

Genomic prediction of ordinal traits via application of the BLASSO and BayesC π methods in simulated data

Geraldo Magela da Cruz Pereira ^{1*}, Andrew de Paula Ribeiro ², Sebastião Martins Filho ¹

¹Universidade Federal de Viçosa (UFV), Departamento de Estatística, Avenida Peter Henry Rolfs, s/nº, Campus Universitário, CEP 36.570-900 Viçosa, MG, Brasil.

²UFV, Departamento de Fitotecnia, Avenida Peter Henry Rolfs, s/nº, Campus Universitário, CEP 36.570-900 Viçosa, MG, Brasil.

*Corresponding author. E-mail: geraldompereira@gmail.com

ABSTRACT. This paper aims at evaluating the use of BLASSO and BayesC π methods for the genomic prediction of ordinal traits, studying factors that influence the performance of the models, and if there is a difference in the ranking of individuals. Genotypic and phenotypic information from a simulated population of 4,100 animals, genotyped by 10k markers (QTL-MAS Workshop) were used. 3,000 animals were used for estimation of the predictive ability and bias accessed through 5-fold cross-validation with five repetitions. The other animals were used as a population of selection. One ANOVA and the Ryan-Einot-Gabriel-Welch test were performed to verify, respectively, which factors influence significantly the genomic prediction and if there is a statistical difference between the models. The results show that the four main factors significantly ($p < 0.05$) affect the predictive ability of GEBVs (genomic estimated breeding values), and that heritability and the number of categories are the most influential factors. Only for ordinal trait 2, with a density of 9k, significant differences ($p < 0.05$) were observed between the predictive ability of the methods. In general, the BayesC π method proved to be more efficient in the identification of relevant SNPs and in the ranking of individuals. Finally, there is a slight superiority of the BayesC π method for the genomic prediction of ordinal traits.

Keywords: Genomic selection, Bayesian methods, ANOVA, Ranking, Molecular markers

DOI: <http://dx.doi.org/10.33837/msj.v3i1.1112>

Received Dec. 7, 2019. Accepted Feb. 17, 2020.

Associate editor: Ana P. S. Siqueira

INTRODUCTION

Methods for genomic prediction have been extensively evaluated for continuous traits in real (Croiseau et al., 2012, Resende Junior et al., 2012) and simulated (Atefi et al., 2016) data. However, it is in the interest of plants and animals breeding to determine the genetic variants associated with production, disease or resistance traits, which are not of a continuous and normally distributed nature (González-Recio & Forni, 2011).

It is known that several traits of economic interest and importance in animal and plant production are measured in an ordinal scale. Kizilkaya et al. (2014) commented that there are phenotypic traits of low heritability that are ordinal in nature, such as susceptibility, resistance to a disease and calving difficulty. In plant breeding, ordinal scores are also

observed for disease resistance or susceptibility. Montesinos-López et al. (2015) used genomic wide selection techniques to predict resistance to gray leaf spot, measured on an ordinal scale, in maize lines, considering three environments.

Methods for genomic prediction of categorical or ordinal traits have been proposed in recent times, as extensions of the methods used for continuous response. Wang et al. (2013) extended the Bayesian methods, BayesA and BayesB, proposed by Meuwissen et al. (2001), and the BayesC π method proposed by Habier et al. (2011), for prediction of genomic estimated breeding values (GEBV) of threshold traits, which originated the BayesTA, BayesTB and BayesTC π methods. The authors have shown that threshold methods are more accurate than their contemporaries when applied to predict genomic breeding values for threshold traits. Pérez & Campos (2014) proposed a class of models, called Bayesian generalized linear regression (BGLR), that allows modeling normal, binary, ordinal or censored traits. The approaches encompassed by BGLR, present the same standard theoretical model, differing in the a priori distribution assumed for the effect of the markers.

Copyright © The Author(s).

This is an open-access paper published by the Instituto Federal Goiano, Urutaí - GO, Brazil. All rights reserved. It is distributed under the terms of the Creative Commons Attribution 4.0 International License.



The accurate estimation of genomic breeding values is of extreme importance in breeding programs for the selection of genetically superior individuals and that they will act as progenitors in the following generations. According to Montesinos-Lopez et al. (2015), the performance of genomic prediction methods is affected by model choice, training population size, marker density, the heritability of the trait, the amplitude of linkage disequilibrium, among other factors. Although the influence of these factors on predictive accuracy for continuous traits has been extensively addressed in the literature, few studies have evaluated how these factors influence the genomic prediction of ordinal traits.

