2014. Assigning unknown parent groups to reduce bias in genomic evaluations of final score in US Holsteins. *J Dairy Sci* 97(9):5814-5821. doi: 10.3168/jds.2013-7821
- VanRaden, P. M. 2008. Efficient methods to compute genomic predictions. *J Dairy Sci* 91(11):4414-4423. doi: 10.3168/jds.2007-0980

- VanRaden, P. M., M. E. Tooker, J. B. Cole, G. R. Wiggans, and J. H. Megonigal, Jr. 2007. Genetic evaluations for mixed-breed populations. *J Dairy Sci* 90(5):2434-2441. doi: 10.3168/jds.2006-704
- Westell, R. A., R. L. Quaas, and L. D. Van Vleck. 1988. Genetic Groups in an Animal Model. *Journal of Dairy Science* 71(5):1310-1318. doi: [https://doi.org/10.3168/jds.S0022-0302\(88\)79688-5](https://doi.org/10.3168/jds.S0022-0302(88)79688-5)

TABLES

Table 3.1. Descriptive statistics of lines 1 and 2

Trait¹	Line 1		Line 2	
	Records	Mean(SD²)	Records	Mean(SD²)
BW	349,697	26.00(3.42)	201,070	31.22(4.20)
CY	23,115	28.52(2.32)	26,458	24.80(3.22)
MT	397,834	1.10(0.28)	409,095	1.13(0.33)
LH	49,954	2.80(0.69)	36,632	1.83(1.07)

¹BW = Body Weight; CY = Carcass Yield; MT = Mortality; LH = Leg Health

²SD = Standard deviation

Table 3.2. Heritability (diagonal), genetic correlation (above-diagonal) for each trait

Traits¹	BW	CY	MT	LH
BW	0.19	0.24	-0.06	0.13
CY		0.45	0.03	0.02
MT			0.002	0.07
LH				0.18

¹BW = Body Weight; CY = Carcass Yield; MT = Mortality; LH = Leg Health

Table 3.3. Number of animals belonging to each genetic group

Scenario	
UPG1/MF1 ¹	UPG2/MF2 ²
2,147	2,129
2,236	2,145
	2,154
	2,236

¹Unknown-Parent Groups 1, Metafounders 1: UPG and MF definitions are based on animal's line of origin; ² Unknown-Parent Groups 2, Metafounders 2: UPG and MF definitions are based on animal's line of origin separated by sex.

Table 3.4. Number of focal animals for each line and trait

	Body Weight	Carcass Yield	Mortality	Leg Health
¹ Origin 1	23,217	3,029	23,285	3,075
² Origin 2	26,460	5,207	26,590	5,220
³ Combined	49,677	8,236	49,875	8,295

¹Origin 1: evaluation of animals from line 1 only; ²Origin 2: evaluation of animals from line 2 only; ³Combined: Animals from lines 1 and 2 are evaluated together with crossbred animals removed.

FIGURES

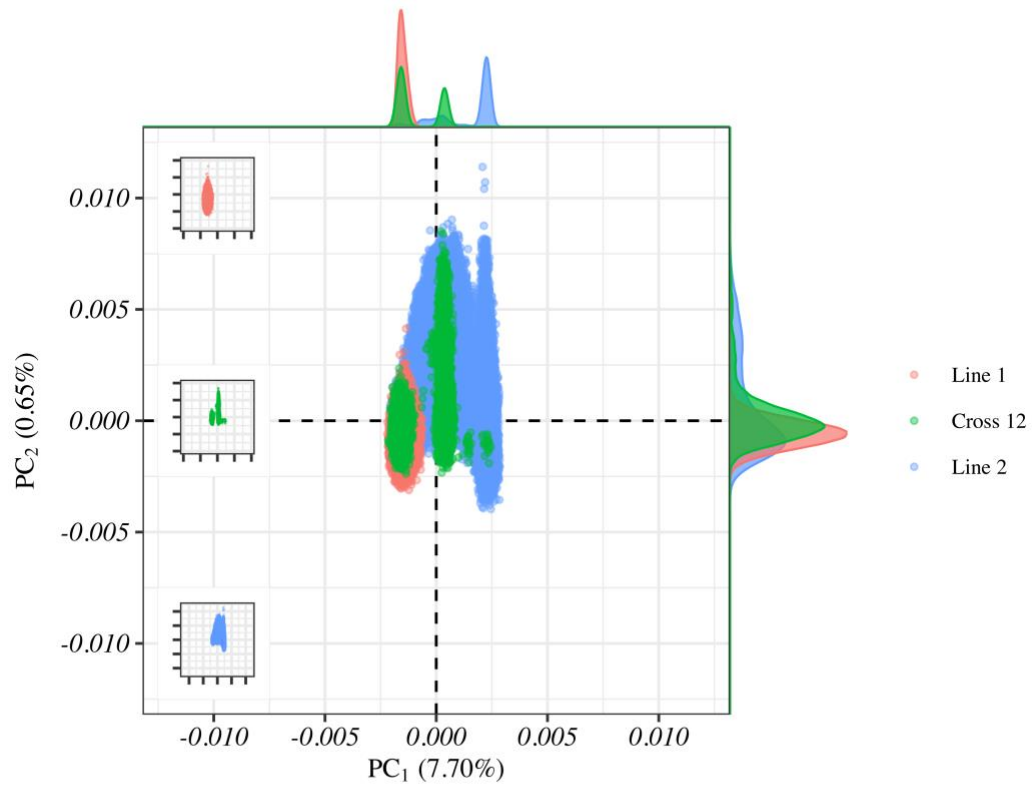


Figure 3.1. Principal component analysis (PCA) of the genomic relationship matrix

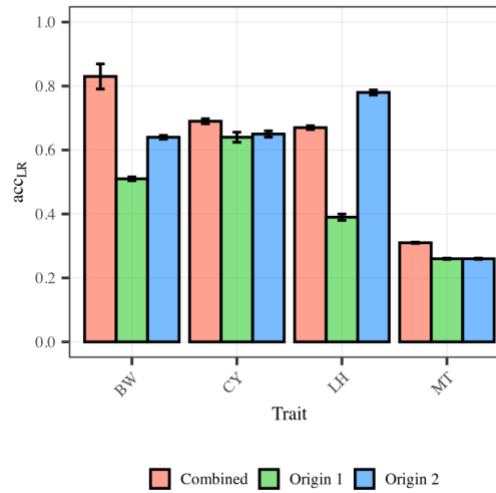


Figure 3.2. Prediction accuracy (acc_{LR}) with standard errors on the y-axis between Combined (both lines evaluated together without crossbred animals), Origin 1 (evaluation of animals from line 1 only) and Origin 2 (evaluation of animals from line 2 only) with M1/NOUPG (M1: Model with no account for animal's line of origin, NOUPG: no use of UPG or MF in the model) of BW (Body Weight), CY (Carcass Yield), LH (Leg Health) and MT (Mortality) on the x-axis

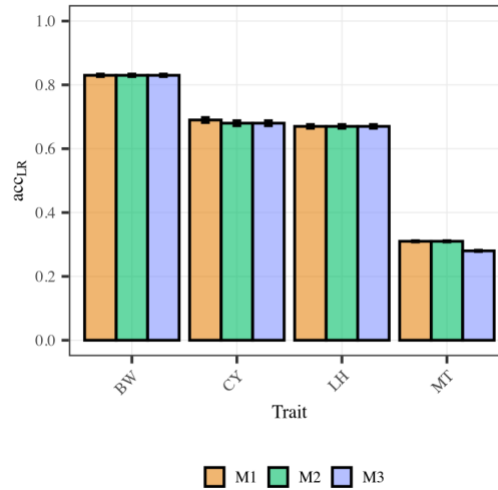


Figure 3.3. Prediction accuracy (acc_{LR}) with standard errors on the y-axis of M1 (Model with no account for animal's line of origin), M2 (Model with a vector of fixed effect based on animal's line of origin included in the model) and M3 (Contemporary groups and sex effects defined based on animal's line of origin) with NOUPG (no use of UPG or MF in the model) of BW (Body Weight), CY (Carcass Yield), LH (Leg Health) and MT (Mortality) on the x-axis

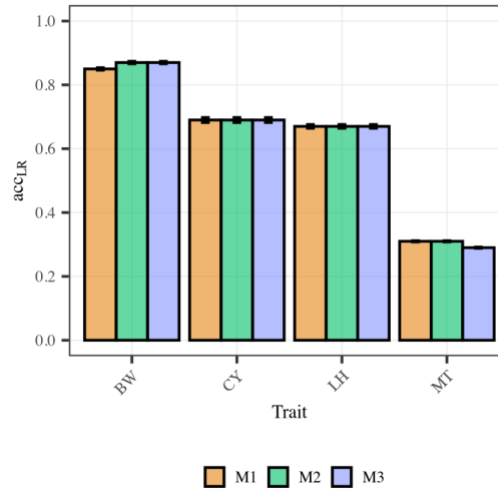


Figure 3.4. Prediction accuracy (acc_{LR}) with standard errors on the y-axis of M1 (Model with no account for animal's line of origin), M2 (Model with a vector of fixed effect based on animal's line of origin included in the model) and M3 (Model with contemporary groups and sex effects defined based on animal's line of origin) with UPG1 (UPG based on animals' line of origin) of BW (Body Weight), CY (Carcass Yield), LH (Leg Health) and MT (Mortality) on the x-axis

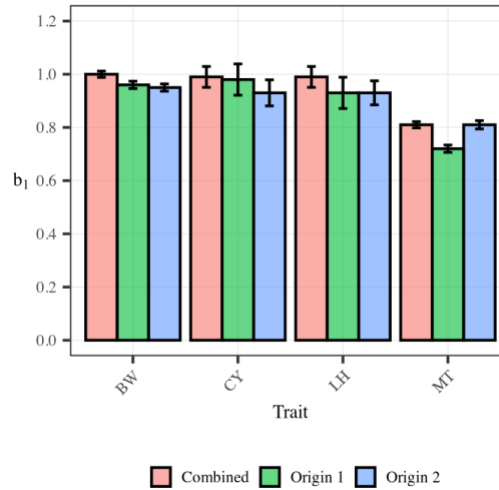


Figure 3.5. Dispersion (b_1) with standard errors on the y-axis levels between Combined (both lines evaluated together without crossbred animals), Origin 1 (evaluation of animals from line 1 only) and Origin 2 (evaluation of animals from line 2 only) with M1/NOUPG (M1: Model with no account for animal's line of origin, NOUPG: no use of UPG or MF in the model) of BW (Body Weight), CY (Carcass Yield), LH (Leg Health) and MT (Mortality) on the x-axis

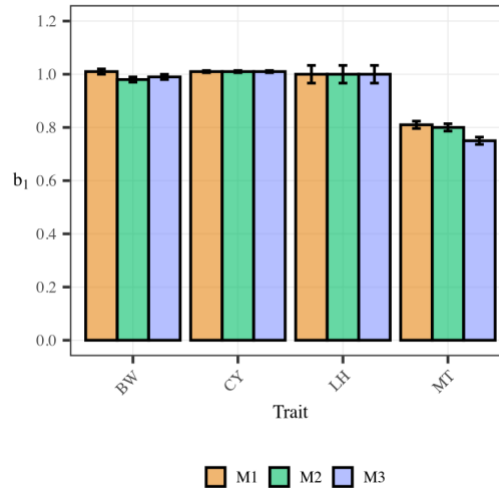


Figure 3.6. Dispersion levels (b_1) with standard errors on the y-axis of M1 (Model with no account for animal's line of origin), M2 (Model with a vector of fixed effect based on animal's line of origin included in the model) and M3 (Model with contemporary groups and sex effects were defined based on animal's line of origin) with MF2 (UPG and MF based on animal's line of origin and separated by sex) of BW (Body Weight), CY (Carcass Yield), LH (Leg Health) and MT (Mortality) on the x-axis

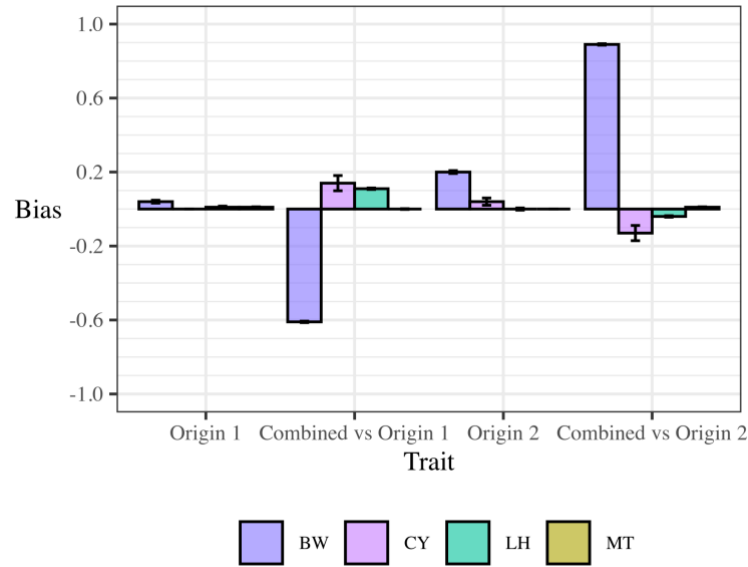


Figure 3.7. Bias with standard errors between Combined (both lines evaluated together without crossbred animals) results from partial dataset vs. Origin 1 (evaluation of animals from line 1 only) and Origin 2 (evaluation of animals from line 2 only) results from whole dataset of BW (Body Weight), CY (Carcass Yield), LH (Leg Health) and MT (Mortality) on the x-axis

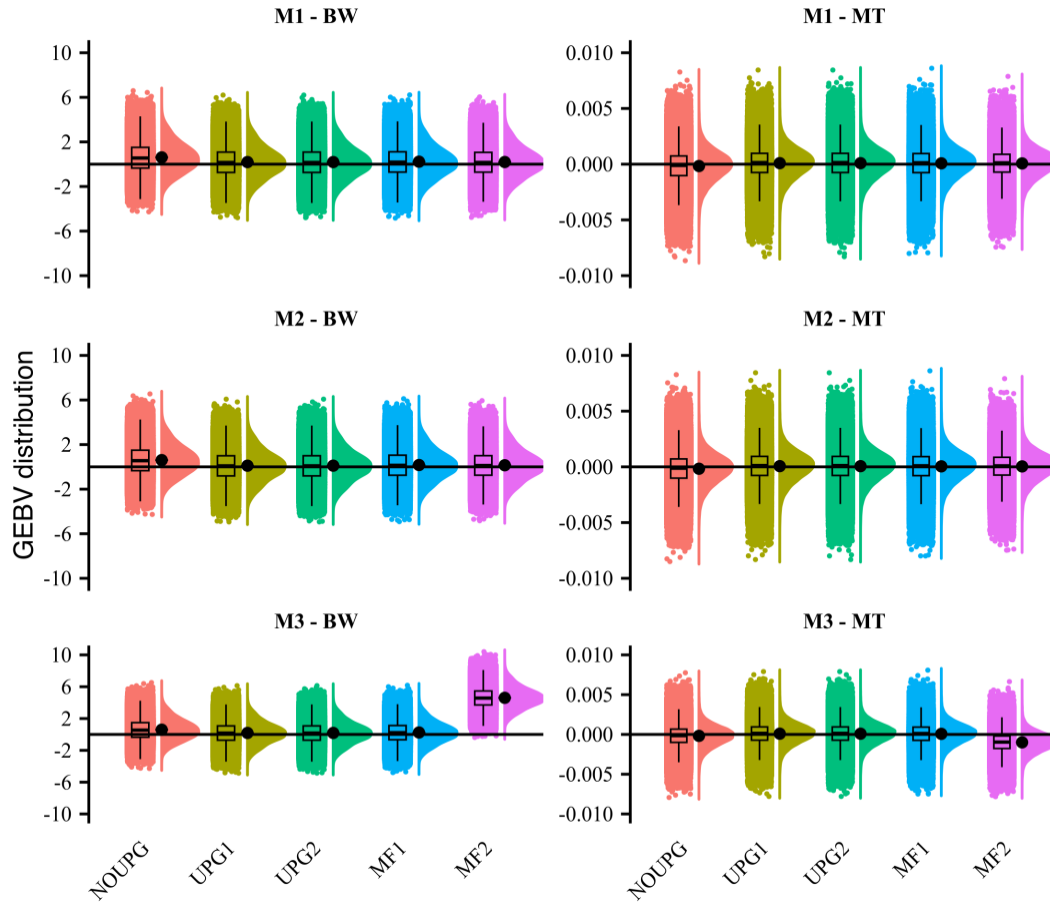


Figure 3.8. Distribution of Body weight (BW) and Mortality (MT) GEBV (Genomic estimated breeding values) for each model and scenario (NOUPG: No UPG or MF in the model, UPG1/MF1: UPG and MF based on animals' line of origin, UPG2/MF2: UPG and MF based on animal's line of origin and separated by sex; M1: no account for any line of origin, M2: Vector of fixed effect based on animal's line of origin included in the model, M3: Contemporary groups and sex effects were defined based on animal's line of origin)

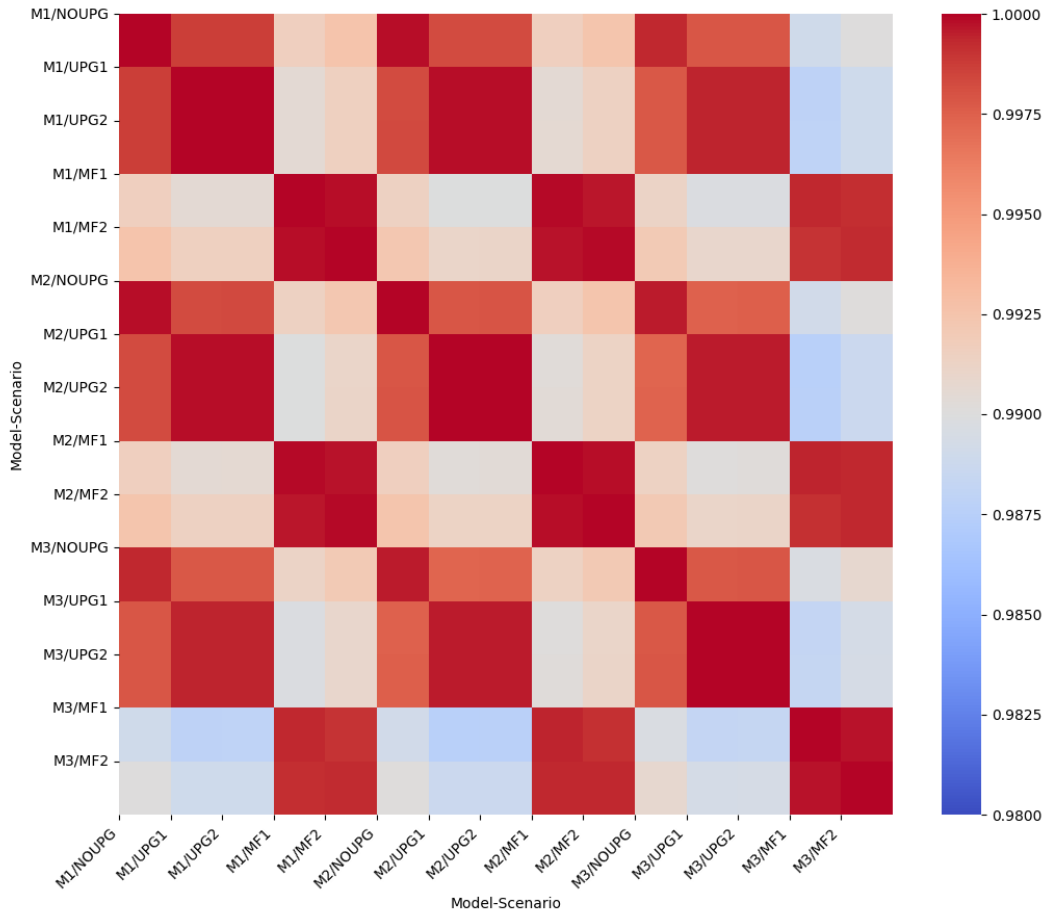


Figure 3.9. Spearman Correlation plot of all models and scenarios for Body weight (NOUPG: No UPG or MF in the model, UPG1/MF1: UPG and MF based on animals' line of origin, UPG2/MF2: UPG and MF based on animal's line of origin and separated by sex; M1: no account for any line of origin, M2: Vector of fixed effect based on animal's line of origin included in the model, M3: Contemporary groups and sex effects were defined based on animal's line of origin)

CHAPTER 4

ALL-BREED SINGLE-STEP GBLUP EVALUATIONS FOR FERTILITY TRAITS IN U.S. DAIRY CATTLE²

²Joe-Menwer Tabet, Daniela Lourenco, Fernando Bussiman, Matias Bermann, Ignacy Misztal, Paul M. VanRaden, Zulma G. Vitezica, Andres Legarra. Submitted to Journal of Dairy Science.

ABSTRACT

The U.S. dairy cattle genetic evaluation is currently a multistep process, including multibreed traditional BLUP estimations followed by single-breed SNP effects estimation. Single-step GBLUP (ssGBLUP) combines pedigree and genomic data for all breeds in one analysis. Unknown parent groups (UPG) or metafounders (MF) can be used to address missing pedigree information. Fertility traits are notably difficult to evaluate due to low heritabilities, changing management, and a higher recent emphasis on selection to move in a favorable direction. We assessed bias, dispersion, and accuracy of fertility traits in all-breed U.S. dairy cattle using pedigree-based BLUP (PBLUP) and ssGBLUP with UPG or MF; with 5% or 10% residual polygenic effect (RPE). Validation methods included the LR (Linear Regression) method and comparison of early and late deregressed proofs for Holstein and Jersey breeds. By comparing MF or UPG in PBLUP, we observed similar results in terms of bias, dispersion, and correlations between early and recent predictions. When genomics was used, ssGBLUP with MF and 10% RPE consistently outperformed other models regarding bias, dispersion, and correlations. Compared with multistep results, ssGBLUP with MF and 10% RPE showed less bias and increased correlations but slightly overdispersed estimates. Overall, genomic prediction of fertility traits using ssGBLUP was accurate and unbiased, more so with MF than with UPG.

