



Alternative measures to evaluate the accuracy and bias of genomic predictions with censored records

Geraldo Magela da Cruz Pereira^{1*}, Sebastião Martins Filho², Renata Veroneze³, Luiz Fernando Brito⁴, Vinícius Silva dos Santos⁵ and Leonardo Siqueira Glória⁶

¹Departamento de Estatística, Universidade Federal de Lavras, Trevo Rotatório Professor Edmir Sá Santos, s/n, Cx. Postal 3037, 37203-202, Lavras, Minas Gerais, Brazil. ²Departamento de Estatística, Universidade Federal de Viçosa, Viçosa, Minas Gerais, Brazil. ³Departamento de Zootecnia, Universidade Federal de Viçosa, Viçosa, Minas Gerais, Brazil. ⁴Department of Animal Sciences, Purdue University, West Lafayette, Indiana, USA. ⁵Departamento de Estatística, Universidade Federal do Acre, Rio Branco, Acre, Brazil. ⁶Departamento de Zootecnia, Universidade Estadual do Norte Fluminense, Campos dos Goytacazes, Rio de Janeiro, Brazil. *Author for correspondence. E-mail: geraldomcpereira@gmail.com

ABSTRACT. This study aimed to propose and compare metrics of accuracy and bias of genomic prediction of breeding values for traits with censored data. Genotypic and censored-phenotypic information were simulated for four traits with QTL heritability and polygenic heritability, respectively: C1: 0.07-0.07, C2: 0.07-0.00, C3: 0.27-0.27, and C4: 0.27-0.00. Genomic breeding values were predicted using the Mixed Cox and Truncated Normal models. The accuracy of the models was estimated based on the Pearson (PC), maximal (MC), and Pearson correlation for censored data (PCC) while the genomic bias was calculated via simple linear regression (SLR) and Tobit (TB). MC and PCC were statistically superior to PC for the trait C3 with 10 and 40% censored information, for 70% censorship, PCC yielded better results than MC and PC. For the other traits, the proposed measures were superior or statistically equal to the PC. The coefficients associated with the marginal effects (TB) presented estimates close to those obtained for the SLR method, while the coefficient related to the latent variable showed almost unchanged pattern with the increase in censorship in most cases. From a statistical point of view, the use of methodologies for censored data should be prioritized, even for low censoring percentages.

Keywords: genomic selection; statistical modeling; simulation; mixed Cox model; truncated normal model.

Received on November 11, 2021.

Accepted on June 1, 2022.

Introduction

With the availability of high-density genomic marker panels, genomic selection (Meuwissen, Hayes, & Goddard, 2001) has become a powerful tool in genetic improvement programs, especially for the prediction of breeding values for complex traits. Many methods have been proposed to enable successful implementation of genomic selection. In most of these methods, the marker information is incorporated into prediction models, in which the phenotypes are regressed considering a set of markers as explanatory variables in a regression model (Campos, Hickey, Pong-Wong, Daetwyler, & Calus, 2013).

Traits that have censored records (e.g., age at first calving and slaughter, longevity) are frequently evaluated in animal breeding programs (Santos et al., 2015; Costa et al., 2019; Oliveira, Miller, Brito, & Schenkel, 2021). In the presence of censoring, the value for a given trait is greater or less than a certain threshold, or belongs to an interval, and it is not possible to observe an exact value for response. Santos et al. (2015) evaluated the trait time-to-slaughter in pigs, in which the censoring was characterized by the presence of animals that did not obtain the minimum weight necessary to be slaughtered (censored observations). Another example of censoring in animal production is the selection of genetically superior individuals for disease resistance, which would result in a longer survival time for the animals (Alemu et al., 2016; Palaiokostas, Ferraresso, Franch, Houston, & Bargelloni, 2016; Vallejo et al., 2016). In these studies, a common approach is to challenge individuals during the test period and the individuals that do not die during the period of the experiment result in censored data.

Various methods for performing genomic prediction of traits with censored data have been proposed over the past years. For instance, Kärkkäinen and Sillanpää (2013) proposed a hierarchical Bayesian model (Bayesian threshold model) for analyses of data with binary, ordinal scale, or with censored records. Pérez

and Campos (2014) proposed a class of models, called Bayesian Generalized Linear Regression (BGLR), which also allow to model normal, binary, ordinal, or censored responses. Furthermore, Santos et al. (2015) were the pioneers in using Mixed Cox model for genomic prediction of traits with censored phenotypes. Comparisons between several models in genomic selection are commonly performed using accuracy and prediction bias (e.g., Teissier et al., 2020; Araujo et al., 2022; Massender et al., 2022). Genomic accuracy can be defined as the Pearson correlation (PC) between the true breeding values (TBV) and the genomic enhanced breeding values (GEBV). Genomic bias can be calculated by adjusting a simple linear regression model between TBV and GEBV. However, these measures may not be adequate when the phenotypes are partially observed, that is, censored observations are present. According to Li, Gillespie, Shedden, and Gillespie (2018), when traditional estimators for complete data are applied to censored data, they reduce the accuracy of the estimates and can introduce bias. According to the authors, approaches that incorporate incomplete observations are generally more statistically powerful. Thus, the use of the PC to assess the predictive ability in the framework of censored data may result in misleading conclusions.

There is a need for evaluating alternative measures to PC and linear regression for the assessment of genomic accuracy and bias, respectively, in the context of censored data. In this context, the main objectives of this study were to 1) propose and compare alternative metrics of accuracy and bias in genomic prediction with censored data; and, 2) compare the Mixed Cox and Truncated Normal models under different scenarios of censoring and genetic architecture of the traits.

Material and methods

Simulated data

The phenotypic and genotypic information used in this study were simulated using the QMSim software (Sargolzaei & Schenkel, 2009). The parameters used in the simulation process were similar to those used by Brito, Neto, Sargolzaei, Cobuci, and Schenkel (2011). The historical population was made up of 1,000 generations with an initial size of 2,000 individuals each. Then a genetic bottleneck was created, so that in the next 1,020 generations, the size of the population gradually decreased from 2,000 to 1,500, in order to create initial linkage disequilibrium (LD). Subsequently, the population was expanded by random mating individuals from the last generation of the historical population, generating eight more generations, with five offspring per female. Finally, a recent population was created, obtained by mating 120 males with 6,000 females, randomly selected from the last generation of the expanded population, generating another ten generations, with one offspring per female. The proportion of male individuals was kept equal to 0.5 in the expanded and recent populations.