This paper aims at evaluating the use of BLASSO and BayesC π methods for the genomic prediction of ordinal traits. Specifically, the objective was to evaluate: which factors significantly influence the performance of the models; whether the predictive abilities of the models assessed in each scenario differ statistically by the Ryan - Einot - Gabriel - Welch test, and if there is a difference in the ranking of individuals by estimating the Kendall's coefficient of concordance.

MATERIAL AND METHODS

In this study we used simulated data from the 16th QTL-MAS Workshop (Usai et al., 2014) composed of 4,100 animals, genotyped by 10,000 biallelic SNPs (single nucleotide polymorphisms) markers distributed by five chromosomes. The database is made up of genotypic, phenotypic, genealogical information and true breeding values (TBV). Of the

10,000 markers, 9,038 passed the quality control based on the lowest allele frequency of 5% and the call rate of 5%. Random samples of markers were taken from the genotype set to obtain different marker densities (3k, 4.5k, 6k, and 9k). Samples were taken sequentially, so all SNPs contained in the set of 6k markers are also part of the 9k, and so on.

The data were simulated in five generations (G0, G1, G2, G3, and G4), so that each generation was formed by 20 males and 1,000 females, given that in the generation zero only for the males the genotypic information was registered. As the simulated phenotypic traits are relative to milk production, only females were phenotyped (G1, G2, and G3), which totaled 3,000 animals. In the fourth generation, comprised of 1,020 animals, of which 20 are males and 1,000 are females, the animals were only genotyped, constituting themselves a population of selection (Usai et al., 2014).

The phenotypic traits 1, 2 and 3 were simulated, respectively, with heritability equal to 0.35, 0.35 and 0.50. These traits were expressed only by females, and correspond to milk and fat yields, and fat content (Usai et al., 2014). Only the traits 1 and 3 were used in this study, because they present greater differences between their heritabilities. The continuous responses of these traits were transformed into ordinals with two (1 or 2), three (1, 2 or 3) or four (1, 2, 3 or 4) ordinal categories by dividing the continuous responses into quantiles (Figure 1). After the transformation, trait 1 and 3 were renamed as ordinal trait 1 and ordinal trait 2.

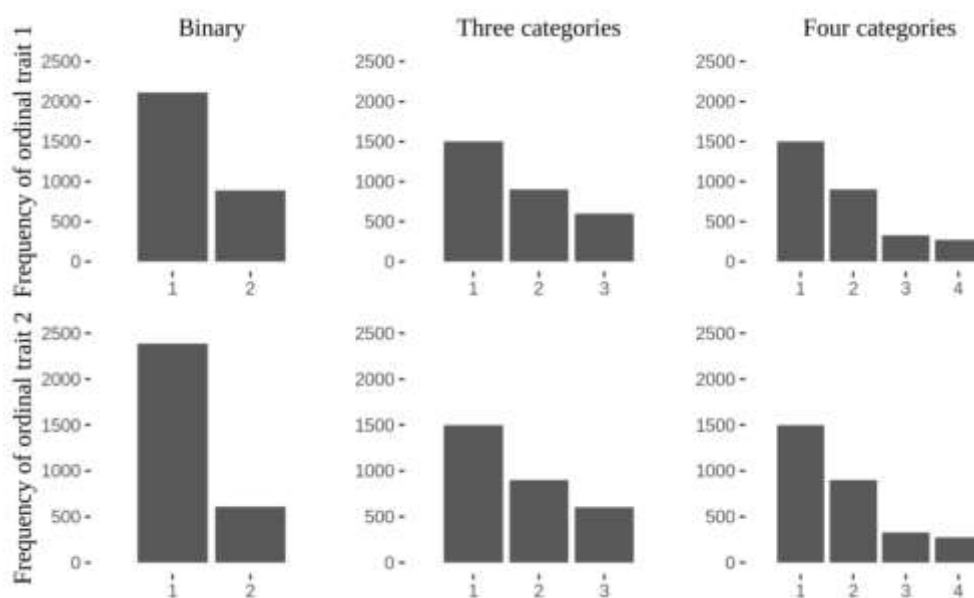


Figure 1. Frequencies obtained for ordinal traits 1 and 2 after the categorization of the continuous response in two, three or four ordinal categories.