INTRODUCTION

The current genetic evaluation system for U.S. dairy cattle follows a multistep process. Initially, breeding values are estimated using an all-breed traditional Best Linear Unbiased Prediction (BLUP). Subsequently, these estimated breeding values (EBV) and their reliabilities

are used to create deregressed proofs and obtain a within-breed estimation of Single Nucleotide Polymorphism (SNP) effects, which in turn are used to compute Direct Genomic Values. The final Genomic Estimated Breeding Values (GEBV) (VanRaden, 2008) combine the EBV and the Direct Genomic Value, including a 10% residual polygenic effect from the initial pedigree-based BLUP. However, this multistep process has some drawbacks when using results from the traditional BLUP method, which is known to be biased due to the influence of genomic preselection (Patri and Ducrocq, 2011). Single-step genomic BLUP (ssGBLUP) predicts genomic breeding values by incorporating all available information, namely phenotypic, pedigree, and genomic data, in a single analysis (Legarra et al., 2009; Aguilar et al., 2010; Christensen and Lund, 2010). The single-step method is, in principle, unbiased and more accurate than BLUP (Masuda et al., 2018; Jibrila et al., 2021; Cesarani et al., 2022).

An important detail in implementing single-step GBLUP in dairy cattle evaluations involves how to include Unknown Parent Groups (UPG) (Quaas, 1988; Tsuruta et al., 2014, 2019). The correct inclusion of UPG without double counting is complex (Misztal et al., 2013; Masuda et al., 2022; Strandén et al., 2022). A different method is Metafounders (MF) (Legarra et al., 2015), which address the incompatibility between the pedigree relationship matrix $\mathbf{A}$ and the genomic relationship matrix $\mathbf{G}$ while simultaneously modeling the differences within and across populations.

In this study, we aimed to apply ssGBLUP to fertility traits used in the genetic evaluation of U.S. dairy cattle. The traits under consideration include Daughter Pregnancy Rate (DPR), Cow Conception Rate (CCR), Heifer Conception Rate (HCR), and Early First Calving (EFC). These traits present a challenging case for evaluation due to their low heritability and the diverse range of records involved in phenotype determination, such as calving interval, insemination dates, and

pregnancy verification (VanRaden et al., 2004, 2014). The primary objectives of this study are twofold: first, to assess the accuracy and potential biases of the single-step GBLUP method when applied to the specific case of all-breed U.S. fertility traits, and second, to explore various options for modeling missing parents, utilizing either UPG or MF. We also evaluate the fraction of pedigree-based residual polygenic effect included in the predictions. Finally, we compare the results of these methods with those of the current multistep evaluation system. This comprehensive study aims to gain insights into the effectiveness and suitability of implementing ssGBLUP methodology in fertility traits for U.S. dairy cattle.

MATERIALS AND METHODS

Data

We used domestic data from the official December 2022 multibreed U.S. dairy genetic evaluation for fertility traits, which were provided by the Council on Dairy Cattle Breeding (Bowie, MD). The phenotypic file consisted of all breeds (Ayrshire, Brown Swiss, Guernsey, Jersey (JE), Holstein (HO), Milking Shorthorn, and crossbreds) records for CCR, HCR, DPR, and EFC. DPR predicts the percentage of non-pregnant cows that will become pregnant during each 21-day period compared to the breed base, and it is a function of Days Open. HCR and CCR measure the percentage of inseminated heifers or cows that successfully become pregnant with each service, modelling the success or failure of each reported insemination. EFC measures the age at first calving. For further details on these traits, refer to VanRaden et al. (2004, 2014). We will focus on DPR and CCR as these are the most relevant traits for the industry. The records for DPR and EFC started in 1960, whereas data collection for CCR and HCR started in 2000 – however, DPR and CCR are strongly genetically correlated traits (Legarra and VanRaden, 2023),

expected to have similar genetic trends. Table 1 provides an overview of the traits. The pedigree file included 93,417,440 animals, with 417 UPG based on breed, birth year, and pathway (sex of the animal with missing parent, sex of its ancestor, and foreign/local origin).

The CDCB database contains over 7 million genotypes, all imputed to a standard set of 78,964 selected single nucleotide polymorphisms (SNP), following the CDCB procedures outlined in <https://uscdcb.com/genomic-evaluations/>. Of these genotyped animals, only those deemed "useful" were considered. This involves two considerations. First, genotypes of culled animals do not need to be included, with an exception that we will describe later. This was shown by Jibrila et al. (2021, 2023) using simulation, and it can also be analytically proven as follows. For a model of the kind $\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}\mathbf{u} + \mathbf{e}$ with $Var(\mathbf{u}) = \mathbf{H}\sigma_u^2$ and $Var(\mathbf{e}) = \mathbf{R}$, single-step GBLUP can be presented using selection index principles (Henderson, 1984) as $\sigma_u^2 \mathbf{H}\mathbf{Z}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\hat{\mathbf{b}})$ with $\hat{\mathbf{b}} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}$ and $\mathbf{V} = \sigma_u^2 \mathbf{Z}\mathbf{H}\mathbf{Z}' + \mathbf{R}$ (this is a single trait proof but readily applies to multiple trait and more complex models). Thus, the elements in $\mathbf{H}$ (the ssGBLUP relationship matrix) of animals without records can be safely discarded, identical to Henderson (1977). Another intuitive explanation is that highly selected young bulls exhibit positive Mendelian Sampling term, which is well accounted for by the difference between progeny and parents' genotypes.

The proof above requires that the matrix $\mathbf{H}$ is "as complete as possible", i.e., it is not affected when genotypes of culled animals with no records and no progeny with records are discarded. However, this is not necessarily true. Consider an animal genotyped without records, and not an ancestor of animals with records. This animal can be ignored in the BLUP evaluation because $\mathbf{A}$ among animals in records is not modified by its presence in the pedigree. However, if one or both parents are not genotyped, this animal's genotype will modify (improve) the row and column in $\mathbf{H}$ of the non-genotyped parent(s). When one or both parents are not genotyped,

Shabalina et al. (2017) and Jibrila et al. (2021) observed bias and reduction in accuracy when genotypes of such animals, informative on $\mathbf{H}$ but not on records, were ignored. However, when both parents are genotyped, the genotypes of these animals are not useful, as shown analytically above and also by Jibrila et al. (2021). We can reasonably hope that selected bulls in the last decade or so are offspring of genotyped sire and dam, in which case the matrix $\mathbf{H}$ is not modified by adding culled half-sibs of the bull, and there should be little or no bias. We, therefore, preferred to use a simple rule rather than try to catch genotyped descendants of ungenotyped parents.

Therefore, genotypes of animals with records or their ancestors were selected, whereas the rest were ignored, e.g., we ignored a foreign genotyped bull with no domestic genotyped daughters in records or a domestic genotyped heifer with no record yet. Out of the 7 million genotypes, 2,041,543 genotyped animals were thus extracted. Table 2 presents the number of genotyped animals per breed.

Estimation of pedigree-based or genomic breeding values:

An all-breed, multi-trait repeatability animal model incorporating all four traits was used to predict pedigree-based EBV or genomic-based GEBV. The pedigree-based methods were pedigree-based BLUP (PBLUP) with UPG and MF (PBLUP_{UPG}, PBLUP_{MF}). The genomic prediction methods were ssGBLUP with UPG or MF fitted using either 5% (ssGBLUP_{UPG5} or ssGBLUP_{MF5}) or 10% (ssGBLUP_{UPG10} or ssGBLUP_{MF10}) pedigree-based polygenic effect (also called residual polygenic effect). The 5% and 10% fractions are, respectively, the default in the blupf90 software suite (Misztal et al., 2014b) and the standard at CDCB MultiStep (MS) genomic prediction. In summary, the different models were PBLUP_{UPG}, PBLUP_{MF}, ssGBLUP_{UPG5}, ssGBLUP_{UPG10}, ssGBLUP_{MF5}, and ssGBLUP_{MF10}.

The results obtained from these models were compared with statistics from the official results of the dairy genetic evaluation provided by CDCB, which employs a traditional all-breed BLUP followed by within-breed GBLUP using deregressed proofs. However, due to model changes, these statistics are one year delayed, as will be explained later.

When MF is used to model unknown parents, both PBLUP and ssGBLUP use an inverse relationship matrix denoted as $\mathbf{A}_\Gamma^{-1}$ Legarra et al. (2015), which requires the estimation of a matrix $\mathbf{\Gamma}$ of relationships among metafounders. The MF are the same as UPG (e.g., 417), with a different method that uses a relationship matrix $\mathbf{\Gamma}$. The $\mathbf{\Gamma}$ matrix was estimated similarly to Wicki et al. (2023) or Legarra and VanRaden (2023) using base allele frequencies inferred during the imputation process at CDCB. We did use the adjustment of the genetic covariance across traits ($\mathbf{G}_0$) to consider $\mathbf{\Gamma}$, as $\mathbf{G}_{0(\text{related})} = \frac{\mathbf{G}_0}{k}$ for $k = \left(1 + \frac{\overline{\text{diag}(\mathbf{\Gamma})}}{2} - \bar{\mathbf{\Gamma}}\right)$ (Legarra et al., 2015) with a value of $k = 0.91$. This is an approximation, as it is not possible to exactly scale variance components to an all-breed situation when breeds are assumed to be related.

When using MF, genomic relationships are created assuming allele frequencies of 0.5 (Christensen, 2012; Legarra et al., 2015), i.e., as crossproducts of the genotype format (-1,0,1) times a scaling factor $\frac{2}{m}$, for m the number of SNPs, then “blended” with a fraction (5% or 10% as described) of $\mathbf{A}_{\Gamma 22}$. Compatibility is automatic because $\mathbf{A}_\Gamma$ is created using $\mathbf{\Gamma}$ as estimated for MF. When using UPG, observed allele frequencies are used for both the crossproduct and scaling of $\mathbf{G}$. Then the matrix is “tuned” to resemble $\mathbf{A}$ (Christensen et al., 2012) then “blended” with $\mathbf{A}_{22}$.

For BLUP and ssGBLUP, we used random UPG with *a priori* covariances equal to the identity matrix multiplied ($\otimes$ Kronecker product) by the genetic covariance matrix, resulting in a matrix $\mathbf{A}_\Sigma^*$ (Masuda et al., 2022). For BLUP, this implies the usual inverse with UPG ($\mathbf{A}^*$) plus a

block-structure in the corresponding block to UPG ($\mathbf{A}_{\Sigma}^* = \mathbf{A}^* + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \Sigma^{-1} \end{bmatrix}$). For ssGBLUP with UPG, we used the “Altered QP model” (Masuda et al., 2022), which is equivalent to adjusting J-factors, i.e., covariates accounting for different allele frequencies across populations, a special case where (in their notation) $\mathbf{Q}_2 = \mathbf{Q}_c$ (Strandén et al., 2022; Himmelbauer et al., 2024).

Choice of core animals for APY:

Given the large number of genotyped animals, the algorithm for proven and young (APY) (Misztal et al., 2014a) was employed for this study to efficiently obtain the inverse of $\mathbf{G}$. Core animals representing genotypes from five pure breeds and crossbreds were included in the evaluation: Ayrshire, Brown Swiss, Guernsey, Holstein, Jersey, and crossbred genotyped animals. Choice of the core can lead to changes in the GEBV (Misztal et al., 2020; Edel et al., 2022). Thus we tried to avoid a truly random sample, given that our data set is structured by breed and sex. The core was selected in the following manner. In the case of Ayrshire and Guernsey, all genotyped animals born in or after 1990 were chosen due to the low number of genotyped animals in those breeds. Based on previous studies (Cesarani et al., 2022), the desired core size was 15K for Jersey and Holstein and 5K for Brown Swiss and crossbreds. Then, we proceeded as follows. First, core animals must be born in or after 1990 and have both parents known. Second, we chose bulls with more than 100 daughters in the pedigree for Brown Swiss, Jersey, and crossbreds and more than 500 daughters in the pedigree for Holstein. Last, we added genotyped cows to the core until (approximately) the desired size. Cows were added if $\text{modulo}(\text{anim}_{\text{key}}, n) = 0$ where anim_{key} is the internal, unique ID used at CDCB and $n = \text{round}\left(\frac{N_G - B}{C}\right)$, where n is the nearest integer to the ratio of the difference between the total number of genotypes of a breed (N_G) and the number of selected bulls (B) of the same breed, and the number core animals (C) and $\text{round}(\cdot)$ is the mathematical rounding function. This ensures that the process is repeatable, provided that no new

information is added. The integer key anim_{key} is roughly correlative with the inclusion of new animals (i.e., in time), so our method should give an approximately balanced sampling of genotyped cows per year.

The reason for requiring a minimum number of daughters with records is to include both sires and cows in the core. Sires represent the next generation but are highly selected; cows represent the current population and are less selected. The combination of both should provide a good representation of the population's genetic diversity. Table 2 presents the number of core animals per breed and sex finally used to create an approximated inverse of the genomic relationship matrix ($\mathbf{G}_{\text{APY}}^{-1}$).

Validation methods

To evaluate the performance of BLUP and ssGBLUP, we employed two validation methods based on (Genomic) Estimated Breeding Value ((G)EBV): Linear Regression (LR) (Legarra and Reverter, 2018) and Improved Genomic Validation (IGV) (VanRaden, 2021). These validation approaches enable the estimation of different metrics to assess bias, dispersion, and accuracy of the estimated breeding values derived from ssGBLUP and pedigree BLUP (PBLUP). To ensure fair comparisons, we expressed all (G)EBVs as deviations from the same base, which is the average (G)EBV of cows with records born in 2015.

Both validation methods involve using a whole dataset, which we have already described, and a partial dataset, which excluded records with calving dates during or after 2019 (but we did not remove genotypes or pedigree). We ran the same models for evaluations for this partial data set. Then the respective (G)EBV predicted using “whole” (with subindex “w”) or “partial” datasets (with subindex “p”) are compared. For the validation process, we identified focal individuals (or validation candidates) as bulls with more than 100 daughters with records for DPR in the whole

dataset and no daughters with records in the partial dataset. To simplify the process, the same set of validation bulls was selected for DPR and CCR. In total, there were 1,891 Holstein and 303 Jersey focal bulls, all born between 2012 and 2019, with the majority falling within the 2014 to 2018 range. Given the limited number of bulls meeting these criteria, no validation set was established for the other breeds.

The first validation method (LR) involves estimating three statistics for the validation bulls. Bias μ_{pw} was estimated as the difference in means of (G)EBV in the partial and whole datasets for the focal animals; estimations were expressed in genetic standard deviations of each trait. Positive values imply an overestimation of early (G)EBV, while negative values imply an underestimation. Dispersion (b_1) is measured as the regression coefficient of the (G)EBV from the whole dataset on the (G)EBV of the partial dataset. If b_1 is equal to 1, there is not over or under-dispersion. As a measure of accuracy, we computed the Pearson correlation between the (G)EBV of the validation animals between the whole and the partial dataset, whose expected value is the ratio $\frac{acc_p}{acc_w}$ where acc_p is the accuracy of the validation candidates from the partial dataset and acc_w the accuracy of the same animals but from the whole dataset. Thus, assuming that the later prediction is highly accurate, higher correlation values mean that the “early” prediction with partial data is more accurate.

The second validation method (IGV) involves computing the deregressed (G)EBV (d(G)EBV) from the difference between (G)EBV in the whole and partial datasets (VanRaden, 2021). The resulting d(G)EBV is analogous to a record (not shrunk), and it only incorporates the information from the gain in reliability between the partial and whole datasets. The calculation of d(G)EBV is based on the following equation (VanRaden, 2021):

$$d(G)EBV = (\widehat{G})EBV_p + \frac{(\widehat{G})EBV_w - (\widehat{G})EBV_p}{Rel_{diff}} \quad [4.1]$$

where: Rel_{diff} is the reliability of the extra information of the whole compared to the partial datasets. The value of Rel_{diff} can be calculated by transforming the two (G)EBV reliabilities, “whole” and “partial”, to equivalent daughter contribution (EDC), and then the difference between the two EDCs (whole – partial) is transformed back into reliability. Another equivalent method to directly calculate Rel_{diff} is (Harris and Johnson, 1998):

$$Rel_{diff} = \frac{Rel_w - Rel_p}{Rel_w \times Rel_p + 1 - 2Rel_p} \quad [4.2]$$

in which Rel_w and Rel_p are the reliabilities of (G)EBV computed from the whole and partial datasets.

After obtaining the $d(G)EBV$ and Rel_{diff} of the focal individuals, the $d(G)EBV$ are then used in the following linear regression:

$$d(G)EBV = b_0 + b_1 \times (\widehat{G})EBV_p + e \quad [4.3]$$

Where b_0 and b_1 are the unknown parameters to be estimated. The $d(G)EBV$ in (3) are weighted by Rel_{diff} , so when there is no gain in reliability, that animal gets zero weight. The expected value of b_1 is 1. Additionally, we calculated standard errors of the metrics as described by Bermann et al. (2024).

To compare with existing procedures, the CDCB provided the results of these metrics using their national MS evaluation method, split into first a BLUP prediction (MS_{BLUP}) and second a genomic prediction ($MS_{Genomic}$) done within breed from de-regressed proofs from the MS_{BLUP} . These metrics were obtained for bulls chosen in the same manner as the validation animals in our study. However, there was a difference in the time span due to changes in edits and models from

2018 to 2022, rendering the respective proofs not directly comparable. Thus, the comparison was conducted one year later. We selected bulls with 0 daughters in the evaluation of December 2019 and at least 100 daughters with DPR records in December 2023. In addition, bulls were chosen to not have MACE information, i.e., foreign bulls with 0 domestic daughters in records but with foreign daughters in December 2019 were excluded from the validation data set. The number of selected bulls was 2,160 and 331 for Holstein and Jersey, respectively. The results from the MS methods contained domestic information, with no foreign progeny records; therefore, no MACE data.

Genetic trends

Genetic trends (means of (G)EBV per year of birth for cows with records) were obtained via ssGBLUP and PBLUP from both whole and partial datasets for CCR and DPR in HO and JE. This visualization serves to illustrate the method's ability to accurately capture the response to selection of a particular trait. Furthermore, it offers insight into the behavior of the estimates over a specific timeframe, potentially highlighting any distortions in the curve's behavior or indicating the homogeneity of genetic progress. The genetic base year was set to cows with records born in 2015 (Norman et al., 2020). In addition, to further understand disparities between methodologies (MF and UPG), we plotted solutions (using the whole dataset) of the UPG or MF for “unknown sires of foreign dams” with ssGBLUP (MF10 and UPG10).

RESULTS

Validation of UPG and MF

The estimated validation metrics for bias, dispersion, and correlation for both methods (LR and IGV) are shown when using ssGBLUP and PBLUP with MF or UPG. Table 3 presents the

results of bias in terms of the difference in the means of (G)EBVs between the partial and whole datasets, expressed in the standard genetic deviation of the trait.

A comparison between PBLUP_{UPG} and PBLUP_{MF} revealed a reduction in bias (difference closer to zero) using MF for both breeds, with a decrease ranging from 0.09 to 0.17 in favor of MF except in the case of Jersey with DPR with a difference of 0.03 in favor of UPG. In the case of ssGBLUP_{MF5} and ssGBLUP_{UPG5}, there was a reduction in bias of 0.03 with MF in DPR and no change in CCR for Holstein. However, using MF in ssGBLUP with a 5% residual polygenic effect did not improve results for the Jersey breed, resulting in an increase of 0.11 between the two methods. Setting the residual polygenic effect to 10% (ssGBLUP_{UPG10} and ssGBLUP_{MF10}) led to a significant decrease in bias with MF (i.e., in MF5), whereas in the case of UPG, the same fraction (UPG10) resulted in opposite, but sizeable, bias. This suggests that UPG may be unfairly capturing management trends, which would indicate that the bias originates from the phenotypes among other factors (CDCB 2019: <https://queries.uscdcb.com/News/Fertility%20Evaluations%20April%202019.pdf>) and that genomics appears to help palliate this bias.

Table 4 presents b_1 in ssGBLUP and PBLUP for both genetic group definitions and validation methods LR and IGV. Results from traditional pedigree BLUP showed inflation in both cases (MF and UPG), as evidenced by regression coefficients lower than 1 in both methods and for all traits). The lowest b_1 for PBLUP was observed with CCR in Jersey, PBLUP_{MF} yielded 0.69 with IGV and 0.76 with LR, while PBLUP_{UPG} produced 0.62 and 0.72, respectively. Although MF decreased the inflation levels for PBLUP, they were still farther than ssGBLUP from the expected value of 1, perhaps due to not accounting for genomic preselection.

Notably, upon adding genomic information, an immediate increase in the regression coefficient towards 1 was observed. In the case of ssGBLUP, and considering UPG and MF, both LR and IGV regression coefficients calculations produced still inflated estimates ($b_1 < 1$) with the highest levels of inflation observed in ssGBLUP_{UPG5}/ssGBLUP_{MF5} with 0.87 (IGV) in CCR Jersey. Both validation methods, IGV and LR, yielded similar results in HO and JE breeds with ssGBLUP, but the values changed in PBLUP depending on the breed. When MF or UPG included more polygenic fractions (5 to 10), there was an increase in the regression coefficient values, approaching 1. Although LR consistently yielded higher coefficients than IGV, the relative rankings across models were very similar.