The genome was simulated with 52,885 single nucleotide polymorphisms (SNPs) and 88 Quantitative Trait Loci (QTL), distributed over 29 chromosomes, making up a total genome size of 2,740 cM. The SNPs were equally spaced and randomly distributed across the genome. The QTL effects were sampled from a gamma distribution with a parameter equal to 0.4 and a scale parameter determined by the QMSim software according to the simulated genetic variance of the trait. A minor allele frequency (MAF) of 0.10 was assumed for markers and QTL, with a mutation rate of 1×10^{-5} , for markers and QTL in the historical population. The genome parameters were chosen to mimic the cattle (*Bos taurus*) genome. In the simulation process we sought to represent the trait age at first calving. To characterize the base population, four phenotypic traits were simulated. The first two traits, C1 and C2, were simulated considering QTL heritability equal to 0.07 and 0.00, respectively, and heritability and phenotypic variance, according to the values estimated by Costa et al. (2019) via linear model - threshold, respectively, equal to 0.07 and 17.64. The other two traits were considered to have total heritability equal to 0.27 and QTL heritability equal to 0.27 (C3) and 0.00 (C4).

The censored observations were obtained by choosing a phenotypic value (C), as a threshold for censoring on the right, thus, phenotypic values greater than C, were considered as censored observations. Since Y_c is the vector of phenotypes simulated by QMSim, from Y_c and C two other phenotypic variables were created. First, $y_c = \min(Y_c, C)$, where C was calculated as $C = Q(Y_c, \text{probability} = 1 - P\%)$, in which P% is the proportion of censored observations and Q is the quantile of the normal distribution; and let δ be a censoring indicator variable, which assumes the value 1, if $Y_c \leq C$ or 0, if $Y_c > C$. The censored variable was constructed in order to obtain phenotypes with 10, 40, and 70% of censored observations. Thus, 12 scenarios were evaluated in this study, obtained by two heritability values, two QTL heritabilities, and three proportions of censored data. The entire simulation process was repeated 10 times.

Statistical methods

Only individuals from generations eight, nine, and 10 of the recent population were used for the genomic prediction analyses, with each generation comprising 6,000 individuals, totaling 18,000 animals. The genotypic information of these animals was used to obtain the genomic kinship matrix (G). As the simulated trait is restricted to females, only their phenotypic information was used to adjust the models. Thus, females of generations eight and nine were considered as the training population and females of generation 10 as the validation population.

The Mixed Cox and the Truncated Normal models were used to predict the additive genetic values (a) of each individual. The Mixed Cox model was described according to Ripatti and Palmgren (2000) and Therneau, Grambsch, and Pankratz (2003) as follows:

$$\lambda(t) = \lambda_0(t)\exp\{\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{a}\}, \quad (1)$$

where: $\lambda_0(t)$ is the base failure rate function; \mathbf{X} and \mathbf{Z} are the incidence matrices for fixed and random effects, respectively; $\boldsymbol{\beta}$ is a vector of fixed effects and \mathbf{a} is a vector of additive genetic effects, in which a normal distribution with mean zero and covariance matrix $G\sigma_a^2$ was assumed, with σ_a^2 being the additive genetic variance. The partial likelihood function for the Mixed Cox model can be described as (Giolo & Demétrio, 2011):

$$L = \int \text{PL}(\boldsymbol{\beta}, \mathbf{a}) \frac{1}{\sqrt{2\pi|\mathbf{G}|}} \exp\left[-\frac{1}{2}\mathbf{a}'\mathbf{G}^{-1}\mathbf{a}\right] d\mathbf{a}, \quad (2)$$

where PL is the partial likelihood function of Cox proportional hazards model and the other terms as previously defined. As the integral of equation (2) does not have a closed form, Ripatti and Palmgren (2000) suggested the use of the Laplace approximation to obtain the logarithm of the likelihood function, and to overcome the problems arising from the multidimensional integral calculation. The Mixed Cox model was adjusted via the “coxme” package (Therneau, 2020) from the R software (R Development Core Team, 2020).

The Truncated Normal model was adjusted via the Reproductive Kernel Hilbert Spaces (RKHS) regression model of the BGLR package (Pérez & Campos, 2014) of the R software (R Development Core Team, 2020), considering the Kernel matrix (K), as being equal to the genomic kinship matrix (G). According to Pérez and Campos (2014), the adjustment of linear models, in the presence of censored phenotypic traits, is done considering that the phenotypes are sampled from a truncated normal distribution. The following mixed linear model was fitted: $\mathbf{y}_c = \mathbf{1}\boldsymbol{\mu} + \mathbf{Z}\mathbf{a} + \mathbf{e}$, where: \mathbf{y}_c is the vector of observed phenotypes; $\boldsymbol{\mu}$ is the intercept vector; \mathbf{Z} is the incidence matrix that relates phenotypes to the random animal effect; \mathbf{a} is the vector of additive genetic values, with distribution $N(0, G\sigma_a^2)$, with σ_a^2 being the additive genetic variance and \mathbf{e} the vector of residues with distribution $N(0, I\sigma_e^2)$. For the additive genetic variance and for residual variance, scaled inverse chi-squared distributions were assumed, with scale parameter and degree of freedom, chosen as indicated by Pérez and Campos (2014).

The genomic relationship matrix G was used in the Mixed Cox and Truncated Normal models. G was calculated as (VanRaden, 2008): $\mathbf{G} = \mathbf{W}\mathbf{W}'/2\sum_{j=1}^m p_j q_j$, where \mathbf{W} is a matrix of order $n \times m$, composed of the elements: $w_{ij}=0-2p_j=-2p_j$, $w_{ij}=1-2p_j=q_j-p_j$ and $w_{ij}=2-2p_j=2q_j$, associated to the genotypes, mm, Mm, and MM, respectively, with p_j being the allele frequency of the SNP M at locus j , and $q_j=1-p_j$ the allele frequency of m at locus j . The Gibbs sampler algorithm was used for the analyses, considering 160,000 iterations, the first 20,000 of which were discarded (burn-in), and with a spacing of 10 between the samples. The convergence analysis was performed using the Geweke criterion implemented in the BOA package (Smith, 2007) of the R software (R Development Core Team, 2020).

Metrics used to evaluate the models

The Mixed Cox and the Truncated Normal models were adjusted considering the phenotypic information, by the double (y_c, δ) and genotypic information by the genomic relationship matrix (G). The models were also adjusted to the complete data \mathbf{Y}_c , considering a scenario where the percentage of censoring was equal to zero. Accuracy (Ac) and bias ($b_{(GEBV, y)}$) are the two most used measures to compare genomic prediction models. To estimate Ac and bias considering the GEBVs obtained for the phenotypic variable with complete observations (\mathbf{Y}_c), we calculated PC between the enhanced genomic breeding values (\widehat{GEBVs}) and TBV, or that is, $Ac = \text{cor}(\widehat{GEBV}, \text{TBV})$. The bias was obtained by estimating the slope of the linear regression of \mathbf{Y}_c as a function of \widehat{GEBVs} , or by the expression: $b_{(GEBV, TBV)} = \text{cov}(\widehat{GEBV}, \text{TBV})/\sigma_{GEBV}^2$. As in practice these measures are also used

for incomplete phenotypes, in a scenario where TBVs are unknown, the bias and accuracy were also estimated considering y_c , that is, $Ac = \text{cor}(\widehat{\text{GEBV}}, y_c) / \sqrt{h^2}$ and $b_{(\widehat{\text{GEBV}}, y_c)}$. The values $r_{(\widehat{\text{GEBV}}, \text{TBV})}$ and $b_{(\widehat{\text{GEBV}}, \text{TBV})}$ were used as a reference to compare alternative measures of accuracy and bias in the presence of incomplete observations. In this study, we proposed two alternative measures to estimate accuracy: the maximal correlation (Gebelein, 1941; Rényi, 1959; Breiman & Friedman, 1985) and the Pearson correlation for censored data (Li, Gillespie, Shedden, & Gillespie, 2018).

