Received: 18 October 2019 | Revised: 9 January 2020 | Accepted: 11 January 2020

DOI: 10.1111/jbg.12468

ORIGINAL ARTICLE

Journal of Animal Breeding and Genetics WILEY

Improving genomic prediction accuracy for meat tenderness in Nellore cattle using artificial neural networks

Fernando Brito Lopes[1,2] | **Cláudio U. Magnabosco**[1] | **Tiago L. Passafaro**[3] | **Ludmilla C. Brunes**[4] | **Marcos F. O. Costa**[5] | **Eduardo C. Eifert**[1] | **Marcelo G. Narciso**[5] | **Guilherme J. M. Rosa**[3,6] | **Raysildo B. Lobo**[7] | **Fernando Baldi**[1]

[1]Department of Animal Science, São Paulo State University (UNESP), Jaboticabal, Brazil

[2]Embrapa Cerrados, Brasilia, Brazil

[3]Department of Animal Sciences, University of Wisconsin-Madison, Madison, WI, USA

[4]Department of Animal Science, Federal University of Goiás (UFG), Goiânia, Brazil

[5]Embrapa Rice and Beans, Santo Antônio de Goiás, Brazil

[6]Department of Biostatistics and Medical Informatics, University of Wisconsin-Madison, Madison, WI, USA

[7]National Association of Breeders and Researchers (ANCP), Ribeirão Preto, Brazil

Correspondence
Fernando Brito Lopes, Department of Animal Science, São Paulo State University (UNESP), Prof. Paulo Donato Castelane, Jaboticabal, SP, 14884-900, Brazil.
Email: camult@gmail.com

Funding information
Fundação de Amparo à Pesquisa do Estado de São Paulo, Grant/Award Number: 2017/03221-9

Abstract

The goal of this study was to compare the predictive performance of artificial neural networks (ANNs) with Bayesian ridge regression, Bayesian Lasso, Bayes A, Bayes B and Bayes Cπ in estimating genomic breeding values for meat tenderness in Nellore cattle. The animals were genotyped with the Illumina Bovine HD Bead Chip (HD, 777K from 90 samples) and the GeneSeek Genomic Profiler (GGP Indicus HD, 77K from 485 samples). The quality control for the genotypes was applied on each Chip and comprised removal of SNPs located on non-autosomal chromosomes, with minor allele frequency <5%, deviation from HWE ($p < 10^{-6}$), and with linkage disequilibrium >0.8. The FImpute program was used for genotype imputation. Pedigree-based analyses indicated that meat tenderness is moderately heritable (0.35), indicating that it can be improved by direct selection. Prediction accuracies were very similar across the Bayesian regression models, ranging from 0.20 (Bayes A) to 0.22 (Bayes B) and 0.14 (Bayes Cπ) to 0.19 (Bayes A) for the additive and dominance effects, respectively. ANN achieved the highest accuracy (0.33) of genomic prediction of genetic merit. Even though deep neural networks are recognized to deliver more accurate predictions, in our study ANN with one single hidden layer, 105 neurons and rectified linear unit (ReLU) activation function was sufficient to increase the prediction of genetic merit for meat tenderness. These results indicate that an ANN with relatively simple architecture can provide superior genomic predictions for meat tenderness in Nellore cattle.

KEYWORDS

animal breeding, Bayesian regression models, deep learning, genomic selection, Zebu

1 | INTRODUCTION

Brazil has the second largest commercial beef cattle herd in the world, occupying a prominent position in the global beef market. Zebu cattle (*Bos indicus*) comprise more than 80% of the beef cattle in Brazil, and the vast majority of these animals are from Nellore breed, given their tolerance to tropical climate and high resistance to ectoparasites (Baldi et al., 2012). Despite their advantages for production in tropical environments, Zebu cattle tend to produce tougher meat than *Bos taurus* breeds (Reverter et al., 2003; Thrift & Thrift, 2002). Brazilian beef quality tends to be highly variable, due to a broad diversity of production environments and management systems, and to a large proportion of Zebu animals in

the national herd. Some studies have shown the potential for improving the meat tenderness in Zebu cattle (Bonilha et al., 2008; Pereira et al., 2015; Smith et al., 2007). Clearly, the use of tropically adapted Zebu cattle such as the Nellore breed is essential for the Brazilian beef industry, but long-term success in the global market will depend on their response to selection for improved meat quality.

Beef eating quality is the result of different factors such as tenderness, juiciness and flavour (Koohmaraie, Kent, Shackelford, Veiseth, & Wheeler, 2002). Such factors together contribute to the consumer's opinion about the meat palatability (Jeremiah, 1982). Therefore, genetic improvement of meat and carcass traits in beef cattle is important to fulfil consumers' demands and to add value and obtain better prices on the global market. Some efforts have been made to improve beef quality in animal breeding programmes, such as using pedigree information and phenotypic measurement (Magnabosco et al., 2016). Meat quality is an economically important trait and must be considered in animal breeding programmes in order to better determine the return on investment in the livestock. At present, meat tenderness cannot be assessed in vivo, since it requires harvesting the animal to measure this trait. Thus, identifying animals that produce tender meat is laborious, time-consuming and expensive. These aspects decrease the opportunities for large-scale progeny testing, and they are a major limitation for genetic improvement in meat quality. Considering the difficulty and high cost for measuring meat tenderness, genomic selection (GS) approach could maximize the prediction accuracy of genetic merit and increase the number of evaluated animals, being beneficial for livestock (Magnabosco et al., 2016).

Different genomic prediction methods have been developed and used to improve breeding efficiency and genetic gains. In GS methods, markers are simultaneously fitted in models to explore linkage disequilibrium (LD) between markers and quantitative trait loci, to capture most of the relevant variation of the genome. The prediction accuracy of GS depends not only on the genetic architecture of the trait and the number of markers and their level of LD with relevant loci, but also on the statistical model employed. Several models have been developed for genomic prediction, including ridge regression, G-BLUP, Bayes A, Bayes B, Bayes C, Bayesian Lasso, Bayes R and reproducing kernel Hilbert space regression (De los Campos, Hickey, Pong-Wong, Daetwyler, & Calus, 2013; Meuwissen, Hayes, & Goddard, 2001; Moser, Tier, Crump, Khatkar, & Raadsma, 2009; Su, Christensen, Janss, & Lund, 2014). These methods differ in the assumptions of the distribution and variances of marker effects, for example, for ridge regression method, all markers are assumed to have the same variance and their effects follow a normal distribution.