Phenotypic traits of ordinal nature can assume C possible values $y_i \in \{1, \dots, C\}$, to $C = 2$ have the binary case. For ordinal response, the probit link function is used to link each category to the linear predictor, by means of $P(y_i = C) = \Phi(\eta_i - \gamma_i) - \Phi(\eta_i - \gamma_{C-1})$, where: $\Phi(\cdot)$ is the cumulative normal distribution function, $\boldsymbol{\eta} = \boldsymbol{\mu} + \mathbf{X}\boldsymbol{\beta}$ is the linear predictor and γ_i the threshold parameter. The probit link function is implemented via data augmentation, through the insertion of the latent variable $l_i = \eta_i + \varepsilon_i$ and the measurement model $y_i = C$ if $\gamma_{C-1} < l_i < \gamma_C$ (Pérez & Campos, 2014).

The BayesC π method proposed by Habier et al. (2011), has the characteristic of assuming that the effects of the markers have homogeneous variances. This method allows some markers to have zero effects, i.e., it excludes from the analyses marker that is not associated with any gene. The following distribution is assumed for the effect of the markers

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

with priori distributions given by $\sigma_a^2 \sim \chi^{-2}(v_a, S_a)$ and $\pi \sim \text{Beta}(b_1, b_2)$, respectively for the additive genetic variance and for the proportion of markers with null effects. The hyper-parameters v_a , S_a , b_1 and b_2 were defined according to Pérez & Campos (2014).

Initially proposed by Park & Casella (2008), the BLASSO (Bayesian Least Absolute Shrinkage and Selection Operator) method was adapted for use in genomic wide selection by Campos et al. (2009). Unlike the BayesC π method, the BLASSO method assumes heterogeneity of variances for the effects of the markers, in addition, to perform the selection of variables indirectly through a strong shrinkage of effects close to zero. The method assumes that the effects of markers follow a double exponential distribution, or equivalently, a mixture of normal distributions. The priori distributions assumed by the model are given by: $\beta_j | \tau_{\beta_j}^2, \sigma_e^2 \sim N(0, \tau_{\beta_j}^2 \sigma_e^2), \tau_{\beta_j}^2 | \lambda \sim \text{Exp}(0.5\lambda^2)$, $\lambda^2 \sim \text{Gamma}(r, s)$ and $\sigma_e^2 \sim \chi^{-2}(v_e, S_e)$, being the hyper-parameters r , s , v_e and S_e , defined a priori according to Pérez & Campos (2014).

The models were compared by 5 - fold cross-validation with five repetitions, through their predictive ability ($r_{GEBV, TBV}$) and bias ($b_{GEBV, TBV}$). The predictive ability was obtained by the Pearson correlation between the genomic estimated breeding values (GEBV) and the true breeding values (TBV) given by the equation $r_{GEBV, TBV} = \text{COR}(GEBV, TBV)$. The bias, obtained through the estimation of the slope coefficient, of the regression of true breeding values on genomic estimated breeding values, can be calculated by: $b_{GEBV, TBV} = \text{COV}(GEBV, TBV) / \sigma_{TBV}^2$. Since the genomic breeding values and true breeding values are not on the same scale, the estimated values for the bias

were transformed to the original scale, using the scale change suggested by Wang et al. (2017), given by: $b_{GEBV, TBV} / \sigma_T$, where σ_T is the true standard deviation threshold.

In the cross-validation process, subjects were equally divided into five subgroups. Four of these groups were used as training populations, in which the genomic prediction models were adjusted and used to estimate the effects of the markers. The fifth group was used as a validation population, in this group were accessed the predictive ability and bias of the models. The procedure was repeated five times so that each of the five groups was part of the validation population only once. The cross-validation process was repeated five times, then mean values for predictive ability and bias were stored for each repetition.

In order to evaluate the significance of the effects of the main factors, heritability, model, the number of categories and density of the markers, a factorial design was used, with five replications, consisting on forty-eight treatments. The assumptions of normality and homogeneity of variances were tested respectively by the Shapiro-Wilk and Bartlett tests. Then, the Ryan-Einot-Gabriel-Welch test (Hsu, 1996) for multiple comparisons was applied at 5% probability.