The maximal correlation is based on the determination of transformations, possibly nonlinear for two variables, subject to the mean zero and variance one, in order to maximize the PC between these two transformations (Feizi, Makhdoumi, Duffy, Kellis, & Medard, 2017). Thus, since X and Y are two real random variables, the maximal correlation between them is defined by: $\rho^*(X, Y) = \max_{f, g} \rho(f(X), g(Y))$, where: ρ is the PC coefficient, and the maximum is taken from the set of measurable functions $f, g: \mathbb{R} \rightarrow \mathbb{R}$, with $0 < \text{Var} f(X) < \infty$ and $0 < \text{Var} g(Y) < \infty$ (Breiman & Friedman, 1985; Blázquez & Miño, 2014). The algorithm proposed by Breiman and Friedman (1985), implemented in the “acepack” package (Spector et al., 2016) of the R software, was used to estimate the maximal correlation.

A specific measure was also used for the correlation of censored data. Li et al. (2018) described the implementation of the profile maximum likelihood method to estimate PC for bivariate data with censoring or missing values. The authors proposed a general model, for the case of interval censored, which includes, right censored, left censored, exact observed values, and data with missing observations. The univariate likelihood function is defined by:

$$L(\theta) = \prod_{i=1}^{k_1} F_{\theta}(x_i^{\text{upper}}) \times \prod_{i=k_1+1}^{k_2} [F_{\theta}(x_i^{\text{upper}}) - F_{\theta}(x_i^{\text{lower}})] \times \prod_{i=k_2+1}^{k_3} [1 - F_{\theta}(x_i^{\text{lower}})] \times \prod_{i=k_3+1}^n f_{\theta}(x_i),$$

The authors considered that each case belongs to the interval $(x_i^{\text{lower}}, x_i^{\text{upper}}]$. For right, left, interval censoring, exact values and missing observations, inequalities are considered, respectively: $-\infty < x_i^{\text{lower}} < x_i^{\text{upper}} = \infty$, $-\infty = x_i^{\text{lower}} < x_i^{\text{upper}} < \infty$, $-\infty < x_i^{\text{lower}} < x_i^{\text{upper}} < \infty$, $-\infty < x_i^{\text{lower}} = x_i^{\text{upper}} < \infty$ and $-\infty = x_i^{\text{lower}} < x_i^{\text{upper}} = \infty$, where: k_1 represents the number of observations left censored with detection limit x_i^{upper} , $i=1, \dots, k_1$; followed by k_2-k_1 interval censored observations with limits $(x_i^{\text{lower}}, x_i^{\text{upper}}]$, $i=k_1+1, \dots, k_2$; k_3-k_2 right censored values in x_i^{lower} , $i=k_2+1, \dots, k_3$ and $n-k_3$ exact values x_i , $i=k_3+1, \dots, n$ (Li et al., 2018).

For right censored data, the authors mentioned that the univariate likelihood function can be written in terms of the density function, for exact observed values, and the survival function ($S_{\theta}(x) = P(X > x)$) for values censored on the right. In this case, θ represents a vector of parameters, which includes the correlation coefficient. Pearson correlation for censored data (PCC) was obtained using the “clikcorr” package (Li et al., 2018) of the R software.

The expected accuracy (Ace) was estimated based on G and the additive matrix associated with the pedigree relationship matrix (A), according to the equation proposed by Wientjes, Veerkamp, and Calus (2013):

$$\text{Ace} = \sqrt{\frac{N_p h^2}{N_p h^2 + M_e}},$$

where: N_p is the number of individuals in the training population, genotyped and phenotyped, h^2 is the heritability of the trait and M_e is the effective number of loci, obtained by: $M_e = 1/\text{Var}(\mathbf{D})$, where $\mathbf{D} = \mathbf{G} - \mathbf{A}$.

The bias was estimated by the slope of the simple linear regression and the Tobit regression considering the latent variable and the marginal effects associated with censored and uncensored observations, by the regression of phenotypes in GEBVs. The Tobit regression model or censored regression was defined as:

$$y_i = \begin{cases} y_i^* = \mathbf{x}_i \boldsymbol{\beta} + \varepsilon_i, \text{ se } y_i^* < t_c \\ t_c, \text{ se } y_i^* \geq t_c \end{cases},$$

with: $i=1, 2, \dots, N$, N is the number of individuals and $\varepsilon_i \sim N(0, \sigma^2)$, y_i^* is an unobserved latent variable, y_i is the observed dependent variable, x_i a vector of regressor variables, $\boldsymbol{\beta}$ is a vector of unknown parameters and t_c is the censoring time. To obtain the marginal effect on the expected value for y , associated with censored and non-censored observations, the slope $\boldsymbol{\beta}$ was multiplied by the probability that the response variable is greater than or equal to the censoring time, that is, $\partial E(y|x) / \partial x_j = \beta_j \times \Phi(t_c - \bar{\mathbf{x}}' \boldsymbol{\beta} / \sigma)$, where: Φ is the cumulative distribution function of the standard normal (Tobin, 1958; Long, 1997). As a measure to evaluate the deviation

between the estimated values of accuracy, based on phenotypes and GEBVs and based on TBVs and GEBVs, the measure of relative difference, was calculated as: $\Delta = 100 \times \left| \frac{Ac_{(GEBV, y_c)} - Ac_{(GEBV, TBV)}}{Ac_{(GEBV, TBV)}} \right|$.

Results and discussion

The estimated values for the accuracy of the models, based on PC, MC, and PCC in different scenarios are shown in Table 1. Positive and negative values were observed, respectively, for the accuracy of the Truncated Normal and Mixed Cox models, when estimated via PC. According to Hou et al. (2009), the sign of these values was due to the fact that, in the Mixed Cox model, the risk of occurrence of the event is modeled, whereas in the Truncated Normal model, the time until the occurrence of the event is directly modeled. Thus, the GEBVs predicted by the Mixed Cox model and the Truncated Normal model, were inversely proportional, and presented different scales, since the GEBVs estimated by the Mixed Cox model were related to the risk of occurrence of the event of interest (Hou et al., 2009; Santos et al., 2015). As the MC can only assume values in the range from zero to one, positive values of similar magnitude were perceived for both models in most of the evaluated scenarios.

Table 1. Estimated mean values for the accuracy (Ac) of the Mixed Cox and Truncated Normal models, considering the incomplete phenotypes (y_c), for 10, 40, and 70% of censoring, using Pearson correlation (PC), maximal correlation (MC), and Pearson correlation for censored data (PCC), as measures of correlation between phenotypes and genomic enhanced breeding values (GEBVs) and for Pearson and maximal correlation between true breeding values (TBVs) and GEBVs, with 0% censoring.