Genomic prediction techniques (De los Campos et al., 2013; Gianola, De Los Campos, Hill, Manfredi, & Fernando, 2009; Meuwissen et al., 2001; Pérez, de los Campos, Crossa, & Gianola, 2010) have shown some improvement on prediction accuracy of estimated breeding values for meat tenderness (Magnabosco et al., 2016; Stone et al., 2005). However, there is still a component of genetic variance that has not been captured yet by these methods (Legarra, Lourenço, & Vitezica, 2018). In this study, we investigate the potential of deep neural network (DNN) to capture non-linear relationships between markers in order to better predict the genetic merit and improve the prediction accuracy for meat tenderness. DNN is an artificial neural network (ANN) with multiple hidden layers. Traditionally, ANN has been implemented with a fully connected architecture with a single hidden layer. Innovations in the field of ANN allowed the construction of ANN with multiple hidden layers and different architectures, providing a powerful tool for pattern recognition. Such evolution in the field of ANN led to exploration of architectures with more than two hidden layers, emerging a new era for DNN applications. DNN is also related to different architectures such as convolution neural networks, often used for image analysis, and recurrent neural networks, frequently applied to time series data. DNN is capable to capture non-linear relationship between markers and phenotypes without using strong assumptions a priori. DNN has been shown to improve the predictive performance of several tasks in image analysis, speech recognition and computational biology (Angermueller, Pärnamaa, Parts, & Stegle, 2016). However, such potential approach has not been widely explored in the context of GS. Therefore, this study was carried out to compare the predictive performance of ANN with some Bayesian regression models (Bayesian Ridge Regression, Bayesian Lasso, Bayes A, Bayes B and Bayes Cπ) in estimating genomic merit for meat tenderness in Nellore cattle.

2 | MATERIALS AND METHODS

2.1 | Ethics statement

The collection of phenotypic information is not categorized as an experiment, since the interventions are related to farming practices, according to the law No 11.794 (8 October 2008; Brazilian Constitution), which lays down procedures for the scientific use of animals. Hence, this study was not submitted to an ethics committee, considering that a data set from a commercial production system was used.

2.2 | Study design

This research was conducted in a joint effort between the OB Ranch, Embrapa (Brazilian Agricultural Research

Corporation), the University of Wisconsin-Madison and the University of California-Davis, which aimed at the genetic characterization and GS of Nellore cattle for meat tenderness. A total of 22 sires representing the main Nellore bloodlines were selected and mated with 552 Polled Nellore cows. The mattings were performed based on the probability of identical genes by descent (Lacy, 1989) to create an experimental contrasting population where the animals in each group had extremes for meat tenderness. Thus, to identify animals with higher and lower genetic potential to produce tender meat, the mattings were based on the relationship structure among animals connected to individuals with genetic and phenotypic information for meat tenderness (Magnabosco et al., 2016). Progeny from these mating was raised on pasture, finished in feedlot for 3 months and harvested at 22 months of age. Warner–Bratzler shear force (WBSF) was determined on samples of the *Longissimus* muscle after ageing for 7 days at 4°C (Wheeler, Cundiff, Shackelford, & Koohmaraie, 2010).

The study was conducted over multiple years so that parents of animals from previous harvests could be evaluated based on their progeny performance. Therefore, the data set was composed of 113, 94, 142, 79, 4, 75, 63 and 5 animals born in 2002, 2003, 2006, 2010, 2011, 2013, 2014 and 2015, respectively. Exploratory analysis was performed to verify data consistency and to evaluate the significance of environmental effects. Hence, the fixed effects consisted of season and year of birth of the animals and sex. The birth season was defined as rainy (animals born between October and March) or dry (animals born between April and September) seasons. In addition, the harvest group was fitted as fixed effect.

2.3 | Estimation of genetic parameters

Variance–covariance components and heritability were estimated considering a pedigree-based animal model, $\mathbf{y} = \mathbf{X\beta} + \mathbf{Zu} + \mathbf{e}$, where $\mathbf{y}$ is the vector of observations (meat tenderness); $\boldsymbol{\beta}$ is a vector of fixed effects (season and year of birth of the animals, sex and harvest group); $\mathbf{X}$ is the incidence matrix associating $\boldsymbol{\beta}$ with $\mathbf{y}$; $\mathbf{u}$ is a vector of random direct additive genetic effects; $\mathbf{Z}$ is the incidence matrix associating $\mathbf{u}$ with $\mathbf{y}$; and $\mathbf{e}$ is a vector of random residual effects. It was assumed that $E[\mathbf{y}] = \mathbf{X\beta}$; and the direct additive genetic and residual effects were normally distributed with zero mean and $Var(\mathbf{u}) = \mathbf{A} \otimes \mathbf{S_a}$ and $Var(e) = \mathbf{I} \otimes \mathbf{S_e}$, in which $\mathbf{A}$ is the numerator relationship matrix, $\mathbf{S_a}$ is the additive genetic covariance matrix, $\mathbf{S_e}$ is the residual covariance matrix, $\mathbf{I}$ is an identity matrix, and $\otimes$ is the direct product. The analysis was performed using the restricted maximum likelihood method with the REMLF90 software (Misztal et al., 2018).

2.4 | SNP quality control (QC)

The animals were genotyped with either the Illumina Bovine HD Bead Chip (HD, 777K from 90 samples) or the GeneSeek Genomic Profiler (GGP Indicus HD, 77K from 485 samples). The quality control (QC) filters were performed separately for each chip. Markers were removed if they fell into the extremes of QC variables including Hardy–Weinberg proportions (HWP), missing proportion (call rate) and minor allele frequency (MAF). Extreme deviation from HWP is typically used to identify gross genotyping errors; a low call rate indicates poor genotype probe performance and low genotyping accuracy; and markers with low MAF are more prone to error, as fewer samples would be within a genotype cluster and most clustering-based calling algorithms do not perform well with rare alleles (Neale & Purcell, 2008). Thus, SNPs were excluded if they significantly deviated from HWP ($p < 1e\text{-}6$; Fisher exact test) or had MAF < 5% or call rate <90%. Markers were also excluded if they were located in non-autosomal regions or had the same genomic coordinates, that is mapped to the same positions (just the replicates were removed), and if they were in LD > 0.8. After these QC edits, 219,863 SNP markers remained for further analysis.

2.5 | Imputation

The FImpute program v. 2.2 (Sargolzaei, Chesnais, & Schenkel, 2014) was used for imputation from the GGP chip to the HD chip. This program uses deterministic methods to combine family and population information. The imputation is based on overlapping sliding windows and assumes that individuals are related to some degree. Overlapping of windows allows for consistency of haplotype phases across windows. As pedigree information was available, FImpute was run using both family and population-based algorithms, with its own default parameters. To assess the accuracy of imputation, extra analyses (results not shown) with over 900 animals genotyped in HD chip, including the animals used in this study, their parents and other related animals were performed. Different numbers of markers were masked to mimic the GGP chip and then imputed to the HD. The overall correlation between the true genotypes and their imputed genotypes was higher than ~0.98.

2.6 | Bayesian alphabet

Genomic prediction models were fitted using five Bayesian specifications: Bayesian ridge regression, Bayesian Lasso, Bayes A, Bayes B and Bayes Cπ. For these methods,

assuming there are p SNPs, the adjusted phenotypic value y_i of individual i can be then described by the following model:

$$y_i = \mu + \sum_{j=1}^{p} x_{ij}a_j + e_i$$

where μ is an overall constant; x_{ij} is a genotype indicator variable for individual i at locus j; a_j for $j = 1, 2, \ldots p$ is the genetic effect (additive or dominance) of the j-th SNP; and e_i is the residual associated to the observation on individual i. The vector of residuals **e** was assumed to be distributed as $\boldsymbol{e} \sim N(\mathbf{0}, I\sigma_e^2)$, where σ_e^2 is the residual variance.

