



# Comparison of single-step methods for genomic prediction of age at first calving in dairy buffaloes

Jessica Cristina Gonçalves dos Santos<sup>1</sup>, Francisco Ribeiro de Araújo Neto<sup>2</sup> , Gabriela Stefani Fernandez<sup>1</sup>, Daniel Jordan de Abreu Santos<sup>1</sup>, Felipe Pereira Cunha<sup>2</sup>, Rusbel Raul Aspilcueta-Borquis<sup>3</sup> and Humberto Tonhati<sup>1</sup>

<sup>1</sup>Faculdade de Ciências Agrárias e Veterinárias de Jaboticabal – UNESP, Jaboticabal, São Paulo, Brazil; <sup>2</sup>Instituto Federal de Ciência e Tecnologia Goiano – IFGoiano, Rio Verde, Goiás and <sup>3</sup>Universidade Tecnológica Federal do Paraná – UFTPR, Dois Vizinhos, Paraná

## Animal Research Paper

**Cite this article:** dos Santos JCG, de Araújo Neto FR, Fernandez GS, de Abreu Santos DJ, Cunha FP, Aspilcueta-Borquis RR, Tonhati H (2024). Comparison of single-step methods for genomic prediction of age at first calving in dairy buffaloes. *The Journal of Agricultural Science* **162**, 377–383. <https://doi.org/10.1017/S0021859624000364>

Received: 1 October 2023  
Revised: 5 May 2024  
Accepted: 17 June 2024  
First published online: 14 November 2024

**Keywords:** Bayesian inference; *Bubalus bubalis*; genomic selection; reproductive traits

**Corresponding author:**  
F. R. Araujo Neto;  
Email: [francisco.neto@ifgoiano.edu.br](mailto:francisco.neto@ifgoiano.edu.br)

### Abstract

The age at first calving (AFC) is an important trait to be considered in breeding programmes of dairy buffaloes, where new approaches and technologies, such as genomic selection, are constantly applied. Thus, the objective of this study was to compare the predictive ability of different genomic single-step methods using AFC information from Murrah buffaloes. From a pedigree file containing 3320 buffaloes, 2247 cows had AFC records and 553 animals were genotyped. The following models were performed: pedigree-based BLUP (PBLUP), single-step GBLUP (ssGBLUP), weighted single-step GBLUP (WssGBLUP), and single-step Bayesian regression methods (ssBR-BayesA, BayesB $\pi$ , BayesC $\pi$ , Bayes-Lasso, and BayesRR). To compare the methodologies, the accuracy and dispersion of (G)EBVs were assessed using the LR method. Accuracy estimates for the genotyped animals ranged from 0.30 (PBLUP) to 0.39 (WssGBLUP). Predictions with the traditional model (PBLUP) were very dispersed from what was expected, while BayesC $\pi$  (0.99) and WssGBLUP (1.00) obtained the lowest dispersion. The results indicate that the use of genomic information can improve the genetic gain for AFC by increasing the accuracy and reducing inflation/deflation of predictions compared to the traditional pedigree-based model. In addition, among all genomic single-step models studied, WssGBLUP and single-step BayesA were the most advantageous methods to be used in the genomic evaluation of AFC of buffaloes from this population.

### Introduction

Age at first calving (AFC) is an important indicator trait of reproductive efficiency that is genetically monitored by the Milk-Recording Buffalo Program in Brazil (Aspilcueta-Borquis *et al.*, 2022). It is well known, however, that the genetic progress obtained by direct selection of AFC can be very slow due to the low heritability magnitude of this trait in Buffaloes (Camargo *et al.*, 2015). Since genomic selection is more accurate than traditional selection, especially for low-heritability traits (Calus *et al.*, 2008), genomic evaluation is the best strategy for predicting breeding values for AFC. Furthermore, Camargo *et al.* (2015) and Araujo Neto *et al.* (2020a) have reported some specific genomic regions highly associated to this trait in GWAS studies in this population, indicating that differential shrinkage models, such as Bayesian alphabet and weighted single-step GBLUP models, would be a reasonable alternative to be explored.

The single-step GBLUP (ssGBLUP) method has become the main methodology used for genomic evaluations in dairy buffaloes (Araujo Neto *et al.*, 2020a, 2020b; Cesarani *et al.*, 2021; Lázaro *et al.*, 2021; Araujo Neto *et al.*, 2022) given the ease of integrating the relationship matrices, based on pedigree (A) and genomic (G) information, without major changes in the mixed model equations (Misztal *et al.*, 2020). Moreover, this methodology allows assigning different weights for each marker, originating the weighted single-step GBLUP method – WssGBLUP (Wang *et al.*, 2012; Zhang *et al.*, 2016).

Subsequently, Fernando *et al.* (2014) proposed single-step methodologies combined to Bayesian regression methods (ssBR), as an option to ssGBLUP. Nonetheless, few studies have been developed comparing the results of the different single-step methodologies (Lee *et al.*, 2017; Gao *et al.*, 2018; Zhou *et al.*, 2018). Based on this, it becomes interesting to perform similar analyses in buffalo databases to assist in decision making in breeding programmes that employ genomic technology. Thus, this work was designed with the aim of comparing the predictive ability of different single-step methods using AFC information from Murrah buffaloes.