The methodologies cited above were used to predict GEBVs of individuals in the selection population. At this point, the independent validation process was adopted, considering the 3,000 genotyped and phenotyped females, to estimate the effects of the markers. Then, the estimated effects were used to predict genomic breeding values of the individuals belonging to the selection population, through the expression $GEBV_i = \mathbf{x}_i \hat{\boldsymbol{\beta}}$, where: \mathbf{x}_i represents the genotypic profile of individual i and $\hat{\boldsymbol{\beta}}$ the vector of estimated effects for the markers. Subsequently, individuals were organized in decreasing order of TBV (only the 300 largest TBVs were considered), this ranking was compared with the ranking suggested by each of the models through the Kendall's coefficient of concordance (Kendall & Babington, 1939).

All analyses were performed in software R (R Core Team, 2018), with the help of the BGLR (Pérez & Campos, 2014), DescTools (Signorell et al., 2019) and Agricolae (Mendiburu, 2019) packages. In Bayesian models, the Markov chains were run with 60,000 cycles of the Gibbs sampler, with the first 15,000 cycles being discarded as burn-in and a thin of 3.

RESULTS AND DISCUSSION

The significance of the influence of factors, heritability, number of SNPs, model and number of categories, on the prediction of GEBVs was evaluated by means of an analysis of variance (ANOVA) (Table 1). The ANOVA results show that the four main factors and two

interactions ($h^2 \times \text{nSNP}$ and $\text{Model} \times \text{nSNP}$) significantly ($p < 0.05$) affect the prediction ability of GEBVs.

Among the factors evaluated, heritability is responsible for the greater variability in predictive ability (higher MS), followed by the number of categories, model used and the number of SNPs. Although the number of categories and the density of markers can be controlled by the researcher, heritability is inherent in the trait and cannot be easily controlled. Atefi et al. (2016) obtained results similar to those found in this study when evaluating the accuracy of parametric and semi-parametric models over the generations for continuous traits. According to the authors, the heritability and density of markers emerged as the most influential factors in the prediction accuracy of genomic estimated breeding values, also, a significant interaction ($p < 0.05$) between heritability and number of SNPs was observed.

The values for predictive ability and bias, obtained through 5-fold cross-validation for the BLASSO and BayesC π methods, varied according to the scenario evaluated, respectively, belonging to the ranges [0.639, 0.818] and [1.010, 1.497] (Table 2). When comparing the predictive ability of the methods within the number of categories, it was observed that in the binary case no significant differences between the models were observed for the two ordinal traits evaluated, considering the different marker densities. Already in the case of three or four categories, for the density of 9k, the BayesC π method proved to be significantly superior to the BLASSO method for ordinal trait 2. In all the scenarios evaluated, consistently higher values for predictive ability were observed for the BayesC π method, however, these values were not statistically different from those obtained for the BLASSO. Regarding the bias, we can see that in the binary case BayesC π was the least biased method. For three or four categories, BLASSO showed the smallest bias most of the time.

In the context of continuous and normally distributed responses variables, several papers

explored these methods for prediction in real and simulated data. Croiseau et al. (2012) evaluated productive traits of bulls from the database composed of genomic information of bulls from five European countries, totaling 50 scenarios. The authors showed that, in most of the evaluated scenarios (28), the BayesC π and BLASSO methods presented similar predictive ability. In 20 of these scenarios, the BayesC π method presented a predictive ability superior to the BLASSO method. Resende Junior et al. (2012), comparing several methodologies regarding the accuracy of genomic prediction for 17 pine traits, corroborated with the results presented in this study, showing that for 11 of these traits, the BLASSO and BayesC π methods presented the same predictive ability. The BayesC π method was more accurate than BLASSO for only five traits.

In most of the evaluated scenarios, the increase in heritability and the number of categories implied a significant increase in the predictive ability of the methods. In general, the gain in predictive ability, with an increase in heritability was more remarkable for the density of 9k. For two or three categories, increasing the predictive ability of BLASSO and BayesC π methods with heritability variation was significant at almost all marker densities. Most of the time, the change in the number of categories of binary for three categories was accompanied by an increase in the predictive ability of the models, this increase was less noticeable with the change from three to four categories, this difference being not significant in none of the scenarios. The estimated values for the bias were lower for the trait with higher heritability. These results are corroborated by Wang et al. (2013), which showed that increases in the number of categories of the phenotypic variable and heritability are accompanied by increases in prediction accuracy of genomic breeding values, BayesA, BayesB and BayesC π methods.

Table 1. Output of Analysis of variance for the predictive ability, considering the four main factors: heritability (h^2), model, number of categories (nCategories) and number of SNPs (nSNP), and their respective interactions.