The parameterization of the SNP matrix **X** was reformulated in terms of breeding values and dominance deviations, according to the marker genotypes at a locus j (Falconer & McKay, 1996). It is important to address to the fact that each effect was fitted in distinct model.

$$\text{Additive effect:}\boldsymbol{X} = \begin{cases} \text{if AA, then } 2-2p \to 2q \\ \text{if AB, then } 1-2p \to q-p \\ \text{if BB, then } 0-2p \to -2p \end{cases}$$

$$\text{Dominance effect:}\boldsymbol{X} = \begin{cases} \text{if AA, then } 0 \to -2q^2 \\ \text{if AB, then } 1 \to 2pq \\ \text{if BB, then } 0 \to -2p^2 \end{cases}$$

In the Bayesian ridge regression method, an independent Gaussian prior with common variance is assigned to each regression coefficient, so that $p(a_1, a_2, \ldots, a_p|\sigma_a^2) = \prod_{j=1}^{p} N(a_j|0, \sigma_a^2)$. The variance parameter σ_a^2 is treated as unknown, and a scaled inverse chi-squared is specified as prior density, $p\left(\sigma_a^2\right) = \chi^{-2}\left(\sigma_a^2|df_a, S_a\right)$, with degrees of freedom, df_a and scale parameter S_a Similarly, a scaled inverse chi-squared prior density is assumed also for the residual variance σ_e^2, that is $p(\sigma_e^2) = \chi^{-2}(\sigma_e^2|df_e, S_e)$.

In the Bayesian Lasso regression (Park & Casella, 2008), the conditional prior distribution of each marker effect $p(a_j|\tau_j^2, \sigma_e^2)$ follows a zero-mean Gaussian with marker-specific prior variance, independent from each other, so that $p(a_j|\tau_j^2, \sigma_e^2) = \prod_{j=1}^{p} N\left(a_j|0, \tau_j^2\sigma_e^2\right)$. This prior induces marker-specific shrinkage of effect estimates, whose extent depends on τ_j^{-2}. The variance parameters (τ_j^{-2}) are assigned exponential IID priors, $p(\tau_1^2, \tau_2^2, \ldots, \tau_K^2|\lambda) = \prod_{j=1}^{K} \text{Exp}(\tau_j^2|\lambda)$, and a *Gamma* distribution is assumed as the prior of the λ square regularization parameter, that is $p(\lambda^2) = \text{Gamma}(r, \theta)$, as suggested by Park and Casella (2008). Under these settings, it is shown that the marginal prior of each marker effect, $p(a_j|\lambda) = \int N(a_j|0, \sigma_e^2\tau_j^2)\,\text{Exp}(\tau_j^2|\lambda^2)\partial\tau_j^2$, is *Double Exponential*. Lastly, the residual variance (σ_e^2) is specified a scaled inverse chi-squared prior density, with degrees of freedom df_e, and scale parameter S_e.

The Bayes A method (Meuwissen et al., 2001; Pérez & de los Campos, 2014) assumes that the conditional prior distribution of a marker effect a_j is assumed to be Gaussian with null mean and marker-specific variance $\sigma_{a_j}^2$, independent from each other. The variance associated with the effect of each marker is assigned an IID scaled inverse chi-square prior distribution, $p(\sigma_{a_j}^2) = \chi^{-2}(\sigma_{a_j}^2|df, S^2)$, where df and S^2 are known degrees of freedom and scale parameters, respectively. With these specifications, the marginal prior distribution of each marker effect, $p(a_j|df, S^2) = \int N(a_j|0, \sigma_{a_j}^2)\chi^{-2}(\sigma_{a_j}^2|df, S^2)\partial\sigma_{a_j}^2$, is shown to be a t-distribution, that is $p(a_j|df, S^2) = t(0, df, S^2)$ (Rosa, Padovani, & Gianola, 2003).

In the Bayes B method, it is assumed that some genetic markers have zero effect and that only a few loci contribute with some genetic variance (Meuwissen et al., 2001). Conditional on the marker-specific variances $\sigma_{a_j}^2$, non-null marker effects are assumed Gaussian, $N(a_j|0, \sigma_{a_j}^2)$, such that the distribution of marker effects can be described with the following mixture model:

$$p(a_j|\sigma_{a_j}^2, \pi) = \begin{cases} 0 & \text{with probability } \pi \\ N(0, \sigma_{a_j}^2) & \text{with probability } (1-\pi) \end{cases}$$

where π is the proportion of markers with null genetic effects. As in the Bayes A, when $\pi = 0$, a scaled inverse chi-square prior distribution is assumed for the marker variances, that is $p(\sigma_{a_j}^2) = \chi^{-2}(\sigma_{a_j}^2|df, S^2)$, so that marginally, after integrating $\sigma_{a_j}^2$ out, the prior of marker effects takes the following form:

$$p(a_j|\pi) = \begin{cases} 0 & \text{with probability } \pi \\ t(0,\ df, S^2) & \text{with probability } (1-\pi) \end{cases}$$

The Bayes Cπ method (Habier, Fernando, Kizilkaya, & Garrick, 2011) is similar to the Bayes B approach, except that a prior distribution is specified for π, and a Gaussian prior with a common variance is assumed for each of the non-null marker effects. The inclusion (or exclusion) of each marker in the model is modelled by an indicator variable δ_j, which is equal to 1 if the marker j is fitted in the model and is zero otherwise. The common effect variance is sampled from a full-conditional posterior, which is a scaled inverse chi-square with degrees of freedom $\widetilde{df}_a = df_a + m^t$ and scale $\widetilde{S}_a^2 = \left(df_a S_a^2 + \sum_{j=1}^{p} a_j^2\right) / \widetilde{df}_a$, where m^t is the number of markers fitted with non-zero effects in iteration t.

2.7 | Deep learning

The deep learning process can be summarized as following: the information from the input layer is transformed by non-linear ways through a single (ANN) or multiple (DNN) hidden layers before the final output is computed at the last layer. The number of hidden layers defines the depth of the neural network, while the number of neurons in its layers defines its width. Training of a neural network relies on an iterative process performing forward and backward passes (i.e., epochs) in order to minimize some loss function (e.g., *MSE*), learn the weights and the biases of the inputs. In the forward pass, activation functions are employed to weights to retrieve the output at each layer and predict the output. The backward pass starts by calculating the derivatives of the error function between the predicted outputs and the real outputs. Then, the derivatives are propagated backwards updating the weights and computing new error terms for that layer. This process is repeated for each layer until the input layer is reached (Goodfellow, Bengio, & Courville, 2016; Waldmann, 2018).

