

ORIGINAL ARTICLE OPEN ACCESS

Genomic Prediction Ability for Novel Profitability Traits Using Different Models in Nelore Cattle

Leticia Silva Pereira¹  | Cláudio Ulhôa Magnabosco² | Guilherme Rosa³ | Nedenia Bonvino Stafuzza⁴ | Tiago Zanett Albertini⁵ | Minos Carvalho⁵ | Raysildo Barbosa Lobo⁶ | Elisa Peripolli⁷ | Eduardo da Costa Eifert² | Fernando Baldi⁸

¹Department of Animal Science, Federal University of Goiás, Goiânia, GO, Brazil | ²Embrapa Cerrados, Planaltina, Brazil | ³Department of Animal Science, University of Wisconsin-Madison, Madison, Wisconsin, USA | ⁴Department of Animal Science, Animal Science Institute (IZ), São Paulo's Agency for Agribusiness Technology (APTA), Sertãozinho, São Paulo, Brazil | ⁵@Tech—Innovation Technology for Agriculture, Piracicaba, SP, Brazil | ⁶National Association of Breeders and Researchers (ANCP), Ribeirão Preto, SP, Brazil | ⁷Department of Animal Science, São Paulo State University—Júlio de Mesquita Filho (UNESP), Access Way Prof. Paulo Donato Castelane, Jaboticabal, SP, Brazil | ⁸Department of Animal Science, Faculty of Animal Science and Food Engineering, University of São Paulo, Pirassununga, São Paulo, Brazil

Correspondence: Leticia Silva Pereira (leticiavet@gmail.com)

Received: 25 April 2025 | **Revised:** 7 September 2025 | **Accepted:** 13 September 2025

Funding: This work was supported by the Fundação de Amparo à Pesquisa do Estado de Goiás (202110267000441).

Keywords: genomic selection | multi-trait model | phenotypes of profit | prediction accuracy

ABSTRACT

The aim of this study was to assess the accuracy, bias and dispersion of genomic predictions for accumulated profitability (APF) and profit per kilogram of liveweight gain (PFT) in Nelore cattle using different prediction approaches. The dataset consisted of 3969 phenotypic records for each trait. The pedigree harboured information from 38,930 animals born between 1998 and 2016, including 2691 sires and 19,884 dams. A total of 2449 animals were genotyped using the Clarifide Nelore 3.0 SNP panel. Nine models for genomic prediction were evaluated: a linear animal model was applied to estimate genetic parameters and perform the genomic single-trait best linear unbiased prediction (ST_{ss}—default). Additionally, a two-trait (ssGBLUP TT_{W450} and TT_{DMI}), three-trait (TTT_{CAR}) and multi-trait ssGBLUP (MT_{ss}) were tested. Finally, two models employing the weighted linear (ST_{sswl1} and ST_{sswl2}) and non-linear (ST_{sswnl1} and ST_{sswnl2}) single-step genomic approach (WssGBLUP) were used to predict genomic breeding values (GEBV). The ability to predict future performance was assessed by calculating the correlation between GEBV and adjusted phenotypes. The average prediction accuracy of the GEBV models ranged from 0.345 to 0.665 for PFT and from 0.425 to 0.603 for APF. The predictive capability of the MT_{ss} model (0.665) was significantly higher than that of the other models for PFT, except for the TTT_{CAR} model (0.604), which also showed an improvement in predictive performance. For APF, the MT_{ss} (0.561) and TT_{W450} (0.556) models demonstrated improved genomic prediction accuracy compared to the other models. In general, the single trait ssGBLUP (ST_{ss}—default) models and the non-linear weighting approach did not enhance prediction accuracy for either trait. For the phenotypic prediction ability of PFT, the linear WssGBLUP models ST_{sswl1} (0.65) and ST_{sswl2} (0.70), TT_{W450} (0.64) and ssGBLUP-M (0.66) demonstrated the highest prediction accuracies. Similar results were observed for the phenotypic prediction ability of APF for both models. However, the linear WssGBLUP models ST_{sswl1} (0.84) and ST_{sswl2} (0.94) provided higher prediction performance compared to the two-, three- and multi-trait models. The results indicate that the multi-trait model achieved better predictive ability for the novel traits PFT and APF. Multi-trait genomic selection may yield greater genetic gains than other models for these forthcoming economically important traits in breeding programmes.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Author(s). *Journal of Animal Breeding and Genetics* published by John Wiley & Sons Ltd.

1 | Introduction

Beef cattle breeding is an indispensable component in the sustainability of the livestock industry worldwide (Mueller and Van Eenennaam 2022). Considerable advances in the tools and methods used for genetic improvement have contributed to the increased rate of genetic change in recent years (Van Eenennaam 2017). Additive genetic components were separated from environmental components using more sophisticated approaches, and objective measures replaced subjective ones (Van Eenennaam et al. 2014). Among these advances, genomics has emerged as the strategy with the greatest potential to accelerate genetic progress and productivity in livestock production (Weigel et al. 2010). It should be noted that genomic selection provides new opportunities as a starting point for selecting emerging traits, most of which have reference populations with a limited number of animals, due to the costs and challenges of practical phenotyping measurements in large volumes (Calus, Berry, et al. 2013; Calus, De Haas, et al. 2013), especially for complex traits with low heritability (Alvarenga et al. 2020).

Even 10 years after the introduction of genomics, the development of novel resources and their use is still ongoing (VanRaden 2020). Despite this progress, it is becoming increasingly challenging to meet both the market and livestock producers' needs, as well as consumer expectations (Merks et al. 2012). Livestock farming faces new challenges regarding sustainability, considering its three main components: social, environmental and economic (Boichard and Brochard 2012). In the current scenario, known as the 'phenomics era', access to more robust phenotypes has become necessary to address these expectations (Boichard et al. 2016). Therefore, new traits should be included as targets in selection objectives (Merks et al. 2012) with the help of precision or large-scale tools (Seidel et al. 2020).

Given this, two proposed traits become strategic in addressing these issues, especially for the finishing phase, where feedlot operators face challenges such as management and input costs, carcass quality and market volatility. These traits are the accumulated profitability in the feedlot (AFP) and the profit per kilogram of live weight gain (PFT). Both traits are directly collected during the finishing phase in confined animals and are interpreted as indicators of an animal's genetic merit for profitability, based on performance and carcass traits, as well as economic data at both input and output stages of the finishing phase. In other words, these traits can be included to help in the early identification and selection of animals with high potential for economic and productive performance in the feedlot, while also contributing to financial and environmental sustainability.

Multiple methods have been suggested to estimate Genomic Estimated Breeding Values (GEBVs), such as Single-Step Genomic Best Linear Unbiased Prediction (ssGBLUP) (Aguilar et al. 2010; Christensen and Lund 2010). The ssGBLUP model incorporates phenotype, pedigree and genomic information for individuals without genotypes but with phenotypic information, or individuals with only genotypic information through a pedigree-based relationship matrix (A) with the genomic relationship matrix (G) in a hybrid matrix (H) (Legarra et al. 2009; Christensen and Lund 2010). However, this model assumes that

the effects of single nucleotide polymorphisms (SNPs) have equal variance, which in biological terms may not be the most appropriate assumption (Meuwissen et al. 2001; VanRaden 2008; Goddard and Hayes 2009). Consequently, the weighted single-step GBLUP (WssGBLUP) method proposed by Zhang et al. (2016), is an extension of ssGBLUP that assigns different weights to the SNPs used in calculating the G matrix.

Univariate analyses have been primarily used in genomic prediction models (Mehrban et al. 2021). However, this approach may not adequately capture the complex interactions between the analysed traits, as it does not account for the flow of information between them through the available information on genetic (co)variances (Gaire et al. 2022). (Co)variances arise from pleiotropy and linkage disequilibrium, resulting from the complex interactions of quantitative traits (Lynch and Walsh 1998). Therefore, the multi-trait model has been shown to be more consistent in integrating information and identifying the effects of associations between traits, resulting in more accurate GEBV predictions than univariate analysis models (Calus and Veerkamp 2011; Guo et al. 2014; Wang et al. 2017). It is assumed that the multi-trait approach could enhance the predictive capacity of GEBVs, especially for traits with low heritability or a limited number of records, through genetically correlated traits (Song et al. 2019).