Source of variation	Degrees of Freedom	Sum of Squares	Mean Squares	F-value	p-value ⁽¹⁾
h^2	1	0.149771	0.149771	251.0565	<0.0001
Model	1	0.039008	0.039008	65.3878	<0.0001
nCategories	2	0.224710	0.112355	188.3366	<0.0001
nSNP	3	0.011496	0.003832	6.4233	0.00036
$h^2 \times \text{Model}$	1	0.002020	0.002020	3.3853	0.06732
$h^2 \times \text{nCategories}$	2	0.000784	0.000392	0.6573	0.51943
Model \times nCategories	2	0.000744	0.000372	0.6236	0.53709
$h^2 \times \text{nSNP}$	3	0.009952	0.003317	5.5609	0.00111
Model \times nSNP	3	0.006287	0.002096	3.5127	0.01629
nCategories \times nSNP	6	0.006451	0.001075	1.8022	0.10055
$h^2 \times \text{Model} \times \text{nCategories}$	2	0.001057	0.000528	0.8858	0.41405
$h^2 \times \text{Model} \times \text{nSNP}$	3	0.001824	0.000608	1.0191	0.38534
$h^2 \times \text{nCategories} \times \text{nSNP}$	6	0.000427	0.000071	0.1193	0.99401
Model \times nCategories \times nSNP	6	0.002193	0.000366	0.6128	0.71992
$h^2 \times \text{Model} \times \text{nCategories} \times \text{nSNP}$	6	0.000404	0.000067	0.1130	0.99483
Residual	192	0.114540	0.000597	-	-

Table 2. Mean values for the predictive ability and bias of the BayesC π and BLASSO methods, estimated by 5-fold cross-validation in a simulated population, for the prediction of genomic breeding values associated to two ordinal traits, considering different SNP densities and phenotype categorization in two, three or four categories⁽¹⁾.

nSNP	Model	Measure	Binary		Three categories		Four categories	
			h_1^2	h_2^2	h_1^2	h_2^2	h_1^2	h_2^2
3k	BL	$r_{GBV,TBV}$	0.665Aa	0.700Aa	0.708Aa	0.760Ba	0.714Aa	0.766Ba
		$b_{GBV,TBV}$	1.458	1.038	1.452	1.053	1.455	1.048
	BC	$r_{GBV,TBV}$	0.682Aa	0.725Aa	0.723Aa	0.780Ba	0.728Aa	0.784Ba
		$b_{GBV,TBV}$	1.420	1.052	1.497	1.097	1.491	1.084
4.5k	BL	$r_{GBV,TBV}$	0.656Aa	0.691Aa	0.704Aa	0.757Ba	0.710Aa	0.763Ba
		$b_{GBV,TBV}$	1.419	1.030	1.446	1.042	1.461	1.048
	BC	$r_{GBV,TBV}$	0.667Aa	0.729Ba	0.720Ba	0.786Ca	0.725Ba	0.792Ca
		$b_{GBV,TBV}$	1.403	1.032	1.471	1.049	1.467	1.063
6k	BL	$r_{GBV,TBV}$	0.649Aa	0.671Aa	0.701Ba	0.738Ba	0.707Ba	0.743Ba
		$b_{GBV,TBV}$	1.419	1.045	1.450	1.040	1.446	1.064
	BC	$r_{GBV,TBV}$	0.655Aa	0.689Aa	0.729Ba	0.754Ba	0.732Ba	0.758Ba
		$b_{GBV,TBV}$	1.369	1.01	1.468	1.045	1.457	1.051
9k	BL	$r_{GBV,TBV}$	0.639Aa	0.680Aa	0.695Ba	0.753Ca	0.701Ba	0.760Ca
		$b_{GBV,TBV}$	1.443	1.039	1.484	1.047	1.431	1.042
	BC	$r_{GBV,TBV}$	0.640Aa	0.728Ba	0.741Ba	0.814Cb	0.746Ba	0.818Cb
		$b_{GBV,TBV}$	1.408	1.012	1.471	1.022	1.458	1.043

⁽¹⁾nSNP, number of SNPs; GBV, genomic estimated breeding values; TBV, true breeding values; $r_{GBV,TBV}$, predictive ability; $b_{GBV,TBV}$, bias; h_1^2 , heritability of trait 1; h_2^2 , heritability of trait 2; BL, BLASSO and BC, BayesC π . Means followed by equal letters, uppercase in the lines and lowercase in the columns, do not differ by the Ryan-Einot-Gabriel-Welch multiple comparison test, at 5% probability.