$Ac = cor(GEBV, y_c) / \sqrt{h^2}$												
P%	C1						C2					
	Mixed Cox model			Truncated Normal model			Mixed Cox model			Truncated Normal model		
	PC	MC	PCC	PC	MC	PCC	PC	MC	PCC	PC	MC	PCC
10%	-0.30	0.32	-0.30	0.30	0.34	0.31	-0.13	0.19	-0.13	0.12	0.16	0.12
	Aa	Aa	Aa	Aa	Ba	Aa	Aa	Aa	Aa	Aa	Aa	Aa
40%	-0.27	0.29	-0.29	0.27	0.31	0.30	-0.10	0.16	-0.11	0.10	0.16	0.11
	Aa	Bba	Ba	Aa	Bba	Ba	Aa	Aa	Aa	Aa	Aa	Aa
70%	-0.20	0.26	-0.27	0.21	0.25	0.27	-0.09	0.16	-0.12	0.07	0.15	0.10
	Ab	Ba	Ba	Ab	Bb	Ba	Aa	Aa	Aa	Aa	Ba	ABa
$Ac = cor(GEBV, TBV)$												
P%	C1						C2					
	Mixed Cox model			Truncated Normal model			Mixed Cox model			Truncated Normal model		
	PC	MC	PCC	PC	MC	PCC	PC	MC	PCC	PC	MC	PCC
0%	-0.38	0.38	-0.38	0.41	0.41	0.41	-0.14	0.15	-0.14	0.13	0.14	0.13
$Ac = cor(GEBV, y_c) / \sqrt{h^2}$												
P%	C3						C4					
	Mixed Cox model			Truncated Normal model			Mixed Cox model			Truncated Normal model		
	PC	MC	PCC	PC	MC	PCC	PC	MC	PCC	PC	MC	PCC
10%	-0.44 ^a	0.44 ^a	-0.45 ^a	0.46 ^b	0.46 ^b	0.46 ^b	-0.19	0.20	-0.19	0.20	0.21	0.20
	Aa	Aba	Ba	Aa	Aba	Aa	Aa	Aa	Aa	Aa	Ba	Aa
40%	-0.40 ^a	0.41 ^a	-0.43 ^a	0.41 ^b	0.43 ^b	0.44 ^b	-0.18	0.20	-0.20	0.19	0.20	0.21
	Ab	ABb	Bb	Ab	Bb	Bb	Aa	Aa	Aa	Aa	Aa	Aa
70%	-0.31 ^a	0.33 ^a	-0.39 ^a	0.32 ^b	0.34 ^b	0.41 ^a	-0.16	0.17	-0.20	0.16	0.18	0.20
	Ac	Bc	Cb	Ac	Bc	Cb	Aa	ABa	Ba	Ab	Ab	Ba

P%: Percentage of censoring; C1 – $h^2 = 0.07$ and h^2 of QTL = 0.07; C2 – $h^2 = 0.07$ and h^2 of QTL = 0; C3 – $h^2 = 0.27$ and h^2 of QTL = 0.27 and C4 – $h^2 = 0.27$ and h^2 of QTL = 0. Means followed by the same letter, uppercase in the lines and lowercase in the columns, do not differ statistically by the paired t test at 5% probability with Bonferroni correction. The significant differences between the Mixed Cox and Truncated Normal models are indicated by the presence of the superscript letters a and b.

The reliability of the expected genomic prediction for heritabilities of 0.07 and 0.27, according to the theoretical formula of accuracy, was 0.41 ± 0.05 and 0.65 ± 0.06 . Thus, the correlation measures that showed values closer to what was expected, were considered the best measures. In scenarios in which the QTL heritability was equal to the total heritability of the trait (C1 and C3), it was observed that in most cases, the MC and PCC measures presented statistically equal estimates, which are different from those obtained for PC, or only PCC differed significantly from PC. The only exceptions occurred for C3 with 70% censoring, in which PCC was statistically superior to PC and MC. The MC, estimated based on the values predicted by the Truncated Normal model, for C1 with 10% censoring was shown to be statistically superior to the other measures. With the QTL heritability equal to zero, few significant differences were observed between the accuracy measures. For the trait C2, the only significant difference found was between MC and PC, for the Truncated Normal model with 70% censoring. For the trait C4 with 70% censoring, considering the GEBVs obtained by the Truncated Normal and Mixed Cox models, it was noted that the PCC measure differed statistically from the PC measure, with the estimates obtained by the MC and PC measures being statistically equal. With 10% censoring, MC was statistically superior to the other metrics, only for the Truncated Normal

model. In general, the estimates obtained by PC were lower than those estimated for MC and PCC, and these in turn, presented similar values in most scenarios.

Several correlation measures have been proposed over the years with the aim of quantifying the dependency between two variables, in a more general context, in which the dependency can be linear or non-linear, considering different distributions and types of dependency relationship. Reshef et al. (2011) presented a measure of dependence between two continuous variables, called the maximal information coefficient (MIC), which, with sufficient sample size, is capable of capturing functional and non-functional associations. When comparing the MIC with the maximal (Rényi, 1959; Breiman, & Friedman, 1985), Pearson (Galton, 1888; Pearson, 1920) and Spearman (Spearman, 1904) correlation measures, the authors noted that for the dependency relationship with linear noise, these measures showed similar performance, with determination coefficients equal to 1. For non-linear noise, such as cubic, exponential, categorical, and parabolic, the MIC and the maximal correlation showed similar results. When sinusoidal noises were considered, the MIC was more appropriate than all the other measures compared, including mutual information via KDE - kernel density estimators (Moon, Rajagopalan, & Lall, 1995) and via Kraskov estimator (Kraskov, Stögbauer, & Grassberger, 2004).

Santos, Takahashi, Nakata, and Fujita (2014) carried out a study to compare statistical methods used to identify the type of dependency relationship between random variables. In this study, the following measures were evaluated, including correlation [Pearson, Spearman, Kendall (Kendall, 1938)], distance correlation (Szekely, Rizzo, & Bakirov, 2007), mutual information via KDE, and MIC (Reshef et al., 2011). Simulated and real data were used, and comparisons between methods were performed using the ROC curve (Receiver Operating Characteristic). The authors showed that Pearson, Spearman, and Kendall correlations can detect only linear or monotonic nonlinear relationships, strictly increasing or decreasing, and that these three measures have a similar performance. The other methods, on the other hand, were able to identify linear relationships, as well as non-monotonic and non-functional relationships. According to the authors, if the hypotheses of linearity or monotonicity are satisfied, the application of the Spearman and Kendall correlations should be preferred, since they are able to identify linear and monotonic relations with high power.