**Materials and methods**

**Dataset**

The phenotypic dataset consisted of 2,314 AFC records of Murrah buffalo cows born between 1995 and 2017. Contemporary groups (CG) were formed considering animals born on the same farm, year and season of birth, which was divided into two (dry and rainy season). Animals with records outside the range between  $\pm 3$  standard deviations from their CG averages and animals belonging to GCs with less than five individuals were removed.

From a pedigree file containing 3320 buffaloes, 553 animals (539 dams and 14 sires) were genotyped with the Axiom Buffalo Genotyping Array 90 K (Iamartino *et al.*, 2017). Only SNPs present on autosomal chromosomes (BBU1-BBU24 referenced by UOA\_WB\_1 genome assembly), as well as those with call rate >95%, MAF >3% and significance level for Hardy Weinberg equilibrium test was  $10^{-5}$ , were remained in the analysis. All samples had a call rate >90%. The database description (phenotypic and genotypic information) after the consistency step is in Table 1.

**PBLUP, ssGBLUP and WssGBLUP models**

This set of methodologies is based on the use of relationship matrices between individuals in mixed model equations, with the differences consisting of the type of information used (pedigree or genomic) and the way these matrices are constructed. The model using pedigree based BLUP (PBLUP) can be described as:

$$y = X\beta + Z\alpha + e$$

where,  $y$  is the vector of phenotypic records (AFC);  $\beta$  is the fixed effect vector (CG),  $X$  is the incidence matrix that associates  $\beta$  with  $y$ ;  $\alpha$  is the vector of additive genetic effects,  $Z$  is an incidence matrix associating  $\alpha$  with  $y$  and,  $e$  is the vector of residuals. The following assumptions were made:

$$E(y) = X\beta \quad E(\alpha) = 0 \quad E(e) = 0$$

$$\text{Var}(\alpha) = A\sigma_\alpha^2 \quad \text{Var}(e) = I_n\sigma_e^2$$

where  $\sigma_\alpha^2$  and  $\sigma_e^2$  represent the additive genetic and residual variances;  $A$  and  $I_n$  are the pedigree relationship and an identity matrices, respectively.

The single step GBLUP (ssGBLUP) method is an extension of the PBLUP model, in which the pedigree-based ( $A$ ) and genomic ( $G$ ) relationship matrices are combined into a single matrix ( $H$ ), as described by Aguilar *et al.* (2010). So, its inverse can be obtained as:

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

where  $A^{-1}$  and  $G^{-1}$  are the inverse matrices of  $A$  and  $G$  respectively, and  $A_{22}^{-1}$  is the inverse of the section of  $A$  related to the genotyped animals. The  $G$  matrix was obtained according to VanRaden (2008) as:

$$G = \frac{ZDZ'}{2 \sum p_i(1 - p_i)}$$

**Table 1.** Description of age at first calving dataset used in this study

Description	Number of records	
	Full dataset	Partial dataset
Total of animals with records	2247	1596
Total of genotyped animals	539	539
Total of genotyped animals with records	509	420
Total of contemporary groups	101	73
Mean (months)	37.34	36.97
Standard Deviation (months)	5.52	5.07

where  $p_i$  is the minor allele frequency (MAF) of SNP  $i$ ,  $Z$  is a matrix relating genotypes of each locus centred by allele frequencies, and  $D$  is a diagonal matrix of weights for SNP variances, (with elements  $d_i = \text{SNP } i \text{ weight}$ ). In ssGBLUP model is set  $D = I$ , so that the weight of all SNPs is equal to 1.

In WssGBLUP method (Wang *et al.*, 2012) markers are assigned with different weights using an iterative process described with the following steps:

1. Set  $D^{(t)} = I$ , when  $t = 1$
2. The genomic relationship matrix is setup for  $t$  as  $G^{(t)} = ZD^{(t)}Z' / (2 \sum p_i(1 - p_i))$
3. The Genomic Breeding Value (GEBV or  $\hat{\alpha}$ ) for  $t$  are obtained;
4. The GEBV  $\hat{\alpha}^{(t)}$  are converted to SNP effects ( $\hat{u}$ ) as  $\hat{u}^t = D^{(t)}Z'(G^{(t)})^{-1}\hat{\alpha}^{(t)}$
5. The weight of the  $i$ th SNP (the  $i$ th element of  $D$  or  $d_i$ ) is calculated;
6. SNP weights are normalized in  $D^{(t+1)}$  to have constant genetic variances of SNP effects;
7. Loop to step 2 three times and exit.