Therefore, this study aimed to evaluate the predictive capacity of new traits related to profitability in feedlots in Nelore cattle. Nine models were evaluated, including best single-step unbiased linear genomic prediction (ssGBLUP) and weighted single-step linear and nonlinear genomic approaches (WssGBLUP) under one-, two-, three- and multiple-trait models. The analysis of these models focused on bias, dispersion and accuracy of genetic values to compare the profitability-related traits.

2 | Materials and Methods

This study was exempt from evaluation by the Animal Ethics Committee (CEUA), in accordance with Law No. 11,794 of 08/10/2008 and Normative Resolution No. 51 of 05/19/2021 from the National Council for Animal Experimentation Control (CONCEA) as all the analyses were conducted using pre-existing databases.

2.1 | Data Source

The dataset used in this study was provided by the National Association of Breeders and Researchers (ANCP, São Paulo, Brazil) in collaboration with @Tech (Innovation Technology for Agriculture LLC—Piracicaba, São Paulo, Brazil). The ANCP provided the pedigree information and genotypes, and novel feedlot profitability-related phenotypes were provided by @Tech. For additional information about the company, please consult the official website: <https://techagr.com/beeftreader>. The animals belonged to 26 different herds located in the Southeast, Northeast and Midwest regions of Brazil. The pedigree contained information from 38,930 animals, born between 1998 and 2016, comprising 2691 sires and 19,884 dams.

2.2 | Phenotyping

2.2.1 | Novel Traits

The new phenotypes analysed were accumulated feedlot profitability (AFP) and feedlot profit per kilogram of liveweight gain (PFT). AFP is the accumulated feedlot profitability in monetary units in the 80-day period of feedlot and the PFT is the profitability per 15 kg of weight gained in feedlot. @Tech's algorithms are designed to make full use of this data and can be collected for up to 150 days, longer than the standard 80-day data collection period in our study. The BeefTrader Decision Support System (Albertini et al. 2017) generates the profitability phenotypes used by the commercial tool Livestock Profit Tool (LPT) to identify the most profitable animals in the herd. The system uses animal growth modelling to define the best day for animals to leave the feedlot.

The BeefTrader algorithm uses animal traits as input variables (gender, breed, body condition score, age, initial weight, initial date, factors that impact on growth dynamics), daily weights individually collected through a weighing sensor (daily basis), and information on the nutritional composition of the diets (Albertini et al. 2017; Biase et al. 2022). The records of the dry matter intake for obtaining the new phenotypes were collected from animals participating in feed efficiency trials, following the same guidelines as Mendes et al. (2020), as mentioned in the section on feed efficiency traits. Based on this information, adjusted for local conditions, weight prediction was carried out in two steps: based on the biology of each animal and with the nutritional data and animal daily weight profile (observed or predicted), it is possible to estimate an optimal growth function for the animals (Step 1); from there, a dynamically adjusted linear or non-linear regression is performed using the least squares method on the weights to fit the predicted growth curve (Step 2).

From the predicted growth curve, it is possible to find other variables required by the model, including animal performance in terms of growth and composition of gain, as well as economic and environmental factors (Gionbelli et al. 2016; Albertini et al. 2017). The equations for calculating phenotypes are presented below:

Accumulated arroba:

$$(\text{sbw} \times \text{cdf} / 100) / 15 \quad (1)$$

where sbw is the shrunk body weight (kg)—96% of the body weight. cdf is the carcass dressing (%).

The carcass dressing used was 55.34% for females and 58.55% for entire males. These are real values from the company's customer database and are close to those found in the literature, which range from 54.9% to 60.6% (Arcanjo et al. 2024; Anaruma et al. 2020; Rodrigues et al. 2003; Moreira et al. 2003).

To assess the profit obtained by meat producers, it is common to use the unit of measurement 'arroba' in Brazil. In this study, the 'arroba' unit is defined as equivalent to 15 kg, following the standard practice in the national livestock industry. Therefore,

for the purposes of this study, the term 'arroba' in this equation will be used to represent the profit obtained per each 15 kg of meat produced.

Arroba × Gain:

$$\text{accumulated kilograms} = \text{accumulated arroba} [d] - \text{initial gain} [1] \quad (2)$$

where accumulate arroba = accumulated kilograms (each 15kg) on a specific day. [d] = final day of the period to be considered, in this case 80. [1] = first day of the feedlot.

Daily cost

$$\text{DMI} \times \text{diet price (kg)} + \text{feedlot daily overhead} \quad (3)$$

where DMI = dry matter intake (kg), diet_price_kg = diet cost (\$/kg), feedlot_daily_overhead = non-feed cost (\$).

Daily revenue:

$$\text{arroba price} \times \text{arroba gain} \quad (4)$$

where arroba_price = price of the arroba (\$/15 kg), arroba gain = (referrig to each 15 kg / day).

Daily profit:

$$\text{revenue daily} - \text{cost daily} \quad (5)$$

where revenue daily = daily revenue (\$), cost daily = daily cost (\$) and total revenue, cost and profit

$$\sum_{\text{time}} \text{daily revenue} \quad (6)$$

$$\sum_{\text{time}} \text{daily cost} \quad (7)$$

$$\sum_{\text{time}} \text{daily profit} \quad (8)$$

where daily revenue = see Equation (4). Daily cost = see Equation (3). Daily profit = see Equation (5).

Cost and Profit per kilograms:

$$\text{total cost} / \text{arroba gain}_t$$

$$\text{total profit} / \text{arroba gain}_t$$

where total profit = see Equation (8). Total cost = see Equation (7).

arroba gain_t = each 15 kg gain over time, see Equation (2).

2.2.2 | Standardisation of Costs and Arroba Pricing

2.2.2.1 | Food Cost. Even considering the effect of the batch (animals evaluated by farm) in the analyses, all common foods between batches, especially among farms, had

their prices standardised to set up the food cost (e.g., for corn silage, the price was always the same for the different lots, and so for all common foods in the diet). Based on the cost of natural matter (as feed) and the percentage of dry matter (DM), from the measurement of each animal's daily individual intake, the food cost for everyone was imputed over the 80-day evaluation period. It is important to note that after the adaptation period, there were 80 days of data collection on weight, DM intake (DMI), and food and non-food costs (operational cost), all individual, to obtain the measure of accumulated profit and profitability per arroba gained by the evaluated animal.

2.2.2.2 | Non-Food Cost (Operational Cost). The non-feed cost was also set at the same value for all evaluated batches with the aim of standardising this cost source in the process, and it is a source that doesn't affect the animals' performance.

2.2.2.3 | Price Paid per Kilograms (Arroba). The arroba price for all batches was standardised to the prices at the time of data collection, with the aim of ensuring that the revenue per arroba was equal for all animals. The prices followed those indicated by the Center for Advanced Studies in Applied Economics (CEPEA, <https://www.cepea.esalq.usp.br/br/indicador/boi-gordo.aspx>)—University of São Paulo (USP).

2.2.3 | Growth

A growth trait considered in this study was adjusted weight at 450 days of age (W450, kg). The standardised weight was calculated using linear regression, considering the average daily gain assessed between 405 and 495 days of age for the variable W450 (Negreiros et al. 2022).

2.2.4 | Carcass

The carcass traits considered in this study were ribeye area (REA, cm²) and rump fat thickness (RFT, mm). Carcass phenotypes were obtained through ultrasound images of the *Longissimus dorsi* muscle, taken between the 12th and 13th ribs (REA) and in the rump region, between the ilium and ischium at the intersection of the *Gluteus medius* and *Biceps femoris*

muscles (RFT). These measurements were conducted using the ALOKA 500 V equipment with a 3.5 MHz linear probe.