Table 5 presents the correlations between (G)EBV derived from the whole and partial datasets for each scenario. PBLUP (MF and UPG) yielded low correlations, ranging between 0.49 and 0.65, which indicated that partial EBV was not a good predictor of EBV of the validation sires when the phenotypes of their daughters were added. Correlations within PBLUP_{MF} were consistently higher than within PBLUP_{UPG}, although the differences were small (0.01 – 0.02), rendering it insignificant in practice. Similar to the regression coefficients, there was a large increase in the correlation when genomic information was included. With ssGBLUP_{MF5}, the lowest correlation observed was 0.80 (CCR in Jersey), and the highest with ssGBLUP_{MF10} was in DPR for Holstein with a value of 0.90.

Overall, the results demonstrated high correlations, indicating that GEBV of young bulls from early evaluations without daughter information are highly correlated with the GEBV when daughter information is added. Implementing MF (whether MF5 or MF10) led on average to small increases in the correlation compared to UPG (UPG5 or UPG10), with minor differences among the four models (UPG5, UPG10, MF5, and MF10) in Holstein for both traits. Increasing the

residual polygenic effect to 10% notably benefitted the Jersey breed in terms of correlations of GEBV, resulting in more accurate predictions with an increase of 0.08 in CCR with ssGBLUP_{MF10} and 0.04 with ssGBLUP_{UPG10}.

Validation of ssGBLUP, PBLUP, and Multistep

In this study, we also compared the results derived from ssGBLUP (UPG5, UPG10, MF5, and MF10) and PBLUP (MF and UPG) with results obtained from the MS results (pedigree-based predictions MS_{BLUP} and within-breed genomic predictions MS_{Genomic}) provided by CDCB. Both validation methods were applied to MS_{Genomic} and MS_{BLUP}.

Table 6 presents the validation metrics provided by the CDCB. Results from PBLUP (Tables 3, 4, and 5) and CDCB MS_{BLUP} were expected to be different. First, CDCB uses UPG, not MF. Second, in the MS_{BLUP} process, if insufficient data is associated with a UPG for accurate estimation, this UPG is automatically merged with the subsequent UPG. In contrast, with PBLUP using the blupf90 programs, such merging does not occur, and there is no genetic group redefinition. We, therefore, observed some discrepancies between PBLUP (MF and UPG) and MS_{BLUP}. A difference in the level of bias existed between PBLUP (MF and UPG) and MS_{BLUP}, with MS_{BLUP} yielding less biased results for CCR in HO than PBLUP, and comparable but opposite biases for JE (biases for MS_{BLUP} are -0.02, 0.13, 0.21, 0.27 whereas for PBLUP_{UPG} are -0.40, -0.12, -0.14, -0.05, always in the order HO_{CCR}, HO_{DPR}, JE_{CCR} and JE_{DPR}). We do not have a clear explanation for these differences.

Regarding ssGBLUP, as observed earlier, all four analyses (UPG5, UPG10, MF5, and MF10) yielded biased estimates (Table 3). Bias also exists with the MS method (MS_{Genomic}, Table 6). Interestingly enough, ssGBLUP_{MF10} yielded the lowest level of bias (approaching zero) among the five genomic predictions (ssGBLUP with UPG5, UPG10, MF5, MF10, and MS_{Genomic}). There

was a drop in bias for HO from 0.25 (CCR) and 0.25 (DPR) in $MS_{Genomic}$ to -0.08 (CCR) and 0.02 (DPR) in $ssGBLUP_{MF10}$. For JE, bias decreased from $MS_{Genomic}$ to $ssGBLUP_{MF10}$, although there were still slight biases, e.g., in $ssGBLUP_{MF10}$ with 0.03 for CCR and 0.05 for DPR compared to $MS_{Genomic}$ with 0.14 and 0.41, respectively.

In terms of dispersion, as anticipated, genomic estimates of CDCB were nearly unbiased (slope of $MS_{Genomic}$ nearly 1) for both LR and IGV. Still, there was inflation in traditional pedigree BLUP with slopes of MS_{BLUP} of 0.93 (IGV) and 0.94 (LR) for Holstein CCR and 0.90 (IGV) and 0.92 (LR) in DPR but lower than 0.80 (IGV and LR) for Jersey for both traits. This is similar to the pattern observed in Table 4, with slopes close to 0.9 – 1.0 for $ssGBLUP$ but much lower for PBLUP. Among the four $ssGBLUP$ models (UPG5, UPG10, MF5, and MF10), the b_1 closest to 1 was observed in $ssGBLUP_{MF10}$, with the highest b_1 values for both validation methods (LR and IGV), reaching 0.97 for DPR in Jersey. These results indicated an overdispersion of breeding values with $ssGBLUP_{MF10}$ and, overall, in $ssGBLUP$. In contrast to the official evaluation results, the $ssGBLUP$ regression coefficient detected inflation (Table 4). For example, in $MS_{Genomic}$, no over or under-dispersion was observed in CCR for HO and either CCR or DPR for JE, with the regression coefficients varying close to 1 for both LR and IGV. However, in $ssGBLUP_{MF10}$, the same results varied between 0.91 and 0.97. Interestingly, in the case of DPR in HO, deviations from 1 in the slope were detected in both $ssGBLUP_{MF10}$ and $MS_{Genomic}$. Early proofs from $ssGBLUP_{MF10}$ were over-dispersed (IGV = 0.94 and LR = 0.95), whereas in $MS_{Genomic}$, early proofs were under-dispersed (IGV = 1.12 and LR = 1.10).

Correlations of (G)EBV from validation animals between whole and partial datasets were similar across our results and CDCB results, with correlations around 0.50 for MS_{BLUP} and around 0.80 for $MS_{Genomic}$. $PBLUP_{MF}$ had slightly higher correlations than MS_{BLUP} , with the most increase

in correlation (0.08 points, MF5 and MF10) observed in DPR with JE. We would expect that $PBLUP_{UPG}$ and MS_{BLUP} would yield similar results since both utilize the same UPG definitions; however, in MS_{BLUP} there is UPG redefinition when a given UPG has insufficient information. There were minor differences, with the highest difference in DPR with Jersey, where $PBLUP_{UPG}$ yielded a 0.63 correlation and $MS_{BLUP} = 0.57$. Concerning genomics, as noted earlier, $ssGBLUP_{MF10}$ exhibited the highest correlations across all traits and breeds. Compared to $MS_{Genomic}$, $ssGBLUP_{MF10}$ demonstrated a general improvement in accuracy, ranging between 0.01 to 0.14, with the most notable increase observed in DPR for JE (0.72 vs. 0.86).

Genetic Trends

Genetic trends for CCR and DPR in Holsteins and Jerseys were visualized by plotting the estimated (G)EBV mean per birth year of cows with phenotypes for DPR. Figure 1 depicts the analyses conducted with $ssGBLUP_{MF10}$ and $PBLUP_{MF}$. Each panel contains one breed, one trait, two methods (BLUP and $ssGBLUP$) and two data sets (partial and whole). Only results with MF will be presented here since similar outcomes were produced with UPG. In all cases, (G)EBV were centered at 0 for animals born in 2015.

In the case of HO, the four trends generally exhibited a consistent direction with minor variations at the beginning and the end of the studied period. The trend for $ssGBLUP_{MF10}$ with the whole dataset was consistently higher from 2000-2019 than $ssGBLUP_{MF10}$ with the partial dataset, and $PBLUP_{MF}$ with the partial dataset had the lowest trend. However, there were small differences in the very last years (2015-2019), depending on the method/dataset combination. The trend for Jersey DPR in the last five years (2015-2019) was essentially flat. For other traits/breeds, $PBLUP_P$ (pedigree-based BLUP with partial dataset) generally showed the lowest values, i.e., the lowest

genetic progress since 2015, and (somewhat unexpectedly) the fastest genetic progress was not always ssGBLUP_{MF10}, for which we discuss possible reasons in the Discussion.

Figure 2 compares the genetic trend (solutions) of UPG10 and MF10 of unknown sires of foreign dams for both HO and JE breeds concerning CCR with ssGBLUP_{MF10} and the whole data set. We show these solutions as an illustration of the virtues of using MF when UPG are hard to estimate. The trend represented by MF10 (depicted in yellow) exhibits a smoother, less heterogeneous behavior, providing a more straightforward interpretation for viewers. In contrast, the trends associated with UPG display more noise over time, particularly evident in the case of Jersey UPG,. The main difference here is that UPG and MF are fit as random effects, but MF are correlated to the neighboring ones, whereas UPG are just shrunk towards 0. This underscores the significance of having sufficient information to accurately estimate the effects of UPG, as inaccurate estimates lead to bias. Consequently, in April 2024, the CDCB changed its definition of UPG by merging foreign UPG (like the one in Figure 2) with domestic ones (<https://uscdcb.com/april-2024-evaluation-changes-whats-new/>).

Concerning computational resources, we used preGSf90 to build the genomic matrices and blup90iod3 for solving, using Preconditioned Conjugate Gradients iterations (Lourenco et al., 2022). Convergence was declared when the average squared difference between the right and left-hand sides of the Mixed Model Equations was smaller than 10^{-12} . Models using ssGBLUP took around 476 rounds to converge with a total computing time of 19 hours for MF and 14 hours for UPG and 8 threads. This does not include the time to create the genomic matrices. Setting up $\mathbf{G}_{APY}^{-1}$ in preGSf90 (Misztal et al., 2014b) took around 5 hours, using about 700 GB memory and 36 CPU. PBLUP models took around 430 rounds to converge, with a total computing time of 7.5 hours for MF and 6 hours for UPG. Other processes were running simultaneously, so these times are not

good benchmarks and are provided as a rough guide. All computations were performed in a Linux server (x86_64) with 1.5 TB of RAM and an Intel Xeon Gold 6354 (3.00 GHz) processor (36 computing cores).

DISCUSSION

Choice of core animals for APY

APY requires a choice of core animals, and this choice is not without consequences. Including an animal in the core/non-core sets affects its GEBV (Misztal et al., 2020; Edel et al., 2022). The objective of this work was not to formally propose and compare the results of a core choice versus other choices. Still, we do propose a deterministic algorithm to choose a core, and this deserves some discussion. First, with our proposal, the pool of sires (with daughters in records) in the core should increase at a slow pace, whereas cows in the core will change with each evaluation.

Given that either there is a large body of genotyped animals (e.g., Holstein and Jersey) or essentially all genotyped animals are in the core (other populations), we expect few global changes from run to run. However, there's the possibility that those changes could affect small subsets of animals in some categories, such as young genomic bulls (Edel et al., 2022). Edel et al. (2022) observed more pronounced changes when the number of phenotyped offspring of young genomic bulls was larger, and the fact that we include all bulls with phenotyped offspring above a certain threshold should minimize these changes. We believe our choice of the core, admittedly heuristic, is a good compromise between fully random choice and choosing only proven sires as was initially considered. Further, choosing within breed, and then both bulls and cows, should provide a good sample of genetic diversity in the population. This is an exploratory proposal, and its consequences

were not fully explored in this work. The quality of this proposal must be evaluated in the medium run, as other compromises which are often made in routine genetic evaluation.

Comparing multistep and single-step GBLUP

Single-step GBLUP has emerged as the preferred method for the genetic evaluation of genotyped and ungenotyped livestock. There is a general tendency of countries to shift national evaluations of dairy cattle to single-step, with (to our knowledge, and at this moment of July 2024) the Czech Republic, France, Uruguay, and the Netherlands running single-step evaluations. Many other countries are on the verge of implementing it. Numerous studies in dairy cattle have explored the possibility of implementing single-step GBLUP, with authors such as Alkhoder and Liu (2021), Cesarani et al. (2021), Himmelbauer et al. (2021) reporting benefits (higher accuracy and less bias).

Our study compared the results estimated by ssGBLUP to those obtained by CDCB using a MS approach. Regarding, b_1 , values in ssGBLUP are well within the range (0.9-1.1) currently accepted by the Interbull genomic validation test (Sullivan, 2023). We observed more inflated estimates with ssGBLUP than MS. This inflation was consistent across both validation methods, IGV and LR. The overdispersion was palliated when the pedigree-based part was increased from 5% to 10% (ssGBLUP_{UPG10}/ssGBLUP_{MF10}), which agrees with previous findings, e.g., Liu et al. (2011). The comparison of ssGBLUP and MS is not entirely fair, as our study compared data until 2018 (partial) with data until 2022 (whole), whereas the CDCB numbers are from evaluations until 2019 and 2023, leading to different sets of bulls being used. The number of bulls CDCB used to calculate these means was 1,489 (vs. 1,891 for ssGBLUP) in Holsteins and 270 (vs. 303 for ssGBLUP) in Jerseys. Additionally, a recent study (Mota et al., 2024) using similar CDCB data obtained slopes of the genomic validation test (IGV) for GEBVs in MS with values for CCR and

DPR ~ 1 for Holstein (similar to ours) but ~ 0.80 for Jersey (whereas we obtained ~ 1 in Table 6). Differences were because of data editing for bulls (in our study, a stricter criterion of more than 100 daughters with records is imposed) and also because Mota et al. (2024) used truncated MACE GEBV, i.e., including MACE data from 4 years ago. Therefore, it is essential to note that the outcomes are influenced by the subset of bulls used for validation.

Regarding bias, we observed levels closer to zero with $ssGBLUP_{MF10}$ than $MS_{Genomic}$. Increasing the residual polygenic effect from 5% ($ssGBLUP_{UPG5}/ssGBLUP_{MF5}$) to 10% ($ssGBLUP_{UPG10}/ssGBLUP_{MF10}$) resulted in less bias (approaching zero with 10% residual polygenic effect in the case of $ssGBLUP_{MF10}$). These findings align with prior studies (Liu et al., 2011; Gao et al., 2012; Su et al., 2012; Alkhoder and Liu, 2021). Pedigree-based BLUP was biased in all cases, including our PBLUP or MS_{BLUP} . Results from $PBLUP_{UPG}$ and MS_{BLUP} should theoretically align, as both analyses employ the same prediction method and UPG definitions in their models. However, differences in these findings are likely due to different UPG handling, stemming from insufficient information available to accurately estimate their effects, particularly in traits with low heritability (Bradford et al., 2019). In the case of MS_{BLUP} , when some UPG lack sufficient data, some of them are combined to avoid increasing bias; this process does not automatically occur in the case of blupf90 programs. Moreover, when comparing GEBV correlations between partial and whole datasets for $ssGBLUP$ and $MS_{Genomic}$, the highest correlations were obtained with $ssGBLUP_{MF10}$, suggesting that this is the most accurate among the tested methods, with the additional benefit of being nearly unbiased.

Model validation between UPG and MF

The second objective of this study was to compare the effect of implementing UPG or MF. $PBLUP_{MF}$ slightly outperformed $PBLUP_{UPG}$ regarding bias and dispersion, consistent with

previous studies (Garcia-Baccino et al., 2017; Bradford et al., 2019; Kudinov et al., 2020, 2022). Correlations between late and early proofs followed a similar trend, with ssGBLUP_{MF10} yielding the highest correlations, between 0.86 and 0.90. ssGBLUP provided early predictions that are therefore rather accurate, given that late proofs are progeny-based. In pedigree models (PBLUP_{MF} and PBLUP_{UPG}), EBVs with models using MF were also better correlated with the later proofs than models with UPG.

One of the main challenges in applying UPG or MF is the high number of UPG often defined in the population, which leads to poor estimates, especially in old genetic groups. MF palliates this problem as it shares information across groups. Still, a careful consideration of UPG or MF assignment results in less biased estimation (Tsuruta et al., 2019; Kudinov et al., 2020).

Related to the use of UPG and MF, we did an all-breed ssGBLUP, which is unusual but not unique (e.g., Kudinov et al. 2020). This implies that SNP effects are defined simultaneously within and across breeds, so-called “uniquely defined” (Stuber and Cockerham, 1966) or “common genetic” (Christensen et al., 2015) approaches and that they are picking up breed differences. This does not seem to particularly harm the genomic prediction as it is generally similar to MS_{Genomic} predictions. We can expect that Holstein's predictions would “dominate” the Jersey ones (as there are many more genotyped Holsteins than Jersey). This seems to be confirmed by the fact that moving from 5% to 10% residual polygenic effect (i.e., less weight on genomics) increases the correlation for Jersey (Table 5). This deserves further investigation in other trait groups.

Genetic trends for all models

Plotting genetic trends of traits (Figure 1) provides a valuable means of evaluating the method's ability to accurately capture the response to selection. Both MF and UPG models show

similar trends. Genomic and non-genomic models exhibit comparable trends until the last few years, consistent with findings by Meyer et al. (2018) and Kudinov et al. (2020).

The method/dataset yielding the apparent highest genetic trend in 2015-2019 varies across breeds and traits. It appears that there are two opposing effects. Because of changes in insemination practices, there is some bias not adequately accounted for in the data entering into CDCB, and this phenotypic bias leads to an upward bias in "early" (partial) predictions (e.g., <https://uscdcb.com/wp-content/uploads/2018/07/Addressing-Variability-in-Fertility-Evaluations-3.pdf>). On the contrary, PBLUP is unable to account for genomic selection, resulting in a downward bias (Patry and Ducrocq, 2011). These two conflicting biases may explain the lack of a consistent ranking of methods across traits and breeds for recent genetic progress.

Furthermore, we plotted the estimated solutions of UPG and MF genetic groups (Figure 2). The MF solutions exhibited a smoother trend than UPG, as in Macedo et al. (2020) and Legarra and VanRaden (2023). The erratic behavior of UPG estimates may be attributed to a lack of records attributable to parent groups, leading to less accurate estimates (Tsuruta et al., 2014), and as mentioned before, it led to a redefinition of groups at CDCB.

CONCLUSIONS

This study found that ssGBLUP is a viable option for evaluating fertility traits in multibreed U.S. dairy cattle. For the same definition of genetic groups, fitting MF resulted in slightly better genomic predictions compared to UPG, offering smoother trends for group solutions, similar or better correlations and slopes than UPG, and less bias in terms of average differences. Compared to the multistep evaluation, ssGBLUP produced less biased estimates than BLUP, with slight inflation observed in the estimates of young, genotyped bulls.

ACKNOWLEDGMENTS

The authors thank the Council on Dairy Cattle Breeding (CDCB, Bowie, MD) for providing data access and participating in the research. The authors would also like to acknowledge the contribution of dairy producers who supplied the data through their participation in the Dairy Herd Improvement program. This project was financed by partially funded by Agriculture and Food Research Initiative Competitive Grant no. 2020- 67015-31030 from the US Department of Agriculture's National Institute of Food and Agriculture (Washington, DC). Authors thank reviewers for very detailed feedback.

REFERENCES

- Aguilar, I., I. Misztal, D.L. Johnson, A. Legarra, S. Tsuruta, and T.J. Lawlor. 2010. Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *J Dairy Sci* 93:743–752. doi:10.3168/jds.2009-2730.
- Alkhoder, H., and Z. Liu. 2021. Application of a single-step SNP BLUP model to conformation traits of German Holsteins. *Interbull Bulletin* 30–40.
- Bermann, M., A. Legarra, A.A. Munera, I. Misztal, and D. Lourenco. 2024. Confidence intervals for validation statistics with data truncation in genomic prediction. *Genetics Selection Evolution* 56:18. doi:10.1186/s12711-024-00883-w.
- Bradford, H.L., Y. Masuda, P.M. VanRaden, A. Legarra, and I. Misztal. 2019. Modeling missing pedigree in single-step genomic BLUP. *J Dairy Sci* 102:2336–2346. doi:10.3168/jds.2018-15434.
- Cesarani, A., D. Lourenco, S. Tsuruta, A. Legarra, E.L. Nicolazzi, P.M. VanRaden, and I. Misztal. 2022. Multibreed genomic evaluation for production traits of dairy cattle in the United States using single-step genomic best linear unbiased predictor. *J Dairy Sci* 105:5141–5152. doi:10.3168/jds.2021-21505.
- Cesarani, A., Y. Masuda, S. Tsuruta, E.L. Nicolazzi, P.M. VanRaden, D. Lourenco, and I. Misztal. 2021. Genomic predictions for yield traits in US Holsteins with unknown parent groups. *J Dairy Sci* 104:5843–5853. doi:10.3168/jds.2020-19789
- Christensen, O.F. 2012. Compatibility of pedigree-based and marker-based relationship matrices for single-step genetic evaluation. *Genetics Selection Evolution* 44:37. doi:10.1186/1297-9686-44-37.

- Christensen, O.F., A. Legarra, M.S. Lund, and G. Su. 2015. Genetic evaluation for three-way crossbreeding. *Genetics Selection Evolution* 47. doi:10.1186/s12711-015-0177-6.
- Christensen, O.F., and M.S. Lund. 2010. Genomic prediction when some animals are not genotyped. *Genetics Selection Evolution* 42:2. doi:10.1186/1297-9686-42-2.
- Christensen, O.F., P. Madsen, B. Nielsen, T. Ostersen, and G. Su. 2012. Single-step methods for genomic evaluation in pigs. *animal* 6:1565–1571. doi:10.1017/S1751731112000742.
- Edel, C., E.C.G. Pimentel, R. Emmerling, and K.-U. Götz. 2022. 338. A critical aspect when using APY inversion with Single-Step GBLUP. Pages 1416–1419 in *Proceedings of 12th World Congress on Genetics Applied to Livestock Production (WCGALP)*. Wageningen Academic Publishers, The Netherlands.
- Gao, H., O.F. Christensen, P. Madsen, U.S. Nielsen, Y. Zhang, M.S. Lund, and G. Su. 2012. Comparison on genomic predictions using three GBLUP methods and two single-step blending methods in the Nordic Holstein population. *Genetics Selection Evolution* 44:8. doi:10.1186/1297-9686-44-8.
- Garcia-Baccino, C.A., A. Legarra, O.F. Christensen, I. Misztal, I. Pocrnic, Z.G. Vitezica, and R.J.C. Cantet. 2017. Metafounders are related to F_{st} fixation indices and reduce bias in single-step genomic evaluations. *Genetics Selection Evolution* 49:34. doi:10.1186/s12711-017-0309-2.
- Harris, B., and D. Johnson. 1998. Approximate Reliability of Genetic Evaluations under an Animal Model. *J Dairy Sci* 81:2723–2728. doi:10.3168/jds.S0022-0302(98)75829-1.
- Henderson, C.R. 1977. Best Linear Unbiased Prediction of Breeding Values Not in the Model for Records. *J Dairy Sci* 60:783–787. doi:10.3168/jds.S0022-0302(77)83935-0.
- Henderson, C.R. 1984. *Applications of Linear Models in Animal Breeding*.