Deebani and Kachouie (2020) evaluated the performance of Pearson, Spearman correlation measures, distance, and maximal correlations and MIC, considering simulated data similar to those used by Reshef et al. (2011), these being extended with the addition of different noise levels: none, low (5%), medium (20%), and high (40%). For the linear relationship between the variables, with no or low noise, the five methods had scores close to or equal to one. With increasing noise, all measures of correlation decreased. The maximum correlation was superior to the MIC and slightly superior to the Pearson, Spearman, and distance measurements correlations. For the exponential, monotonic and strictly increasing relationship, without noise, scores equal to one were obtained by the maximum correlation, Spearman correlation and by the MIC. With the increase in noise, the maximal correlation showed consistently higher scores than the other measures. Spearman correlation yielded higher scores than Pearson correlation only in cases with no or low noise. In the other two scenarios these measures were similar.

Pearson correlation is a measure used to quantify the linear dependence between two variables. This measure is often used in genomic selection to estimate the accuracy of predictive models, and its application is appropriate when the variables have a linear relationship and are normally distributed. The Spearman and Kendall correlations are measures that go beyond quantifying the linear dependence between two variables, they measure the monotonic association between the variables. These measures do not have the assumption of linearity, and they can also be used for variables on an ordinal scale, which make them more appropriate, when the phenotype is survival time. MC does not show preference for linear or monotonic relationships, and does not require any assumptions regarding the distribution of the data. MC can completely characterize the independence between two variables, since a correlation equal to zero implies the independence of the variables. Finally, for the identification of non-linear or non-functional dependency relationships, methods based on ranking or information theory proved to be more appropriate (Santos, Takahashi, Nakata, & Fujita, 2014).

The highest mean values in absolute value for accuracy were observed for the percentage of censoring of 10%. With the increase in the percentage of censoring, there was a loss of predictive ability, which can be explained by the reduction of phenotypic information used for the prediction of GEBVs. Only for trait C3, this reduction in accuracy was significant, in the other scenarios, no significant differences were observed, or the difference was significant only between the classes with 10% and 40% and the class of 70%. This fact was observed for the three measures, the reduction in accuracy being more evident for PC. These results are in agreement with those obtained by Kärkkäinen and Sillanpää (2013), who, when considering the censoring

percentages of 20, 50, and 80%, observed that the accuracy of the Bayesian threshold model, calculated by PC, gradually decreases with increasing the percentage of censoring.

Pearson correlation was most affected by the increase in the percentage of censoring, due to the fact that it did not use the information present in the censored observations for estimation. When using PC and MC for continuous response, the phenotypes of the censored observations were considered to be exact, which implied a reduction in the information used. Although MC does not make use of censored observations, it is more flexible than PC, which makes it possible to model more complex relationships. According to Li et al. (2018), partially observed bivariate data appear in different forms. If one of the variables presents completely missing observations, although these observations do not provide direct information about the correlation coefficient, they are informative as to the marginal distributions. In the presence of censored observations, and considering that none of the variables are completely missing, the observations contribute information about the correlation, and the marginal distributions.

Several correlation measures have been proposed to estimate the correlation in the presence of variables with censored observations on the left, right, or interval. Newton and Rudel (2007) proposed a correlation measure based on maximum likelihood (ML) estimators for data analysis with different points of censoring on the left. This measure was compared with Kendall's tau-b correlation coefficient (Oakes, 1982) for censored data, and with PC, Spearman and Kendall's tau correlation, considering censored observations equal to half the detection limit or as missing observations. The authors showed via a simulation study that for sample sizes greater than or equal to 100, the correlation measure based on ML whose standard deviations were estimated separately, presented the best results, which can be used for censoring proportions ranging from 60 to 90%. This performance was dependent on the parametric value of the correlation. Most of the measures evaluated were highly biased towards the proportion of censoring ranging from 0 to 90%.

Pearson correlation for censored data estimated via profiled likelihood proposed by Li et al. (2018) is a general correlation measure that can be applied to data with left, right, interval, and missing data censoring, assuming a normal or Student's t distribution for the data. The authors evaluated the performance of this correlation measure in a simulation study considering different sample sizes, percentages of censoring and underlying distributions. For data generated from the normal bivariate distribution, it was shown that the probability of covering the confidence interval provides satisfactory values for censoring on the right with fixed or random points, and that the proposed estimator is not biased for random censoring on the right. If the underlying marginal distributions have tails that are heavier than that of the normal distribution, and the correlations are high, the use of the bivariate Student t distribution provides better coverage probabilities than those obtained with the bivariate normal distribution.

In traits C1 and C3, all genetic variance was explained by the QTL, whereas traits C2 and C4 were polygenic in nature. The three measures were equally useful regarding the discernment between the models, that is, the average efficiency of the Truncated Normal model relative to the Mixed Cox model, presented close values for the three measures of accuracy. For traits C1 and C3, these values were on average 3 and 4%, respectively. Only for C2, the Mixed Cox model was better than the Truncated Normal model. In this case, the relative efficiency of the Truncated Normal method was on average 6% lower than that of the Mixed Cox model. For trait C4, also polygenic, the previous result was not repeated, the Truncated Normal model was 6% more efficient than the Mixed Cox model.

The accuracy calculated based on the GEBVs and TBVs, were lower than the expected accuracy, and, in most cases, greater than the estimated accuracy based on the incomplete phenotypes. In studies involving censored real data, the only information available about the trait is the phenotype and the censoring status. In these cases, to assess the accuracy of the models, PC between the phenotypic values and the GEBVs is generally estimated. When analyzing simulated data, TBVs are available, and when correlating TBVs with GEBVs, we can compare the predicted with the simulated values. In this study, by correlating the GEBVs with TBVs and incomplete phenotypes, it was possible to measure how much the estimated correlations with incomplete phenotypes differed from those estimated with TBVs, by calculating the measure of relative difference (Δ).

Table 2 shows the values obtained for the relative difference in each of the scenarios. Most of the time, for the Mixed Cox and Truncated Normal models, the results showed that PCC had a smaller relative difference when compared with PC and MC. In all scenarios, the relative difference in the PCC and MC measures were smaller than those obtained for PC. The calculated average relative differences were 12.45, 13.95 and 12.68%, respectively, for PCC, MC, and PC, with 10% censored. With the increase in censored, the average difference

associated with PC increased to 22.78 and 37.88%, respectively, to 40 and 70% of censored, while for PCC the average relative difference increased to 16.80 and 20.63%. The MC presented values of average relative distance similar to those obtained for PCC, with a larger difference found for 10% of censored. The Mixed Cox and Truncated Normal models are models known to perform better for polygenic traits, then for traits that are influenced by genes with moderate to large effects. This fact probably justifies a smaller relative difference for the traits C2 and C4. In general, the accuracy values estimated based on the PCC measure, were closer to the values estimated based on the TBVs, than those estimated with PC and MC.

Table 2. Estimated mean values for the relative difference (Δ) in percentage, considering the estimated accuracy based on incomplete phenotypes (y_c) and true breeding values (TBV), based on Pearson (PC), maximal (MC), and Pearson correlation for censored data (PCC) correlations, considering the genomic breeding values estimated by the Mixed Cox and Truncated Normal models, for phenotypic traits with 10, 40, and 70% censored.