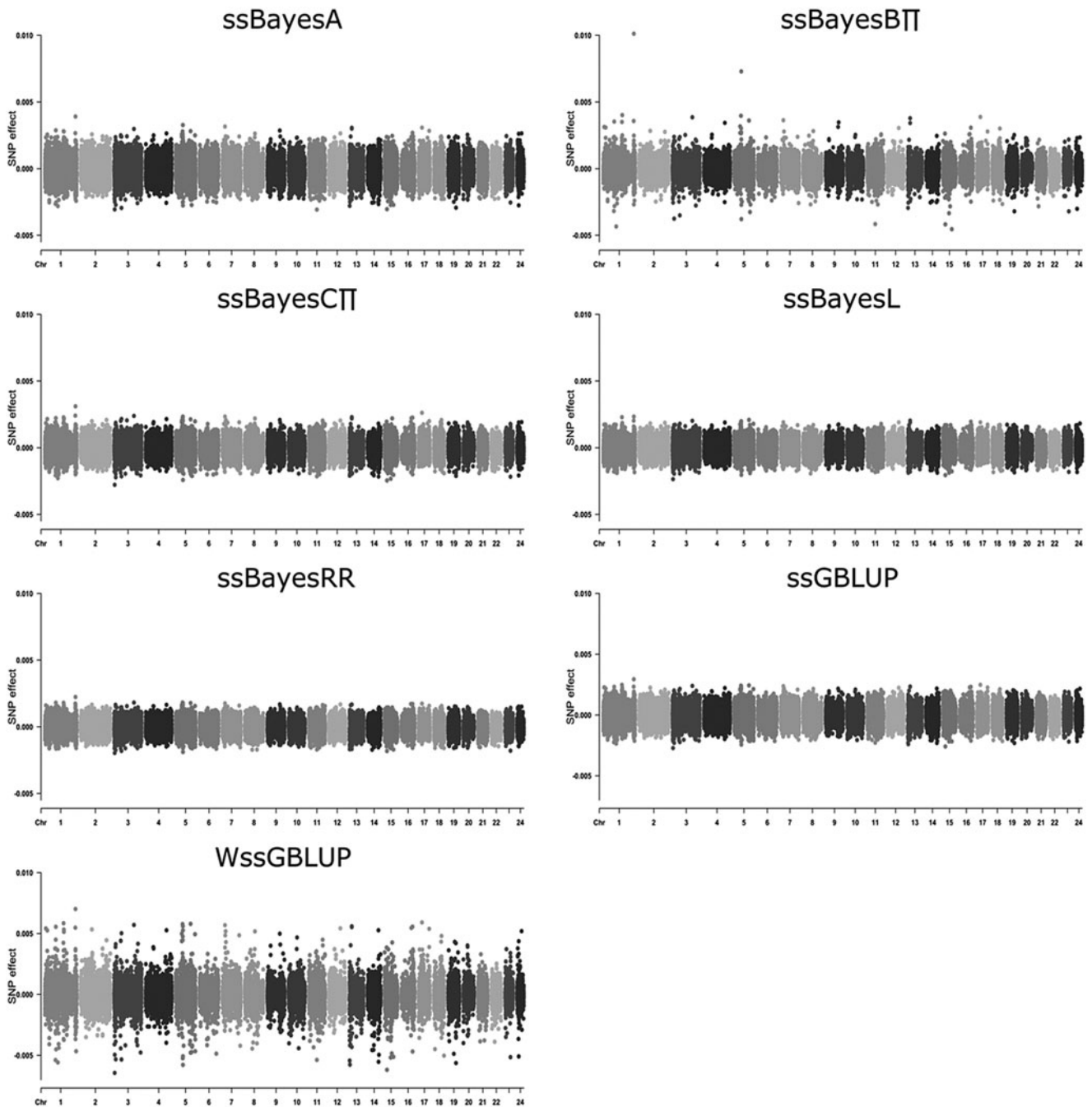
For the methods mentioned above, the program BLUPF90+ (Misztal *et al.*, 2014) was used for the estimation of the variance components and for the validation step. The estimation of the effects of the markers and the calculation of the weighting (WssGBLUP) were performed with the program POSTGSF90 (Misztal *et al.*, 2014). The heritability estimates for AFC were obtained from these estimated variance components.

**Single-step Bayesian regression models (ssBR)**

The ssBR models combine all available data from genotyped and non-genotyped animals and use imputed marker covariates for animals that are not genotyped (Fernando *et al.*, 2014), can be described as:

$$\begin{bmatrix} y_n \\ y_g \end{bmatrix} = \begin{bmatrix} X_n \\ X_g \end{bmatrix} \beta + \begin{bmatrix} Z_n & 0 \\ 0 & Z_g \end{bmatrix} \begin{bmatrix} M_n \alpha + \epsilon \\ M_g \alpha \end{bmatrix} + e$$

where vectors and matrices for non-genotyped animals are denoted with the subscript  $n$  and for genotyped animals with the subscript  $g$ . Thus,  $y_n$  and  $y_g$  represent the vectors of phenotypic observations;  $\beta$  is the vector of systematic effects of GC (equivalent to the fixed effect in frequentist methods);  $X_n$  and  $X_g$  design matrices for the fixed effects;  $Z_n$  and  $Z_g$  represent