- Himmelbauer, J., H. Schwarzenbacher, and C. Fuerst. 2021. Implementation of single-step evaluations for fitness traits in the German and Austrian Fleckvieh and Brown Swiss populations. *Interbull Bulletin* 82–89.
- Himmelbauer, J., H. Schwarzenbacher, C. Fuerst, and B. Fuerst-Waltl. 2024. Exploring unknown parent groups and metafounders in single-step genomic BLUP: Insights from a simulated cattle population. *J Dairy Sci.* doi:10.3168/jds.2024-24891.
- Jibrila, I., G. de Jong, and M.P.L. Calus. 2023. Genomic models account for genomic preselection by correctly estimating Mendelian sampling terms of preselected animal. *Interbull Bulletin* 41–46.
- Jibrila, I., J. Vandenplas, J. ten Napel, R.F. Veerkamp, and M.P.L. Calus. 2021. Avoiding preselection bias in subsequent single-step genomic BLUP evaluations of genomically preselected animals. *Journal of Animal Breeding and Genetics* 138:432–441. doi:10.1111/jbg.12533.
- Kudinov, A.A., M. Koivula, G.P. Aamand, I. Strandén, and E.A. Mäntysaari. 2022. Single-step genomic BLUP with many metafounders. *Front Genet* 13. doi:10.3389/fgene.2022.1012205.
- Kudinov, A.A., E.A. Mäntysaari, G.P. Aamand, P. Uimari, and I. Strandén. 2020. Metafounder approach for single-step genomic evaluations of Red Dairy cattle. *J Dairy Sci* 103:6299–6310. doi:10.3168/jds.2019-17483
- Legarra, A., I. Aguilar, and I. Misztal. 2009. A relationship matrix including full pedigree and genomic information. *J Dairy Sci* 92:4656–4663. doi:10.3168/jds.2009-2061.

- Legarra, A., O.F. Christensen, Z.G. Vitezica, I. Aguilar, and I. Misztal. 2015. Ancestral relationships using metafounders: Finite ancestral populations and across population relationships. *Genetics* 200:455–468. doi:10.1534/genetics.115.177014.
- Legarra, A., and A. Reverter. 2018. Semi-parametric estimates of population accuracy and bias of predictions of breeding values and future phenotypes using the LR method 01 Mathematical Sciences 0104 Statistics. *Genetics Selection Evolution* 50. doi:10.1186/s12711-018-0426-6.
- Legarra, A., and P.M. VanRaden. 2023. Effect of modelling unknown parent groups and metafounders on the historical genetic trend of fertility traits. *Interbull Bulletin* 26–27.
- 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. *Genetics Selection Evolution* 43:19. doi:10.1186/1297-9686-43-19.
- Lourenco, D., S. Tsuruta, I. Aguilar, Y. Masuda, M. Bermann, A. Legarra, and I. Misztal. 2022. 366. Recent updates in the BLUPF90 software suite. Pages 1530–1533 in *Proceedings of 12th World Congress on Genetics Applied to Livestock Production (WCGALP)*. Wageningen Academic Publishers, The Netherlands.
- Macedo, F.L., O.F. Christensen, J.-M. Astruc, I. Aguilar, Y. Masuda, and A. Legarra. 2020. Bias and accuracy of dairy sheep evaluations using BLUP and SSGBLUP with metafounders and unknown parent groups. *Genetics Selection Evolution* 52:47. doi:10.1186/s12711-020-00567-1.
- Masuda, Y., P.M. VanRaden, I. Misztal, and T.J. Lawlor. 2018. Differing genetic trend estimates from traditional and genomic evaluations of genotyped animals as evidence of preselection bias in US Holsteins. *J Dairy Sci* 101:5194–5206. doi:10.3168/jds.2017-13310.

- Masuda, Y., P.M. VanRaden, S. Tsuruta, D.A.L. Lourenco, and I. Misztal. 2022. Invited review: Unknown-parent groups and metafounders in single-step genomic BLUP. *J Dairy Sci* 105:923–939. doi:10.3168/jds.2021-20293.
- Meyer, K., B. Tier, and A. Swan. 2018. Estimates of genetic trend for single-step genomic evaluations. *Genetics Selection Evolution* 50. doi:10.1186/s12711-018-0410-1.
- Misztal, I., A. Legarra, and I. Aguilar. 2014a. Using recursion to compute the inverse of the genomic relationship matrix. *J Dairy Sci* 97:3943–3952. doi:10.3168/jds.2013-7752.
- Misztal, I., S. Tsuruta, D. Lourenco, M. Yutaka, I. Aguilar, A. Legarra, and Z. Vitezica. 2014b. Manual for Blupf90 Family of Programs. Accessed May 14, 2024. http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90_all7.pdf.
- Misztal, I., S. Tsuruta, I. Pocrnic, and D. Lourenco. 2020. Core-dependent changes in genomic predictions using the Algorithm for Proven and Young in single-step genomic best linear unbiased prediction. *J Anim Sci* 98. doi:10.1093/jas/skaa374.
- Misztal, I., Z.G. Vitezica, A. Legarra, I. Aguilar, and A.A. Swan. 2013. Unknown-parent groups in single-step genomic evaluation. *Journal of Animal Breeding and Genetics* 130:252–258. doi:10.1111/jbg.12025.
- Mota, R.R., P.G. Sullivan, E.L. Nicolazzi, T. McWhorter, A. Legarra, and P.M. VanRaden. 2024. Genomic Validation Software: USA Update Including Truncated MACE. Accessed June 3, 2024. https://interbull.org/static/web/5_3_Mota_Interbull2024.pdf.
- Norman, H.D., P. VanRaden, and G. Wiggans. 2020. April 2020: Genetic Base Change. Accessed June 3, 2024. https://www.uscdcb.com/wp-content/uploads/2020/02/Norman-et-al-Genetic-Base-Change-April-2020-FINAL_new.pdf.

- Patry, C., and V. Ducrocq. 2011. Evidence of biases in genetic evaluations due to genomic preselection in dairy cattle. *J Dairy Sci* 94:1011–1020. doi:10.3168/jds.2010-3804.
- Quaas, R.L. 1988. Additive Genetic Model with Groups and Relationships. *J Dairy Sci* 71:1338–1345. doi:10.3168/jds.S0022-0302(88)79691-5.
- Shabalina, T., E.C.G. Pimentel, C. Edel, L. Plieschke, R. Emmerling, and K.-U. Götz. 2017. Short communication: The role of genotypes from animals without phenotypes in single-step genomic evaluations. *J Dairy Sci* 100:8277–8281. doi:10.3168/jds.2017-12734.
- Strandén, I., G.P. Aamand, and E.A. Mäntysaari. 2022. Single-step genomic BLUP with genetic groups and automatic adjustment for allele coding. *Genetics Selection Evolution* 54:38. doi:10.1186/s12711-022-00721-x.
- Stuber, C., and C. Cockerham. 1966. Gene effects and variances in hybrid populations.
- Su, G., P. Madsen, U.S. Nielsen, E.A. Mäntysaari, G.P. Aamand, O.F. Christensen, and M.S. Lund. 2012. Genomic prediction for Nordic Red Cattle using one-step and selection index blending. *J Dairy Sci* 95:909–917. doi:10.3168/jds.2011-4804.
- Sullivan, P.G. 2023. Updated Interbull software for genomic validation tests. *Interbull Bulletin* 7–16.
- Tsuruta, S., D.A.L. Lourenco, Y. Masuda, I. Misztal, and T.J. Lawlor. 2019. Controlling bias in genomic breeding values for young genotyped bulls. *J Dairy Sci* 102:9956–9970. doi:10.3168/jds.2019-16789.
- Tsuruta, S., I. Misztal, D.A.L. Lourenco, and T.J. Lawlor. 2014. Assigning unknown parent groups to reduce bias in genomic evaluations of final score in US Holsteins. *J Dairy Sci* 97:5814–5821. doi:10.3168/jds.2013-7821.

- VanRaden, P.M. 2008. Efficient Methods to Compute Genomic Predictions. *J Dairy Sci* 91:4414–4423. doi:10.3168/jds.2007-0980.
- VanRaden, P.M. 2021. Improved genomic validation including extra regressions. *Interbull Bulletin* 65–69.
- VanRaden, P.M., A.H. Sanders, M.E. Tooker, R.H. Miller, H.D. Norman, M.T. Kuhn, and G.R. Wiggins. 2004. Development of a National Genetic Evaluation for Cow Fertility. *J Dairy Sci* 87:2285–2292. doi:10.3168/jds.S0022-0302(04)70049-1.
- VanRaden, P.M., M.E. Tooker, J.R. Wright, C. Sun, and J.L. Hutchison. 2014. Comparison of single-trait to multi-trait national evaluations for yield, health, and fertility. *J Dairy Sci* 97:7952–7962. doi:10.3168/jds.2014-8489.
- Wicki, M., J. Raoul, and A. Legarra. 2023. Effect of subdivision of the Lacaune dairy sheep breed on the accuracy of genomic prediction. *J Dairy Sci* 106:5570–5581. doi:10.3168/jds.2022-23114.

TABLES

Table 4.2. Descriptive statistics of the studied traits

Trait¹	Minimum	Maximum	Average	SD²	N³	h^{2*}
CCR	-15.64	123.03	52.44	39.25	35,212,583	1.60%
HCR	-24.90	130.42	70.60	38.53	11,587,364	1.00%
DPR	0.00	100.00	45.51	34.23	89,672,054	1.40%
EFC	-2273.00	-519.00	-787.04	101.15	35,412,656	2.70%

¹CCR: Cow Conception Rate, HCR: Heifer Conception Rate, DPR: Daughter Pregnancy Rate, EFC: Early First Calving.

²Standard deviation; ³Number of records; *Heritability.

Table 4.3. Number of core animals per breed and sex and total number of genotypes across breed

Breed	# of Genotypes	# of Sires (Dams) in Core
Ayrshire	1,608	311 (1,175)
Brown Swiss	9,560	611 (4,313)
Guernsey	3,561	219 (3,258)
Holstein	1,669,795	6,890 (8,113)
Jersey	300,976	3,186 (11,883)
Crossbreds	56,528	141 (4,616)

Table 4.3. Estimated bias, in genetic standard deviations, of (G)EBV comparing whole and partial datasets for ssGBLUP with MF/UPG at 5% and 10% residual polygenic effect (ssGBLUP_{UPG5}/ssGBLUP_{MF5} and ssGBLUP_{UPG10}/ssGBLUP_{MF10}), and traditional pedigree BLUP with MF and UPG (PBLUP_{MF} and PBLUP_{UPG})

Breed ^{Trait}	Bias					
	³ ssGBLUP				³ PBLUP	
	MF10	MF5	UPG10	UPG5	MF	UPG
¹ HO _{CCR}	-0.08	-0.40	0.15	-0.40	-0.23	-0.40
¹ HO _{DPR}	0.02	-0.13	0.22	-0.16	0.01	-0.12
² JE _{CCR}	0.03	-0.30	0.20	-0.19	-0.05	-0.14
² JE _{DPR}	0.05	-0.33	0.19	-0.22	0.08	-0.05

¹HO_{CCR}; HO_{DPR}: CCR and DPR in Holstein

²JE_{CCR}; JE_{DPR}: CCR and DPR in Jersey

³Standard errors: HO_{CCR} and HO_{DPR}: ssGBLUP $\leq$ 0.10, PBLUP $\leq$ 0.05; JE_{CCR}: and JE_{DPR}: ssGBLUP $\leq$ 0.24, PBLUP $\leq$ 0.12

Table 4.4. Estimated regression coefficient of (G)EBV from whole on partial datasets for ssGBLUP with UPG/MF at 5% and 10% residual polygenic effect (ssGBLUP_{UPG5}/ssGBLUP_{MF5} and ssGBLUP_{UPG10}/ssGBLUP_{MF10}), and traditional pedigree BLUP with MF and UPG (PBLUP_{MF} and PBLUP_{UPG}) with Linear Regression (LR) Improved Genomic Validation (IGV)

Breed _{Trait}	b₁					
	³ ssGBLUP				³ PBLUP	
	MF10	MF5	UPG10	UPG5	MF	UPG
LR						
¹ HO _{CCR}	0.94	0.95	0.90	0.90	0.93	0.92
¹ HO _{DPR}	0.95	0.93	0.91	0.90	0.81	0.82
² JE _{CCR}	0.94	0.89	0.90	0.88	0.76	0.72
² JE _{DPR}	0.97	0.92	0.95	0.90	0.81	0.79
IGV						
¹ HO _{CCR}	0.93	0.94	0.89	0.89	0.91	0.88
¹ HO _{DPR}	0.94	0.93	0.90	0.89	0.74	0.75
² JE _{CCR}	0.91	0.87	0.87	0.87	0.69	0.62
² JE _{DPR}	0.97	0.91	0.94	0.89	0.74	0.69

¹HO_{CCR}; HO_{DPR}: CCR and DPR in Holstein

²JE_{CCR}; JE_{DPR}: CCR and DPR in Jersey

³Standard errors for LR: HO_{CCR} and HO_{DPR}: ssGBLUP =0.03, PBLUP =0.01; JE_{CCR} and JE_{DPR}: ssGBLUP =0.07, PBLUP =0.02

Table 4.5. Correlation of (G)EBV from whole and partial datasets for ssGBLUP with MF at 5% and 10% residual polygenic effect ($\text{ssGBLUP}_{\text{UPG5}}/\text{ssGBLUP}_{\text{MF5}}$ and $\text{ssGBLUP}_{\text{UPG10}}/\text{ssGBLUP}_{\text{MF10}}$), ssGBLUP with UPG ($\text{ssGBLUP}_{\text{UPG}}$) and traditional pedigree BLUP with MF and UPG (PBLUP_{MF} and $\text{PBLUP}_{\text{UPG}}$)

Breed ^{Trait}	Correlation					
	³ ssGBLUP				³ PBLUP	
	MF10	MF5	UPG10	UPG5	MF	UPG
¹ HO _{CCR}	0.87	0.87	0.86	0.85	0.54	0.53
¹ HO _{DPR}	0.90	0.90	0.88	0.89	0.49	0.50
² JE _{CCR}	0.88	0.80	0.85	0.81	0.56	0.54
² JE _{DPR}	0.86	0.85	0.86	0.86	0.65	0.63

¹HO_{CCR}; HO_{DPR}: CCR and DPR in Holstein

²JE_{CCR}; JE_{DPR}: CCR and DPR in Jersey

³Standard errors: Holstein ≤ 0.02 for all traits and methods; Jersey: ≤ 0.06 for all traits and methods

Table 4.6. LR (Linear Regression) and IGV (Improved Genomic Validation) statistics for multistep predictions with genomics (MS_{Genomic}) and without genomics (MS_{BLUP})

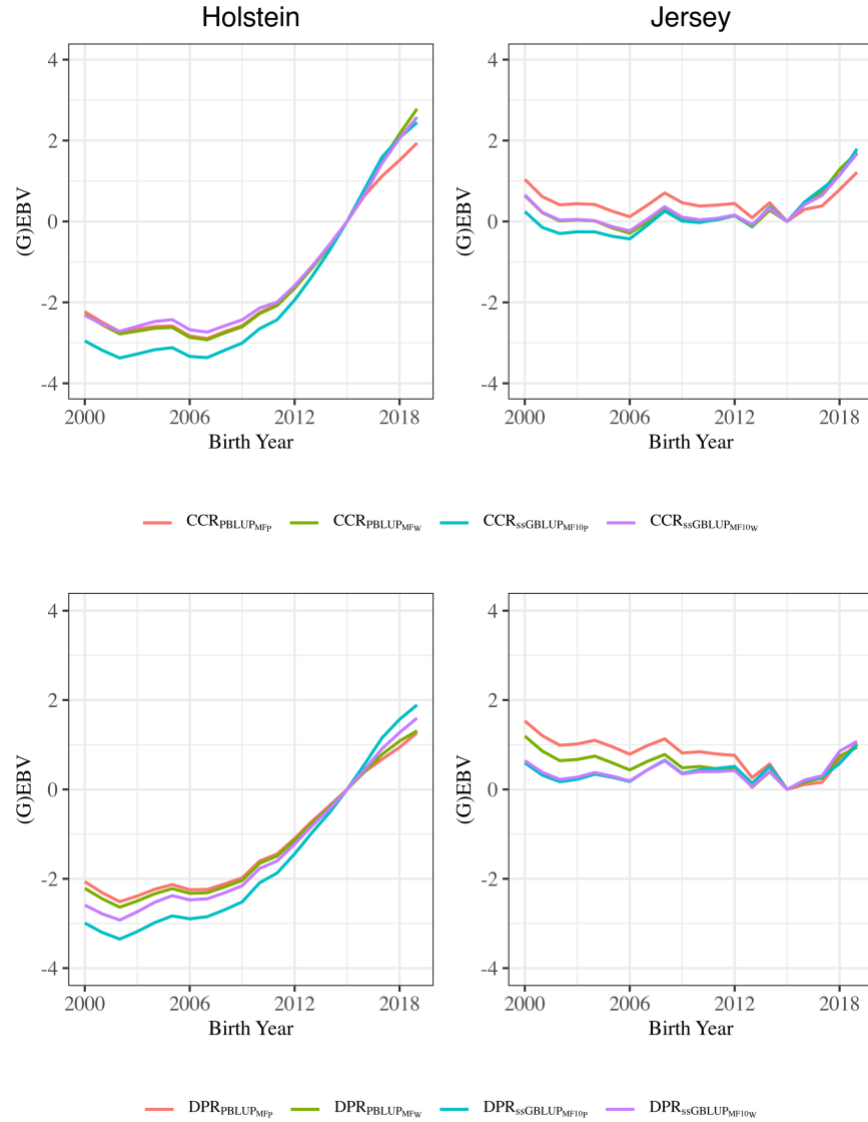
Breed _{Trait}	Bias		Correlation		b_1			
	MS_{Genomic}	MS_{BLUP}	MS_{Genomic}	MS_{BLUP}	MS_{Genomic}		MS_{BLUP}	
					IGV	LR	IGV	LR
¹ HO _{CCR}	0.25	-0.02	0.86	0.51	1.01	1.01	0.93	0.94
¹ HO _{DPR}	0.25	0.13	0.84	0.47	1.12	1.10	0.90	0.92
² JE _{CCR}	0.14	0.21	0.82	0.53	1.02	1.01	0.73	0.78
² JE _{DPR}	0.41	0.27	0.72	0.57	1.00	1.00	0.68	0.74

¹HO_{CCR}; HO_{DPR}: CCR and DPR in Holstein

²JE_{CCR}; JE_{DPR}: CCR and DPR in Jersey

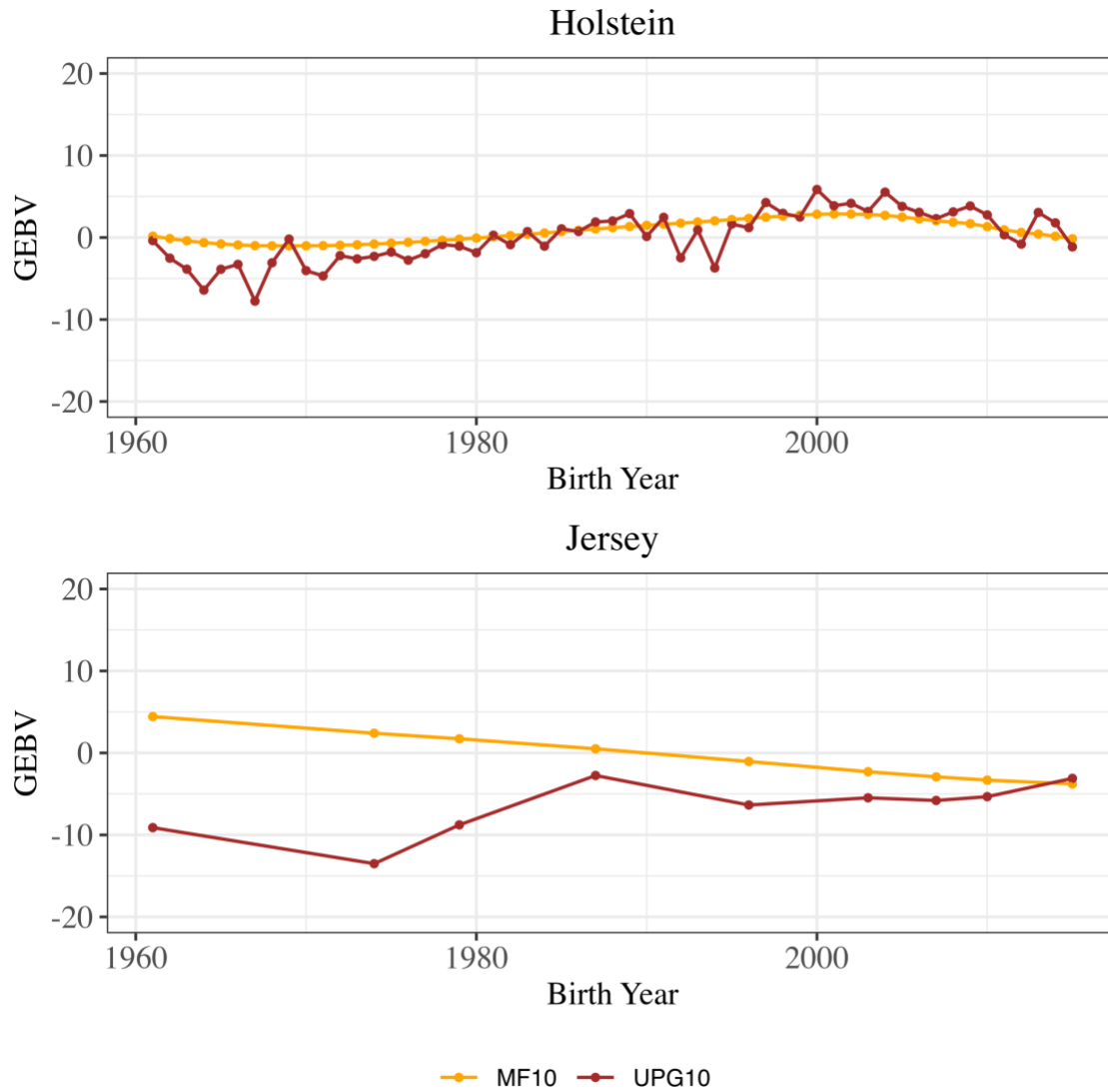
FIGURES

Figure 4.1. Genetic trend of CCR and DPR in Holstein and Jersey using Metafounders in ssGBLUP with 10% residual polygenic effect and PBLUP



$CCR_{PBLUP_{MFw}}$, $DPR_{PBLUP_{MFw}}$, $CCR_{ssGBLUP_{MF10w}}$, $DPR_{ssGBLUP_{MF10w}}$: Mean (G)EBV from the whole dataset with $PBLUP_{MF}$ and $ssGBLUP_{MF10}$ for CCR and DPR in Holstein and Jersey;
 $CCR_{PBLUP_{MFp}}$, $DPR_{PBLUP_{MFp}}$, $CCR_{ssGBLUP_{MF10p}}$, $DPR_{ssGBLUP_{MF10p}}$: Mean (G)EBV from the partial dataset with $PBLUP_{MF}$ and $ssGBLUP_{MF10}$ for CCR and DPR in Holstein and Jersey

Figure 4.2. Estimates of Genetic groups for unknown sires of foreign dams in Holstein and Jersey from the whole dataset for CCR with $ssGBLUP_{MF10}$ and $ssGBLUP_{UPG10}$ ($ssGBLUP$ with MF/UPG and residual polygenic effect = 10%)



CHAPTER 5

APPROXIMATING RELIABILITIES OF INDIRECT PREDICTIONS USING SNP EFFECTS FROM LARGE SSGBLUP EVALUATIONS³

³Joe-Menwer Tabet, Fernando Bussiman, Jorge Hidalgo, Matias Bermann, Alberto Cesarani, Ignacy Misztal, Daniela Lourenco. Submitted to Journal of Dairy Science.