P%	Δ											
	C1						C2					
	Mixed Cox model			Truncated Normal model			Mixed Cox model			Truncated Normal model		
	PC	MC	PCC	PC	MC	PCC	PC	MC	PCC	PC	MC	PCC
10%	21.05	15.79	21.05	24.39	17.07	24.39	7.14	26.67	7.14	7.69	14.29	7.69
40%	28.95	23.68	23.68	34.15	24.39	26.83	28.57	6.67	21.43	23.08	14.29	15.38
70%	47.37	31.58	28.95	48.78	39.02	34.15	35.71	6.67	14.29	46.15	7.14	23.08
	C3						C4					
10%	18.52	18.52	16.67	17.86	19.30	17.86	0	0	0	4.76	0	4.76
40%	25.93	24.07	20.37	26.79	24.56	21.43	5.26	0	5.26	9.52	4.76	0
70%	42.59	38.89	27.78	42.86	40.35	26.79	15.79	15	5.26	23.81	14.29	4.76

P%: Percentage of censoring; C1 - $h^2 = 0.07$ and h^2 of QTL = 0.07; C2 - $h^2 = 0.07$ and h^2 of QTL = 0; C3 - $h^2 = 0.27$ and h^2 of QTL = 0.27 and C4 - $h^2 = 0.27$ and h^2 of QTL = 0.

Table 3 presents the estimated average values for the slope of the simple linear regression model, the marginal effect on censored and uncensored observations and the slope for the latent variable for the Tobit regression (TB), considering all the scenarios evaluated, for the GBVs estimated by mixed Cox and Truncated Normal models. The coefficients estimated for the Mixed Cox and Truncated Normal models differed markedly in terms of the magnitude and sign of the estimates. This fact was due to the inverse relationship, and the difference in scale between the GBVs estimated by the two models. It was also observed that the slope coefficients estimated by the regression of the TBVs in the GEBVs based on the complete data were greater than or equal to the values estimated in the presence of censored observations in all scenarios.

Table 3. Mean values estimated for the slope coefficients obtained by regression of the incomplete phenotypes (y_c), in the genomic enhanced breeding values (GEBV) using Mixed Cox and Truncated Normal models, by simple linear (SLR) and Tobit (TB) regressions, considering different values of total heritability and QTL heritability. For the Tobit regression, the first number represents the marginal effect associated with censored and uncensored observations, and the second number the effect associated with the latent variable.

$b_{(y_c, GEBV)} = \text{reg}(y_c \sim GEBV)$										
P%	C1						C2			
	Mixed Cox Model		Truncated Normal model		Mixed Cox Model		Truncated Normal model			
	SLR	TB	SLR	TB	SLR	TB	SLR	TB	SLR	TB
10%	-4.08	-4.09; -4.48	0.63	0.63; 0.69	-2.26	-2.25; -2.46	0.26	0.26; 0.29		
40%	-1.96	-1.95; -3.09	0.44	0.45; 0.71	-1.20	-1.23; -1.94	0.18	0.19; 0.30		
70%	-0.81	-0.83; -2.51	0.23	0.23; 0.69	-0.95	-0.82; -2.50	0.09	0.09; 0.29		
	C3						C4			
10%	-2.99	-3.03; -3.22	0.87	0.88; 0.94	-2.41	-2.41; -2.56	0.60	0.60; 0.64		
40%	-1.95	-1.94; -2.85	0.64	0.64; 0.94	-1.53	-1.51; -2.20	0.46	0.45; 0.66		
70%	-0.87	-0.87; -2.35	0.35	0.34; 0.93	-0.73	-0.71; -1.88	0.27	0.27; 0.70		
$b_{(TBV, GEBV)} = \text{reg}(TBV \sim GEBV)$										
	C1						C2			
0%	-5.23	-5.23	0.71	0.71	-2.80	-2.80	0.30	0.30		
	C3						C4			
0%	-3.27	-3.27	0.92	0.92	-2.61	-2.61	0.61	0.61		

P%: Percentage of censoring; C1: $h^2 = 0.07$ and h^2 of QTL = 0.07; C2: $h^2 = 0.07$ and h^2 of QTL = 0; C3: $h^2 = 0.27$ and h^2 of QTL = 0.27 and C4: $h^2 = 0.27$ and h^2 of QTL = 0.

The slope coefficient and the coefficient associated with the marginal effect estimated, respectively, by the simple linear regression and by the Tobit regression, presented similar values in all the evaluated scenarios. These coefficients were inversely related to the proportion of censorship for the Truncated Normal

model, and directly related to the proportion of censorship for the Mixed Cox model, that is, in the Truncated Normal model, as the censorship increases, the coefficient decreases, in Mixed Cox model, there was an increase in the value of the coefficient, with an increase in censoring. The slope coefficient estimated based on the latent variable, showed a markedly increasing pattern, with the increase of censorship, only for the Mixed Cox model in scenarios C1, C3, and C4. In general, the coefficients estimated for the Truncated Normal model based on the latent variable suffered less variation with the increase in censoring, while those estimated by linear regression and based on marginal effects were drastically influenced by censoring, mainly with the change of 40 to 70% censoring. These last two coefficients were more suitable for the assessment of bias in the presence of censored observations. This is because with the reduction of phenotypic information used to estimate GBVs, it is expected that the bias of the prediction varies considerably, since the censoring can reach up to 70% of the observations. According to Long (1997), the Tobit regression model makes use of all available information, which includes censored observations. Thus, the model is able to provide consistent estimates, while the linear regression model, when treating all phenotypic values as observed, provides inconsistent estimates for the parameters.

The parameter estimates in the simple linear regression model are obtained via ordinal least squares, with the assumption of linearity in the data, normality, and homogeneity of the residues. The Tobit regression model, on the other hand, is based on maximum likelihood estimators, and has the same assumptions as simple linear regression. Assuming that the assumption of normality is met, Amemiya (1973) presented results regarding the consistency and asymptotic normality of ML estimators. Lumley, Diehr, Emerson, & Chen (2002) showed by means of a simulation study, that for a sufficiently large sample, simple linear regression can present a reasonable performance even for non-normal data. According to Amore and Murtinu (2019), the Tobit regression model is more sensitive to breaking the assumptions of normality and homogeneity of the residuals than the simple linear regression model.

Through a graphic inspection, it was found that in the present study, the residues were not normally distributed, and that the assumption of homogeneity of the residues was not accepted for all repetitions and censorship percentages. Lewis and McDonald (2014) found that for data simulated with normal residues, with a thousand observations, and with 25% censoring, the Tobit regression model was the most efficient among all models, which is less biased than the simple linear regression, and with root mean square error (RMSE) approximately four times smaller than the one obtained for the simple linear regression. For residues with mixed normal and log-normal distributions, the two models showed to be highly biased, whereas for mixed normal the models presented similar values of bias and RMSE and for log-normal, simple linear regression showed lower values of bias and RMSE.