**Figure 1.** Estimates of SNP marker effects (Manhattan plot) for the different methods employed (ssBA, Single step Bayes A; ssBπ, Single step Bayes Bπ; ssBCπ, Single step Bayes Cπ; ssBL, Single step Bayes Lasso; ssBRR, Single step Bayes Ridge Regression), and cluster analysis based on similarity between marker effects.

incidence matrices associated with the genomic values of the animals;  $\alpha$ ,  $\epsilon$  and  $e$ , represent in this order, the vectors of marker effects, imputation residual and model residual;  $M_g$  is the marker matrix for the genotyped animals and  $M_n$  is the imputed marker matrix for the non-genotyped animals. The matrix  $M_n$  can be obtained as:

$$M_n = A_{ng}A_{22}^{-1}M_g$$

where  $A_{ng}$  represents the pedigree-based relationship matrix between genotyped and non-genotyped animals and;  $A_{2g}^{-1}$  represents the inverse of the pedigree-based relationship matrix

between genotyped animals. The following ssBR methods were used: BayesA (ssBA- Meuwissen *et al.*, 2001), BayesBπ (ssBBπ- Meuwissen *et al.*, 2001), BayesCπ (ssBCπ- Habier *et al.*, 2011), Lasso Bayesian (ssBL- Yi and Xu, 2008) and Bayesian ridge regression (ssBRR- Campos *et al.*, 2013), where the assumptions for markers effects are shown in Table 2. The hibayes package (Yin *et al.*, 2022), available for the R program (R Core Team, 2021), was used to perform the ssBR analyses. A total of 350 000 samples and a burn-in period of 150 000 samples were generated. The convergence was evaluated using graphical analysis. The heritability estimates for AFC were also obtained using these models.

**Table 2.** Description of different prior for markers effects used in Bayesian single-step regression methods used for genomic analysis of age at first calving in dairy buffaloes

Model	Presuppositions
Single step Bayes A	$\alpha_j \sim N(0, \sigma_{aj}^2)$ $\sigma_{aj}^2 \sim \chi^{-2}(v, S)$
Single step Bayes $B\pi$	$\alpha_j \sim (1 - \pi)N(0, \sigma_a^2) + \pi\delta_0$ $\sigma_{aj}^2 \sim \chi^{-2}(v, S)$
Single step Bayes $C\pi$	$\alpha_j \sim (1 - \pi)N(0, \sigma_a^2) + \pi\delta_0$ $\alpha_j \sim N(0, \sigma_a^2)$
Single step Bayes Lasso	$\alpha_j \sim N(0, \sigma_{aj}^2)$ $\sigma_{aj}^2 \sim \text{Expon}(\lambda^2/2)$
Single step Bayes Ridge Regression	$\alpha_j \sim N(0, \sigma_a^2)$

### Validation of (G)EBVs

The predictive ability of the methods was assessed using accuracy and dispersion statistics proposed by Legarra and Reverter (2018). This methodology is based on predicting (G)EBV using full ( $\hat{u}_w$ ) and a partial ( $\hat{u}_p$ ) datasets. The partial dataset (reference group) was obtained by omitting phenotypic records of genotyped animals born after 2010 (validation group). Thus, to assess the prediction ability of the models, only genotyped animals from the validation group were used to calculate the following statistics:

$$r_{A\hat{A}} = \sqrt{\frac{\text{cov}(\hat{u}_w, \hat{u}_p)}{(1 - \bar{F})\hat{\sigma}_{g,i}^2}} \quad \text{Dispersion} \quad b_{w,p} = \frac{\text{cov}(\hat{u}_w, \hat{u}_p)}{\text{var}(\hat{u}_p)}$$

where  $\bar{F}$  is the average inbreeding coefficient of the validation individuals and  $\hat{\sigma}_{g,i}^2$  is the estimate of additive genetic variance.

### Results

In this study, the estimated heritability for AFC was 0.17 using BLUP-based methods and 0.16 to 0.18 using ssBR methods and, the estimates of the effects of the markers can be seen in Figure 1. Accuracy estimates for the genotyped animals ranged from 0.30 (ssBRR) to 0.39 (WssGBLUP) (Fig. 2). Estimates of accuracy with low magnitude were observed for non-genotyped animals, with values ranging from 0.29 (ssGBLUP) to 0.33 (ssBA). The PBLUP model presented accuracies of 0.30 and 0.29 for genotyped and non-genotyped animals, respectively.

Regarding the dispersion of the GEBV prediction (Fig. 2), ssBC $\pi$  (0.99) and WssGBLUP (1.00) obtained the best results considering only genotyped animals, while for non-genotyped animals only WssGBLUP (1.02) showed the best value, although it was also the higher than 1. The PBLUP model had the worst result with values straying 35% below from the reference value (1.0).