ABSTRACT

Reliability (REL) is a metric used for assessing breeding values. However, calculating true reliabilities requires the prediction error variance (PEV) obtained from the inverse of the coefficient matrix of the mixed model equations (MME). This process is computationally intensive for large populations, so approximations are often used. Additionally, when individuals have indirect predictions (IP) based on single nucleotide polymorphisms (SNP) effects, they are not included in the MME, making traditional methods unsuitable for computing their reliabilities. We aimed to approximate the reliabilities of IP when SNP effects are calculated in large, genotyped populations. We compared two reliability approximations: 1) REL_{IP} , based on the approximated prediction error covariance (PEC) of SNP effects based on core animals by back solving from the genomic estimated breeding values (GEBV), using single-step GBLUP (ssGBLUP) with the Algorithm for Proven and Young (APY), and 2) using a benchmark reliability, REL_{GS3} , by calculating PEV in a ssGBLUP model using a block sparse inversion of the genomic relationship matrix with APY and effective record contributions as weights. Data from the Council on Dairy Cattle Breeding (CDCB) included 98.9 million milk yield (MY) records from multiple breeds, 49.6 million animals in the pedigree, and 1.88 million genotyped animals. Reliabilities were calculated using both a multi-breed and a single-breed approach. Validation excluded pedigree, phenotype, and progeny data for focal animals born between 2019 and 2021. The reliabilities of IP were compared to those derived from a benchmark analysis that included genotypes of focal animals. In the multi-breed analysis, the intercepts were 0.03 (males) and 0.04 (females) for Holstein (HO), 0.98 (males) and 0.24 (females) for Jersey (JE), and -0.31 (males) and -0.19 (females) for Brown Swiss (BS); the slope values were 0.95 (males) and 0.91 (females) for HO, -0.14 (males) and 0.63 (females) for JE, and 1.11 (males) and 0.99 (females) for BS; the correlations

between both reliabilities were 0.95 (males) and 0.94 (females) for HO, -0.25 (males) and 0.40 (females) for JE, and 0.98 (males) and 0.80 (females) for BS. In the single-breed analysis, correlations improved for both JE (0.99 for both sexes) and BS (0.99 for both sexes). The mean reliabilities in the single-breed evaluation were more consistent than in the multi-breed evaluation, demonstrating that it is possible to approximate reliabilities based on the PEC of SNP effects of core animals in single-breed evaluations. When using multi-breed models, dominant breeds affect the reliabilities, indicating the need to have a within-breed approximation of reliabilities.

INTRODUCTION

To select animals, breeders rely not only on genomic estimated breeding values (GEBV) but also on the reliability accompanying those estimates (Tier and Meyer, 2004). Reliability measures the information used to estimate an animal's breeding value, helping breeders make informed mating and culling decisions within their herd. In the mixed model theory, reliability can be derived from the elements of the inverse of the coefficient matrix, namely from the prediction error variance (PEV) associated with the genetic effect (Henderson, 1984).

Since the introduction of genomic selection, more reliable GEBV have been obtained (VanRaden, 2008; VanRaden et al., 2009), and the cost of genotyping has drastically decreased (Wiggans and Carrillo, 2022). Recently, the Council of Dairy Cattle Breeding (CDCB; Bowie, MD) registered the ten-millionth genotyped animal into their dataset. Managing such a large genomic dataset has created opportunities for developing methods that can handle the data without compromising the estimation of breeding values. Misztal et al. (2014a) presented the algorithm for proven and young (APY), proposing that the dimensionality of the genomic relationship matrix (**G**) is limited to the number of eigenvalues explaining 98% of its variance; therefore, that number

is used to determine the number of core (c) and non-core (n) animals. This strategy decreases the computational cost of inverting $\mathbf{G}$ by using an approximated matrix $\mathbf{G}_{\text{APY}}^{-1}$.

Another strategy to reduce computational cost is removing genotyped animals that do not contribute to the evaluation (Bussiman et al., 2023). Those removed animals, i.e., young, genotyped animals with no records (either own or progeny) or culled animals that are no longer available (e.g., inactive bulls not being used for artificial insemination), can rely on indirect predictions (IP) as their genomic prediction. The IP are calculated from single-nucleotide polymorphisms (SNP) effects which are backsolved from the GEBV of other animals in previous evaluations (or in the reference population). Previous studies have found high correlations between IP and GEBV, demonstrating that IP from an APY ssGBLUP analysis can be accurate (Lourenco et al., 2018; Garcia et al., 2020; Tsuruta et al., 2021).

In the genomic evaluation of large populations, computing the inverse of the left-hand side of the mixed model equations (MME) is computationally unfeasible, and true reliabilities cannot be directly obtained. Bermann et al. (2022b) presented an efficient algorithm to approximate GEBV reliabilities in a single-step genomic best linear unbiased prediction (ssGBLUP) evaluation with APY taking advantage of the sparse structure of $\mathbf{G}_{\text{APY}}^{-1}$; their algorithm does not require inverting the MME and adds weights (based on effective record contributions) to the diagonal of the genomic relationship matrix to approximate the PEV. In a different study, Bermann et al. (2022a) showed the calculation of the prediction error covariance (PEC) of SNP effects depends solely on the genotypes of core animals, making it possible to calculate reliabilities of IP with lower computational cost. Hence, our objective in this study was to assess a reliability approximation method for IP in the U.S. multi-breed dairy cattle evaluation.

MATERIALS AND METHODS

Data and model:

Ethical approval for this study was not needed because the data came from existing databases. Data from the official multi-breed genomic evaluation in the U.S. dairy cattle population was provided by the CDCB, consisting of 98,938,654 records (first five lactations) for milk yield (MY), dating from February 1963 to June 2021. These records covered all breeds used in the official evaluation: Ayrshire (AY), Brown Swiss (BS), Guernsey (GU), Holstein (HO), Jersey (JE), Milking shorthorn (MS) and crossbreds (XX). The genotypic dataset comprised of 1,877,704 genotyped animals, imputed to a standard set of 78,964 SNP (Table 2). The total number of animals in the pedigree was 49,585,192 (tracing back three generations).

A single-trait multi-breed repeatability animal model was applied as follows:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_h\mathbf{h} + \mathbf{Z}_u\mathbf{u} + \mathbf{Z}_p\mathbf{p} + \mathbf{e} \quad [5.1]$$

where: $\mathbf{y}$ is the vector of MY records; $\boldsymbol{\beta}$ is the vector of fixed effects comprised of herd-management, age-parity, inbreeding coefficient covariates, and heterosis, $\mathbf{h}$ is the vector of random herd-by-sire interaction effects, $\mathbf{u}$ is the vector of random animal additive genetic effects, $\mathbf{p}$ is the vector of random permanent environmental effects, and $\mathbf{e}$ is the vector of residuals. $\mathbf{X}$, $\mathbf{Z}_h$, $\mathbf{Z}_u$, $\mathbf{Z}_p$ are incidence matrices relating the phenotypic data to their respective effects. Under multivariate normality assumption, the random effects (co)variance matrix is given by:

$$Var \begin{bmatrix} \mathbf{h} \\ \mathbf{u} \\ \mathbf{p} \\ \mathbf{e} \end{bmatrix} = \begin{bmatrix} \mathbf{I}\sigma_h^2 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{H}\sigma_u^2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}\sigma_p^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}\sigma_e^2 \end{bmatrix} \quad [5.2]$$

where $\mathbf{I}$ is an identity matrix, $\mathbf{H}$ is the realized relationship matrix that combines pedigree and genomic information (Legarra et al., 2009; Aguilar et al., 2010; Christensen and Lund, 2010), and

σ_h^2 , σ_u^2 , σ_p^2 , and σ_e^2 are the variance components of herd-by-sire interaction effects, animal additive genetic effects, permanent environmental effects, and residuals. In addition to the multi-breed, we ran two single-breed, single-trait ssGBLUP models that included only JE or BS populations. These models were run to compare the changes in reliabilities between the multi- and single-breed evaluations.

Computation of genomic and IP reliabilities

Reliabilities were assessed for animals born between 2019 and 2021 (focal animals). Those were genotyped males and females, totaling 3,696 animals for HO, 3,888 for JE, and 842 for BS. These animals had their pedigree, own, and progeny phenotypes excluded from the analyses. Two analyses were performed: one including genotypes for focal animals (REL_{GS3}) used as benchmark reliability and another excluding these genotypes and posteriorly computing their IP with corresponding reliability (REL_{IP}). The REL_{GS3} can be viewed as genomic reliability since the only information focal animals had, was their genotypes, whereas REL_{IP} relied on SNP effects computed without focal animals' genotypes. Due to the number of genotyped animals, the APY was used (Misztal et al. (2014a)). Core animals were randomly selected within each breed with a 50% sex ratio based on Cesarani et al. (2022), except for GU, which lacked sufficient genotyped males, so the 5k threshold was completed by selecting genotyped females. The number of core animals per breed and sex is shown in Table 1. Additionally, focal animals were not chosen as core animals.

Reliabilities with APY were obtained as follows:

1. Constructing $\mathbf{G}_{APY}^{-1}$

For core (c) and non-core (n) animals, $\mathbf{G}_{APY}^{-1}$ (Misztal et al., 2014a) is constructed as:

$$\mathbf{G}_{APY}^{-1} = \begin{bmatrix} \mathbf{G}_{cc}^{-1} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} -\mathbf{G}_{cc}^{-1}\mathbf{G}_{cn} \\ \mathbf{I} \end{bmatrix} \mathbf{M}_{nn}^{-1} [-\mathbf{G}_{nc}\mathbf{G}_{cc}^{-1} \quad \mathbf{I}] \quad [5.3]$$

where $\mathbf{G}_{cc}^{-1}$ is the inverse of $\mathbf{G}$ for core animals, $\mathbf{G}_{cn}$ is the genomic relationship matrix relating core and non-core animals, and $\mathbf{M}_{nn} = \text{diag}(\mathbf{G}_{nn} - \mathbf{G}_{nc}\mathbf{G}_{cc}^{-1}\mathbf{G}_{cn})$.

2. Saving $\mathbf{G}_{APY}^{-1}$ and $\mathbf{A}_{22}^{-1}$:

Using preGSf90 (Aguilar et al., 2014), the $\mathbf{G}_{APY}^{-1}$ and $\mathbf{A}_{22}^{-1}$ (inverse of pedigree-based relationship matrix for genotyped animals) were saved.

3. Calculating REL_{GS3} :

The algorithm implemented in accf90GS3 calculates pedigree reliabilities based on Tier and Meyer (2004) and genomic reliabilities based on Bermann et al. (2022b). It requires $\mathbf{G}_{APY}^{-1}$ and $\mathbf{A}_{22}^{-1}$ to obtain the PEV as $\text{diag}(\mathbf{D} + \mathbf{G}_{APY}^{-1}\lambda)^{-1}$, hence step 2 was used to save time.

4. Repeat step 1, excluding genotypes for focal animals, saving $\mathbf{G}_{APY}^{-1}$ and $\mathbf{A}_{22}^{-1}$.

5. Obtain GEBV with BLUP90IOD3, reading saved matrices from step 4.

6. Back-solve SNP effects from GEBV:

Using APY, SNP effects ($\hat{\mathbf{a}}$) depend only on core animals (Bermann et al., 2022a):

$$\hat{\mathbf{a}}|\hat{\mathbf{u}}_c = (1 - \alpha)b \frac{1}{2 \sum p_l(1 - p_l)} \mathbf{Z}_c' \mathbf{G}_{cc}^{t-1} \hat{\mathbf{u}}_c \quad [5.4]$$

where subscript c refers to core animals; $\hat{\mathbf{u}}_c$ is the vector of GEBV; p_l is the l^{th} allele frequency; $\mathbf{Z}_c$ is the centered gene content matrix; $\mathbf{G}_{cc}^{t-1}$ is the inverse of the genomic relationship matrix after tuning and blending; that is, $\mathbf{G}_{cc}^t = (1 - \alpha)(\mathbf{1}\mathbf{1}'a + b\mathbf{G}_{cc}) + \alpha\mathbf{A}_{22cc}$, with a blending parameter $\alpha = 0.05$ and tuning parameters a and b computed as in Vitezica et al. (2011).

7. Compute the variance of SNP effects:

Once $\hat{\mathbf{a}}$ is obtained, its variance is computed as Bermann et al. (2022b):

$$\text{Var}(\hat{\mathbf{a}}) = (1 - \alpha)b \frac{1}{2 \sum p_l(1 - p_l)} \mathbf{Z}'_c \mathbf{G}_{cc}^{t-1} (\mathbf{G}_{cc}^t \sigma_u^2 - \mathbf{PEC}_{cc}) \mathbf{G}_{cc}^{t-1} \mathbf{Z}_c \frac{1}{2 \sum p_l(1 - p_l)} (1 - \alpha) \quad [5.5]$$

Where:

$\mathbf{PEC}_{cc}$ is the prediction error covariance based on the core animals calculated as

$$\mathbf{PEC}_{cc} = (\mathbf{D}_c + \lambda \mathbf{G}^{cc} - \lambda \mathbf{G}^{cn} (\lambda \mathbf{M}_{nn}^{-1} + \mathbf{D}_n)^{-1} \lambda \mathbf{G}^{nc})^{-1} \quad [5.6]$$

$\mathbf{D}_c$ and $\mathbf{D}_n$ are the c and n blocks of the diagonal matrix based on effective record

contributions and $\lambda = \frac{\sigma_e^2}{\sigma_u^2}$,

8. Compute REL_{IP} using Predf90 is calculated as:

$$REL_{IP} = \frac{\mathbf{Z}_f \text{Var}(\hat{\mathbf{a}}) \mathbf{Z}'_f}{\text{Var}(\mathbf{u})} \quad [5.7]$$

Where:

$\mathbf{Z}_f$ is the centered gene content matrix of focal animals, and $\text{Var}(\mathbf{u})$ is the true breeding value variance.

To evaluate the relationship between REL_{GS3} and REL_{IP} , the following linear regression was used:

$$REL_{GS3_i} = b_0 + b_1 REL_{IP_i} + e_i \quad [5.8]$$

where REL_{GS3_i} is the reliability of GEBV for the i^{th} animal; REL_{IP_i} is the reliability of IP for the i^{th} animal; b_0 is the intercept; b_1 is the slope; and e_i is a random residual term. Provided that REL_{GS3} and REL_{IP} are in the same scale, $E(b_0) = 0$, and if REL_{IP} is neither over- nor under-dispersed, $E(b_1) = 1$. Additionally, we calculated the mean of both reliabilities and the Pearson correlation between the REL_{GS3} and REL_{IP} .

RESULTS AND DISCUSSION

Multi-breed ssGBLUP evaluation

Regression coefficients for males and females show that REL_{IP} are slightly over-dispersed with $b_1 = 0.95$ (males) and 0.91 (females) for HO (Figure 1), with intercepts of -0.02 (males) and 0.03 (females). The REL_{IP} were overestimated due to the blending parameter ($\alpha = 0.05$), a necessary process to invert $\mathbf{G}$. Legarra et al. (2018) derived formulas for blending and tuning parameters when computing IP. However, in this study, we observed that an adjustment or scaling factor is necessary to correct this overestimation of REL_{IP} due to the blending parameter (Liu, 2017). To check this hypothesis, we ran the same analysis with $\alpha = 0.01$ (data not shown) and the average of REL_{IP} changed from 0.94 ($\alpha = 0.05$) to 0.92 ($\alpha = 0.01$) while REL_{GS3} changed from 0.88 ($\alpha = 0.05$) to 0.92 ($\alpha = 0.01$).

Some animals had larger differences between REL_{GS3} and REL_{IP} . However, we found that those were crossbred females considered as pure HO in the pedigree. This could lead to these extreme deviations since they come from populations with different allele frequencies from a pure HO population.

For JE and BS, calculating REL_{IP} based on PEC_{cc} did not yield accurate results (Figure 2). The REL_{IP} was over- or under-dispersed in both sexes and breeds. For JE, b_1 was -0.14 (males) and 0.63 (females); while for BS, it was 1.11 (males) and 0.99 (females). The scale of REL_{IP} differed in both breeds; with b_0 of -0.31 (BS) and 0.98 (JE) in males and -0.19 (BS) and 0.24 (JE) in females. This highlights the need to consider breed differences in allele frequencies when approximating the PEV based on core animals in multi-breed evaluations. In a simulated population, Lourenco et al. (2016) compared breeding values calculated from either breed-specific allele frequencies or allele frequencies of all breeds together and found that the best approach was constructing $\mathbf{G}$ based on within-breed allele frequencies. In our case, these large differences could result from different population structures, where assuming uniform allele frequencies across

populations is incorrect. The REL_{IP} could differ largely when uniform allele frequencies are assumed across all breeds, even if GEBV differences are negligible from a prediction standpoint.

Table 3 shows correlations between REL_{GS3} and REL_{IP} from the multi-breed evaluation. For HO, correlations were 0.95 (males) and 0.94 (females). For JE, they were -0.25 (males) and 0.40 (females), and for BS, they were 0.99 (males) and 0.80 (females). These results were expected, as Figures 1 and 2 show reliabilities that did not completely align, even when their correlations are high (i.e., for HO and BS). Previous studies (Lourenco et al., 2018; Garcia et al., 2020) showed high correlations between GEBV and IP estimations (around 0.99) when calculated based on core animals. However, no comparison of their reliabilities was done. For JE, the regression coefficient (-0.14 for males and 0.63 females), as well as correlations, suggest that the approximation of $Var(\hat{\mathbf{a}})$ based on PEC_{cc} may be affected by population structure and selection, ultimately affecting SNP effects and their variance.

In the multi-breed evaluation, allele frequencies seem to be dominated by the breed with more genotypes (HO in our case), leading to inflated and over-dispersed reliabilities in the breeds with fewer genotypes. Similarly, Cesarani et al. (2022) showed that GEBV can be inflated in multi-breed evaluations when there is no balanced breed representation in the core set.

Single-breed ssGBLUP evaluation

We ran single-breed evaluations to assess REL_{IP} without the influence of breed differences. For BS, REL_{IP} improved after removing the genotypes from other breeds (Figure 3), showing that the PEC approximation based on core animals works well. The plot reveals some outliers: few males had REL_{IP} values around 0.10 while REL_{GS3} was 0. These 32 males are the progeny of genotyped sires and ungenotyped dams, highlighting that REL_{IP} depends on an animal's relationship with the core set. These males were offspring of young sires with limited

information (5-10 progenies), explaining their low reliability (Bermann et al., 2022a). Excluding these animals, b_1 and b_0 improved (from 1.47 and -0.23 to 1.10 and -0.07, respectively). Similarly, for females with overestimated REL_{IP} ($n = 11$), removing them changed b_1 from 1.16 to 1.03 and b_0 from -0.08 to -0.02.

In Table 3, the correlations between multi-breed and single-breed evaluations for BS show an improvement for females, increasing from 0.80 in the multi-breed evaluation to 0.93 in the single-breed evaluation. Additionally, the means of REL_{GS3} and REL_{IP} in Table 4, both sexes were more similar in the single-breed evaluation compared to the multi-breed evaluation. In the single-breed evaluation, the differences were 0.05 for males and 0.01 for females, whereas in the multi-breed, the differences were 0.22 for males and 0.20 for females. This demonstrates that excluding the genotypes of other breeds allowed for a better approximation of reliabilities.