For continuous phenotypic traits, the predictive ability of the methods is expected to increase, or remain constant, with an increasing number of markers (Grattapaglia & Resende, 2011). As shown in Table 1, marker density is the main factor that least influences the predictive ability of genomic breeding values, so it is expected that few significant differences will be found. For the binary case, the increase of the density of the markers did not imply significant gains in the predictive ability, for the two ordinal traits. In the case of three or four categories, only for ordinal trait 2, a significant difference was observed with the change in the number of markers from 6k to 9k for the BayesC π method.

The estimated values for the effects of the markers were plotted for each of the methods, considering the density of 9k markers and the three forms of categorization of the phenotype for the two ordinal traits (Figure 2 and 3). Identifying markers with large effects is an important step in genomic selection, because, by the position of these markers, it is possible to check the existence of QTLs associated with the trait. When comparing the effects estimated for the markers by BLASSO and BayesC π methods, it is noticed that in general, the position of the most relevant SNPs was similar for the two methods.

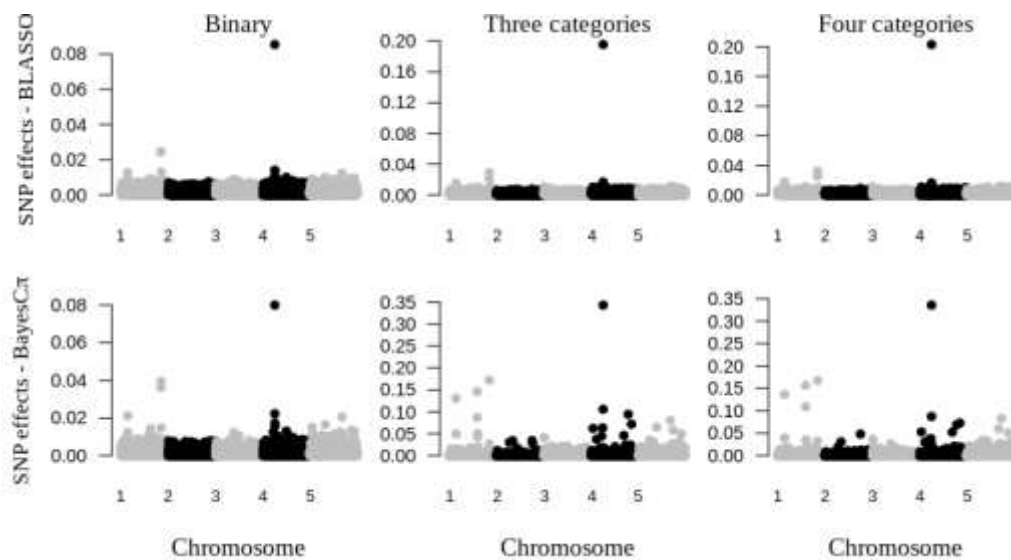


Figure 2. Manhattan plot of the effects of markers, estimated by the BLASSO and BayesC π methods, for the ordinal trait 1 categorized in two, three or four categories.