The learning process was performed in three steps: DNN was employed using different number of neurons (i.e., 5, 35, 70, 105, 150 neurons). The number of neurons that gave the highest prediction accuracy in the previous step was then used to feed the first hidden layer followed by a reduction in the number of neurons in the second hidden layer; the third step was similar to the second one, where the number of neurons in the first and second hidden layers was kept, while decreasing and increasing them in the third hidden layer. However, to decide the number of neurons, hidden layers and activation function, a random discrete strategy was employed to perform a random search of all the combinations of the hyperparameters, where the stop metric was the *MSE* and the maximum number of epochs was 10,000. Rectifier and Maxout (both with and without dropout) activation function were tested (results not shown), using up to four hidden layers (HL). Empirical research has been employed to determine the optimal number of neurons (Masters, 1993; Moradi & Hariri-Ardebili, 2019; Ozturan, Kutlu, & Ozturan, 2008). In our study, it was uses 5% of the number of SNP as input layer (*nI*, # of input), while the number of neurons in the output layer was equal to a single node (*n0*). We limited *nI* to 5% because of the limited number of observations; more than that caused convergence problems with our data. A quadratic loss function was used in the list of hyperparameters. Hence, the number of neurons in each hidden layer was set as following:

1 hidden layer: $\mathrm{HL} = (nI/n0)^{1/2}$;
2 hidden layers: $\mathrm{HL} = [n0 \times r^2; n0 \times r]$, where $r = (nI/n0)^{1/3}$;
3 hidden layers: $\mathrm{HL} = [n0 \times r^3;\ n0 \times r^2; n0 \times r]$, where $r = (nI/n0)^{1/4}$;
4 hidden layers: $\mathrm{HL} = [n0 \times r^4;\ n0 \times r^3; n0 \times r^2; n0 \times r]$, where $r = (nI/n0)^{1/5}$.

The activation function used to train DNN was the rectified linear unit (ReLU), not only because it is faster to learn then sigmoid and hyperbolic tangent functions (Glorot, Bordes, & Bengio, 2011), but also because of its better performance during the random grid search. The back-propagation step consists of updating network weights. During the learning process, predicted and observed values of meat tenderness were compared to compute a loss, which was backward propagated through the network to compute gradients and update weights. The loss function is typically optimized by using gradient-based descent which involves derivatives to find the direction of the gradient. DNNs are likely to quickly overfit the training populations. To help in overcoming such issue, a model can be used through simulations with different network architectures by randomly dropping out nodes during training in parallel. Such technique is a regularization approach called dropout (Srivastava, Hinton, Krizhevsky, Sutskever, & Salakhutdinov, 2014), which is very useful for error modelling generalization. The ADADELTA was the adaptive learning rate algorithm used in all ANN (Zeiler, 2012). This learning rate automatically combines the benefits of learning rate annealing and momentum training to avoid slow convergence. For reproducibility and fine-tuning of model parameters, these analyses were performed on a single node. Overall, the activation function used was Rectifier linear units with dropout, mean square error as loss function and 10,000 epochs. Overall, the definition of the hyperparameter that resulted in the best architecture of the ANN was based on different number of units (up to three layers) considering the optimum number of inputs in each layer, different activation function and regularization methods.

2.8 | Cross-validation approach

To implement cross-validation for accuracy of GEBV, the observed phenotypic values for all animals were split into training and validation data sets. In practical livestock applications, training occurs on pre-existing animals, whereas the validation population for implementation of GS are usually performed on young animals (selection candidates). Hence, the animals were grouped based on their year of birth: animals born between 2002 and 2013 were used as training population ($n = 507$), while animals born in 2014 and 2015 were used as validation population (68). In addition, most

of the animals in the validation population were progeny of new sires, not used in the reference population. The prediction accuracy was assessed by the correlation between the genomic estimated breeding values and adjusted meat tenderness.

The meat tenderness was adjusted using a linear model considering the fixed effects of the contemporary group and the harvest group:

$$y_{ij} = \mu + \mathrm{CG}_i + \mathrm{HG}_j + e_{ii}$$

where y_{ij} is the observed value for meat tenderness; μ is an overall intercept; CG and HG are fixed effects of the contemporary group and harvest group, respectively; and e_{ij} is a residual term.

2.9 | Computer resources

The BGLR package (Pérez & de los Campos, 2014) was used to implement Bayesian Lasso, Bayesian ridge regression, Bayes A, Bayes B and Bayes Cπ methods. The BGLR defaults priors were used for all models, where $df_0 = 5$ and $S_0 = 3.76$ for all models, and π (*pronIn*) equal to 0.23 and 0.47 for Bayes B and Bayes Cπ, respectively. The Markov Chain Monte Carlo sampling length was 100,000 iterations, with the first 20,000 iterations excluded as burn-in and a thinning interval of 20 cycles. Convergence was checked by visual inspection of trace plots of the residual variance. ANN analyses were performed using the open-source software for big data analysis called H2O (https://CRAN.R-project.org/package=h2o), which can be used through the statistical R environment (2018).

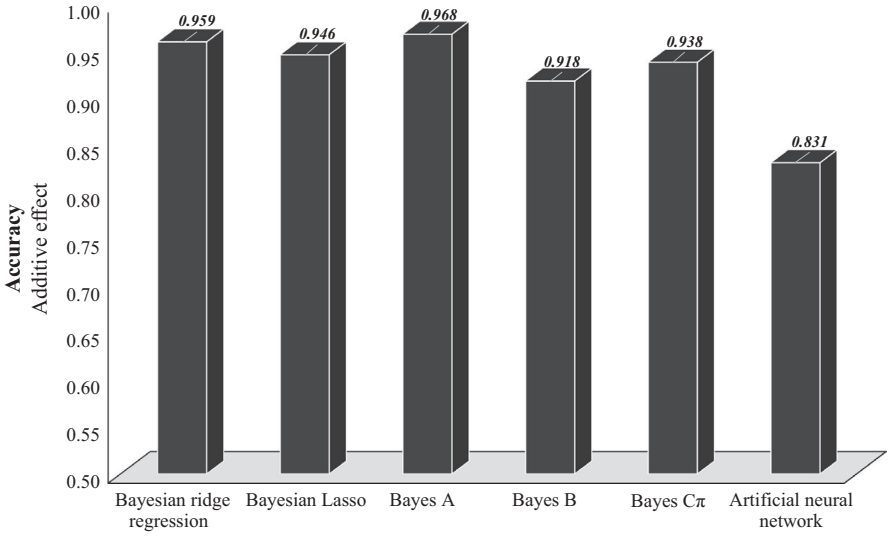

FIGURE 1 Model fitting accuracy estimated as the correlation between meat tenderness and its estimated breeding value in the training population using the additive SNP matrix