For JE, REL_{IP} changed between multi-breed and single-breed evaluations (Figure 4), and correlations improved significantly for both sexes (-0.25 to 0.99 in males and 0.40 to 0.99 in females). However, REL_{IP} remained inflated with b_1 of 1.43 (males) and 1.22 (females) and b_0 equal to 0.42 (males) and -0.23 (females). This may result from incorrect breed assignment of genotypes (e.g., crossbred JE vs. purebred JE) or the fact that these animals are offspring of animals not well connected to the population. Table 4 shows that, like BS, single-breed evaluations produced closer reliabilities, with differences of 0.07 (males) and 0.05 (females), compared to 0.12 and 0.13 in multi-breed evaluations.

Several studies have explored the possibility of performing multi-breed evaluations using reference populations of different breeds. These studies focused on the quality of estimated breeding values. VanRaden et al. (2009) highlighted that estimating SNP effects across breeds

without accounting for breed differences can cause allele frequency changes, particularly for rare alleles. Harris and Johnson (2010) showed that GEBV can be accurately predicted if the reference population representing all breeds is sufficiently large and uses a dense marker chip. Hayes et al. (2009) compared GBLUP evaluation accuracies obtained from the diagonal of $\mathbf{G}$ to realized accuracies in HO and JE populations. They found no improvement in multi-breed GBLUP evaluations, concluding that a large sample of genotypes representing both populations is important. Nowadays, this is less of an issue for most breeds in the U.S. as CDCB now has over 10 million genotyped animals, which presents a challenge for handling the genomic dataset, requiring the development of methods and approximations (Miszta et al., 2014a; Bermann et al., 2022b).

Garcia et al. (2020) showed in an Angus beef population that IP can be obtained using SNP effects from either a core group or an entire genotyped population with ssGBLUP and APY. Similarly, Cesarani et al. (2022) investigated the changes in predictive ability across five breeds with 4 million genotyped animals, finding similar predictive ability for cows and bulls in single- and multi-breed evaluations. This was attributed mainly to having a representative core set of animals from all five breeds and using unknown parent groups to model breed differences. Garcia et al. (2022) showed that IP accuracies could be derived from the variance of SNP effects based on the PEV of GEBV using $\mathbf{G}_{APY}^{-1}$.

In this study, we investigated the approximation of reliabilities based on PEC_{cc} in the U.S. dairy cattle multi-breed evaluation. Correct reliabilities require considering breed differences when constructing $\mathbf{G}_{cc}^{-1}$ and calculating the SNP variances. Allele frequencies were mostly dominated by breeds with larger genotype counts, leading to differences in reliabilities for breeds with fewer genotypes. A potential solution could be using metafounders (MF) (Legarra et al., 2015), but this

remains yet to be investigated. Bermann et al. (2023) developed a method for calculating reliabilities using MF that accounts for base population differences. They showed that reliabilities are identical to those obtained with a single-population model without MF.

In our study, REL_{IP} performed better in the single-breed evaluation. Differences were observed in specific cases, particularly for animals with one or both parents ungenotyped, parents poorly connected to the population, or incorrect breed assignments. Therefore, careful consideration is necessary when determining which animals to include in the evaluation and which should depend on REL_{IP} .

CONCLUSIONS

The results showed that in a multi-breed ssGBLUP evaluation, accounting for breed-specific allele frequencies is important when calculating PEC_{cc} . We demonstrated that the dominant breeds' allele frequencies affect the approximations of reliabilities of less represented breeds. The reliabilities of HO animals in the evaluation, or those calculated based on SNP effects, showed only minor differences, except in certain specific cases. In contrast, single-breed evaluations gave better reliability approximation for JE and BS, yet an overestimation of reliabilities was detected, and further investigation should be made. Moreover, when including a residual polygenic effect in the model, the reliabilities slightly shifted in scale, suggesting the need for adjustments.

ACKNOWLEDGMENTS

The authors thank the Council on Dairy Cattle Breeding (CDCB, Bowie, MD) for providing data access and participating in the research. The authors would also like to acknowledge the contribution of dairy producers who supplied the data through their participation in the Dairy Herd Improvement program.

REFERENCES

- Aguilar, I., I. Misztal, D. L. Johnson, A. Legarra, S. Tsuruta, and T. J. Lawlor. 2010. Hot topic: a unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *J Dairy Sci* 93(2):743-752. doi: 10.3168/jds.2009-2730
- Aguilar, I., I. Misztal, S. Tsuruta, A. Legarra, and H. Wang. 2014. PREGSF90 – POSTGSF90: Computational Tools for the Implementation of Single-step Genomic Selection and Genome-wide Association with Ungenotyped Individuals in BLUPF90 Programs 10. World Congress on Genetics Applied to Livestock Production (WCGALP). American Society of Animal Science, Vancouver, Canada.
- Bermann, M., I. Aguilar, D. Lourenco, I. Misztal, and A. Legarra. 2023. Reliabilities of estimated breeding values in models with metafounders. *Genetics Selection Evolution* 55(1):6. doi: 10.1186/s12711-023-00778-2
- Bermann, M., D. Lourenco, N. S. Forneris, A. Legarra, and I. Misztal. 2022a. On the equivalence between marker effect models and breeding value models and direct genomic values with the Algorithm for Proven and Young. *Genetics Selection Evolution* 54(1):52. doi: 10.1186/s12711-022-00741-7
- Bermann, M., D. Lourenco, and I. Misztal. 2022b. Efficient approximation of reliabilities for single-step genomic best linear unbiased predictor models with the Algorithm for Proven and Young. *Journal of Animal Science* 100(1)doi: 10.1093/jas/skab353
- Bussiman, F., C.-Y. Chen, J. Holl, M. Bermann, A. Legarra, I. Misztal, and D. Lourenco. 2023. Boundaries for genotype, phenotype, and pedigree truncation in genomic evaluations in pigs. *Journal of Animal Science* 101doi: 10.1093/jas/skad273

- Cesarani, A., D. Lourenco, S. Tsuruta, A. Legarra, E. L. Nicolazzi, P. M. VanRaden, and I. Misztal. 2022. Multibreed genomic evaluation for production traits of dairy cattle in the United States using single-step genomic best linear unbiased predictor. *J Dairy Sci* 105(6):5141-5152. doi: 10.3168/jds.2021-21505
- Christensen, O. F., and M. S. Lund. 2010. Genomic prediction when some animals are not genotyped. *Genet Sel Evol* 42(1):2. doi: 10.1186/1297-9686-42-2
- Garcia, A., I. Aguilar, A. Legarra, S. Tsuruta, I. Misztal, and D. Lourenco. 2022. Theoretical accuracy for indirect predictions based on SNP effects from single-step GBLUP. *Genet Sel Evol* 54(1):66. doi: 10.1186/s12711-022-00752-4
- Garcia, A. L. S., Y. Masuda, S. Tsuruta, S. Miller, I. Misztal, and D. Lourenco. 2020. Indirect predictions with a large number of genotyped animals using the algorithm for proven and young. *J Anim Sci* 98(6)doi: 10.1093/jas/skaa154
- Harris, B. L., and D. L. Johnson. 2010. Genomic predictions for New Zealand dairy bulls and integration with national genetic evaluation. *Journal of Dairy Science* 93(3):1243-1252. doi: <https://doi.org/10.3168/jds.2009-2619>
- Hayes, B. J., P. J. Bowman, A. C. Chamberlain, K. Verbyla, and M. E. Goddard. 2009. Accuracy of genomic breeding values in multi-breed dairy cattle populations. *Genetics Selection Evolution* 41(1):51. doi: 10.1186/1297-9686-41-51
- Henderson, C. R. 1984. *Applications of linear models in animal breeding*
- Legarra, A., I. Aguilar, and I. Misztal. 2009. A relationship matrix including full pedigree and genomic information. *J Dairy Sci* 92(9):4656-4663. doi: 10.3168/jds.2009-2061

- Legarra, A., O. F. Christensen, Z. G. Vitezica, I. Aguilar, and I. Misztal. 2015. Ancestral Relationships Using Metafounders: Finite Ancestral Populations and Across Population Relationships. *Genetics* 200(2):455-468. doi: 10.1534/genetics.115.177014
- Legarra, A., D. A. Lourenco, and Z. G. Vitezica. 2018. Bases for genomic prediction. *Short course* 1(1):1-141.
- Liu, Z. 2017. Approximating genomic reliabilities for national genomic evaluation. *Interbull Bulletin* (51)
- Lourenco, D., A. Legarra, S. Tsuruta, D. Moser, S. Miller, and I. Misztal. 2018. Tuning indirect predictions based on SNP effects from single-step GBLUP. *Interbull Bulletin* 52:48-53.
- Lourenco, D. A. L., S. Tsuruta, B. O. Fragomeni, C. Y. Chen, W. O. Herring, and I. Misztal. 2016. Crossbreed evaluations in single-step genomic best linear unbiased predictor using adjusted realized relationship matrices¹. *Journal of Animal Science* 94(3):909-919. doi: 10.2527/jas.2015-9748
- Misztal, I., A. Legarra, and I. Aguilar. 2014. Using recursion to compute the inverse of the genomic relationship matrix. *J Dairy Sci* 97(6):3943-3952. doi: 10.3168/jds.2013-7752
- Tier, B., and K. Meyer. 2004. Approximating prediction error covariances among additive genetic effects within animals in multiple-trait and random regression models. *Journal of Animal Breeding and Genetics* 121(1):77-89. doi: 10.1111/j.1439-0388.2003.00444.x
- Tsuruta, S., D. A. L. Lourenco, Y. Masuda, T. J. Lawlor, and I. Misztal. 2021. Reducing computational cost of large-scale genomic evaluation by using indirect genomic prediction. *JDS Communications* 2(6):356-360. doi: <https://doi.org/10.3168/jdsc.2021-0097>
- VanRaden, P. M. 2008. Efficient methods to compute genomic predictions. *J Dairy Sci* 91(11):4414-4423. doi: 10.3168/jds.2007-0980

- VanRaden, P. M., C. P. Van Tassell, G. R. Wiggans, T. S. Sonstegard, R. D. Schnabel, J. F. Taylor, and F. S. Schenkel. 2009. Invited review: reliability of genomic predictions for North American Holstein bulls. *J Dairy Sci* 92(1):16-24. doi: 10.3168/jds.2008-1514
- Vitezica, Z. G., I. Aguilar, I. Misztal, and A. Legarra. 2011. Bias in genomic predictions for populations under selection. *Genet Res (Camb)* 93(5):357-366. doi: 10.1017/s001667231100022x
- Wiggans, G. R., and J. A. Carrillo. 2022. Genomic selection in United States dairy cattle. *Front Genet* 13:994466. doi: 10.3389/fgene.2022.994466

TABLES

Table 5.1. Distribution of phenotypes and genotypes per breed

Breed	Phenotypes	Genotypes
AY	216,122	12,822
BS	366,892	58,271
GU	637,022	6,964
HO	34,112,088	1,523,968
JE	2,894,825	268,278
MS	57,159	0
XX	571,995	7,401

Abbreviations: AY = Ayrshire; BS = Brown Swiss; GU = Guernsey; HO = Holstein; JE = Jersey; MS = Milking Shorthorn; XX = Crossbreds.

Table 5.2. Number of core animals per breed and sex and total number of genotypes across breed

Breed	Number of core animals	
	Males	Females
AY	2,500	2,500
BS	2,500	2,500
GU	1,490	3,510
HO	7,500	7,500
JE	7,500	7,500
MS	2,500	2,500

Abbreviations: AY = Ayrshire; BS = Brown Swiss; GU = Guernsey; HO = Holstein; JE = Jersey; MS = Milking Shorthorn; XX = Crossbreds.

Table 5.3. Correlation between genomic reliabilities (REL_{GS3}) and reliability of Indirect predictions (REL_{IP}) for Holstein (HO), Jersey (JE), and Brown-Swiss (BS) from multi-breed and single-breed evaluations.

	HO	JE	BS
	Multi-breed		
Male	0.95	-0.25	0.98
Female	0.94	0.40	0.80
	Single-breed		
Male		0.99	0.98
Female		0.99	0.98

Table 5.4. Means of genomic reliability (REL_{GS3}) and reliability of Indirect predictions (REL_{IP}) for Jersey (JE) and Brown-Swiss (BS) from multi-breed and single-breed evaluations.

Multi-breed				
	JE		BS	
	Male	Female	Male	Female
REL_{GS3}	0.85	0.86	0.56	0.64
REL_{IP}	0.97	0.99	0.78	0.84
Single-breed				
REL_{GS3}	0.73	0.81	0.33	0.44
REL_{IP}	0.80	0.86	0.38	0.45

FIGURES

Figure 5.1. Multi-breed reliability for Holstein focal animals with 5% residual polygenic effect.

Abbreviations: REL_{GS3} = reliability of genomic predicted breeding value; REL_{IP} = reliability of indirect prediction.

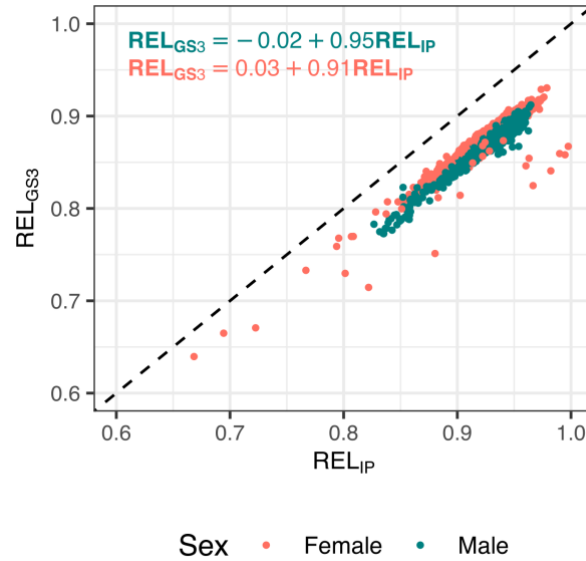


Figure 5.2. Multi-breed reliability for Jersey (JE) and Brown Swiss (BS) focal animals with 5% residual polygenic effect. Abbreviations: REL_{GS3} = reliability of genomic predicted breeding value; REL_{IP} = reliability of indirect prediction.

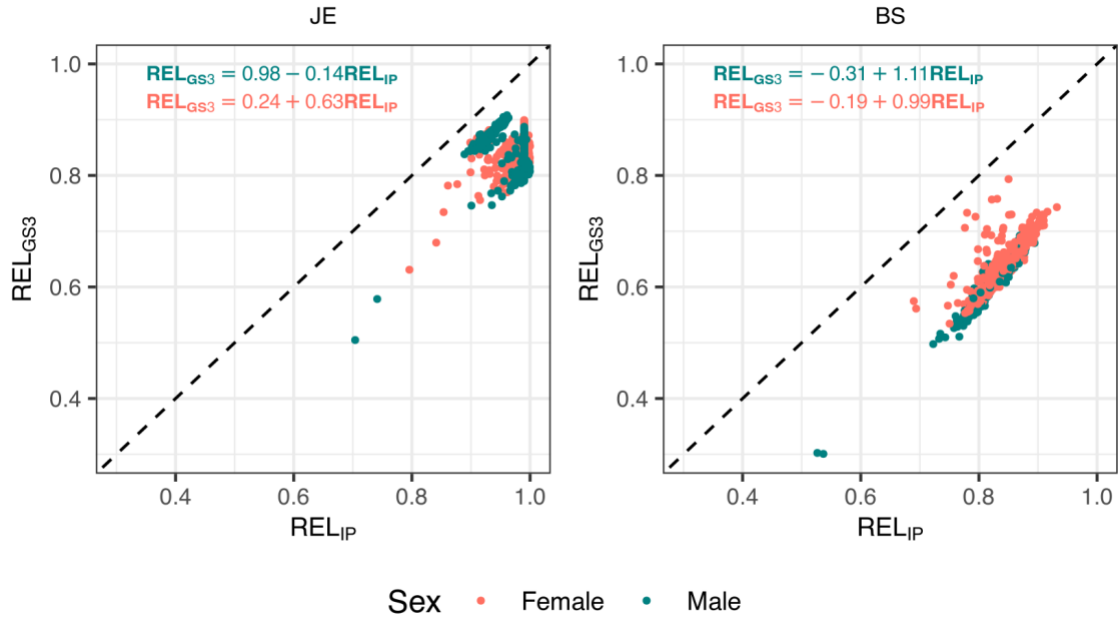


Figure 5.3. Single-breed reliability for Brown Swiss focal animals with 5% residual polygenic effect. Abbreviations: REL_{GS3} = reliability of genomic predicted breeding value; REL_{IP} = reliability of indirect prediction.

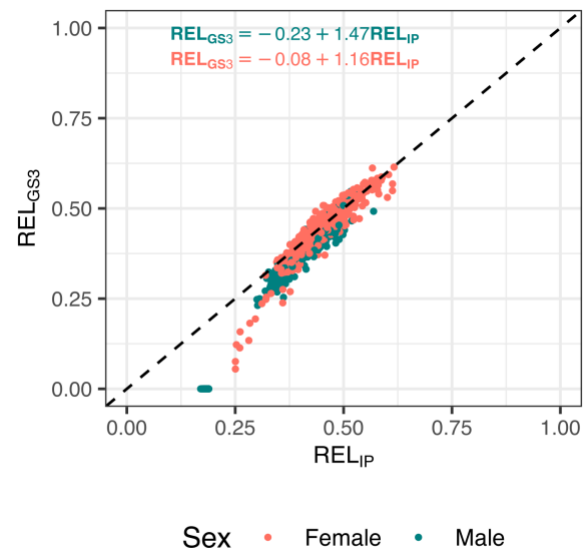
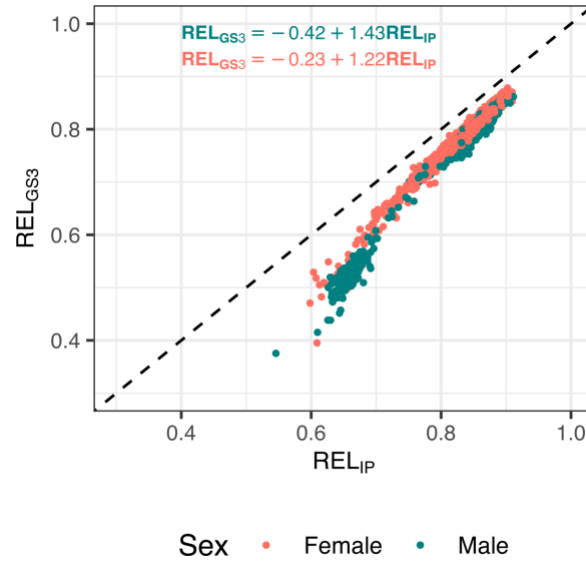


Figure 5.4. Single-breed reliability for Jersey focal animals with 5% residual polygenic effect.

Abbreviations: REL_{GS3} = reliability of genomic predicted breeding value; REL_{IP} = reliability of indirect prediction.



CHAPTER 6

CORRECTING OVERESTIMATION OF APPROXIMATE TRADITIONAL RELIABILITIES WHEN YOUNG GENOMIC BULLS ARE USED IN FEW HERDS

³Joe-Menwer Tabet, Ignacio Aguilar, Matias Bermann, Daniela Lourenco, Ignacy Misztal, Paul M. VanRaden, Zulma G. Vitezica, Andres Legarra. Submitted to Genetics Selection Evolution.

ABSTRACT

Differential treatment of daughters of the same sire within a herd is modelled as the herd-sire effect. Recent changes in management practices may have led to the extensive use of certain bulls in a limited number of herds. In that case, although the effect can be well accounted for in genetic evaluation models, some approximation methods for reliabilities do not consider it correctly, leading to an overestimation of some sires' approximated reliabilities. This study assessed the potential bias of these approximated reliabilities due to the herd-sire effect in both simulated and real dairy cattle records. Two existing methods were tested: Misztal-Wiggans, which includes a specific modification for herd-sire, and Tier-Meyer, which does not. We also modified and tested a Tier-Meyer method considering herd-sire. We observed that in the presence of the herd-sire effect, reliabilities obtained by approximations were overestimated by Tier-Meyer for sires with many daughters in a limited number of herds. This was true even for sires with a large number of daughters. The Misztal-Wiggans method performed correctly. We introduced a modified Tier-Meyer method that weighs the information transmitted by the daughter to the sire as a function of the herd-sire information. As a result, the modified Tier-Meyer method performed well in both simulated and real data. For cows, the inclusion of the herd-sire effect had minimal impact. This study identified possible overestimation of approximated reliabilities of sires with daughters concentrated in a few herds when there is a herd-sire effect. This bias occurs when the herd-sire effect is not correctly modeled in reliability approximation methods. Methods that specifically accounted for the herd-sire effect produced unbiased reliability estimates.

INTRODUCTION

Over the past two decades, the number of U.S. dairy farms has decreased while the number of large farms has grown. For example, farms with fewer than 1,000 cows have declined, while those with 1,000 or more cows have increased (Gillespie et al., 2024). These large herds tend to have systematic management practices, often using a few bulls to simplify their operations. Alongside this trend, genomic selection has led to the increased use of young genomic bulls, which have limited semen straw availability. As a result, some young bulls are exclusively used in a small number of herds instead of the traditional sampling approach (e.g., common around the year 2000) in which sires were more widely assigned across herds. Moreover, the extensive use of a single sire within a herd can result in preferential treatment of that sire's daughters. This can be understood as a (random) herd-sire (HS) effect or, equivalently, as creating an environmental covariance between half-sibs of the same sire within a herd.

Historically, the importance of the HS effect was greater before the widespread use of Artificial Insemination (AI) because natural service sires have daughters only in one herd, resulting in higher environmental covariance. In the traditional sampling approach, AI sires had several offspring distributed in many herds (Van Vleck, 1966; Kelleher et al., 1967). The HS effect has recently regained importance with the recent changes in management practices in large herds, leading to the non-random, less distributed allocation of sires to herds (Jang et al., 2022).