In the full dataset, Pearson's correlation between GEBVs and EBVs (predicted via PBLUP) ranged from 0.92 (ssBRR) to 0.93 (ssGBLUP), while for animals with both genotypic and phenotypic information, the correlation between GEBVs and EBVs ranged from 0.71 (ssBRR) to 0.77 (WssGBLUP). The correlations between all GEBVs predicted with the single-step genomic models were higher than 0.95.

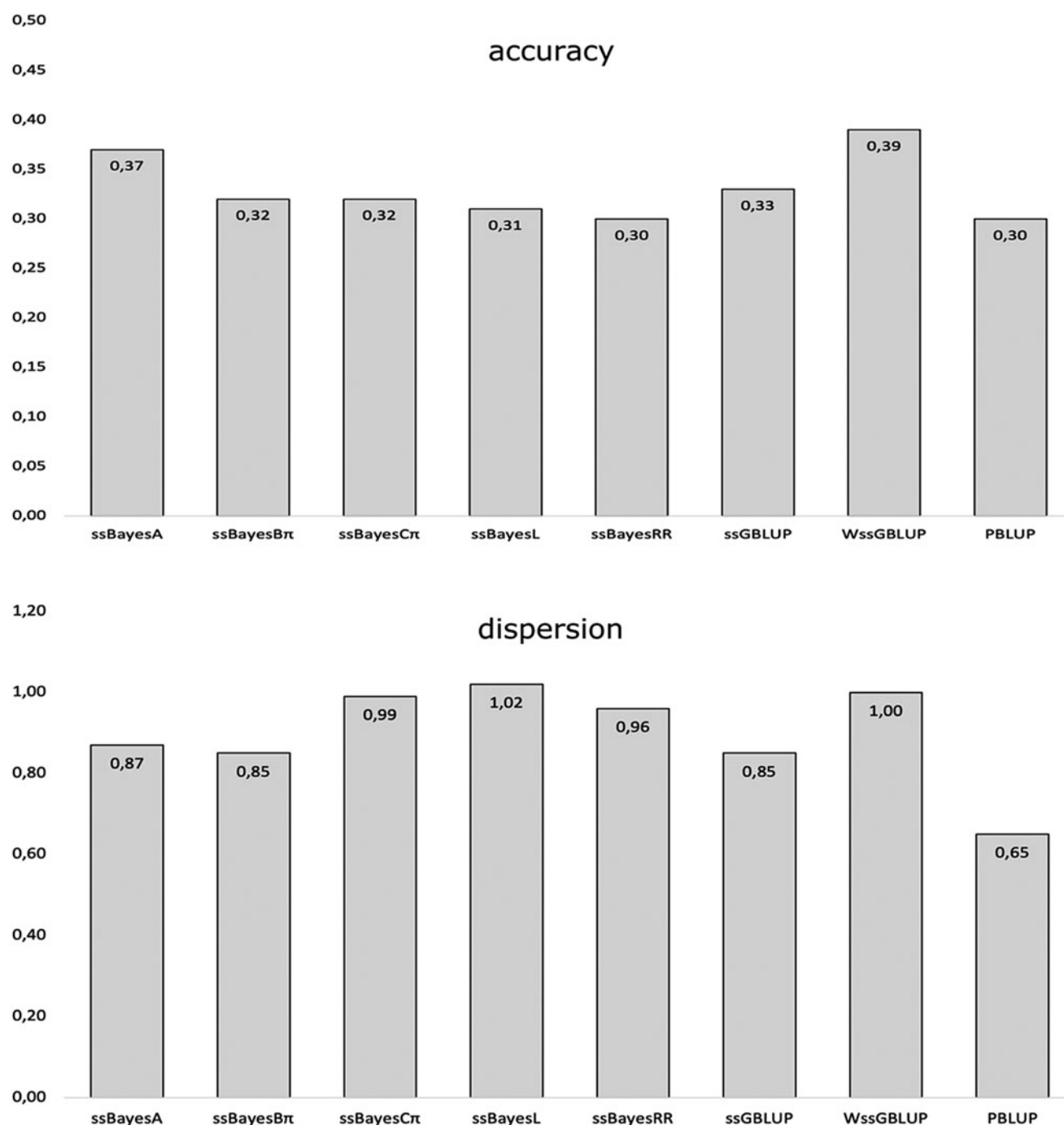
### Discussion

In this study, a heritability estimate obtained for AFC, had magnitude similar to those described in several studies with the buffalo species, with values ranging from 0.13 to 0.19 (Agudelo *et al.*, 2015; Thiruvankadan *et al.*, 2015; Fernandes *et al.*, 2016; Araujo Neto *et al.*, 2020a, 2020b). These results show that despite the greater information provided by the markers in the analysis, there was very little influence on the estimation of the genetic parameters, possibly due to the polygenic nature of the trait. There are few reports in the literature on comparison of genomic prediction for reproductive traits, and it is not yet possible to find studies that compared the predictive ability between different single-step methods for AFC. Moreover, several accuracy measures have been reported by different authors, which makes the comparison between their values even more difficult. For instance, Costa *et al.* (2019), working only with genotyped Nellore breed animals and considering the correlation between GEBV and adjusted AFC ( $r_{GEBV, Y^*}$ ), estimated values with magnitudes similar to those found in our study with BayesC $\pi$  (0.31) and Bayes-Lasso (0.31), while Toghiani *et al.* (2017) reported lower values with BayesA (0.148) and BayesC $\pi$  (0.15), studying the same breed and using the same type of accuracy. In both cited studies, accuracy estimates were not presented for PBLUP model, which does not allow us to fully contrast with our findings. Possibly, for comparison purposes, checking the accuracy increments in relation to PBLUP is the best approach to discuss our results. Thus, all literature discussed about predictability were made based on the difference in performance between methods also used in our study.