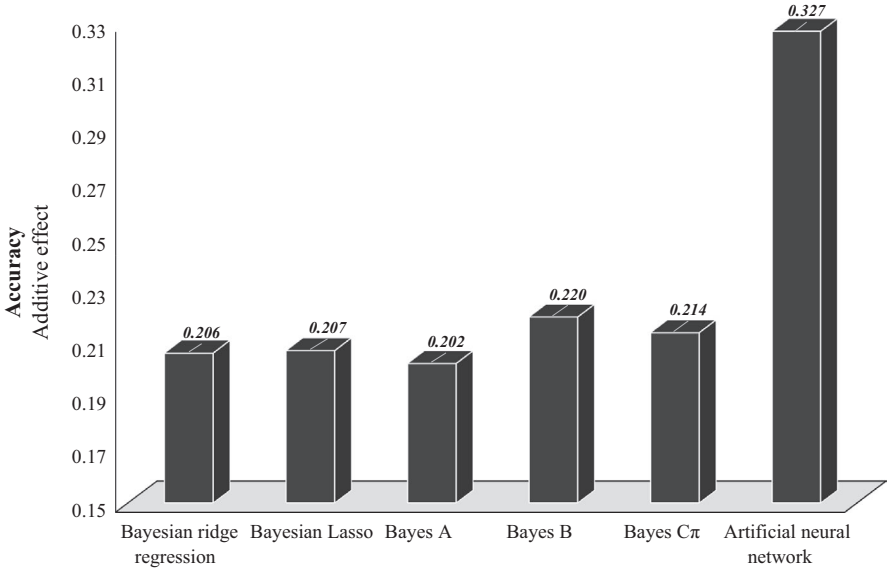

FIGURE 2 Estimates of correlation between meat tenderness and genomic breeding values predicted by Bayesian ridge regression, Bayesian Lasso, Bayes A, Bayes B, Bayes Cπ and artificial neural network models using the additive SNP matrix

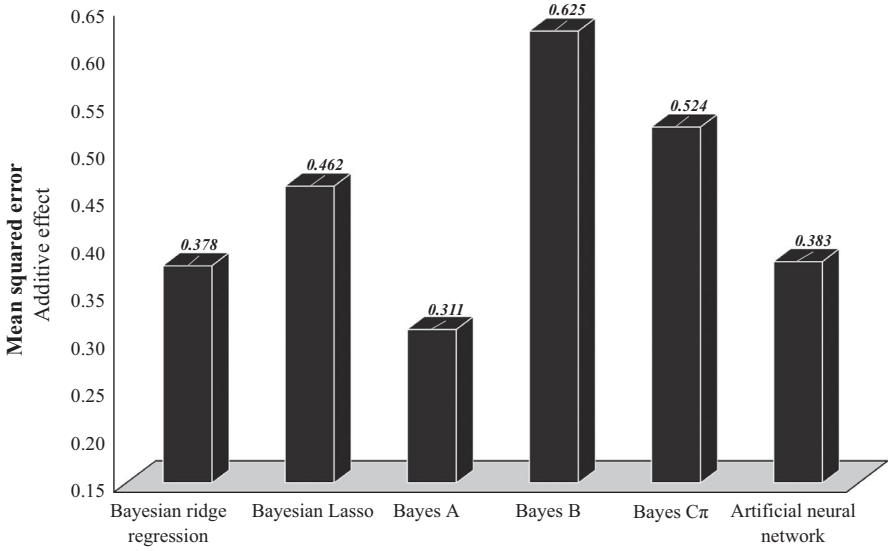

FIGURE 3 Model fitting performance computed as the mean squared error of prediction in the training population using the additive SNP matrix

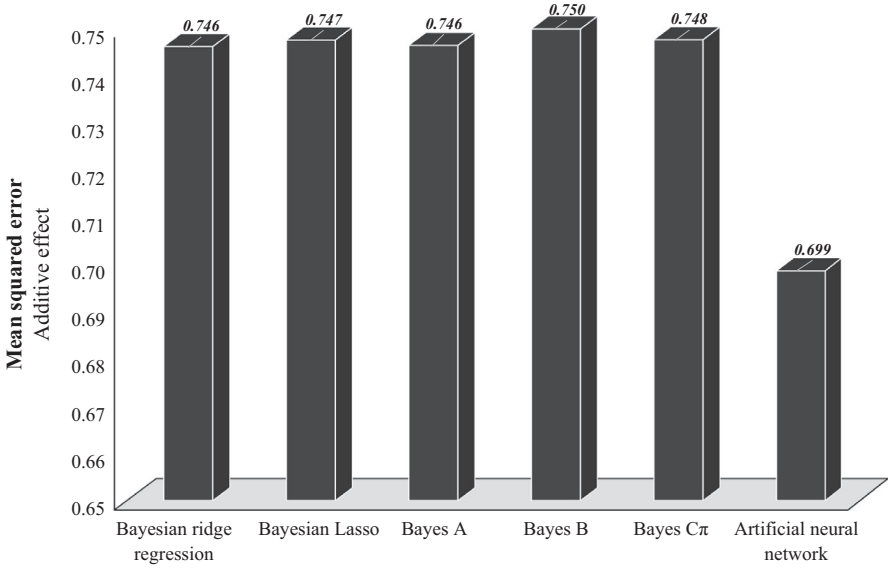

FIGURE 4 Mean squared error of prediction of meat tenderness estimated by Bayesian ridge regression, Bayesian Lasso, Bayes A, Bayes B, Bayes Cπ and artificial neural network models using the additive SNP matrix

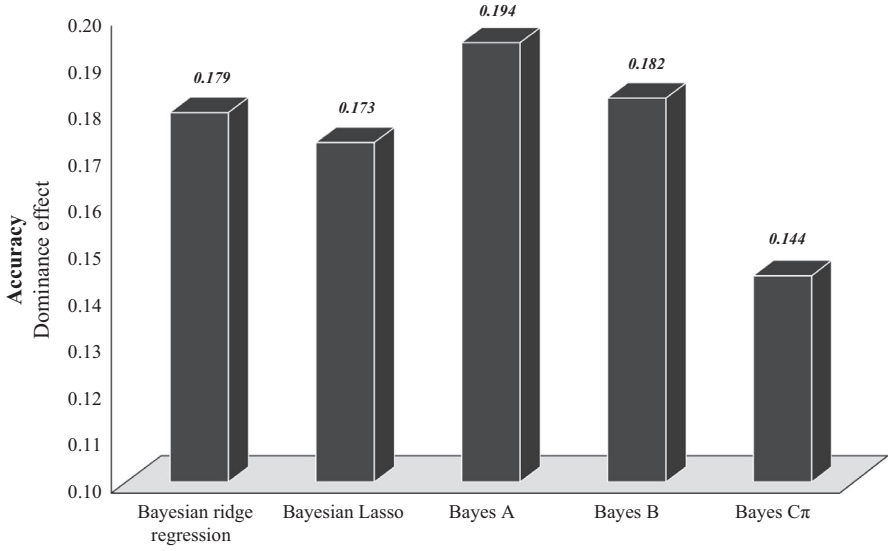

FIGURE 5 Estimates of correlation between meat tenderness and genomic breeding values predicted by Bayesian ridge regression, Bayesian Lasso, Bayes A, Bayes B and Bayes Cπ using dominance SNP matrix

3 | RESULTS

Estimate of additive heritability for meat tenderness was 0.35 (±0.047), indicating that this trait would respond to selection. Although meat tenderness has not been used as a breeding goal, mainly because it is difficult and expensive to measure, our results showed that its additive direct genetic variability is high, allowing selection of best animals as parents of future generations.