The HS effect is often introduced in the genetic evaluation model (Norman, 1974; Wiggans et al., 1988; VanRaden and Wiggans, 1991). Failing to do so can introduce bias and overestimate reliabilities (Meyer, 1987). Additionally, Banos and Shook (1990) noted that ignoring the HS effect overestimates the accuracy in sire evaluations, particularly when sires have daughters in a few herds. Previous methods for genetic evaluation include specific corrections of reliabilities for

the existence of herd-sire interactions (Norman, 1974; Meyer, 1987; Wiggans et al., 1988; VanRaden and Wiggans, 1991).

The overestimation of reliabilities by approximate methods when there is an HS effect is a problem specific to HS, unlike, say, contemporary groups (CG) (Harris and Johnson, 1998; Liu et al., 2004; Tier and Meyer, 2004). This bias in reliabilities is shown in the following example. Consider a “sire model” with two sires and two herds, each with 100 records. Assume that there is no herd effect but there is a preferential treatment HS effect, i.e., daughters of sire 1 are treated differently than daughters of sire 2, with these differences varying by herd. If each sire is assigned to a single herd, separating the sire from the preferential treatment captured by HS becomes impossible, i.e., they are collinear. Conventional approaches for reliability, e.g. (Harris and Johnson, 1998), rely on the number of records of its HS effect (e.g., 100); however, the collinearity of HS (an environmental effect) and sire itself (a genetic effect) is not considered. If each sire is assigned to both herds, then there are two HS effects per sire, the collinearity between HS and sire is broken, and reliability increases with the number of records in the CG.

The genetic evaluation model used by the Council on Dairy Cattle Breeding (CDCB; Bowie, MD) for U.S. dairy cattle includes the HS effect for most traits. However, the reliabilities estimated are an approximation that ignores the off-diagonals that associate HS with sires of daughters within the same herd. So far, recent reliability estimation methods have not addressed the overestimation of reliabilities due to the HS. As a result, there are anecdotal reports of genomic bulls assigned to a few herds whose traditional (non-genomic) reliabilities may be overestimated. For example, the traditional estimated breeding value (EBV) for milk yield of bulls JEUSA000117723269 and JE840003012659218 had, in December 2024, 0.97 and 0.98 traditional reliabilities with respectively 377 daughters in a single herd and 390 daughters in three herds.

These reliabilities are likely overestimated because the model includes HS interaction. Although such cases represent a small fraction of the population, they often involve significant financial investments for their owners. Therefore, it is fair that they should receive unbiased reliability.

The objective of this study was to explore, through simulation and real data, the extent of bias in reliability approximation when the HS effect is included in the model and ignored in the approximation process and provide possible solutions to the bias using approximate methods. Two approximation methods were computed, Misztal and Wiggans (1988) proposed a method for approximating PEV by simplifying computations, but its applicability is limited in complex models. Tier and Meyer (2004) introduced a three-step reliability approximation method adaptable to multi-trait and complex models, but it overlooks the HS effect, potentially causing overestimation. As a byproduct, a modification to the approach by Tier and Meyer (2004) is proposed to account for the HS effects.

MATERIALS AND METHODS

Methods to estimate reliabilities

Existing methods

The reliability of the EBV for the i^{th} animal (EBV_i) can be directly calculated by inverting the coefficient matrix of the Mixed Model Equations (MME) as $REL_{EXACT_i} = 1 - \frac{PEV_i}{a_{ii}\sigma_u^2}$, where PEV_i is the prediction error variance of EBV_i , obtained from the diagonal of the inverse of the MME, a_{ii} is the diagonal element in the numerator relationship matrix (**A**) for animal i , and σ_u^2 is the animal additive genetic variance. The computational costs of inversion are too high for large data sets. Methods have been developed to approximate PEV, either by approximating the diagonal elements of the inverse coefficient matrix (Misztal and Wiggans, 1988; Meyer, 1989) or by

sequentially accumulating information from the animal's own or progeny records, parents, or other relatives (Tier and Meyer, 2004).

Misztal and Wiggans (1988) proposed a method that approximates PEV by ignoring off-diagonals and absorbing the non-genetic model effects (i.e., subclasses, permanent environmental effect (pe)). Their approach, along with a method to consider genomic information in the approximation (Bermann et al., 2022b), is programmed in the software accf90GS2 from the BLUPF90 suite (Lourenco et al., 2022). Using a tailored approach, the HS effect was included in the reliability estimate from the beginning (Misztal and Wiggans, 1988; Wiggans et al., 1988). This approach calculates “herd-sire weights” for each sire by accumulating daughter weights across herds while accounting for the information from each lactation separately, adjusted for the CG effects. However, a limitation of the method in (Misztal and Wiggans, 1988) is that extending to more complex models (e.g., random regression models) and complex missing trait patterns is hard. We call this approximation of reliability REL_{MW} .

Tier and Meyer (2004) developed a three-step approach to approximate PEV by adjusting the diagonal blocks of the coefficient matrix. The first adjustment accounts for the animal's records based on the size of its subclass, followed by accumulating information on progeny and other descendants and, finally, information on parents, ancestors, and other relatives. The advantage of this method is its natural extension to multiple trait and random regression models. However, it does not explicitly account for the HS effect. This method is programmed in the software accf90GS3 from the BLUPF90 suite (still unreleased). We call this approximation of reliability REL_{TMnohs} .

Modified Tier-Meyer to consider herd-sire effect.

The standard Tier-Meyer method does not account for the HS effect in its current form. In this study, we propose a modification to incorporate it. This modification, called REL_{TM} , involves assigning a weight to matrix $\mathbf{E}_l$ of Eq. 6 from Tier and Meyer (2004). This equation sends contributions of information from offspring to parents, so the adjusted equation for animal i and its $l = 1 \dots p_i$ progeny is:

$$\mathbf{E}_i = \frac{1}{3}\mathbf{G}_0^{-1} - \frac{4}{9}\mathbf{G}_0^{-1} \left(\mathbf{D}_i + \sum_{l=1}^{p_i} \mathbf{K}_l \mathbf{E}_l \mathbf{K}_l + \frac{4}{3}\mathbf{G}_0^{-1} \right)^{-1} \mathbf{G}_0^{-1} \quad [6.1]$$

Where $\mathbf{G}_0$ is a matrix with genetic covariances and $\mathbf{D}$ and $\mathbf{E}$ are submatrices corresponding, respectively, to the contribution of the animal's own phenotype (after adjusting for fixed effects) and to the contribution of its progeny. Matrix $\mathbf{K}_l$ is a (number of traits by number of traits) diagonal matrix, different for each animal l , with its (k, k) element equal to $\sqrt{c_k}$. The weight c_k (of cow l) represents the amount of information cow l for trait k gives to her sire, compared to a weight of 1 if there was no HS effect. The weight is first calculated within each herd-sire j for trait k as: $chs_{jk} = \frac{EDC_{jk}}{\sum w_{jk}}$, where the denominator $\sum w_{jk}$ is the weighted sum of the number of records for trait k and herd-sire j , and EDC_{jk} represents the Equivalent Daughter Contributions (EDC) at herd-sire j and trait k . This EDC_{jk} is computed separately for each trait in two steps. First, the formula by VanRaden and Wiggans (1991) gives the maximum possible reliability of sire due to daughters in a single HS j for trait k , $REL_{anim(jk)}^*$:

$$REL_{anim(jk)}^* = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_{hs}^2 + \frac{\sigma_\epsilon^2}{\sum w_{jk}}} \quad [6.2]$$

Where, for trait k , $\sigma_s^2 = 0.25\sigma_u^2$ is the variance component of the sire effect; σ_{hs}^2 is the variance component of the HS effect; $\sigma_\epsilon^2 = \sigma_e^2 + 0.5\sigma_u^2$, assuming the dam's breeding value is known

(VanRaden and Wiggans, 1991). If permanent environment (pe) effects are included, then $\sigma_e^2 = \sigma_e^2 + 0.5\sigma_u^2 + \sigma_{pe}^2$, and $\sum w_{jk}$ is the weighted sum of the number of records for the HS as above. Second, this quantity is transformed into *EDC*:

$$EDC_{(jk)} = \frac{4 - 2h^2}{h^2} \left(\frac{REL_{anim(jk)}^*}{1 - REL_{anim(jk)}^*} \right) \quad [6.3]$$

Which is finally used in $chs_{jk} = \frac{EDC_{jk}}{\sum w_{jk}}$. Because a cow can potentially change herds, the final weight $c_{l(k)}$ for cow l in trait k is an average of the weights chs_{jk} of all the herd-sires j in which the cow l participated for that trait, i.e., if the cow had one lactation in herd A and two lactations in herd B, $c_{l(k)} = \frac{1chs(A)_k + 2chs(B)_k}{3}$.

Finally, the weights are not used when transmitting the information from parents to offspring (Eq. 7 in (Tier and Meyer, 2004)) because they consider the phenotypic information, which is already done in our Eq. 1.

Data and model

All data came from simulated or preexisting databases. Ethical approval for the use of animals was thus deemed unnecessary.

Simulated data

To examine changes in approximated reliabilities due to ignoring the HS effect, we conducted a simulation that generates a data structure mimicking a two-generation U.S. dairy cattle population -parent and offspring- with sires classified into “Odd” and “Normal.” Odd sires have a limit of 500 straws, whose use was deemed in large “quotas” of a few hundred straws by the same producers at a time; as a result, they were limited to 1-2 herds, with a maximum of 500 daughters. Normal sires had 10-600 daughters spread across 1-60 herds. We simulated two types of herds: “Large” with more than 3,000 animals and “Average” with less than 3,000. A total of 3,542 sires

(542 Odd and 3,000 Normal) and 1,000 herds were simulated. All daughters had a unique dam and a single record. As a result, we created a structure consisting of sires with varying reliabilities based on the type and distribution of daughters. Figures 1 and 2 illustrate the distribution of daughters and herds per each sire type. Odd sires were found in 1 to 2 herds with 500 daughters, and Normal sires were distributed across many herds with a more homogenous distribution of their daughters.

We considered a single-trait animal model with the herd as a fixed effect, HS effect with variance $\mathbf{I}\sigma_{hs}^2$, where $\sigma_{hs}^2 = 5$, an animal additive genetic effect with a mean zero and variance $\mathbf{A}\sigma_u^2$ with $\sigma_u^2 = 20$. The residual variance was 75. Reliabilities were obtained either from the inverse of the MME (REL_{EXACT}) or from the approximation methods Misztal-Wiggans (REL_{MW}), which corrects for herd and herd-sire, standard Tier-Meyer (REL_{TMnohs}), which corrects for herd only, and modified Tier-Meyer (REL_{TM}), which corrects for herd and herd-sire.

Real data

We also used data provided by the CDCB. A subset of records for Jersey (JE) animals was extracted. The first data included 7,824,163 milk yield records from 2,967,998 cows, 3,863,135 animals in the pedigree, and 687,889 HS; this will be referred to as DataFull. However, this file proved too large to obtain exact reliabilities by inversion, which we attribute to a large number of fill-ins in the sparse Cholesky decomposition caused by the HS effect. Thus, we further reduced the dataset to records of cows born after 2015, resulting in DataReduced, with 1,605,821 records, 776,390 cows in records, 1,725,100 animals in the pedigree, and 88,035 HS. In both cases, the pedigree spanned as many ancestors as possible.

We ran a single-trait repeatability model with CG and age-parity as fixed effects. The HS ($\sigma_{hs}^2 = 73 \times 10^5$) and animal additive genetic ($\sigma_u^2 = 21 \times 10^6$) effects were modeled as random effects defined earlier, with pe random effect with a mean zero and variance $\mathbf{I}\sigma_{pe}^2$, where $\sigma_{pe}^2 = 29 \times 10^6$ the variance component of the pe effect. The residual variance was 47×10^6 . Computation of reliabilities ignored Unknown Parent Groups in all cases, as typically assumed.

Exact reliabilities REL_{EXACT} were obtained using blupf90+, REL_{MW} are computed by accf90GS2, and REL_{TMnohs} and REL_{TM} are computed by accf90GS3, all programs from the BLUPF90 suite (Lourenco et al., 2022).

Reliability comparison

For the simulated dataset and DataReduced, we compared REL_{EXACT} versus REL_{MW} , REL_{TM} and REL_{TMnohs} . In the case of DataFull, REL_{EXACT} was not available because it was impossible to invert the MME due to their large size, so we compared REL_{MW} versus REL_{TM} and REL_{TMnohs} .

The comparisons were detailed by bull category. In the simulated data, bulls are classified as “Odd” and “Normal”. With the real JE data, it is necessary to identify sires that may have many daughters concentrated in only a few herds to categorize them as “Odd” or “Normal”. In that case, we calculated a measure that we called the h-index, analogous to the Google Scholar h-index (Wikipedia, 2024, November 12). A sire with an h-index of n has at least n daughters in n herds. For instance, a sire with a daughter count of 10, 7, 4, and 2 across 4 herds has an h-index of 3 as it has at least 3 daughters in 3 herds. Sires with many daughters concentrated only in a few herds were classified as “Odd” with an h-index ≤ 5 . In contrast, sires with a more even distribution of daughters across multiple herds were classified as “Normal” if they had an h-index ≥ 10 , implying at least 100 daughters with records. Results for sires between h-indexes of 5 and 10 were

inconclusive in preliminary inspection, and we did not use them for diagnostics. For example, consider two sires, each with 200 daughters. Sire 1 has 100, 50, 23, 20, 3, 2, 1, and 1 daughters across 8 herds, giving an h-index of 4. This indicates that most of its daughters are concentrated in a few herds, resulting in a larger number of half-siblings within each herd. Sire 2, on the other hand, has 15 daughters in each of 11 herds and 5 daughters in each of other 7 herds, resulting in an h-index of 11, leading to less influence of preferential treatment in the sire proof.

RESULTS AND DISCUSSION

For simulated data, Figure 3 and Table 1 display the results of the true and approximated reliabilities of Odd bulls. Both REL_{MW} and REL_{TM} correctly estimated reliabilities with a slight overestimation for REL_{TM} , whereas REL_{TMnohs} overestimated reliabilities by more than 0.40 because this method does not explicitly account for the HS effect. Table 1 presents the correlations and slope values of the regression of REL_{EXACT} on the approximated reliabilities. The correlations were 0.99 for both REL_{MW} and REL_{TM} and 0.48 for REL_{TMnohs} . Including the HS, the correlations increased. Additionally, the modified Tier-Meyer improved the slope of the regression of REL_{EXACT} on approximated reliabilities from 0.85 (REL_{TMnohs}) to 1.03 (REL_{TM}).

Table 1 shows Normal sires' true and approximated reliabilities. We can see a slight overestimation with REL_{TM} . We see correlations of 0.99 across all three cases. This shows that when sires are well spread among herds, the correction for HS is less important. However, REL_{TMnohs} ($b_1 = 1.22$) were under-dispersed compared to REL_{TM} and REL_{MW} with b_1 values close to 1.

REL_{TM} slightly overestimated reliabilities, likely due to only accounting for first-order relationships within the same sub-classes when evaluating observations with respect to the animal and its ancestors (Tier and Meyer, 2004).

For DataReduced, results in Figure 4 and Table 2 show exact and approximated reliabilities for Odd and Normal sires according to the h-index; other sires are excluded from the figure for clarity. Approximated reliabilities REL_{TM} and REL_{MW} were only slightly under-estimated for Odd sires, whereas REL_{TMnohs} clearly over-estimated reliabilities. The correlations and slopes of REL_{EXACT} regressed on REL_{TM} or REL_{MW} were close to 1, while REL_{TMnohs} showed lower correlation and higher dispersion. For Normal sires, all approximated reliabilities were quite similar to REL_{EXACT} , except for the slopes. Again, we confirm that for Odd sires, the correction of reliabilities for HS is important, but less so for Normal sires.

In DataReduced, we also studied cow reliabilities. Figure 5 shows the reliabilities of cows with records, indicating no impact of considering the HS effect on reliability estimates. Table 3 shows correlations exceeding 0.81 and regression slopes around 1 for REL_{MW} , REL_{TM} and REL_{TMnohs} .

In DataFull, we were unable to obtain REL_{EXACT} . Based on previous results, we considered REL_{MW} as a good approximation to REL_{EXACT} . Figure 6 and Table 4 compare REL_{TM} and REL_{TMnohs} to REL_{MW} for both Odd and Normal sires in DataFull. The results show that REL_{TM} and REL_{MW} had similar means across both types of sires. The modified Tier-Meyer method, REL_{TM} , had dispersions of 0.87 and 1.10 for Odd and Normal sires, respectively, but ignoring the HS in REL_{TMnohs} had regression coefficients far from 1. The differences were more pronounced for Odd sires, with correlations for REL_{TM} of 0.85 with REL_{MW} and 0.45 for REL_{TMnohs} .

The bulls in our example had REL_{TM} and REL_{TMnohs} of 0.71 and 0.97 (JEUSA000117723269), and 0.65 and 0.96 (JE840003012659218) in DataFull. Values for REL_{TMnohs} agree with CDCB's traditional reliabilities. Correcting for the HS effect as in REL_{TM} likely reduced the overestimation, although we do not have REL_{EXACT} values to compare.

This study confirmed with both real and simulated data that ignoring the specifics of the HS effect in reliability approximation leads to overestimated reliabilities for sires with daughters concentrated in few herds, even if the number of daughters is large (Kelleher et al., 1967; Norman, 1974; Misztal and Wiggans, 1988). This could lead to incorrect selection decisions and different appraisals of risk for sires' owners. In the past, the inclusion of HS in reliability approximation was highly emphasized (Meyer, 1987; Wiggans et al., 1988; Banos and Shook, 1990), but then it had less importance due to the use of AI sires, and recently, this effect has regained importance for certain sires. Decades ago, Kelleher et al. (1967) suggested that using AI sires across herds would make HS less relevant for reliability approximation. Here, we confirm that the effect is still relevant even with AI because some sires are not spread across herds.

Exact reliabilities are difficult to compute due to the computational demands of inverting the MME, so approximation methods are needed to calculate PEV. The method by Misztal and Wiggans (1988) performed well, though its application to complex models (e.g., random regression) is not straightforward. We found that the more general method by Tier and Meyer (2004) did not explicitly address the HS effect, resulting in overestimated reliabilities for sires. We modified the Tier-Meyer method by assigning a weight (c) based on HS counts of records, to reduce the amount of information the cow transfers back to its sire. By doing so, we compensated for the overestimation of reliabilities, aligning estimates with the true values. With the

modification, both methods (Misztal-Wiggans and modified Tier-Meyer) yielded similar results and close to exact values obtained by inversion.

CONCLUSIONS

This study evaluated the bias in approximated reliabilities due to ignoring the HS effect in two approximated methods used for computing reliabilities. We identified excessively high approximated reliabilities, particularly when sires have daughters concentrated in a few herds, and the HS effect was not correctly accounted for. The existing Misztal-Wiggans algorithm for approximated reliabilities considers the HS well. We modified the Tier-Meyer method to weigh down the reliability of a sire, accounting for the loss of information due to the descendants of a sire being in a few HS classes. As a result, the modified Tier-Meyer algorithm performed well and eliminated the bias.

ACKNOWLEDGMENTS

The authors thank the Council on Dairy Cattle Breeding (CDCB, Bowie, MD) for providing data access and participating in the research. The authors would also like to acknowledge the contribution of dairy producers who supplied the data through their participation in the Dairy Herd Improvement program. This material is based upon research supported by the Chateaubriand Fellowship of the Office for Science & Technology of the Embassy of France in the United States.