Considering the percentage of accuracy increment in relation to PBLUP predictions in only genotyped animals, we observed three groups of models: increments less than 5% – ssBL and ssBRR; increments close 10% – ssBB $\pi$ , ssBC $\pi$  and ssGBLUP and; increments greater than 25% – ssBA and WssGBLUP. Estimates of accuracy increments using SNP markers with magnitudes similar to those we have found for AFC, have been described in the literature for productive traits in buffaloes (Cesarani *et al.*, 2021; Herrera *et al.*, 2021).

For milk, fat and protein yield, greater increases in accuracy ( $r_{GEBV, Y^*}$ ) are presented by Herrera *et al.* (2021) (13.04 to 76.47%) and Cesarani *et al.* (2021) (15.28 to 33.33%). Aspilcueta-Borquis *et al.* (2015), studying the same population of this study and ssGBLUP model, reported similar increase in the average accuracy (based on prediction error variance) of (G)EBVs, which ranged from 8.52 to 12.05% for several dairy traits. This gain in accuracy observed with the use of genomic information is due to the additional capture of both Mendelian sampling variations and relationships between animals (Christensen *et al.*, 2012; Cesarani *et al.*, 2020). The similarity in predictive performance between the different single-step genomic methods, however, could be related to the reduced number of animals evaluated, the small subset of genotyped animals analysed, and the polygenic nature of the AFC (Calus, 2010; Campos *et al.*, 2013).

The proximity between ssBA and GRM-based model observed in this study (by the magnitude of estimated accuracies) may be explained by the polygenic nature of the trait and the robustness of ssBA method for different genetic architectures. Zhou *et al.* (2018), studying the effects of QTL number on the accuracy prediction with simulated data, also verified that ssBA method was robust and performed similarly to ssGBLUP in scenarios with a large number of QTL.



**Figure 2.** Estimates of accuracy and dispersion of genetic and genomic values using different single-step methods (ssBA, Single step Bayes A; ssB $\pi$ , Single step Bayes B $\pi$ ; ssBC $\pi$ , Single step Bayes C $\pi$ ; ssBL, Single step Bayes Lasso; ssBRR, Single step Bayes Ridge Regression) for the study of age at first calving in dairy buffaloes (genotyped animals).

The dispersion criterion is related to the degree of inflation or deflation of predictions, which was measured in our validation set as the slope of the regression of the GEBV obtained with the full dataset on GEBV from the partial dataset. Our results showed that the inclusion of genomic information reduces the dispersion of prediction compared to the traditional evaluations (PBLUP), regardless of the single-step model considered. The same improvement in prediction slopes by using genomic information was reported by Lázaro *et al.* (2021) studying several milk-related traits with random regression models in the same buffalo

population, while other authors, such as Gao *et al.* (2018) also analysing dairy traits in Finnish red dairy cattle, report no benefit. Moreover, the GEBVs predicted with WssGBLUP did not disperse in relation to what was expected for the subset of genotyped animals and presented little deflation for the prediction of non-genotyped animals. The low dispersion of GEBVs predicted with WssGBLUP compared to the predictions from other genomic models is probably due to the self-fitting of the SNP weights to the dataset (weighting steps), which regulates the scale of the additive effect to the full dataset previously (*a priori*).

The results indicate that the use of genomic information can improve the genetic gain for AFC by increasing the accuracy and reducing inflation/deflation of predictions of this trait compared to the traditional pedigree-based model. In addition, among all genomic single-step models studied, WssGBLUP and ssBA were the most advantageous models to be used for the genomic evaluation of AFC of buffaloes from this population.

**Authors' contributions.** Jessica Cristina Gonçalves dos Santos: Methodology, Analysis, Writing – original draft.; Francisco Ribeiro de Araujo Neto: Conceptualization, Methodology, Statistical analysis, Writing – review; Gabriela Stefani Fernandez: Database; Daniel Jordan de Abreu Santos: Writing – review & editing; Felipe Pereira Cunha: Writing editing; Rusbel Raul Aspilueta-Borquis: Statistical analysis; Humberto Tonhati: Conceptualization, Database, Supervision.