2.2.5 | Feed Efficiency

The feed efficiency trait considered in this study was DMI. This trait was measured using the Intergado and GrowSafe electronic systems. Feed efficiency tests followed the guidelines established by Mendes et al. (2020) for assessing individual feed intake in beef cattle using both electronic systems. Animals were housed in either collective or individual pens and underwent to a 21-day adaptation period, followed by a valid 70-day testing phase. During this period, each animal's average weight was recorded either manually every 14 days or through automated weighing platforms (Intergado). Daily dry matter intake (DMI, kg/day) was calculated as the mean of all valid individual daily intake values electronically recorded by the Intergado and GrowSafe systems during the test period.

2.3 | Statistical and Quality Control Analyses

The contemporary groups (CGs) were formed based on the effect of farm, year and season of birth (dry season: March to August, and rainy season: September to February), management group and sex. For the feed efficiency trait, the feed efficiency test identification was also considered in forming the CGs. Phenotypic quality control excluded records that deviated 3.5 standard deviations from the overall mean of the CGs and those with fewer than four records. Descriptive statistics for APF, PFT, carcass, growth and feed efficiency related traits used in the single, two-, three- and multi-trait analyses after quality control are summarised in Table 1.

2.4 | Genotyping

Genotypes were provided by the National Association of Breeders and Researchers (ANCP), Ribeirão Preto, Brazil. A total of 2449 animals were genotyped using the low-density panel (Clarifide Nelore 3.0). Genotype quality control (QC) was carried out by the PREGSF90 program (Aguilar et al. 2014), excluding both animals and SNPs from the dataset with call rates < 0.90. Additionally, SNPs with a minor allele

TABLE 1 | Descriptive statistics for profitability, growth, feed efficiency and carcass-related traits in Nelore cattle.

TRAIT	N	MEAN	SD	MIN	MAX	CV (%)	CG
APF (\$)	3969	157.71	67.18	-17.98	420.06	42.60	252
PFT (\$/kg)	3969	36.65	9.54	-13.33	51.23	26.04	252
W450 (kg)	55,052	289.96	63.23	119.00	592.00	21.81	2211
DMI (kg/day)	11,169	8.17	2.06	3.18	20.66	25.23	251
REA (cm ²)	37,091	57.55	12.85	20.45	114.97	22.33	1559
RFT (mm)	37,003	4.35	2.74	0.13	24.39	62.94	1559

Abbreviations: APF, accumulated profitability; BW450, weight at 450 days of age; CG, number contemporary group; CV, coefficient of variation; DMI, dry-matter intake; MAX, maximum; MIN, minimum; N, number animals with records; PFT, profit per kilogram of liveweight gain; REA, rib eye area; RFT, rump fat thickness; SD, standard deviation.

frequency (MAF) < 0.05, Mendelian conflicts > 1%, monomorphic SNPs with redundant positions, SNPs deviating from Hardy–Weinberg equilibrium expectations (0.15), and those located on non-autosomal chromosomes were also excluded. After QC, the dataset included 2449 genotyped animals and 35,658 SNPs for analysis.

2.5 | Genomic Prediction Models

Genomic prediction models for APF and PFT were performed using nine genomic models applying the ssGBLUP methodology, which included a single-trait model, two- and three-trait models and a multi-trait model. Additionally, in the case of the single-trait model, both the weighted linear and non-linear single-step genomic approach (WssGBLUP) were applied in the analyses. The analyses were performed using the BLUPF90 family (Miszta et al. 2002), and the general model can be defined as:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e} \quad (9)$$

where \mathbf{y} is a vector of phenotypic records; $\boldsymbol{\beta}$ is a vector of fixed effects; \mathbf{X} is a design matrix associating $\boldsymbol{\beta}$ with \mathbf{y} ; \mathbf{u} is a vector of random effects of the direct additive genetic effects; \mathbf{Z} is the incidence matrix associating \mathbf{u} with \mathbf{y} ; \mathbf{e} is the residual effect. Assumptions for residual effects are described below:

$$\mathbf{e} \sim N(\mathbf{0}, I\sigma_e^2),$$

where σ_e^2 is the residual variance, and I is an identity matrix with a dimension equal to the number of animals with records.

The description of the genomic prediction models is presented in Table 2.

2.5.1 | Single-Trait Prediction Model

The ssGBLUP single-trait model using information from both genotyped and non-genotyped phenotype information and using both marker and pedigree information for genetic evaluations was performed. The ssGBLUP is a modification of the BLUP model, in which the inverse of the numerator relationship matrix A^{-1} is replaced by H^{-1} (Aguilar et al. 2010), which is given by:

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

where \mathbf{H} is the relationship coefficient matrix between the animals; \mathbf{A} is the (numerator) additive relationship matrix; \mathbf{A}_{22}^{-1} is a partition of \mathbf{A} corresponding to the genotyped animals and \mathbf{G}^{-1} is the genomic relationship matrix described by VanRaden (2008) as $\mathbf{G} = \mathbf{Z}\mathbf{Z}'$, where:

$$\mathbf{Z} = (\mathbf{M} - \mathbf{P}) / \left[2 \sum_{j=1}^K p_j (1 - p_j) \right]^{1/2}$$

in which \mathbf{M} is the matrix of K SNP genotype for each animal, and \mathbf{P} is the matrix of frequency of the second allele p in the locus j (p_j) multiplied by two.

TABLE 2 | Genomic prediction models for feedlot profitability-related traits in Nelore cattle.

Model	Description
<i>Single-trait</i>	
ST_ss (default)	ssGBLUP based on genotypic records (default)
ST_ssw1	ssGBLUP weighting the diagonal of D matrix with the λ values obtained in the first iteration of the WssGWAS for nonlinear model
ST_ssw2	ssGBLUP weighting the diagonal of D matrix with the λ values obtained in the second iteration of the WssGWAS for linear model
ST_sswn1	ssGBLUP weighting the diagonal of D matrix with the λ values obtained in the first iteration of the WssGWAS for nonlinear model
ST_sswn2	ssGBLUP weighting the diagonal of D matrix with the λ values obtained in the 2st iteration of the WssGWAS for nonlinear model
<i>Two-trait</i>	
TTT_CAR	
TT—W450	ssGBLUP based on genotypic records + BW450 records
TT_DMI	ssGBLUP based on genotypic records + DMI records
<i>Three-trait</i>	
TTT_CAR	ssGBLUP based on genotypic records + carcass records
<i>Multi-trait</i>	
MT_ss	ssGBLUP based on genotypic records + W450, DMI, REA and RFT records

2.5.2 | Two-, Three- and Multi-Trait Prediction Model

In the analysis of two-, three- and multi-trait models, carcass (REA and RFT), growth (W450) and feed efficiency (DMI) traits were identified as those genetically correlated with APF and PFT in the feedlot and were used as predictors in the genomic prediction models. The covariances and genetic correlations are presented in Table S1. Further details about the datasets used in this study are described in Pereira et al. (2025). Regarding the construction of the models, for the two-trait analysis, the traits W450 and DMI were considered alongside APF and PFT in the feedlot, resulting in the following combinations: APF-W450, APF-DMI, PFT-W450 and PFT-DMI. The three-trait analysis combined carcass (CAR) traits (CAR: REA and RFT) with APF and PFT, resulting in APF-CAR and PFT-CAR combinations. In the multi-trait analysis, the traits W450, DMI, REA and RFT were considered in the APF and PFT models.