Model prediction accuracies, computed as the correlation between meat tenderness and the predicted genomic breeding value in the training population, were higher for the Bayesian alphabet models, ranging from 0.92(±0.049) for the Bayes B to 0.97(±0.031) for Bayes A, and lower for the ANN (0.83 ± 0.069) (Figure 1). As discussed, the models that presented the highest additive accuracy of prediction in the training population showed the lowest additive prediction accuracy in the validation population. The highest additive accuracy in the training set was achieved by the Bayes A model (0.97 ± 0.031), and this model showed the lowest predictive accuracy (0.20 ± 0.120) in the validation set. Overall, correlations between meat tenderness and genomic breeding values due to additive effect for Bayesian ridge regression, Bayesian LASSO, Bayes A, Bayes B and Bayes Cπ models were 0.21 ± 0.120, 0.21 ± 0.120, 0.20 ± 0.121, 0.22 ± 0.120 and 0.21 ± 0.120, respectively (Figure 2).

The mean squared error (*MSE*) was employed as a measure of the overall fit achieved with each model. In the training population, the additive *MSEs* varied considerably (Figure 3) and were lower than the validation population, which were very similar among the Bayesian alphabet models, and smaller for the ANN (Figure 4). Estimates of correlation between meat tenderness and genomic breeding values due to dominance effect for Bayesian ridge regression, Bayesian LASSO, Bayes A, Bayes B and Bayes Cπ models were 0.18 ± 0.122, 0.17 ± 0.121, 0.19 ± 0.121, 0.18 ± 0.120 and 0.14 ± 0.122, respectively (Figure 5). Although the additive prediction accuracies were very close across the Bayesian regression models, for the dominance effects, such predictions have varied and the Bayes A model, which uses Student t-distribution for the marker effects, was the best model for recovering the dominance variance.

4 | DISCUSSION

Although in the current study a relatively small number of animals was used, the design of a highly segregating population (Magnabosco et al., 2016) was fundamental to ensure increased genetic variability of meat tenderness, which contributed to achieve such good prediction accuracy. To identify animals with higher and lower genetic potential to produce tender meat, sires and dams were grouped according to their genetic similarity in a way that one group presented progeny with high WBSF and the other group low WBSF. Due to the limited sample size, estimates of EBV showed very low accuracy overall. Sires and dams were grouped based on their relatedness followed by their progeny performance for meat tenderness. Hence, animals within groups were genetically similar and somewhat different between groups. In addition, on average, sires and dams were only selected and mated if they both have had progeny with either tender or tough meat.

Some simulation studies have suggested that variable selection methods (e.g., Bayes B and Bayes Cπ) could outperform other methods such as Bayesian ridge regression. Nonetheless, the performance of Bayesian ridge regression has often been comparable to those of variable selection methods when analysing real data (Clark, 2011; Habier, Fernando, & Dekkers, 2007; Meuwissen et al., 2001). A possible explanation could be related to the distribution of QTL effects, because most of economically important traits are generally affected by several QTL with small effects. It has also been reported that methods with conceptual differences reached very similar predictive abilities, as well as methods with different assumptions have presented similar performance (Hayes, Bowman, Chamberlain, Verbyla, & Goddard, 2009; Heslot, Yang, Sorrells, & Jannink, 2012; Resende, Silva, Lopes, & Azevedo, 2012).

The correlation between meat tenderness and the predicted genomic breeding (model fitting accuracy – "training population") obtained by ANN was lower than the Bayesian models. Typically, higher values (close to 1) might indicate overfitting in the training set, when it occurs it is expected a lower prediction accuracy in the validation population, concomitant with a higher accuracy in the training population. The overall fit of the models to the data, assessed by the *MSE*, has favoured the ANN method over the Bayesian regression methods. The lower *MSE* verified with the use of ANN can be a attributed to the use of non-linear models to predict the outcome (Dai, Huo, & Wang, 2011). ANN is a potential technique to be employed in animals' genetic evaluation. It becomes even more important when environmental effects may mask the genomic merit of selected candidates, leading to a wrong selection. To get around this, ANN can decrease these environmental effects by backpropagation of the error, which is a technique with great potential for prediction of genetic value (Peixoto, Bhering, & Cruz, 2015). Thus, ANN may assist in selecting the best animals based on their genomic prediction, while reducing the mean squared error of prediction and trying to maximize the prediction accuracy in the validation population.

A few studies have considered the use of ANN methods for prediction of complex traits, and it has been suggested that such approach can achieve good prediction accuracy (Gianola, Okut, Weigel, & Rosa, 2011; González-camacho, Crossa, Pérez-rodríguez, Ornella, & Gianola, 2016;

González-Camacho et al., 2012; Okut, Gianola, Rosa, & Weigel, 2011; Perez-Rodriguez et al., 2012). However, there has not been consistent evidence indicating that they can outperform linear models in terms of predictive ability. A reasonable explanation for such dubious performance is that most of the available studies have been based on small population size and limited numbers of SNP markers. Remarkably, in our study the predictive ability of the ANN was higher than the Bayesian regression models. The use of the rectified linear unit activation function (Glorot et al., 2011) in combination with the regularization approach applying a dropout (Srivastava et al., 2014) of 50% of the SNP markers might has helped in the generalization of the ANN, thus improving its predictive ability.

As pointed out by Azevedo et al. (2015), Gaussian, Student and Double Exponential distributions are likely good approximations of the true distributions of genetic effects. Hence, it is more important to test these distributions when predicting genomic breeding values, because it might reveal which prior distribution is more adequate and/or robust. Remarkably, ANN achieved the highest prediction accuracy (0.33). Summing up the additive and dominance accuracies, the predictive ability of Bayesian regression models ranged from 0.36 (Bayes Cπ) to 0.40 (Bayes B). Therefore, this might assist in explaining the superiority of ANN over the Bayesian regression models. The ANN ability of capturing non-linear relationships might have made it possible to capture some of the genetic variance that is due to additive and non-additive genetic effects. It is important to address the fact that in real data, it is difficult to precisely differentiate additive from non-additive genetic effects, as their true effects are not known. However, the results showed that the Bayesian linear models and ANN capture 35% and 56% of the additive genetic variability, respectively. Such additive genetic variability might not be all due to linear, but also non-linear effects. Hence, ANN might be capturing additive genetic effects, as well as some non-additive genetic effects. In addition, due to the DNN/ANN ability to capture non-linear associations, it might be useful for phenotypic prediction.

Deep neural networks are widely recognized to delivery more accurate predictions in different application. In our study, however, an ANN with a single hidden layer and 105 neurons was enough to increase the predictive ability for meat tenderness, which might be explained by the fact that ANN can capture non-linear relationships such as non-additive genetic effects. The use of DNN/ANN needs to be carefully evaluated, mainly regarding the non-additive effects that may be captured. In cases where non-additive effects assist in improving the predictive ability of a model, breeding programmes that use breeding values (additive genetic values) may not benefit from the use of such model models, because they are capturing also interactions between genes, which might comprise non-heritable genetic variation that changes due to recombination and other factors. In addition, the use of non-additive effects can potentially increase the power of GS in cross-bred populations. Such effects, in addition to additive effects, can increase the predictive ability of economic important and complex traits. Esfandyari, Sørensen, and Bijma (2015) have also addressed the fact that non-additive effects, such as dominance and epistasis effects, might also impact the additive genetic effect.