Due to the difference in scale existing between the GBVs obtained by the Mixed Cox and Truncated Normal models, the bias is not a good measure to directly compare these models, which does not affect the objective of the study that the evaluation of an alternative measure to estimate the bias in the context of censored data. Santos et al. (2015) used Spearman correlation as a measure of accuracy for data with censored observations, and the Kappa coefficient as a measure for assessing the agreement of the ranking of individuals according to the GBVs estimated by the Mixed Cox model and by the linear mixed model with complete or imputed data.

Conclusion

From a statistical point of view, the proposed methodologies are preferred because they deal adequately with the presence of incomplete observations in the data. The results showed that the proposed measures of correlation presented estimates significantly higher than those presented by PC, mainly for the trait with QTL heritability equal to 0.27. The coefficient associated with the marginal effect of the TB presents values close to those obtained by SLR. The coefficient associated with the latent variable in the TB, in most cases, had little variation with changes in censoring proportions, while SLR is extremely affected by the presence of censored observations.

Acknowledgements

We would like to thank Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for providing a scholarship to the first author.

References

- Alemu, S. W., Calus, M. P. L., Muir, W. M., Peeters, K., Vereijken, A., & Bijama, P. (2016). Genomic prediction of survival time in a population of brown laying hens showing cannibalistic behavior. *Genetics Selection Evolution*, *48*(68), 1-10. DOI: <https://doi.org/10.1186/s12711-016-0247-4>
- Amemiya, T. (1973). Regression analysis when the dependent variable is truncated normal. *Econometrica*, *41*(6), 997-1016. DOI: <https://doi.org/10.2307/1914031>
- Amore, M. D., & Murtinu, S. (2019). Tobit models in strategy research: Critical issues and applications. *Global Strategy Journal*, *11*(3), 331-355. DOI: <https://doi.org/10.1002/gsj.1363>
- Araujo, A. C., Carneiro, P. L., Alvarenga, A. B., Oliveira, H. R., Miller, S. P., Retallick, K., & Brito, L. F., (2022). Haplotype-based single-step GWAS for yearling temperament in American Angus cattle. *Genes*, *13*, 17. DOI: <https://doi.org/10.3390/genes13010017>
- Blázquez, F. L., & Miño, B. S. (2014). Maximal correlation in a non-diagonal case. *Journal of Multivariate Analysis*, *131*(C), 265-278. DOI: <https://doi.org/10.1016/j.jmva.2014.07.008>
- Breiman, L., & Friedman, J. H. (1985). Estimating optimal transformations for multiple regression and correlation. *Journal of the American statistical Association*, *80*(391), 580-598. DOI: <https://doi.org/10.1080/01621459.1985.10478157>
- Brito, F. V., Neto, J. B., Sargolzaei, M., Cobuci, J. A., & Schenkel, F. S. (2011). Accuracy of genomic selection in simulated populations mimicking the extent of linkage disequilibrium in beef cattle. *BMC Genetics*, *12*(80), 1-10. DOI: <https://doi.org/10.1186/1471-2156-12-80>
- Campos, G., Hickey, J. M., Pong-Wong, R., Daetwyler, H. D., & Calus, M. P. L. (2013). Whole genome regression and prediction methods applied to plant and animal breeding. *Genetics*, *193*(2), 327-345. DOI: <https://doi.org/10.1534/genetics.112.143313>
- Costa, E. V., Ventura, H. T., Veroneze, R., Silva, F. F., Pereira, M. A., & Lopes, P. S. (2019). Bayesian linear-threshold censored models for genetic evaluation of age at first calving and stayability in Nellore cattle. *Livestock Science*, *230*(103833). DOI: <https://doi.org/10.1016/j.livsci.2019.103833>
- Deebani, W., & Kachouie, N. N. (2020) Monte Carlo ensemble correlation coefficient for association detection. *Communications in Statistics - Simulation and Computation*, *51*(12), 7095-7109. DOI: <https://doi.org/10.1080/03610918.2020.1823413>
- Feizi, S., Makhdoumi, A., Duffy, K., Kellis, M., & Medard, M. (2017). Network maximal correlation. *IEEE Transactions on Network Science and Engineering*, *4*(4), 229-247. DOI: <https://doi.org/10.1109/TNSE.2017.2716966>
- Galton, F. (1888). Co-relations and their measurement, chiefly from anthropometric data. *Proceedings of the Royal Society of London*, *45*(273-279), 135-145. DOI: <https://doi.org/10.1098/rspl.1888.0082>
- Gebelein, H. (1941). Das statistische problem der correlation als variation und eigenwertproblem und sein zusammenhang mit der ausgleichrechnung. *Zeitschrift für angewandte Mathematik und Mechanik*, *21*(6), 364-379. DOI: <https://doi.org/10.1002/zamm.19410210604>
- Giolo, S. R., & Demétrio, C. G. B. (2011). A frailty modeling approach for parental effects in animal breeding. *Journal of Applied Statistics*, *38*(3), 619-629. DOI: <https://doi.org/10.1080/02664760903521492>
- Hou, Y., Madsen, P., Labouriau, R., Zhang, Y., Lund, M. S., & Su, G. (2009). Genetic analysis of days from calving to first insemination and days open in Danish Holsteins using different models and censoring scenarios. *Journal Dairy Science*, *92*(3), 1229-1239. DOI: <https://doi.org/10.3168/jds.2008-1556>
- Kärkkäinen, H. P., & Sillanpää, M. J. (2013). Fast Genomic Predictions via Bayesian G-BLUP and Multilocus Models of Threshold Traits Including Censored Gaussian Data. *G3*, *3*(9), 1511-1523. DOI: <https://doi.org/10.1534/g3.113.007096>
- Kendall, M. (1938). A new measure of rank correlation. *Biometrika*, *30*(1/2), 81-93. DOI: <https://doi.org/10.1093/biomet/30.1-2.81>
- Kraskov, A., Stögbauer, H., & Grassberger, P. (2004). Estimating mutual information. *Physical Review E*, *69*(6), 1-16. DOI: <https://doi.org/10.1103/PhysRevE.69.066138>
- Lewis, R. A., & McDonald, J. B. (2014). Partially Adaptive Estimation of the Censored Regression Model. *Econometric Reviews*, *33*(7), 732-750. DOI: <https://doi.org/10.1080/07474938.2012.690691>