The two-, three- and multi-trait models were applied to estimate genomic breeding values for traits through the realised matrix (\mathbf{H}) as follows (adapted from Guo et al. 2014):

$$\begin{matrix} y_1 \\ \vdots \\ y_n \end{matrix} = \begin{bmatrix} I_1 & 0 \\ \vdots & \vdots \\ 0 & I_N \end{bmatrix} \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_n \end{bmatrix} + \begin{bmatrix} Z_1 & 0 \\ \vdots & \vdots \\ 0 & Z_n \end{bmatrix} \begin{bmatrix} g_1 \\ \vdots \\ g_n \end{bmatrix} + \begin{bmatrix} e_1 \\ \vdots \\ e_n \end{bmatrix}$$

where y is the vector that includes each of the n type traits. In these two-, three and multi-trait model, it was assumed that the genomic effects $\sim N(\mathbf{0}, G \otimes \mathbf{H})$ and the residuals $(e) \sim N(\mathbf{0}, R \otimes \mathbf{I})$, where \otimes is the Kronecker direct product; \mathbf{G} and \mathbf{R} are the genetic and residual covariance matrices, respectively; and \mathbf{I} is an identity matrix. The assumed co-variance structure is:

$$\text{Var} \begin{bmatrix} g \\ e \end{bmatrix} = \begin{bmatrix} G \otimes H & 0 \\ 0 & I \sigma_e^2 \end{bmatrix}$$

2.5.3 | Single-Trait WssGBLUP Linear and Nonlinear

Before performing the weighted genomic prediction, a single-step weighted GWAS (WssGWAS) was conducted to identify SNPs and their respective weights. For the ssGWAS, a single-trait animal model was applied. The effects and variances of the SNPs were estimated following the methodology proposed by Wang et al. (2012). In this methodology, SNP effects are derived from the genomic values estimated by the ssGBLUP model, with SNP weights being iteratively updated. The iterative process increases the weights of SNPs with large effects and decreases those with smaller effects, effectively regressing them towards the mean. Equation (9) was employed to construct the WssGBLUP model. For the derivation of SNPs effects and weights, the animal effect was decomposed into genotyped animals (a_g) and not genotyped (a_n), as described by Wang et al. (2012). The animal effect of the genotyped animals is a function of the SNP effects (Wang et al. 2012):

$$a_g = Z_g u$$

where Z_g represents the relationship matrix of the genotypes of each locus, and u is a vector of the SNPs effects. The variance of animal effects was assumed as:

$$\text{var}(a_g) = \text{var}(Z_g u) = Z_g D Z_g' \sigma_u^2 = G^* \sigma_a^2$$

where D is a diagonal matrix of weights for variances of SNP variances ($D=I$ for GBLUP), σ_u^2 is the variance of the additive genetic effect obtained from each SNP when the same variance is assumed for all SNPs, σ_a^2 is the additive genetic variance, and G^* is the weighted genomic relationship matrix.

The ratio of covariance of additive genetic (a_g) and SNPs (u) effects is:

$$\text{var}_u^{a_g} = \begin{bmatrix} Z D Z' & Z D' \\ D Z' & D \end{bmatrix} \sigma_u^2$$

Sequentially:

$$G^* = \frac{\text{var}(a_g)}{\sigma_a^2} = \frac{\text{var}(Zu)}{\sigma_a^2} = Z D Z' \lambda$$

where λ is a normalising constant described by VanRaden et al. (2009) as:

$$\lambda = \frac{\sigma_u^2}{\sigma_a^2} = \frac{1}{\sum_{i=1}^m 2p_i(1-p_i)}$$

where m is the number of SNPs and p_i is the frequency of the second allele in the i th SNP. The SNP effects can be described by Wang et al. (2012):

$$\hat{u} = \lambda D Z' G^{*-1} \hat{a}_g = D Z' [Z D Z']^{-1} \hat{a}_g$$

The estimated SNP effects can be used to calculate the variance of each individual SNP (Zhang et al. 2010), which can be used as different weighting for each SNP:

$$\sigma_{u,i}^2 = \hat{U}_i^2 2p_i(1-p_i)$$

where $\sigma_{u,i}^2$ is the j SNP weight (equivalent to j SNP variance); \hat{u} is a vector of estimated j SNP effect; and p is the allele frequency of j SNP.

Two strategies were used to weight SNPs and perform genomic prediction: linear and non-linear methodology. Weighted ssGWAS is an iterative process with several steps, considering t as the iteration number; the steps are (Wang et al. 2012):

1. Let $D=I$ in the first step.
2. Calculate $G = Z_g D Z_g' \lambda$.
3. Calculate GEBVs for the entire data set using the ssGBLUP.
4. Convert GEBVs to SNP effects (\hat{u}): $\hat{u} = \lambda D Z' (Z_g D Z_g' \lambda)^{-1} \hat{a}_g$, where \hat{a}_g is the GEBVs of animals which were also genotyped.
5. Calculate weight for each SNP for linear model (Zhang et al. 2010): $d_i = \hat{a}_i^2 2p_i(1-p_i)$ (default)

For the non-linear weighted model, using similar approach to VanRaden (2008), SNP-specific weights were calculated as:

$$d_i = 1.125 \frac{|\hat{a}_i|}{\sigma(\hat{a}_i)} - 2$$

6. Normalise SNP weights to maintain the total genetic variance constant.
7. Back to step 2.

The analysis involved two iterations of the WssGBLUP model. The first iteration used an identity matrix to weight SNP effects, and the second used the D matrix derived from the SNP solutions estimated in the first step. This second iteration led to higher GEBV accuracies in the preliminary analysis. The SNP solutions were estimated using the POSTGSF90 software (Aguilar et al. 2014), and the genomic association analyses were performed using the BLUPF90 software family (Misztal et al. 2002), incorporating the genomic information as outlined by Aguilar et al. (2010). The results were presented as the proportion of the additive genetic variance explained by windows of 10 adjacent SNPs, as below:

$$\frac{\text{Var}(a_i)}{\sigma_a^2} \times 100 = \frac{\text{Var}\left(\sum_{j=1}^{10} Z_g c\right)}{\sigma_a^2} \times 100$$

where a_i is the genetic value of the i th region that consists of continuous 10 SNPs; σ_a^2 is the additive genetic variance; and c is the marker effect of the i th SNP within the i th region.

Manhattan plots of the first and second iterations of linear and nonlinear WssGWAS, along with the respective percentages of additive genetic variance explained by the 10-SNP windows for APF and PFT are presented in [Supporting Information](#).

2.6 | Prediction Accuracy, Bias and Dispersion

To conduct the prediction analyses, the dataset was divided into whole (w , training) and partial (p , validation) subsets. The validation subset was created by removing the phenotypic records of the target validation animals from the dataset, ensuring that evaluations were initially performed with the partial dataset. In this process, the partial GEBV (GEBV p) is calculated based on relatives and genomic information. Subsequently, the phenotypes of the validation animals are included, and the whole genomic breeding value (GEBV w) is achieved. Animals born in 2022 and 2023 from six herds were selected as the target validation group, representing young selection candidates. This validation methodology simulates the practical scenario in genomic selection, where the prediction of GEBV for younger animals (i.e., selection candidates) is based on phenotypic and genotypic information from proven or older animals within the total population. Table 3 presents the descriptive statistics of the traits for both the whole and partial datasets.

Prediction accuracy, bias and dispersion were calculated according to the methodology proposed by Legarra and Reverter (2018). Prediction accuracy (acc_p^2) represents the squared population accuracy of the GEBV estimated from a partial dataset. This method uses the covariance between GEBVs obtained from partial and whole datasets ($\rho_{Covw,p}^2$) which is modelled as a function of the reliability (squared accuracy) of the partial-data GEBV, as shown below::

$$acc_p^2 = \rho_{Covw,p}^2 = \sqrt{\frac{\text{cov}(\hat{u}_w, \hat{u}_p)}{(1 - \bar{F})\sigma_a^2}}$$

where \hat{u}_w is the GEBV estimated using the whole dataset with genomic and phenotypic information; \hat{u}_p is the (G)EBV considering the partial dataset; \bar{F} is the average inbreeding coefficient estimated for the animals included in the validation dataset; and σ_a^2 is the additive genetic variance for APF and PFT.