**Funding statement.** This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

**Competing interests.** None.

**Ethical standards.** Not applicable (the work was developed with a database).

## References

- Agudelo-Gomez D, Pineda-Sierra S and Cerón-Muñoz MF (2015) Genetic evaluation of dual-purpose buffaloes (*Bubalus bubalis*) in Colombia using principal component analysis. *PLoS One* **10**, e0132811.
- Aguilar I, Misztal I, Johnson DL, Legarra A, Tsuruta S and Lawlor TJ (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**, 743–752. <https://doi.org/10.3168/jds.2009-2730>.
- Araujo Neto FR, Takada L, Santos DJA, Aspilueta-Borquis RR, Cardoso DF, Nascimento AV, Leão KM, Oliveira HN and Tonhati H (2020a) Identification of genomic regions related to age at first calving and first calving interval in water buffalo using single-step GBLUP. *Reproduction in Domestic Animals* **55**, 1565–1572. <https://doi.org/10.1111/rda.13811>.
- Araujo Neto FR, Santos DJA, Fernandes Júnior GA, Aspilueta-Borquis RR, Nascimento AV, Seno LO, Tonhati H and Oliveira HN (2020b) Genome-wide association studies for growth traits in buffaloes using the single step genomic BLUP. *Journal of Applied Genetics* **61**, 113–115. doi: 10.1007/s13353-019-00528-5.
- Araujo Neto FR, Santos JCG, Silva Arce CD, Borquis RRA, Santos DJA, Guimarães KC, Nascimento AV, Oliveira HN and Tonhati H (2022) Genomic study of the resilience of buffalo cows to a negative energy balance. *Journal of Applied Genetics* **63**, 379–388. doi: 10.1007/s13353-021-00680-x.
- Aspilueta-Borquis RR, Araujo Neto FR, Santos DJ, Hurtado-Lugo NA, Silva JA and Tonhati H (2015) Multiple-trait genomic evaluation for milk yield and milk quality traits using genomic and phenotypic data in buffalo in Brazil. *Genetics and Molecular Research* **14**, 18009–18017. doi: 10.4238/2015.December.22.27.
- Aspilueta-Borquis RR, Seno LO, Araujo Neto FR, Santos DJA, Hurtado-Lugo NA and Tonhati H (2022) Lifetime productivity: genetic study of longevity and its associations with economically important traits in dairy buffaloes. *Livestock Science* **259**, 104900. <https://doi.org/10.1016/j.livsci.2022.104900>.
- Calus MPL (2010) Genomic breeding value prediction: methods and procedures. *Animal* **4**, 157–164. <https://doi.org/10.1017/S1751731109991352>.
- Calus MPL, Meuwissen THE, De Roos APW and Veerkamp RF (2008) Accuracy of genomic selection using different methods to define haplotypes. *Genetics* **178**, 553–561. <https://doi.org/10.1534/genetics.107.080838>.
- Camargo GMF, Aspilueta-Borquis RR, Fortes MRS, Porto-Neto R, Cardoso DF, Santos DJA, Lehnert SA, Reverter A, Moore SS and Tonhati H (2015) Prospecting major genes in dairy buffaloes. *BMC Genomics* **16**, 872. <https://doi.org/10.1186/s12864-015-1986-2>.
- Campos G, Hickey JM, Pong-Wong R, Daetwyler HD and Calus MPL (2013) Whole-genome regression and prediction methods applied to plant and animal breeding. *Genetics* **193**, 327–345. <https://doi.org/10.1534/genetics.112.143313>.
- Cesarani A, Garcia A, Hidalgo J, Degano L, Vicario D, Macciotta NPP and Lourenco D (2020) Genomic information allows for more accurate breeding values for milk ability in dual-purpose Italian Simmental cattle. *Journal of Dairy Science* **104**, 5719–5727. <https://doi.org/10.3168/jds.2020-19838>.
- Cesarani A, Biffani S, Garcia A, Lourenco D, Bertolini G, Neglia G, Misztal I and Macciotta NPP (2021) Genomic investigation of milk production in Italian buffalo. *Italian Journal of Animal Science* **20**, 539–547. <https://doi.org/10.1080/1828051X.2021.1902404>.
- Christensen OF, Madsen P, Nielsen B, Ostensen T and Su G (2012) Single-step methods for genomic evaluation in pigs. *Animal* **6**, 1565–1571. <https://doi.org/10.1017/S1751731112000742>.
- Costa RB, Irano N, Diaz IDPS, Takada L, Hermisdorff IC, Carneiro R, Baldi F, Oliveira HN, Tonhati H and Albuquerque LG (2019) Prediction of genomic breeding values for reproductive traits in Nelore heifers. *Theriogenology* **125**, 12–17. <https://doi.org/10.1016/j.theriogenology.2018.10.014>.
- Fernandes PB, Marques KO, de Araujo Neto FR, de Oliveira DP, Hurtado-Lugo NA, Aspilueta-Borquis RR and Tonhati H (2016) Genetic-quantitative study of the first-service pregnancy probability of Murrah heifers. *Reproduction in Domestic Animals* **51**, 428–434.
- Fernando RL, Dekkers JC and Garrick DJ (2014) A class of Bayesian methods to combine large numbers of genotyped and non-genotyped animals for whole-genome analyses. *Genetics Selection Evolution* **46**, 50. <https://doi.org/https://doi.org/10.1186/1297-9686-46-50>.
- Gao H, Koivula M, Jensen J, Strandén I, Madsen P, Pitkänen T, Aamand GP and Mäntysaari EA (2018) Short communication: genomic prediction using different single-step methods in the Finnish red dairy cattle population. *Journal of Dairy Science* **101**, 10082–10088. doi: 10.3168/jds.2018-14913.
- Habier D, Fernando RL, Kizilkaya K and Garrick DJ (2011) Extension of the Bayesian alphabet for genomic selection. *BMC Bioinformatics* **12**, 186. <https://doi.org/10.1186/1471-2105-12-186>.
- Herrera JRV, Flores EB, Duijvesteijn N, Moghaddar N and van der Werf JH (2021) Accuracy of genomic prediction for milk production traits in Philippine dairy buffaloes. *Frontiers in Genetics* **12**, 682576. doi: 10.3389/fgene.2021.682576.
- Iamartino D, Nicolazzi EL, Van Tassell CP, Reecy JM, Fritz-Waters ER, Koltes JE, Biffani S, Sonstegard TS, Schroeder SG, Ajmone-Marsan P, Negrini R, Pasquariello R, Ramelli P, Coletta A, Garcia JF, Ali A, Ramunno L, Cosenza G, Oliveira DAA, Drummond MG, Bastianetto E, Davassi A, Pirani A, Brew F and Williams JL (2017) Design and validation of a 90K SNP genotyping assay for the water buffalo (*Bubalus bubalis*). *PLoS One* **12**, e0185220. <https://doi.org/10.1371/journal.pone.0185220>.
- Lázaro SF, Tonhati H, Oliveira HR, Silva AA, Nascimento AV, Santos DJA, Stefani G and Brito LF (2021) Genomic studies of milk-related traits in water buffalo (*Bubalus bubalis*) based on single-step genomic best linear unbiased prediction and random regression models. *Journal of Dairy Science* **104**, 5768–5793. <https://doi.org/10.3168/jds.2020-19534>.
- Lee J, Cheng H, Garrick D, Golden B, Dekkers J, Park K, Lee D and Fernando R (2017) Comparison of alternative approaches to single-trait genomic prediction using genotyped and non-genotyped Hanwoo beef cattle. *Genetics Selection Evolution* **49**, 2. <https://doi.org/10.1186/s12711-016-0279-9>.
- Legarra A and Reverter A (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**, 53. <https://doi.org/10.1186/s12711-018-0426-6>.
- Meuwissen THE, Hayes BJ and Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. *Genetics* **157**, 1819–1829. doi: 10.1093/genetics/157.4.1819.
- Misztal I, Tsuruta S, Lourenço D, Masuda Y, Aguilar I, Legarra A and Vitezica Z (2014) BLUPF90 family of programs. Disponível em: <http://nce.ads.uga.edu/wiki/doku.php>.
- Misztal I, Lourenco D and Legarra A (2020) Current status of genomic evaluation. *Journal of Animal Science* **98**, skaa101. <https://doi.org/10.1093/jas/skaa101>.
- R Core Team (2021) R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.