- Li, Y., Gillespie, B. W., Shedden, K., & Gillespie, J. A. (2018). Profile Likelihood Estimation of the Correlation Coefficient in the Presence of Left, Right or Interval Censoring and Missing Data. *The R Journal*, *10*(2), 159-179. DOI: <https://doi.org/10.32614/RJ-2018-040>
- Long, J. S. (1997). *Regression Models for Categorical and Limited Dependent Variables*. Thousand Oaks, CA: Sage Publications.
- Lumley, T., Diehr, P., Emerson, S., & Chen, L. (2002). The importance of the normality assumption in large public health data sets. *Annual Review of Public Health*, *23*, 151-169. DOI: <https://doi.org/10.1146/annurev.publhealth.23.100901.140546>
- Massender, E., Brito, L.F., Maignel, L., Oliveira, H.R., Jafarikia, M., Baes, C.F., ... Schenkel, F.S., (2022). Single-and multiple-breed genomic evaluations for conformation traits in Canadian Alpine and Saanen dairy goats. *Journal of Dairy Science*, *105*(7), 5985-6000. DOI: <https://doi.org/10.3168/jds.2021-21713>
- 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. DOI: <https://doi.org/10.1093/genetics/157.4.1819>
- Moon, Y., Rajagopalan, B., & Lall, U. (1995). Estimation of mutual information using kernel density estimators. *Physical Review E*, *52*(3), 2318-2321. DOI: <https://doi.org/10.1103/PhysRevE.52.2318>
- Newton, E., & Rudel, R. (2007). Estimating correlation with multiply censored data arising from the adjustment of singly censored data. *Environmental science & technology*, *41*, 221-228. DOI: <https://doi.org/10.1021/es0608444>
- Oakes, D. (1982). A concordance test for independence in the presence of censoring. *Biometrics*, *38*(2), 451-455.
- Oliveira, H. R., Miller, S. P., Brito, L. F., & Schenkel, F. S. (2021). Impact of censored or penalized data in the genetic evaluation of two longevity indicator traits using random regression models in North American Angus cattle. *Animals*, *11*(3). DOI: <https://doi.org/10.3390/ani11030800>
- Palaiokostas, C., Ferraresso, S., Franch, R., Houston, R. D., & Bargelloni, L. (2016). Genomic Prediction of Resistance to Pasteurellosis in Gilthead Sea Bream (*Sparus aurata*) Using 2b-RAD Sequencing. *G3*, *6*(11), 3693-3700. DOI: <https://doi.org/10.1534/g3.116.035220>
- Pearson, K. (1920). Notes on the history of correlation. *Biometrika*, *13*, 25-45. DOI: <https://doi.org/10.1093/biomet/13.1.25>
- Pérez, P., & Campos, G. (2014). Genome-wide regression and prediction with the BGLR statistical package. *Genetics*, *198*(2), 483-495. DOI: <https://doi.org/10.1534/genetics.114.164442>
- R Development Core Team. (2020). *R: a language and environment for statistical computing*. Vienna, AU: R Foundation for Statistical Computing. Retrieved from <https://cran.r-project.org/bin/windows/base/>
- Rényi, A. (1959). On measures of dependence. *Acta Mathematica Hungarica*, *10*, 441-451. DOI: <https://doi.org/10.1007/bf02024507>
- Reshef, D. N., Reshef, Y. A., Finucane, H. K., Grossman, S. R., McVean, G., Turnbaugh, P. J., ... Sabeti, P. C. (2011). Detecting Novel Associations in Large Datasets. *Science*, *334*(6062), 1518-1524. DOI: <https://doi.org/10.1126/science.1205438>
- Ripatti, S., & Palmgren, J. (2000). Estimation of multivariate frailty models using penalized partial likelihood. *Biometrics*, *56*(4), 1016-1022. DOI: <https://doi.org/10.1111/j.0006-341x.2000.01016.x>
- Santos, S. S., Takahashi, D. Y., Nakata, A., & Fujita, A. (2014). A comparative study of statistical methods used to identify dependencies between gene expression signals. *Briefings in Bioinformatics*, *15*(6), 906-918. DOI: <https://doi.org/10.1093/bib/bbt051>
- Santos, V. S., Martins, F. S., Resende, M. D., Azevedo, C. F., Lopes, P. S., Guimarães, S. E., ... Silva, F. F. (2015). Genomic selection for slaughter age in pigs using the Cox frailty model. *Genetics and Molecular Research*, *14*(4), 12616-12627. DOI: <https://doi.org/10.4238/2015.October.19.5>
- Sargolzaei, M., & Schenkel, F. S. (2009). QMSim: A large-scale genome simulator for livestock. *Bioinformatics*, *25*(5), 680-681. DOI: <https://doi.org/10.1093/bioinformatics/btp045>
- Smith, B. J. (2007). boa: An R Package for MCMC Output Convergence Assessment and Posterior Inference. *Journal of Statistical Software*, *21*(11), 1-37. DOI: <https://doi.org/10.18637/jss.v021.i11>
- Spearman, C. (1904). "General intelligence", objectively determined and measured. *The American Journal of Psychology*, *15*(2), 201-292. DOI: <https://doi.org/10.2307/1412107>

- Spector, P., Friedman, J., Tibshirani, R., Lumley, T., Garbett, S., & Baron, J. (2016). *Acepack: ACE and AVAS methods for choosing regression transformations. R package version 1.4.1*. Retrieved from <https://cran.r-project.org/web/packages/acepack/index.html>
- Szekely, G., Rizzo, M., & Bakirov, N. (2007). Measuring and testing dependence by correlation of distances. *The Annals of Statistics*, 35(6), 2769-2794. DOI: <https://doi.org/10.1214/009053607000000505>
- Teissier, M., Larroque, H., Brito, L. F., Rupp, R., Schenkel, F. S., & Robert-Granié, C., (2020). Genomic predictions based on haplotypes fitted as pseudo-SNP for milk production and udder type traits and SCS in French dairy goats. *Journal of Dairy Science*, 103(12), 11559-11573. DOI: <https://doi.org/10.3168/jds.2020-18662>
- Therneau, T. M., Grambsch, P. M., & Pankratz, V. S. (2003). Penalized survival models and frailty. *Journal of Computational and Graphical Statistics*, 12, 156-175. DOI: <https://doi.org/10.1198/1061860031365>
- Therneau, T. M. (2020). Coxme: Mixed Effects Cox Models. *R-package description.*, 1-14. Retrieved from <https://cran.r-project.org/web/packages/coxme/vignettes/coxme.pdf>
- Tobin, J. (1958). Estimation of Relationships for Limited Dependent Variables. *Econometrica*, 26, 24-36. DOI: <https://doi.org/10.2307/1907382>
- Vallejo, R. L., Leeds, T. D., Fragomeni, B. O., Gao, G., Hernandez, A. G., Misztal, I., ... Palti, Y. (2016). Evaluation of genome-enabled selection for bacterial cold-water disease resistance using progeny performance data in rainbow trout: Insights on genotyping methods and genomic prediction models. *Frontiers in Genetics*, 7(96), 1-13. DOI: <https://doi.org/10.3389/fgene.2016.00096>
- VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *Journal of Dairy Sciences*, 91(11), 4414-4423. DOI: <https://doi.org/10.3168/jds.2007-0980>
- Wientjes, Y. C. J., Veerkamp, R. F., & Calus, M. P. L. (2013). The effect of linkage disequilibrium and family relationships on the reliability of genomic prediction. *Genetics*, 193(2), 621-631. DOI: <https://doi.org/10.1534/genetics.112.146290>