The mean of the estimated breeding values was used to evaluate bias, which has an expected value of 0 under an unbiased evaluation:

$$\mu_{wp} = \bar{\hat{u}}_p - \bar{\hat{u}}_w$$

The dispersion was calculated by regressing the EBVs from whole dataset on EBVs from the partial dataset. The $b_{w,p}$ has an expected value $E(b_{w,p}) = 1$, assuming there is no over/under dispersion:

$$b_{w,p} = \frac{\text{cov}(\hat{u}_w, \hat{u}_p)}{\text{var}(\hat{u}_p)}$$

To assess the predictive capacity of the phenotypes across the different models, Pearson's correlation between the GEBV and the adjusted phenotype for the fixed effects (Y_c) was used (Legarra et al. 2008). Phenotypic adjustment was carried out using the PREDICTF90 software from the BLUPF90 suite (Misztal et al. 2002). The correlation between the GEBV and the adjusted phenotype (y^* , defined as the phenotype y corrected for fixed effects) was calculated for individuals in the validation population. This correlation was subsequently divided by the square root of the heritability (h^2) (h^2):

$$ACC_p = \frac{r(\hat{u}_p, y^*)}{\sqrt{h^2}}$$

3 | Results

The prediction accuracy, bias and dispersion of genomic prediction for PFT and APF using different prediction models are described in Table 4. The results revealed that the prediction accuracy of the ST_ssw11 (0.587) and ST_ssw12 (0.528) models was higher than that of the ST_ss model (0.345) for PFT. The WssGBLUP model provided an expressive increase of 59% in prediction accuracy, as shown in Table 4. A similar trend was observed for APF, whose prediction accuracy of the ST_ssw11 (0.575) and ST_ssw12 (0.603) models was higher than the single-trait model ssGBLUP (0.425). On the other hand, in terms of prediction accuracy, the nonlinear models ST_ssw11 (0.365 and 0.443, respectively) and ST_ssw12 (0.365 and

TABLE 3 | Number of animals with phenotypic records in the whole and partial subsets for profitability-related traits in Nelore cattle.

Dataset	Trait	N	Mean	SD	Min	Max	CV (%)	CG
Whole subset	APF	3478	160.26	62.54	-17.98	420.06	39.02	242
	PFT	3478	37.62	8.83	-13.34	51.23	23.47	242
Partial subset	APF	502	139.64	91.75	-6.59	374.61	65.70	28
	PFT	502	29.76	11.39	-3.49	48.57	38.28	28

Abbreviations: APF, accumulated profitability; CG, number contemporary group; CV, coefficient of variation; MAX, maximum; MIN, minimum; N, number animals with records; PFT, Profit per kilogram of liveweight gain; SD, standard deviation.

TABLE 4 | Accuracy, bias and dispersion of genomic predictions for feedlot accumulated profitability and profit per kilogram of liveweight gain in Nelore cattle across different models.

Trait	Model	Accuracy	Bias	Dispersion	
Profit per kilogram of liveweight gain	<i>Single-trait</i>				
		<i>ST_ss</i>	0.345 (0.000)	-0.771 (0.250)	1.124 (0.043)
		<i>ST_ssw1</i>	0.587 (0.000)	-0.792 (0.242)	0.713 (0.013)
		<i>ST_ssw2</i>	0.528 (0.000)	-0.746 (0.234)	0.470 (0.014)
		<i>ST_sswn1</i>	0.365 (0.000)	-0.770 (0.249)	1.111 (0.038)
		<i>ST_sswn2</i>	0.365 (0.000)	-0.770 (0.249)	1.120 (0.038)
	<i>Two-trait</i>				
	<i>TTT_CAR</i>				
		<i>TT_W450</i>	0.523 (0.000)	-1.031 (0.349)	1.063 (0.019)
		<i>TT_DMI</i>	0.372 (0.000)	-1.075 (0.346)	1.122 (0.034)
	<i>Three-trait</i>				
		<i>TTT_CAR</i>	0.604 (0.005)	-1.230 (0.345)	0.969 (0.019)
	<i>Multi-trait</i>				
		<i>MT_ss</i>	0.665 (0.00)	-1.111 (0.344)	0.982 (0.016)
Accumulated profitability	<i>Single-trait</i>				
		<i>ST_ss</i>	0.425 (0.000)	0.002 (0.018)	1.104 (0.045)
		<i>ST_ssw1</i>	0.575 (0.000)	0.049 (0.017)	0.739 (0.023)
		<i>ST_ssw2</i>	0.603 (0.000)	0.033 (0.024)	0.556 (0.022)
		<i>ST_sswn1</i>	0.443 (0.000)	0.001 (0.017)	1.084 (0.040)
		<i>ST_sswn2</i>	0.442 (0.000)	0.001 (0.017)	1.093 (0.040)
	<i>Two-trait</i>				
	<i>TTT_CAR</i>				
		<i>TT_W450</i>	0.556 (0.000)	0.002 (0.015)	1.087 (0.028)
		<i>TT_DMI</i>	0.503 (0.000)	-0.003 (0.016)	1.065 (0.034)
	<i>Three-trait</i>				
		<i>TTT_CAR</i>	0.505 (0.000)	0.096 (0.015)	1.068 (0.032)
	<i>Multi-trait</i>				
		<i>MT_ss</i>	0.5612 (0.000)	0.004 (0.015)	1.049 (0.027)

Abbreviations: λ = proportion of additive genetic variance explained by each SNP. TTT_CAR. MT_ss = ssGBLUP with genotypic, BW450, DMI, REA and RFT records; ST_ss = ssGBLUP (default, based on genotypic records); ST_ssw1 = ssGBLUP weighting the diagonal of the D matrix with λ values from the first iteration of WssGWAS (nonlinear); ST_ssw2 = ssGBLUP with λ from the second iteration (nonlinear); TT_DMI = ssGBLUP with genotypic + DMI records; TT_W450 = ssGBLUP with genotypic + BW450 records; TTT_CAR = ssGBLUP with genotypic + carcass records.

0.442, respectively) showed no improvements compared to the ssGBLUP model for both traits. When incorporating information from correlated traits in the two-, three- and multi-trait models, such as W450 and carcass with PFT, improvements in the accuracy of the GEBVs were observed, reaching 0.604 (TTT_CAR) and 0.523 (TT_W450). Similarly, gains were observed in the predictive capacity for APF with W450, DMI and carcass, with accuracy values of 0.503 (TT_DMI), 0.505 (TTT_CAR) and 0.556 (TT_W450). As shown in the genetic correlations in Table S1, the genomic prediction for PFT indicated

that the two-trait model (TT_DMI) did not significantly improve prediction accuracy, possibly due to low genetic correlations. This may be due to the low genetic association between these traits and the limited amount of DMI data, resulting in lower accuracy of 0.372.

The dispersion values varied between PFT and APF within the models, as shown in Table 4. These values suggest variation in predictive ability, with both inflation and deflation of the estimated GEBVs. Values above one indicate underdispersion

(deflation), while values below one indicate overdispersion (inflation) of the estimated genetic values Himmelbauer et al. (2023). For both traits, the ST_{ss}, ST_{sswn1}, ST_{sswn2}, TTT_CAR, TT_W450 and TT_DMI provided relatively larger slopes for predictions, with values ranging from 1.063 to 1.124 for PFT and from 1.065 to 1.104 for AFP, respectively. The ST_{sswn1} and ST_{sswn2} obtained predictions with the regression coefficients furthest from one, with values of 0.713 and 0.470 for PFT and 0.556 and 0.739 for APFT, respectively. The MT_{ss} models were generally less dispersed (values closer to 1), with values of 0.982 (PFT) and 1.049 (APF).

The prediction accuracies for the fraction of additive genetic variance explaining the phenotype were obtained by dividing the correlation between the partial GEBV and the phenotype adjusted for fixed effects by the square root of the heritability. The ST_{sswn2} presented the highest precision values of the fraction of additive genetic variance that explains the adjusted phenotype for PFT, with a value of 0.70, followed by 0.66 (MT_{ss}), 0.65 (ST_{sswn1}) and 0.64 (TT_W450). Slightly lower predictive ability was observed for the ST_{sswn1} (0.58) and ST_{sswn2} (0.55) models, while the lowest predictive ability values were observed for the TTT_CAR (0.41), TT_DMI (0.38) and ST_{ss} (0.33) models. For APF, weighted linear models followed the same trend, with the ST_{sswn1} (0.84) and ST_{sswn2} (0.94) models showing higher predictive accuracies compared to TT_W450 (0.57), MT_{ss} (0.56), ST_{sswn1} (0.54), ST_{sswn2} (0.54), TT_DMI (0.52), TTT_CAR (0.46) and ST_{ss} (0.46) models. For this dataset, the weighted linear model WssGBLUP was more efficient in predicting the fraction of additive genetic variance responsible for explaining phenotype expression.