REFERENCES

- Banos, G., and G. E. Shook. 1990. Genotype by Environment Interaction and Genetic Correlations Among Parities for Somatic Cell Count and Milk Yield¹. *Journal of Dairy Science* 73(9):2563-2573. doi: [https://doi.org/10.3168/jds.S0022-0302\(90\)78942-4](https://doi.org/10.3168/jds.S0022-0302(90)78942-4)
- Bermann, M., D. Lourenco, and I. Misztal. 2022. Efficient approximation of reliabilities for single-step genomic best linear unbiased predictor models with the Algorithm for Proven and Young. *Journal of Animal Science* 100(1)doi: 10.1093/jas/skab353
- Gillespie, J., E. Njuki, and A. Terán. 2024. Structure, costs, and technology used on U.S. dairy farms. ERR-334, U.S. Department of Agriculture, Economic Research Service.
- Harris, B., and D. Johnson. 1998. Approximate Reliability of Genetic Evaluations Under an Animal Model. *Journal of Dairy Science* 81(10):2723-2728. doi: [https://doi.org/10.3168/jds.S0022-0302\(98\)75829-1](https://doi.org/10.3168/jds.S0022-0302(98)75829-1)
- Jang, S., D. Lourenco, and S. Miller. 2022. Inclusion of sire by herd interaction effect in the genomic evaluation for weaning weight of American Angus. *Journal of Animal Science* 100(3)doi: 10.1093/jas/skac057
- Kelleher, D. J., A. E. Freeman, and J. L. Lush. 1967. Importance of Bull $\times$ Herd-Year-Season Interaction in Milk Production¹. *Journal of Dairy Science* 50(10):1703-1707. doi: [https://doi.org/10.3168/jds.S0022-0302\(67\)87697-5](https://doi.org/10.3168/jds.S0022-0302(67)87697-5)
- Liu, Z., F. Reinhardt, A. Bünger, and R. Reents. 2004. Derivation and Calculation of Approximate Reliabilities and Daughter Yield-Deviations of a Random Regression Test-Day Model for Genetic Evaluation of Dairy Cattle. *Journal of Dairy Science* 87(6):1896-1907. doi: [https://doi.org/10.3168/jds.S0022-0302\(04\)73348-2](https://doi.org/10.3168/jds.S0022-0302(04)73348-2)

- Lourenco, D., S. Tsuruta, I. Aguilar, Y. Masuda, M. Bermann, A. Legarra, and I. Misztal. 2022. 366. Recent updates in the BLUPF90 software suite. In: Proceedings of 12th World Congress on Genetics Applied to Livestock Production (WCGALP). p 1530 - 1533.
- Meyer, K. 1987. Estimates of variances due to sire $\times$ herd interactions and environmental covariances between paternal half-sibs for first lactation dairy production. *Livestock Production Science* 17:95-115.
- Meyer, K. 1989. Approximate accuracy of genetic evaluation under an animal model. *Livestock Production Science* 21(2):87-100. doi: [https://doi.org/10.1016/0301-6226\(89\)90041-9](https://doi.org/10.1016/0301-6226(89)90041-9)
- Misztal, I., and G. R. Wiggans. 1988. Approximation of Prediction Error Variance in Large-Scale Animal Models. *Journal of Dairy Science* 71:27-32. doi: [https://doi.org/10.1016/S0022-0302\(88\)79976-2](https://doi.org/10.1016/S0022-0302(88)79976-2)
- Norman, H. D. 1974. Factors That Should Be Considered in a National Sire Summary Model1. *Journal of Dairy Science* 57(8):955-962. doi: [https://doi.org/10.3168/jds.S0022-0302\(74\)84992-1](https://doi.org/10.3168/jds.S0022-0302(74)84992-1)
- Tier, B., and K. Meyer. 2004. Approximating prediction error covariances among additive genetic effects within animals in multiple-trait and random regression models. *Journal of Animal Breeding and Genetics* 121(1):77-89. doi: 10.1111/j.1439-0388.2003.00444.x
- Van Vleck, L. D. 1966. Environmental Correlations in Sire Evaluation. *Journal of Dairy Science* 49(1):56-60. doi: [https://doi.org/10.3168/jds.S0022-0302\(66\)87786-X](https://doi.org/10.3168/jds.S0022-0302(66)87786-X)
- VanRaden, P. M., and G. R. Wiggans. 1991. Derivation, calculation, and use of national animal model information. *J Dairy Sci* 74(8):2737-2746. doi: 10.3168/jds.S0022-0302(91)78453-

- Wiggans, G. R., I. Misztal, and L. D. Van Vleck. 1988. Implementation of an Animal Model for Genetic Evaluation of Dairy Cattle in the United States. *Journal of Dairy Science* 71:54-69. doi: [https://doi.org/10.1016/S0022-0302\(88\)79979-8](https://doi.org/10.1016/S0022-0302(88)79979-8)
- Wikipedia. 2024, November 12. *h-index*. <https://en.wikipedia.org/wiki/H-index> (Accessed November 14, 2024).

TABLES

Table 6.1. Reliability means, correlation, and regression of REL_{EXACT} on approximated reliabilities of Odd and Normal sires from simulated dataset.

	¹ Odd			Normal		
	Mean	b ₁	Correlation	Mean	b ₁	Correlation
² REL_{EXACT}	0.51			0.86		
REL_{MW}	0.50	1.00	0.99	0.85	0.97	0.99
REL_{TM}	0.58	1.03	0.99	0.87	1.04	0.99
REL_{TMnohs}	0.95	0.85	0.48	0.90	1.22	0.99

¹**Odd:** Sires with many daughters concentrated in a small number of herds; **Normal:** Sires with daughters more evenly distributed across many herds.

² REL_{EXACT} : Exact reliabilities from the inverse of the coefficient matrix with blupf90+; REL_{MW} : Approximated reliabilities based on Misztal-Wiggans; REL_{TM} : Approximated reliabilities based on modified Tier-Meyer; REL_{TMnohs} : Approximated reliabilities based on standard Tier-Meyer.

Table 6.2. Reliability means correlation and regression of REL_{EXACT} on approximated reliabilities of Odd and Normal sires from DataReduced.

	¹ Odd			Normal		
	Mean	b ₁	Correlation	Mean	b ₁	Correlation
² REL_{EXACT}	0.70			0.95		
REL_{MW}	0.65	0.86	0.97	0.98	0.77	0.98
REL_{TM}	0.66	0.90	0.97	0.95	0.89	0.99
REL_{TMnobs}	0.76	0.79	0.64	0.97	1.36	0.90

¹**Odd:** Sires with many daughters concentrated in a small number of herds; **Normal:** Sires with daughters more evenly distributed across many herds.

² REL_{EXACT} : Exact reliabilities from the inverse of the coefficient matrix with blupf90+; REL_{MW} : Approximated reliabilities based on Misztal-Wiggans; REL_{TM} : Approximated reliabilities based on modified Tier-Meyer; REL_{TMnobs} : Approximated reliabilities based on standard Tier-Meyer.

Table 6.3. Correlation and regression of RE_{EXACT} with approximated reliabilities for cows with records in DataReduced

	Mean	b₁	Correlation
¹ RE_{EXACT}	0.50		
REL_{MW}	0.45	1.07	0.91
REL_{TM}	0.43	1.00	0.81
$\text{REL}_{\text{TMnohs}}$	0.43	1.02	0.81

¹ RE_{EXACT} : Exact reliabilities from the inverse of the coefficient matrix with blupf90+; REL_{MW} : Approximated reliabilities based on Misztal-Wiggans; REL_{TM} : Approximated reliabilities based on modified Tier-Meyer; $\text{REL}_{\text{TMnohs}}$: Approximated reliabilities based on Tier-Meyer.

Table 6.4. Reliability means, correlation, and regression of REL_{MW} on approximated reliabilities of Odd and Normal sires from DataFull.

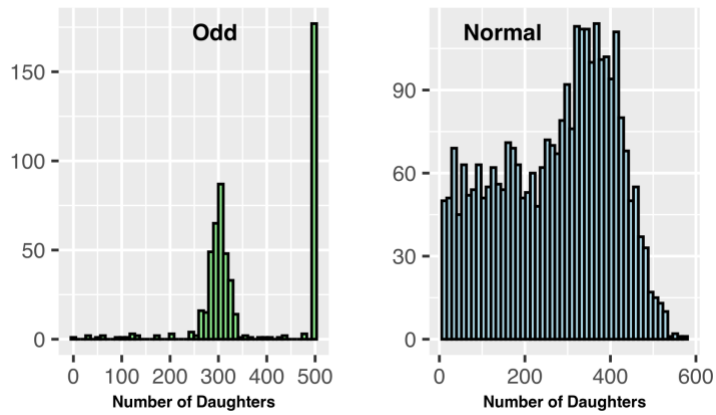
	¹ Odd			Normal		
	Mean	b ₁	Correlation	Mean	b ₁	Correlation
² REL_{MW}	0.60			0.94		
REL_{TM}	0.63	0.87	0.85	0.95	1.10	0.98
REL_{TMnohs}	0.78	0.45	0.31	0.97	1.41	0.85

¹**Odd:** Sires with many daughters concentrated in a small number of herds; **Normal:** Sires with daughters more evenly distributed across many herds.

² REL_{MW} : Approximated reliabilities based on Misztal-Wiggans; REL_{TM} : Approximated reliabilities based on modified Tier-Meyer; REL_{TMnohs} : Approximated reliabilities based on standard Tier-Meyer.

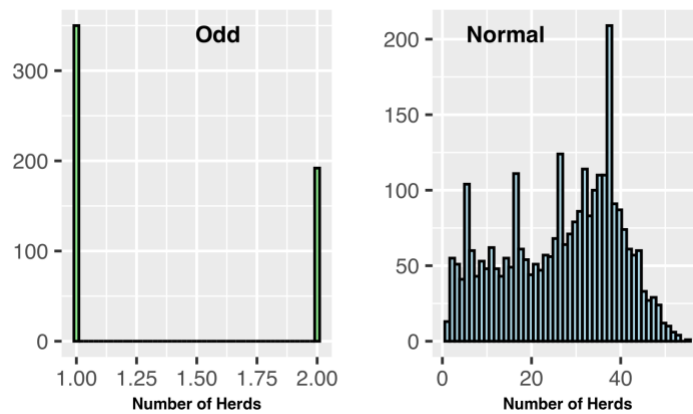
FIGURES

Figure 6.1. Distribution of daughters per sire



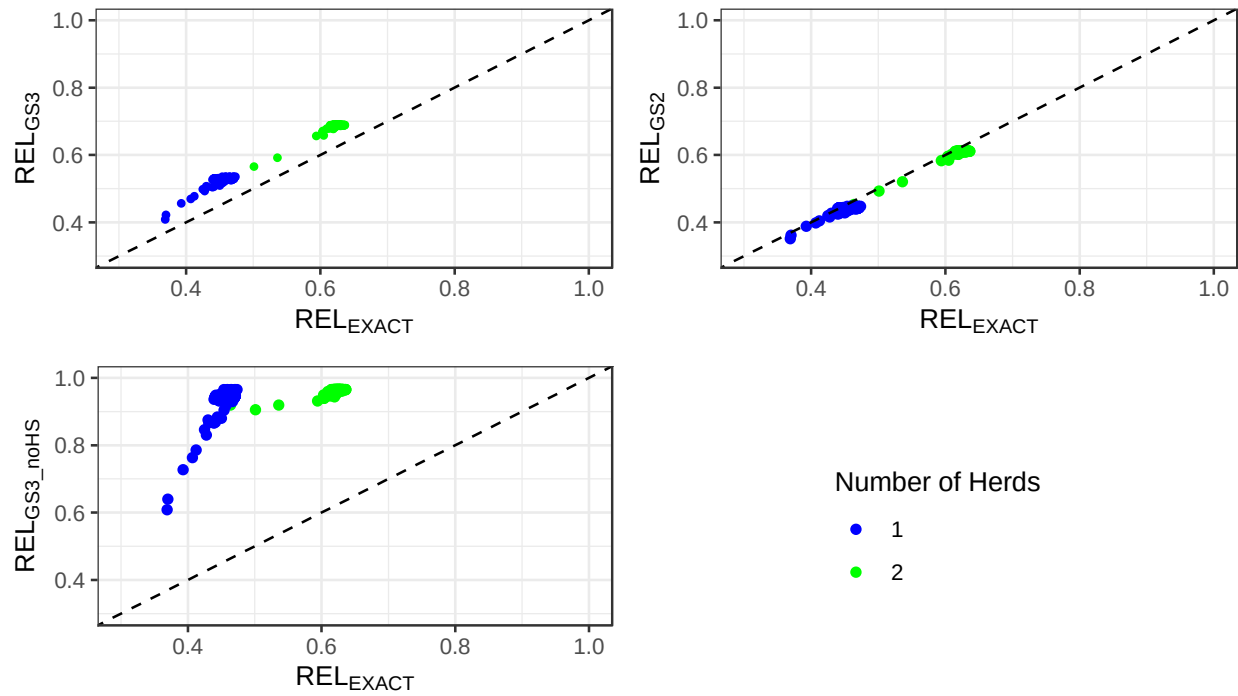
Abbreviations: **Odd:** Sires with many daughters concentrated in a small number of herds;
Normal: Sires with daughters more evenly distributed across many herds

Figure 6.2. Distribution of herds per sire



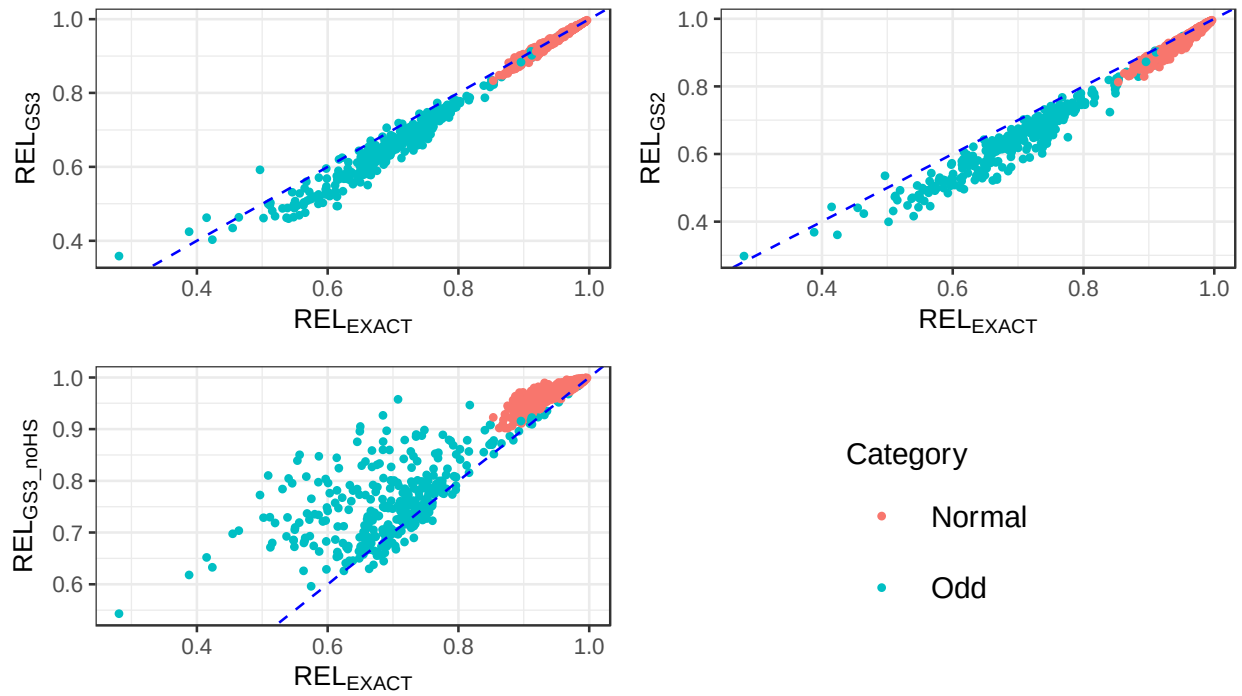
Abbreviations: Odd: Sires with many daughters concentrated in a small number of herds; Normal: Sires with daughters more evenly distributed across many herds

Figure 6.3. Reliabilities of Odd sires from the simulated dataset



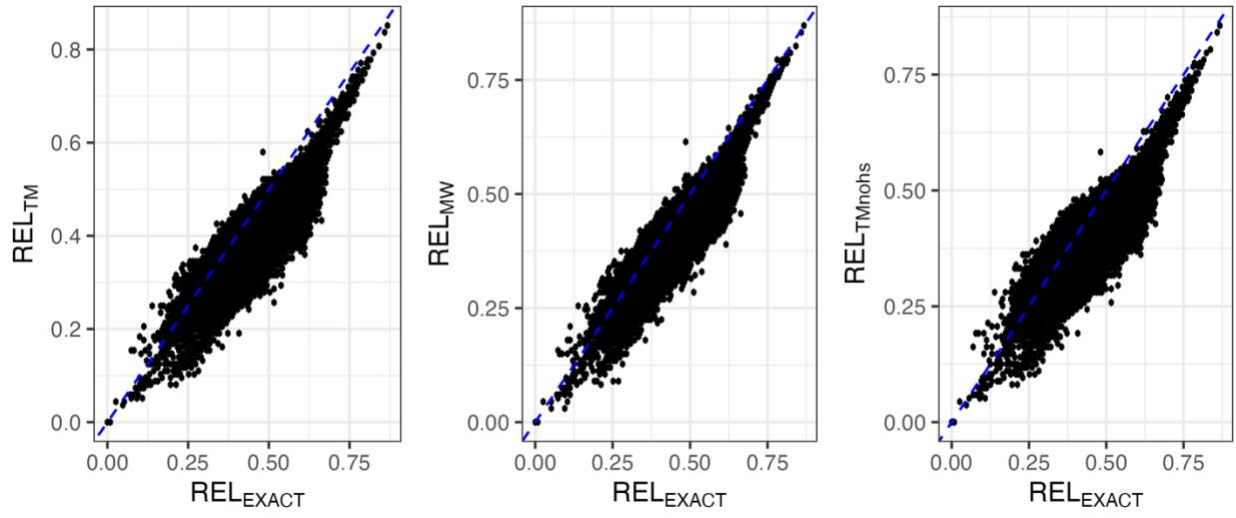
Abbreviations: **Odd:** Sires with many daughters concentrated in a small number of herds; **Normal:** Sires with daughters more evenly distributed across many herds; **REL_{EXACT} :** Exact reliabilities from the inverse of the coefficient matrix with blupf90+; **REL_{MW} :** Approximated reliabilities based on Misztal-Wiggans; **REL_{TM} :** Approximated reliabilities based on modified Tier-Meyer; **REL_{TMnohs} :** Approximated reliabilities based on standard Tier-Meyer

Figure 6.4. Reliabilities of Odd and Normal sires from DataReduced.



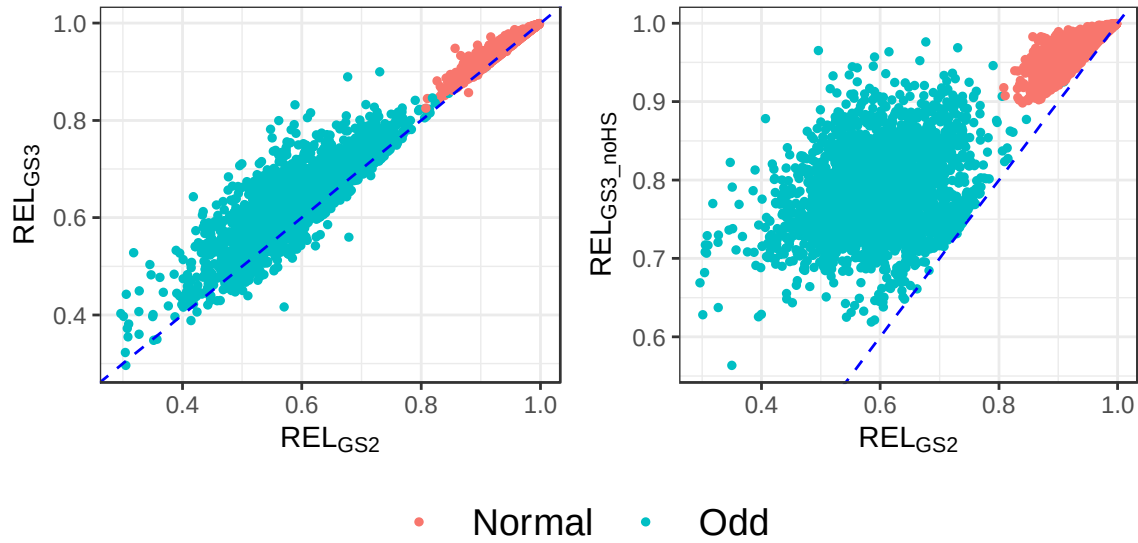
Abbreviations: **Odd:** Sires with many daughters concentrated in a small number of herds; **Normal:** Sires with daughters more evenly distributed across many herds; **REL_{EXACT} :** Exact reliabilities from the inverse of the coefficient matrix with blupf90+; **REL_{MW} :** Approximated reliabilities based on Misztal-Wiggans; **REL_{TM} :** Approximated reliabilities based on modified Tier-Meyer; **REL_{TMnohs} :** Approximated reliabilities based on standard Tier-Meyer

Figure 6.5. Reliabilities of cows with records from DataReduced.



Abbreviations: **REL_{EXACT}** : Exact reliabilities from the inverse of the coefficient matrix with blupf90+; **REL_{MW}** : Approximated reliabilities based on Misztal-Wiggans; **REL_{TM}** : Approximated reliabilities based on modified Tier-Meyer; **REL_{TMnohs}** : Approximated reliabilities based on standard Tier-Meyer

Figure 6.6. Reliabilities of Odd and Normal sires from DataFull.



Abbreviations: **Odd**: Sires with many daughters concentrated in a small number of herds; **Normal**: Sires with daughters more evenly distributed across many herds; **REL_{EXACT}**: Exact reliabilities from the inverse of the coefficient matrix with blupf90+; **REL_{MW}**: Approximated reliabilities based on Misztal-Wiggans; **REL_{TM}**: Approximated reliabilities based on modified Tier-Meyer; **REL_{TMnohs}**: Approximated reliabilities based on standard Tier-Meyer

CHAPTER 7

CONCLUSIONS

Continuous advancements in genomic evaluations and reliability approximations are essential as data availability increases, technology evolves, and breeding goals shift. The transition from Best Linear Unbiased Prediction (BLUP) to the incorporation of genomic information has significantly enhanced genetic gain by increasing accuracy, reducing bias, and lowering computational costs in genetic evaluations.

This dissertation introduced new modeling strategies that account for population differences, evaluated alternative prediction methods to improve the genetic evaluation process, and explored reliability approximation to reduce the computational costs in large-scale genomic evaluations. Additionally, it derived a modification of an existing reliability approximation method to mitigate its overestimation in young genomic bulls.

The findings demonstrated substantial improvements in genomic prediction accuracy, genetic gain, and better-quality reliability approximations. When combining two divergent populations in the same evaluations, results indicated that modeling genetic and non-genetic differences between the populations increased prediction accuracy and prevented the inflation of genomic estimated breeding values (GEBVs).

Applying single-step genomic BLUP (ssGBLUP) to fertility traits in a multi-breed dairy population proved effective in reducing GEBV inflation and capturing the response to selection in the population.

Various strategies can be implemented to decrease the computational cost of genetic evaluations. The study in Chapter 5 demonstrated that one such approach is the selective removal

of certain genotyped animals from the official evaluation. These animals can instead be evaluated using Single Nucleotide Polymorphism (SNP) effects derived from previous evaluations, with their reliabilities approximated based on SNP prediction error (co)variances. This strategy reduces the number of genotypes used in the evaluation and, consequently, lowers computational demands.

Finally, ensuring correct reliability approximations for bulls in a dairy population is critical for making sound selection decisions. The final study showed that introducing a weight into the reliability approximation method, one that considers all the effects included in the evaluation model, effectively reduced the overestimation of reliabilities, ultimately improving the quality of genetic evaluations.