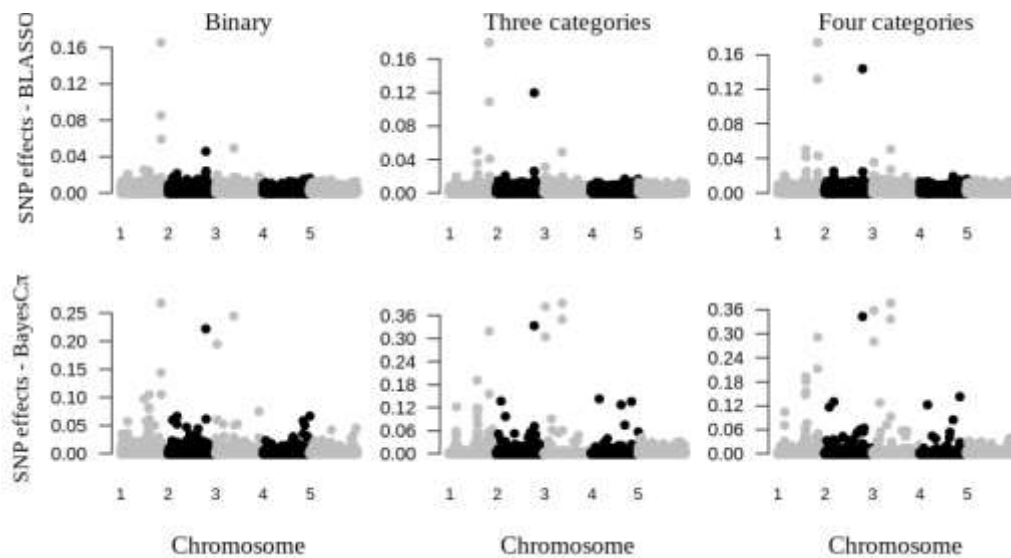


Figure 3. Manhattan plot of the effects of markers, estimated by the BLASSO and BayesC π methods, for the ordinal trait 2 categorized in two, three or four categories.

For the ordinal trait 1, the methods identified an SNP of great effect in the initial position of chromosome four, in the three forms of categorization of the phenotype. The methods suggest the existence of two QTLs, one at the beginning and the other at the end of chromosome one, by the presence of SNPs with moderate effects. For the ordinal trait 2, SNPs of large effects were identified by both methods on chromosome 1. With the increase in the number of categories, the BLASSO method starts to identify a large-effect SNP located at the end of chromosome 2, SNP identified by BayesC π with two categories. The initial part of chromosome 3 presented SNPs with large effects by the BayesC π method, which suggests the existence of a QTL in this region. However, the effects estimated by BLASSO were small. These results are in agreement with those obtained by Grosse-Brinkhaus et al. (2014) who, through Genomic Wide Association Studies (GWAs), identified QTLs for the two traits without categorizing the phenotype, in positions similar to those indicated in Figures 2 and 3.

The positions of the QTLs are more evident when evaluating the effects estimated by the BayesC π method, with three or four categories, than with the effects estimated by the BLASSO method. It can be seen from the graphs that the effect of shrinkage on effects estimates is greater for the BLASSO method than for the BayesC π method. With the variation of the

number of categories, it can be seen that the effects estimated by the methods for the two ordinal traits suffered a stronger shrinkage in the binary case, for three or four categories, few differences were observed in the magnitudes of the estimated effects. The BayesC π method was more sensitive to changes in the number of categories than BLASSO. As the number of categories increases, it is noticed that markers with moderate and small effects are more affected by the shrinkage effect than markers with great effect.

The estimated values for the Kendall's coefficient of concordance (W), for the ranking of the individuals according to the methods and with the TBVs two to two, were significant and greater than 0.65 in all scenarios evaluated (Table 3). The most assertive method is the one that shows the highest agreement with the rankings suggested by the TBVs. A high almost perfect for ordinal trait 1 (0.998) and 0.953 for ordinal trait 2 with two categories. However, with the increase in the number of categories, this agreement decreases, respectively, to 0.94 and 0.88. In all scenarios, the agreement of the BayesC π method, with the ranking suggested by the TBVs, was superior to that obtained by BLASSO. The ranking of individuals assists in the choice of genetically superior individuals, which is of paramount importance in animal and plant genetic improvement.

Table 3. Estimated values for the Kendall's coefficient of concordance for the two ordinal traits, considering the genomic breeding values predicted by the models and true breeding values (TBV), for the ranking of individuals with the 300 largest TBVs⁽¹⁾.

Number of categories	BayesC π vs TBV	BLASSO vs TBV	BayesC π vs BLASSO
Ordinal trait 1			
2	0.687*	0.682*	0.998*
3	0.717*	0.704*	0.937*
4	0.727*	0.706*	0.943*
Ordinal trait 2			
2	0.737*	0.701*	0.953*
3	0.745*	0.727*	0.888*
4	0.739*	0.717*	0.880*

*Significant by chi-square test at 5% probability.

CONCLUSION

Few differences were found between the BLASSO and BayesC π methods, regarding predictive ability. Only for the density of 9k markers, for three or four categories, the BayesC π method proved to be significantly ($p < 0.05$) superior to BLASSO.

Among the evaluated factors, heritability and number of categories appear as factors that most influence the prediction of genomic breeding values.

The BayesC π method is more efficient than the BLASSO method in identifying regions that influence ordinal traits and in the ranking of genetically superior individuals.

ACKNOWLEDGMENTS

To Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) for the scholarship granted.