4 | Discussion

The prediction values and accuracy obtained under the different models support the idea that a model's performance depends on its ability to capture and consider the genetic structure of the traits under study, including the effects of polymorphisms, regardless of their magnitude, whether those with large or small effect (Terakado et al. 2021). Based on the observed results, it can be inferred that the relevant increase in the prediction accuracy of the linear models ST_{sswn1} and ST_{sswn2} compared to the nonlinear model can be attributed to the polygenic nature of the traits under study, in which the additive genetic variation for traits related to feedlot profitability in Nelore cattle is largely explained by numerous small-effect SNPs.

The differences observed in the predictive capacity of GEBVs from the two- and three-trait models can be explained by the traits with the strongest genetic correlation with PFT, such as W450 and carcass-related traits, as well as APF with DMI, W450 and carcass-related traits. Additionally, the number of records contributed by each trait in the models may also have influenced these differences.

Despite the moderate genetic correlation between PFT and carcass-related traits, as well as the lower number of records compared to W450, significant improvements in predictive capacity were observed when carcass traits (TTT_CAR) were

included in the model. Regarding APF, despite the high genetic correlation with W450 (Table S1), the TTT_CAR and TT_DMI models also contributed to its prediction accuracy. The highest prediction accuracy for both traits was obtained in the MT_{ss} model, demonstrating superior predictive ability over the single-trait linear model WssGBLUP and the two-trait model ssGBLUP.

The results presented herein align with those reported by Guo et al. (2014) and Song et al. (2019), in which the two-, three- and multi-trait genomic models proved more effective for traits with a smaller set of phenotypic data. This approach is suitable for genetic improvement programs, as in a reference population, most individuals will often not have phenotypes available for all traits of interest (Guo et al. 2014). Moreover, it is particularly relevant for research on novel traits to be incorporated as selection criteria, considering their costs and the difficulty of measurement due to the varying environments and regions in which different production systems operate. According to Guo et al. (2014), incorporating records of genetically correlated and easily measured traits, as well as multi-trait models, contributes to improvements in the prediction ability of GEBVs. However, the differences between the multi-trait and single-trait models are not likely to significantly differ in terms of prediction accuracy for traits with complete records and high heritability.

Regarding bias, it was observed to be low and followed the same trends for all prediction models of the PFT, except for APF, whose values were close to zero for all models, ranging from -0.003 to 0.096. Different variables may limit the response of prediction accuracy and reduce the bias of genomic predictions, as reported by Lund et al. (2009), Zhang et al. (2011) and Goddard (2009), who noted that the genetic architecture of the traits of interest can be characterised by heritability and the number of associated QTLs. This finding is consistent with the biases observed in our results, which may be attributed to the low genetic variance and heritability of the PFT trait. TTT_CAR. This result suggests a slight tendency towards inflation and more significant variability in predictions.

However, it is important to note that, although these models provided greater predictive ability compared to the others, the dispersion values exhibited greater underdispersion, suggesting probable deflation in the estimated genetic values. When predictions are deflated and greater than one, in practical terms, the difference between the progeny of selected sires is expected to be greater than what is predicted by the GEBV. Conversely, the opposite occurs when the predictor is inflated (Chiaia et al. 2017). Although the predictive capacity values are similar between the multi-trait and two- and three-trait models, the regression coefficient values close to one for the model suggest the better consistency of this model in predicting GEBVs. Therefore, the two-trait model was more effective in predicting the absolute differences between individuals under evaluation.

These results also indicate that incorporating information from weighted regions of GWAS into genomic prediction models through the weighted linear WssGBLUP approach can be an effective strategy to improve the accuracy of predicting the fraction of additive genetic variance for the traits under study. Van den Berg et al. (2016), Raymond et al. (2018) and Peripolli et al. (2024) reported that sequence variants selected from a

GWAS could be used to improve predictive capacity, as well as to predict the fraction of additive genetic variance responsible for explaining the phenotype.

The genetic architecture of a complex trait exerts a considerable influence on genomic prediction (Hayes et al. 2010). Potential genomic regions harbouring genes with informative SNPs for APF and PFT were identified, as illustrated in the Manhattan plots in the supplemental material (Figures S1–S8), with potential candidate genes involved in biological processes such as carbohydrate, protein, and lipid metabolism, immunity and feeding behaviour. Additionally, these genomic regions may be associated with traits related to growth, reproduction, feed efficiency and carcass.

These results suggest that the analyses of two-, three- and multi-trait and nonlinear weighted models did not contribute to the predictive accuracies of the adjusted phenotypes. On the other hand, weighted regions, where weighted linear models used their own QTL information, contributed to a better predictive capacity of the fraction of additive genetic variance explaining the adjusted phenotypes for PFT and APF. This indicates that these models were more efficient in capturing the additive genetic variation that effectively contributes to the phenotypic expression of each trait.

Regarding the models evaluated, although multitrait models displayed better predictive capacity, their use requires the estimation of a large number of (co)variance components. This implies higher computational demands, such as processing time and memory usage, due to the need to estimate all variance and covariance matrices among all the traits under study. According to Meyer (2007), multivariate analyses have been limited by high computational requirements coupled with the difficulty in accurately estimating a large number of covariance components simultaneously. Alternatively, principal component analysis (PCA) has been proposed as a dimensionality reduction strategy, aiming to simplify model complexity (Meyer 2007). It is considered an alternative to reduce both the number of estimated EPDs and the number of parameters required to model the genetic covariance matrix (Kirkpatrick and Meyer 2004).

Similarly, models that use genomic weighting (WssGBLUP) may be related to hyperparameterisation, as these models assume that certain genomic regions contribute unequally to the total genetic variance, which may have contributed to greater bias and dispersion. Furthermore, the weighted model requires steps prior to estimating genomic predictions, such as analysing the weighting of genomic regions, since the weighted genomic matrix increases the complexity of the process and, consequently, requires higher computational demand.

From a practical standpoint, these factors may impair the application of these multi-trait models in large-scale commercial selection programmes, especially for those with more limited computational infrastructure or less availability of phenotypic data and genetically correlated traits. Despite the limitations associated with multi-trait models, it is essential to consider that incorporating information from genetically correlated traits can significantly enhance the predictive accuracy of genomic prediction for novel phenotypes with smaller databases

or traits that are difficult to measure. Therefore, genetic improvement programmes should carefully consider the balance between gains in predictive capacity and the technical and operational feasibility of implementing these models in commercial settings.

5 | Conclusions

The results obtained herein showed that multi-trait models improve the genomic prediction capacity for profitability-related traits in feedlots. Genomic prediction using multi-trait models is beneficial when the trait is complex and the number of phenotypes (i.e., novel traits) is limited. This strategy is more accurate for improving genomic prediction estimates of these traits and integrating them efficiently into genetic improvement programmes. These results provide additional support to breeders in improving management and selection decisions to enhance feedlot profitability operations. The incorporation of weighted regions for GEBV prediction in the single-trait weighted linear model helps to improve predictive capacity; however, more biased predictions and dispersion values with more significant inflation were observed.

From a practical point of view, for commercial herds, feedlots, and calf comparators for rearing and finishing, the weighted single-trait models (ST_sswnl2) were superior in predicting the phenotype. The results obtained for predicting future performance (phenotype) can help producers make earlier and more assertive management decisions, improving production efficiency and economic return per animal in their operations, by prioritising animals with higher genetic merit for profitability-related traits in feedlots.

Author Contributions

Leticia Silva Pereira: conceptualisation, methodology, formal analysis, writing-original draft preparation. Cláudio Uihôa Magnabosco: conceptualisation, methodology, formal analysis, writing and review. Guilherme Rosa: methodology, formal analysis, writing and review. Nedenia Bonvino Stafuzza: conceptualisation, writing and review. Tiago Zanett Albertini: data curation, writing-review and editing. Minos Carvalho: data curation, writing-review and editing. Raysildo Barbosa Lobo: data curation, review. Elisa Peripolli: conceptualisation. Eduardo da Costa Eifert: conceptualisation, writing and review. Fernando Baldi: conceptualisation, methodology, formal analysis, writing-original draft preparation.