5 | CONCLUSION

The ANN has shown the highest prediction accuracy among all GS models evaluated. Although additive prediction accuracies were very close among the Bayesian regression models, for the dominance effects, such predictions have varied, and the Bayes A was the best model for recovering the dominance variance. Overall, these results indicate that an ANN with relatively simple architecture was more efficient in predicting the genomic breeding values for meat tenderness in Nellore cattle.

ACKNOWLEDGEMENTS

The authors thank to São Paulo Research Foundation for providing postdoctoral financial support for the first, grant #2017/03221-9. The authors also thank ANCP (Brazilian Society of Breeders and Researchers) for providing data set and genealogical data.

CONFLICT OF INTEREST

The authors declare that they do not have any conflict of interest.

ORCID

Fernando Brito Lopes https://orcid.org/0000-0003-1292-3331
Ludmilla C. Brunes https://orcid.org/0000-0001-9012-520X

REFERENCES

Angermueller, C., Pärnamaa, T., Parts, L., & Stegle, O. (2016). Deep learning for computational biology. *Molecular Systems Biology*, *12*(878), 1–16. https://doi.org/10.15252/msb.20156651

Azevedo, C. F., Deon, M., Resende, V. D., Fonseca, F., Marcelo, J., Viana, S., … Muñoz, P. (2015). Ridge, Lasso and Bayesian additive-dominance genomic models. *BMC Genetics*, *16*(105), 1–13. https://doi.org/10.1186/s12863-015-0264-2

Baldi, F., de Albuquerque, L. G., Cyrillo, J. N. S. G. D. S. G., Branco, R. H., Junior, B. C. D. O., Mercadante, M. E. Z., … Mercadante, M. E. Z. (2012). Genetic parameter estimates for live weight and daily live weight gain obtained for Nellore bulls in a test station using different models. *Livestock Science*, *144*(1–2), 148–156. https://doi.org/10.1016/j.livsci.2011.11.009

Bonilha, S. F. M., Tedeschi, L. O., Packer, I. U., Razook, A. G., Alleoni, G. F., Nardon, R. F., & Resende, F. D. (2008). Evaluation of carcass characteristics of *Bos indicus* and tropically adapted *Bos taurus* breeds selected for postweaning weight. *Journal of Animal Science*, *86*(8), 1770–1780. https://doi.org/10.2527/jas.2007-0507

Clark, T. E. (2011). Real-time density forecasts from Bayesian Vector Autoregressions with Stochastic Volatility. *Journal of Business and Economic Statistics*, *29*(3), 327–341. https://doi.org/10.1198/jbes.2010.09248

Dai, X., Huo, Z., & Wang, H. (2011). Simulation for response of crop yield to soil moisture and salinity with artificial neural network. *Field Crops Research*, *121*(3), 441–449. https://doi.org/10.1016/j.fcr.2011.01.016

De Los Campos, G., Hickey, J. M. M., Pong-Wong, R., Daetwyler, H. D. D., & Calus, M. P. L.. (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

Esfandyari, H., Sørensen, A. C., & Bijma, P. (2015). A crossbred reference population can improve the response to genomic selection for crossbred performance. *Genetics Selection Evolution*, *47*, 76. https://doi.org/10.1186/s12711-015-0155-z

Falconer, D. S., & McKay, T. (1996). *Introduction to quantitative genetics*. Harlow, UK: Earson Education Limited.

Gianola, D., De Los Campos, G., Hill, W. G., Manfredi, E., & Fernando, R. (2009). Additive genetic variability and the Bayesian alphabet. *Genetics*, *183*(1), 347–363. https://doi.org/10.1534/genetics.109.103952

Gianola, D., Okut, H., Weigel, K. A., & Rosa, G. J. M. (2011). Redicting complex quantitative traits with Bayesian neural networks: A case study with Jersey cows and wheat. *BMC Genetics*, *12*(87).

Glorot, X., Bordes, A., & Bengio, Y. (2011). Deep sparse rectifier neural networks. *International Conference on Artificial Intelligence and Statistics*, 315–323.

González-camacho, J. M., Crossa, J., Pérez-rodríguez, P., Ornella, L., & Gianola, D. (2016). Genome-enabled prediction using probabilistic neural network classifiers. *BMC Genomics*, *17*(208), 1–16. https://doi.org/10.1186/s12864-016-2553-1

González-Camacho, J. M., de Los Campos, G., Pérez, P., Gianola, D., Cairns, J. E., Mahuku, G., … Crossa, J. (2012). Genome-enabled prediction of genetic values using radial basis function. *Theoretical and Applied Genetics*, *125*(4), 759–771.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. Cambridge, UK: MIT Press.

Habier, D., Fernando, R. L., & Dekkers, J. C. M. (2007). The impact of genetic relationship information on genome-assisted breeding values. *Genetics*, *117*(4), 2389–2397. https://doi.org/10.1534/genetics.107.081190

Habier, D., Fernando, R. L. R., Kizilkaya, K., & Garrick, D. J. D. (2011). Extension of the Bayesian alphabet for genomic selection. *BMC Bioinformatics*, *12*(1), 186. https://doi.org/10.1186/1471-2105-12-186

Hayes, B. J., Bowman, P. J., Chamberlain, A. C., Verbyla, K., & Goddard, M. E. (2009). Accuracy of genomic breeding values in multi-breed dairy cattle populations. *Genetics Selection Evolution*, *41*(51), 1–9. https://doi.org/10.1186/1297-9686-41-51

Heslot, N., Yang, H. P., Sorrells, M. E., & Jannink, J. L. (2012). Genomic selection in plant breeding: A comparison of models. *Crop Science*, *52*, 146–160. https://doi.org/10.2135/cropsci2011.06.0297

Jeremiah, L. E. (1982). Consumer preferences regarding cooking methods and times for beef loin steaks in central Alberta. *Journal of Consumer Studies and Home Economics*, *6*(1), 79–86. https://doi.org/10.1111/j.1470-6431.1982.tb00587.x

Koohmaraie, M., Kent, M. P., Shackelford, S. D., Veiseth, E., & Wheeler, T. L. (2002). Meat tenderness and muscle growth: Is there any relationship? *Meat Science*, *62*(3), 345–352. https://doi.org/10.1016/S0309-1740(02)00127-4

Lacy, R. C. (1989). Analysis of founder representation in pedigrees: Founder equivalents and founder genome equivalents. *Zoo Biology*, *8*(2), 111–123. https://doi.org/10.1002/zoo.1430080203

Legarra, A., Lourenço, D. A. L., & Vitezica, Z. G. (2018). *Bases for Genomic Prediction*. Paris, França: INRA, Animal Genetics Department.

Magnabosco, C., Lopes, F., Fragoso, R., Eifert, E., Valente, B., Rosa, G., & Sainz, R. (2016). Accuracy of genomic breeding values for meat tenderness in Polled Nellore cattle. *Journal of Animal Science*, *94*, 2752–2760. https://doi.org/10.2527/jas.2016-0279

Masters, T. (1993). *Practical neural network recipes in C++*. San Diego, CA: Academic Press Professional Inc.