REFERENCES

- Atefi, A., Shadparva, A. A., & Hossein-Zadeh, N. G. (2016). Comparison of whole genome prediction accuracy across generations using parametric and semi parametric methods. *Acta Scientiarum. Animal Sciences*, 38(4), 447-453.
- Campos, G. de los et al. (2009). Predicting quantitative traits with regression models for dense molecular markers and pedigree. *Genetics*, 182(1), 375-385.
- Croiseau, P., Guillaume, F., & Fritz, S. (2012). Comparison of Genomic Selection Approaches in Brown Swiss within Intergenomics. *Interbull Bulletin*, 46, 127-132.
- González-Recio, O., & Forni, S. (2011). Genome-wide prediction of discrete traits using bayesian regressions and machine learning. *Genetics Selection Evolution*, 43(7), 1-12.
- Grattapaglia, D., & Resende, M. D. V. (2011). Genomic selection in forest tree breeding. *Tree Genetics & Genomes*, 7, 241-255.
- Grosse-Brinkhaus, C., Bergfelder, S., & Tholen, E. (2014). Genome wide association analysis of the QTL MAS 2012 data investigating pleiotropy. *BMC Proceedings*, 8(suppl 5), 1-5.
- Habier, D., Fernando, R. L., Kizilkaya, K., & Garrick, D. J. (2011). Extension of the bayesian alphabet for genomic selection. *BMC Bioinformatics*, 12, 1-12.
- Hsu, J. C. (1996). *Multiple comparisons: theory and methods*. London: Chapman and Hall.
- Kendall, M. G., & Babington Smith, B. (1939). The problem of m rankings. *The Annals of Mathematical Statistics*, 10(3), 275-287.
- Kizilkaya, K., Fernando, R. L., & Garrick, D. J. (2014). Reduction in accuracy of genomic prediction for ordered categorical data compared to continuous observations. *Genetics Selection Evolution*, 46(37), 1-13.
- Mendiburu, F. (2019). *Agricolae: Statistical Procedures for Agricultural Research*, R package version 1.3-1. Available at: <<https://cran.r-project.org/web/packages/agricolae/index.htm>>. Accessed on: 16/01/2020.
- 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.
- Montesinos-López, O. A., Montesinos-López, A., Pérez-Rodríguez, P., Campos, G. de los, Eskridge, K., & Crossa, J. (2015). Threshold Models for Genome-Enabled Prediction of Ordinal Categorical Traits in Plant Breeding. *Genomic selection*, 5, 291-300.
- Park, T., & Casella, G. (2008). The Bayesian Lasso. *Journal of the American Statistical Association*, 103(482), 681-686.
- Pérez, P., & Campos, G. de los (2014). Genome-Wide Regression and Prediction with the BGLR Statistical Package. *Genomic selection*, 198, 483-495.
- R Core Team. (2018). R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. Available at: <<https://www.r-project.org/>>. Accessed on: 02/07/2018.
- Resende Junior, M. F. R. et al. (2012). Accuracy of Genomic Selection Methods in a Standard Data Set of Loblolly Pine (*Pinus taeda* L.). *Genomic selection*, 190, 1503-1510.
- Signorell, A. et al. (2019) DescTools: Tools for Descriptive Statistics, R package version 0.99.31. Available at: <<https://cran.r-project.org/web/packages/DescTools/index.html>>. Accessed on: 16/01/2020.
- Usai, M. G., Gaspa, G., Macciotta, N. P. P., Carta, A., & Casu, S. (2014). XVI th QTLMAS: simulated data set and comparative analysis of submitted results for QTL mapping and genomic evaluation. *BMC Proceedings*, 8, 1-9.
- Wang, C.L. et al. (2013). Bayesian methods for estimating GEBVs of threshold traits. *Heredity*, 110, 213-219.
- Wang, C., Li, X., Qian, R., Su, G., Zhang, Q., & Ding, X. (2017). Bayesian methods for jointly estimating genomic breeding values of one continuous and one threshold trait. *Plos one*, 12(4), 1-18.

To cite this paper, use:

Pereira, G. M. da C., Ribeiro, A. de P., & Martins Filho, S. (2020). Genomic prediction of ordinal traits via application of the BLASSO and BayesC π methods in simulated data. *Multi-Science Journal*, 3(1): 1-7. DOI: <http://dx.doi.org/10.33837/msj.v3i1.1112>