Acknowledgements

The authors thank the National Association of Breeders and Researchers (ANCP, Ribeirão Preto, Brazil) and @Tech Company (Piracicaba, Brazil) for the support and dataset available for this research. We acknowledge the Foundation for the Support of Research of the State of Goiás (FAPEG) for the financial support granted to LS Pereira through the scholarship (grant #202110267000441), in collaboration with the Graduate Program in Animal Science at the Federal University of Goiás (UFG). This study is part of the first author's doctoral research. We acknowledge the São Paulo Research Foundation (FAPESP) for the phenotypes generated in the technological projects of the innovation (2022/04423-2). The Article Processing Charge for the publication of this research was funded by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) (ROR identifier: 00x0ma614).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

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." *Journal of Dairy Science* 93, no. 2: 743–752. <https://doi.org/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." In *10. World Congress on Genetics Applied to Livestock Production (WCGALP)*. American Society of Animal Science. <https://hal.inrae.fr/hal-02743809>.
- Albertini, T. Z., A. G. Biase, M. Barbosa, et al. 2017. "BeefTrader: Optimal Economical Endpoint Decision Support System for Feedlots and Meat Packers." *Journal of Animal Science* 95, no. suppl_4: 307. <https://doi.org/10.2527/asasann.2017.627>.
- Alvarenga, A. B., R. Veroneze, H. R. Oliveira, et al. 2020. "Comparing Alternative Single-Step GBLUP Approaches and Training Population Designs for Genomic Evaluation of Crossbred Animals." *Frontiers in Genetics* 11: 507742. <https://doi.org/10.3389/fgene.2020.00263>.
- Anaruma, R. J., L. G. Reis, P. E. de Felicio, et al. 2020. "Castration Age and Growth, Meat Production and Meat Quality of Nellore Male Cattle." *Animal Production Science* 60: 725–731. <https://doi.org/10.1071/AN18460>.
- Aranjo, A. H. M., L. C. V. Ítavo, C. C. Brandão Ferreira Ítavo, et al. 2024. "Carcass and Meat Characteristics of Nellore Young Bulls Fed Diet Using Cottonseed Cake as a Replacer of the Forage Fiber Source." *Scientific Reports* 14: 8027. <https://doi.org/10.1038/s41598-024-58738-9>.
- Biase, A. G., T. Z. Albertini, and R. F. de Mello. 2022. "On Supervised Learning to Model and Predict Cattle Weight in Precision Livestock Breeding." *Computers and Electronics in Agriculture* 195: 106706. <https://doi.org/10.1016/j.compag.2022.106706>.
- Boichard, D., and M. Brochard. 2012. "New Phenotypes for New Breeding Goals in Dairy Cattle." *Animal* 6, no. 4: 544–550. <https://doi.org/10.1017/S1751731112000018>.
- Boichard, D., V. Ducrocq, P. Croiseau, and S. Fritz. 2016. "Genomic Selection in Domestic Animals: Principles, Applications, and Perspectives." *Comptes Rendus Biologies* 339, no. 7–8: 274–277. <https://doi.org/10.1016/j.crvi.2016.04.007>.
- Calus, M. P., and R. F. Veerkamp. 2011. "Accuracy of Multi-Trait Genomic Selection Using Different Methods." *Genetics Selection Evolution* 43: 1–14. <https://doi.org/10.1186/1297-9686-43-26>.
- Calus, M. P. L., D. P. Berry, G. Banos, Y. De Haas, and R. F. Veerkamp. 2013. "Genomic Selection: The Option for New Robustness Traits?" *Advances in Animal Biosciences* 4, no. 3: 618–625. <https://doi.org/10.1017/S2040470013000186>.
- Calus, M. P. L., Y. De Haas, M. Pszczola, and R. F. Veerkamp. 2013. "Predicted Accuracy of and Response to Genomic Selection for New Traits in Dairy Cattle." *Animal* 7, no. 2: 183–191. <https://doi.org/10.1017/S1751731112001450>.
- Chiaia, H. L. J., E. Peripoli, R. M. de Oliveira Silva, et al. 2017. "Genomic Prediction for Beef Fatty Acid Profile in Nellore Cattle." *Meat Science* 128: 60–67. <https://doi.org/10.1016/j.meatsci.2017.02.007>.
- Christensen, O. F., and M. S. Lund. 2010. "Genomic Prediction When Some Animals Are Not Genotyped." *Genetics Selection Evolution* 42: 1–8. <https://doi.org/10.1186/1297-9686-42-2>.
- Gaire, R., M. P. de Arruda, M. Mohammadi, G. Brown-Guedira, F. L. Kolb, and J. Rutkoski. 2022. "Multi-Trait Genomic Selection Can Increase Selection Accuracy for Deoxynivalenol Accumulation Resulting From Fusarium Head Blight in Wheat." *Plant Genome* 15, no. 1: e20188. <https://doi.org/10.1002/tpg2.20188>.
- Gionbelli, M. P., S. C. Valadares Filho, and E. Detmann. 2016. "Adjusting Cattle Body Weight to Physiological and Feeding Conditions." In *Nutrient Requirements of Zebu and Crossbred Cattle? BR-CORTE*, edited by S. C. Valadares Filho, L. F. C. Costa E Silva, M. P. Gionbelli, et al., vol. 1, 3rd ed., 1–14. Suprema.
- Goddard, M. 2009. "Genomic Selection: Prediction of Accuracy and Maximisation of Long-Term Response." *Genetica* 136: 245–257.
- Goddard, M. E., and B. J. Hayes. 2009. "Genomic Selection Based on Dense Genotypes Inferred From Sparse Genotypes." *Proceedings of the Association for the Advancement of Animal Breeding and Genetics* 18: 26–29.
- Guo, G., F. Zhao, Y. Wang, Y. Zhang, L. Du, and G. Su. 2014. "Comparison of Single-Trait and Multiple-Trait Genomic Prediction Models." *BMC Genetics* 15: 1–7. <https://doi.org/10.1186/1471-2156-15-30>.
- Hayes, B. J., J. Pryce, A. J. Chamberlain, P. J. Bowman, and M. E. Goddard. 2010. "Genetic Architecture of Complex Traits and Accuracy of Genomic Prediction: Coat Colour, Milk-Fat Percentage, and Type in Holstein Cattle as Contrasting Model Traits." *PLoS Genetics* 6, no. 9: e1001139. <https://doi.org/10.1371/journal.pgen.1001139>.
- 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, no. 12: 9026–9043. <https://doi.org/10.3168/jds.2023-23575>.
- Kirkpatrick, M., and K. Meyer. 2004. "Direct Estimation of Genetic Principal Components: Simplified Analysis of Complex Phenotypes." *Genetics* 168, no. 4: 2295–2306. <https://doi.org/10.1534/genetics.104.029181>.
- Legarra, A., I. Aguilar, and I. Misztal. 2009. "A Relationship Matrix Including Full Pedigree and Genomic Information." *Journal of Dairy Science* 92, no. 9: 4656–4663. <https://doi.org/10.3168/jds.2009-2061>.
- 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." *Genetics Selection Evolution* 50, no. 1: 53. <https://doi.org/10.1186/s12711-018-0426-6>.
- Legarra, A., C. Robert-Granié, E. Manfredi, and J.-M. Elsen. 2008. "Performance of Genomic Selection in Mice." *Genetics* 180, no. 1: 611–618. <https://doi.org/10.1534/genetics.108.088575>.
- Lund, M. S., G. Sahana, D. J. de Koning, G. Su, and Ö. Carlborg. 2009. "Comparison of Analyses of the QTLMAS XII Common Dataset. I: Genomic Selection." *BMC Proceedings* 3, no. S1: 1–8. <https://doi.org/10.1186/1753-6561-3-S1-S1>.
- Lynch, M., and B. Walsh. 1998. *Genetics and Analysis of Quantitative Traits*. Vol. 1, 535–557. Sinauer.
- Mehrban, H., D. Lee, and N. Ibáñez-Escriche. 2021. "Multi-Trait Single-Step GBLUP Improves Accuracy of Genomic Prediction for Carcass Traits Using Yearling Weight and Ultrasound Traits in Hanwoo." *Frontiers in Genetics* 12: 692356. <https://doi.org/10.3389/fgene.2021.692356>.