Meuwissen, T. H. E., Hayes, B. J., & Goddard, M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, *157*(4), 1819–1829. 11290733

Misztal, I., Tsuruta, S., Lourenco, D. A. L., Masuda, Y., Aguilar, I., Legarra, A., & Vitezica, Z. (2018). Manual for BLUPF90 family programs. University of Georgia. http://nce.ads.uga.edu/wiki/doku.php?id=documentation

Moradi, M. J., & Hariri-Ardebili, M. A. (2019). Developing a library of shear walls database and the neural network based predictive meta-model. *Applied Science*, *9*, 2562.

Moser, G., Tier, B., Crump, R. E., Khatkar, M. S., & Raadsma, H. W. (2009). A comparison of five methods to predict genomic breeding values of dairy bulls from genome-wide SNP markers. *Genetic Selection Evolution*, *41*(56), 1–16. https://doi.org/10.1186/1297-9686-41-56

Neale, B. M., & Purcell, S. (2008). The positives, protocols, and perils of genome-wide association. *American Journal of Medical Genetics, Part B: Neuropsychiatric Genetics*, *147*(7), 1288–1294. https://doi.org/10.1002/ajmg.b.30747

Okut, H., Gianola, D., Rosa, G. J. M., & Weigel, K. A. (2011). Prediction of body mass index in mice using dense molecular markers and a regularized neural network. *Genetic Research*, *93*, 189–201. https://doi.org/10.1017/S0016672310000662

Ozturan, M., Kutlu, B., & Ozturan, T. (2008). Comparison of concrete strength prediction techniques with artificial neural network approach. *Building Research Journal*, *56*, 23–36.

Park, T., & Casella, G. (2008). The Bayesian Lasso. *Journal of the American Statistical Association*, *103*, 681–686. https://doi.org/10.1198/016214508000000337

Peixoto, L. A., Bhering, L. L., & Cruz, C. D. (2015). Artificial neural networks reveal efficiency in genetic value prediction. *Genetics and Molecular Research*, *14*(2), 6796–6807. https://doi.org/10.4238/2015.June.18.22

Pereira, A. S. C., Baldi, F., Sainz, R. D., Utembergue, B. L., Chiaia, H. L. J., Magnabosco, C. U., … Sobral, P. J. A. (2015). Growth performance, and carcass and meat quality traits in progeny of Poll Nellore, Angus and Brahman sires under tropical conditions. *Animal Production Science*, *55*(10), 1295–1302. https://doi.org/10.1071/AN13505

Pérez, P., & de los Campos, G. (2014). Genome-wide regression and prediction with the BGLR statistical package. *Genetics*, *198*(2), 483–495. https://doi.org/10.1534/genetics.114.164442

Pérez, P., de los Campos, G., Crossa, J., & Gianola, D. (2010). Genomic-enabled prediction based on molecular markers and pedigree using the Bayesian linear regression package in R. *The Plant Genome Journal*, *3*(2), 106. https://doi.org/10.3835/plantgenome2010.04.0005

Pérez-Rodríguez, P., Gianola, D., González-Camacho, J. M., Crossa, J., Manès, Y., & Dreisigacker, S. (2012). Comparison between linear and non-parametric regression models for genome-enabled prediction in wheat. *Genes Genomes Genetics*, *2*(12), 1595–1605. https://doi.org/10.1534/g3.112.003665

R Core Team (2018). *R: A Language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from: https://www.R-project.org

Resende, M. D. V., Silva, F. F., Lopes, P. S. P., & Azevedo, C. F. (2012). *Seleção Genômica Ampla (GWS) via Modelos Mistos (REML/BLUP), Inferência Bayesiana (MCMC), Regressão Aleatória Multivariada (RRM) e Estatística Espacial*. Viçosa Brazil: Universidade Federal de Viçosa/Departamento de Estatistica.

Reverter, A., Johnston, D. J., Ferguson, D. M., Perry, D., Goddard, M. E., Burrow, H. M., … Bindon, B. M. (2003). Genetic and phenotypic characterisation of animal, carcass, and meat quality traits from temperate and tropically adapted beef breeds. 4. Correlations among animal, carcass, and meat quality traits. *Australian Journal of Agricultural Research*, *54*(2), 149–158. https://doi.org/10.1071/AR02088

Rosa, G. J. M., Padovani, C. R., & Gianola, D. (2003). Robust linear mixed models with normal/independent distributions and Bayesian MCMC implementation. *Biometrical Journal*, *45*(5), 573–590. https://doi.org/10.1002/bimj.200390034

Sargolzaei, M., Chesnais, J., & Schenkel, F. S. (2014). *FImpute User's guide* (p. 17).

Smith, T., Domingue, J. D., Paschal, J. C., Franke, D. E., Bidner, T. D., & Whipple, G. (2007). Genetic parameters for growth and carcass traits of Brahman steers. *Journal of Animal Science*, *85*(6), 1377–1384. https://doi.org/10.2527/jas.2006-653

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, *15*, 1929–1958.

Stone, R. T., Casas, E., Smith, T. P. L., Keele, J. W., Harhay, G., Bennett, G. L., Snelling, W. M. (2005). Identification of genetic markers for fat deposition and meat tenderness on bovine chromosome 5: Development of a low-density single nucleotide polymorphism map. *Journal of Animal Science*, *83*(10), 2280–2288. https://doi.org/10.2527/2005.83102280x

Su, G., Christensen, O. F., Janss, L., & Lund, M. S. (2014). Comparison of genomic predictions using genomic relationship matrices built with different weighting factors to account for locus-specific variances. *Journal of Dairy Science*, *97*(10), 6547–6559. https://doi.org/10.3168/jds.2014-8210

Thrift, F. A., & Thrift, T. A. (2002). The Issue of carcass tenderness expressed by cattle varying in *Bos indicus* inheritance. *The Professional Animal Scientist*, *18*(3), 193–201. https://doi.org/10.15232/S1080-7446(15)31522-9

Waldmann, P. (2018). Approximate Bayesian neural networks in genomic prediction. *Genetics Selection Evolution*, *50*, 70. https://doi.org/10.1186/s12711-018-0439-1

Wheeler, T. L., Cundiff, L. V., Shackelford, S. D., & Koohmaraie, M. (2010). Characterization of biological types of cattle (Cycle VIII): Carcass, yield, and longissimus palatability traits. *Journal of Animal Science*, *88*(9), 3070–3083. https://doi.org/10.2527/jas.2009-2497

Zeiler, M. D. (2012). *Adadelta: an adaptive learning rate method* (pp. 1–6). Retrieved from https://arxiv.org/pdf/1212.5701.pdf

How to cite this article: Brito Lopes F, Magnabosco CU, Passafaro TL, et al. Improving genomic prediction accuracy for meat tenderness in Nellore cattle using artificial neural networks. *J Anim Breed Genet*. 2020;00:1–11. https://doi.org/10.1111/jbg.12468