- Thiruvenkadan AK, Panneerselvam S and Murali N** (2015) Study on ages at first mating and calving of Murrah buffaloes in hot and humid climate of Tamil Nadu. *India. Indian Journal of Animal Research* **49**, 591–594.
- Toghiani S, Hay E, Sumreddee P, Geary TW, Rekaya R and Roberts AJ** (2017) Genomic prediction of continuous and binary fertility traits of females in a composite beef cattle breed. *Journal of Animal Science* **95**, 4787–4795. doi: 10.2527/jas2017.1944.
- VanRaden PM** (2008) Efficient methods to compute genomic predictions. *Journal Dairy Science* **91**, 4414–4423.
- Wang H, Misztal I, Aguilar I, Legarra A and Muir W** (2012) Genome-wide association mapping including phenotypes from relatives without genotypes. *Genetics Research* **94**, 73–83. doi: 10.1017/S0016672312000274.
- Yi N and Xu S** (2008) Bayesian LASSO for quantitative trait loci mapping. *Genetics* **179**, 1045–1055. doi: 10.1534/genetics.107.085589.
- Yin L, Zhang H, Li X, Zhao S and Liu X** (2022) Hibayes: an R package to fit individual-level, summary-level and single-step Bayesian regression models for genomic prediction and genome-wide association studies. *bioRxiv*. Accessed in December 2022. Available in: <https://doi.org/10.1101/2022.02.12.480230>.
- Zhang X, Lourenco D, Aguilar I, Legarra A and Misztal I** (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>.
- Zhou L, Mrode R, Zhang S, Zhang Q, Li B and Liu JF** (2018) Factors affecting GEBV accuracy with single-step Bayesian models. *Heredity* **120**, 100–109. doi: 10.1038/s41437-017-0010-9.