- Mendes, E. D. M., C. U. de Faria, R. D. Sainz, et al. 2020. *Procedimentos para mensuração de consumo individual de alimento em bovinos de corte*. Associação Nacional de Criadores e Pesquisadores.
- Merks, J. W. M., P. K. Mathur, and E. F. Knol. 2012. "New Phenotypes for New Breeding Goals in Pigs." *Animal* 6, no. 4: 535–543. <https://doi.org/10.1017/S1751731111002266>.
- Meuwissen, T. H. E., B. J. Hayes, and M. E. Goddard. 2001. "Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps." *Genetics* 157: 1819–1829. <https://doi.org/10.1093/genetics/157.4.1819>.
- Meyer, K. 2007. "Análises multivariadas de características de carcaça de gado Angus ajustadas a modelos analíticos de classificação reduzida e fator." *Journal of Animal Breeding and Genetics* 124, no. 2: 50–64. <https://doi.org/10.1111/j.1439-0388.2007.00637.x>.
- Misztal, I., S. Tsuruta, T. Strabel, B. Auvray, T. Druet, and D. H. Lee. 2002. "BLUPF90 and Related Programs (BGF90)." In *Proceedings of the 7th World Congress on Genetics Applied to Livestock Production*, 19–23. Communication no. 28–07.
- Moreira, F. B., N. E. Souza, M. Matsushita, I. N. Prado, and W. G. Nascimento. 2003. "Evaluation of Carcass Characteristics and Meat Chemical Composition of Bos Indicus and Bos Indicus × Bos Taurus Crossbred Steers Finished in Pasture Systems." *Brazilian Archives of Biology and Technology* 46, no. 4: 609–616. <https://doi.org/10.1590/S1516-89132003000400016>.
- Mueller, M. L., and A. L. Van Eenennaam. 2022. "Synergistic Power of Genomic Selection Assisted Reproductive Technologies, and Gene Editing to Drive Genetic Improvement of Cattle." *CABI Agriculture and Bioscience* 3, no. 1: 13. <https://doi.org/10.1186/s43170-022-00080-z>.
- Pereira, L. S., F. Baldi, G. Rosa, et al. 2025. "Genome-Wide Association Study of Novel Feedlot Profitability-Related Traits in Nelore Cattle." *Animal Genetics* 56, no. 3. <https://doi.org/10.1111/age.70018>.
- Peripolli, E., L. Bordin Temp, H. Fukumasu, et al. 2024. "Identification of Candidate Genes and Genomic Prediction for Early Heifer Pregnancy in Nelore Beef Cattle." <https://doi.org/10.2139/ssrn.4829576>.
- Raymond, B., A. C. Bouwman, C. Schrooten, J. Houwing-Duistermaat, and R. F. Veerkamp. 2018. "Utility of Whole-Genome Sequence Data for Across-Breed Genomic Prediction." *Genetics Selection Evolution* 50: 1–12. <https://doi.org/10.1186/s12711-018-0396-8>.
- Rodrigues, V. C., I. F. Andrade, R. T. Freitas, M. C. Bressan, and J. C. Teixeira. 2003. "Carcass and Slaughter Yield of Castrated and Entire Buffaloes and Cattle." *Revista Brasileira De Zootecnia* 32: 663–671. <https://doi.org/10.1590/S1516-35982003000300018>.
- Seidel, A., N. Krattenmacher, and G. Thaller. 2020. "Dealing With Complexity of New Phenotypes in Modern Dairy Cattle Breeding." *Animal Frontiers* 10, no. 2: 23–28. <https://doi.org/10.1093/af/vfaa005>.
- Song, H., J. Zhang, Q. Zhang, and X. Ding. 2019. "Using Different Single-Step Strategies to Improve the Efficiency of Genomic Prediction on Body Measurement Traits in Pig." *Frontiers in Genetics* 9: 730. <https://doi.org/10.3389/fgene.2018.00730>.
- Terakado, A. P. N., R. B. Costa, N. Irano, et al. 2021. "Comparison of Methods for Predicting Genomic Breeding Values for Growth Traits in Nelore Cattle." *Tropical Animal Health and Production* 53, no. 3: 349. <https://doi.org/10.1007/s11250-021-02785-1>.
- Van den Berg, I., D. Boichard, and M. S. Lund. 2016. "Sequence Variants Selected From a Multi-Breed GWAS Can Improve the Reliability of Genomic Predictions in Dairy Cattle." *Genetics Selection Evolution* 48: 1–18. <https://doi.org/10.1186/s12711-016-0259-0>.
- Van Eenennaam, A. L. 2017. "Genetic Modification of Food Animals." *Current Opinion in Biotechnology* 44: 27–34. <https://doi.org/10.1016/j.copbio.2016.10.007>.
- Van Eenennaam, A. L., K. A. Weigel, A. E. Young, M. A. Cleveland, and J. C. Dekkers. 2014. "Applied Animal Genomics: Results From the Field." *Annual Review of Animal Biosciences* 2, no. 1: 105–139. <https://doi.org/10.1146/annurev-animal-022513-114119>.
- VanRaden, P. M. 2008. "Efficient Methods to Compute Genomic Predictions." *Journal of Dairy Science* 91: 4414–4423. <https://doi.org/10.3168/jds.2007-0980>.
- VanRaden, P. M. 2020. "Symposium Review: How to Implement Genomic Selection." *Journal of Dairy Science* 103, no. 6: 5291–5301. <https://doi.org/10.3168/jds.2019-17684>.
- VanRaden, P. M., C. P. Van Tassell, G. R. Wiggans, et al. 2009. "Invited Review: Reliability of Genomic Predictions for North American Holstein bulls." *Journal of Dairy Science* 92, no. 1: 16–24. <https://doi.org/10.3168/jds.2008-1514>.
- Wang, C., X. Li, R. Qian, G. Su, Q. Zhang, and X. Ding. 2017. "Bayesian Methods for Jointly Estimating Genomic Breeding Values of One Continuous and One Threshold Trait." *PLoS One* 12, no. 4: e0175448. <https://doi.org/10.1371/journal.pone.0175448>.
- Wang, H., I. Misztal, I. Aguilar, A. Legarra, and W. M. Muir. 2012. "Genome-Wide Association Mapping Including Phenotypes from Relatives Without Genotypes." *Genetics Research* 94, no. 2: 73–83. <https://doi.org/10.1017/s0016672312000274>.
- Weigel, K. A., C. P. Van Tassell, J. R. O'Connell, P. M. VanRaden, and G. R. Wiggans. 2010. "Prediction of Unobserved Single Nucleotide Polymorphism Genotypes of Jersey Cattle Using Reference Panels and Population-Based Imputation Algorithms." *Journal of Dairy Science* 93, no. 5: 2229–2238. <https://doi.org/10.3168/jds.2009-2849>.
- Zhang, X., D. Lourenco, I. Aguilar, A. Legarra, and I. Misztal. 2016. "Weighting Strategies for Single-Step Genomic BLUP: An Iterative Approach for Accurate Calculation of GEBV and GWAS." *Frontiers in Genetics* 7: 151. <https://doi.org/10.3389/fgene.2016.00151>.
- Zhang, Z., J. Liu, X. Ding, P. Bijma, D.-J. de Koning, and Q. Zhang. 2010. "Best Linear Unbiased Prediction of Genomic Breeding Values Using a Trait-Specific Marker-Derived Relationship Matrix." *PLoS One* 5, no. 9: e12648. <https://doi.org/10.1371/journal.pone.0012648>.
- Zhang, Z., Q. Zhang, and X. D. Ding. 2011. "Advances in Genomic Selection in Domestic Animals." *Chinese Science Bulletin* 56: 2655–2663. <https://doi.org/10.1007/s11434-011-4632->

Supporting Information

Additional supporting information can be found online in the Supporting Information section.