Iranian Society of Animal Science

Journal of Livestock Science and Technologies

Shahid Bahonar University of Kerman

ISSN: 2322-3553 (Print) ISSN: 2322-374X (Online)

Paper type: Original Research

Assessing the predictive performance of the Bagging algorithm for genomic selection

*Farhad Ghafouri-Kesbi**

Department of Animal Science, Faculty of Agriculture, Bu-Ali Sina University, Hamedan, Iran

*Corresponding author,
E-mail address:
f.ghafouri@basu.ac.ir

Received: 13 Jul 2024,
Received in revised form: 08 Sep 2024,
Accepted: 24 Sep 2024,
Published online: 26 Sep 2024,

ORCID
Farhad Ghafouri-Kesbi
0000-0002-2219-055X

Abstract The aim of the present study was to compare the predictive performance of the Bagging algorithm with other decision tree-based methods, including Regression Tree (RT), Random Forest (RF) and Boosting in genomic selection. A genome including ten chromosomes for 1,000 individuals on which 10,000 single nucleotide polymorphisms (SNP) were evenly distributed was simulated. QTL effects were assigned to 10% of the polymorphic SNPs, with effects sampled from a gamma distribution. Predictive performance measures including accuracy of prediction, reliability and bias were used to compare the methods. Computing time and memory requirements of the studied methods were also measured. In all methods studied, the accuracy of genomic evaluation increased following increase in the heritability level from 0.10 to 0.50. While RT was the most efficient user of time and memory, it was not recommended for genomic selection due to its poor predictive performance. The obtained results showed that the predictive performance of Bagging was equal to RF and higher than RT and Boosting. However, it required significantly higher computational time and memory requirements. Considering the overall performance, Bagging was recommended for genomic selection, especially when due to the size and structure of the genomic data, the use of RF is limited.

Keywords: gamma distribution, genomic selection, heritability, regression tree, SNP

Introduction

In recent decades, researchers have attempted to identify the genes or influential chromosomal segments underlying the economic traits for use as molecular genetic information to improve the economic traits in livestock (Budhlakoti et al., 2022). The outcome of these attempts was the marker-assisted selection (MAS) (Fernando and Grossman, 1989) which is an indirect selection process, where individuals for a particular trait of interest are selected based on the known markers linked to the trait (Dekkers, 2004), however, MAS resulted in small increases in genetic improvement compared to the conventional BLUP (Spealman and Garrik, 1997). With the successful completion of the human genome sequencing and advances in polymorphism genotyping technology, including the low cost, and high speed and accuracy, it became possible to select for single nucleotide polymorphism at the genome scale, which was termed genomic selection (GS) (Meuwissen et al., 2001). In GS, a revolutionary advance in the world of animal breeding, the individual effect of each marker is estimated, and by summing all the marker effects, the genomic estimated breeding value (GEBV) of each individual is estimated (Meuwissen et al., 2001). There are certain

factors such as the size of the reference population, genetic diversity in the reference population, heritability of the trait, influence of genotype by environment (G×E) interaction, density of markers, and genetic relationship between reference and breeding populations, which can influence the accuracy of the genomic prediction (Budhlakoti et al., 2022). Besides these factors, the model of genomic selection can also affect the accuracy of GS (Howard et al., 2014; Ashoori et al., 2021). The choice of model is an essential factor in implementing GS, where several parametric and non-parametric models can be used for this purpose (Budhlakoti et al., 2022). Decision-tree based methods, subclass of machine learning approach, are able to model both the additive and epistatic effects for genomic selection (Legarra and Reverter, 2018), including the Regression tree (RT), Random Forest (RF) and Boosting (Howard et al., 2014; Ashoori et al., 2021; Ahmadi et al., 2021). Another method that is in the same family as RF is the Bagging. Bagging, an abbreviation for Bootstrap Aggregating, is a machine learning ensemble strategy for enhancing the reliability and precision of the predictive models (Hastie et al., 2009). Bagging is a kind of regular ensemble classifier technique in which several predictors are made independently and combined using some model averaging methods such as weighted average or majority vote (Jafarzadeh et al., 2021). Although, RF and Boosting have been widely used for genomic selection (Oguto et al., 2011; Howard et al., 2014; Ashoori et al., 2021), little efforts have been made to implement Bagging in the context of genomic selection in livestock. Gianola et al. (2014) combined Bagging and GBLUP (BagBLUP) for estimating the marker effects and reported that bagging can ameliorate the predictive performance of GBLUP and make it more robust against overfitting. Accordingly, Sahebalam et al. (2022) compared BagBLUP with GBLUP and Bayesian LASSO and reported that when the data are stable, the parametric methods (GBLUP and Bayesian LASSO) provided higher prediction accuracy compared to BagBLUP. In this study, we compare the predictive performance of Bagging implemented in the *ipred* package (Peters et al., 2023) with other decision tree-based methods.

Materials and methods

Data simulation

Using the package *hypred* (Technow, 2013) in R (R Development Core Team, 2023), the genome and population were simulated. At first, a base population including 50 males and 50 females was simulated and by using random mating for 50 generations, the LD between the marker and QTL was established. LD was calculated using the r^2 statistic of Hill and Robertson (1968) as follows:

$$r^2 = D^2/\mathrm{freq}(A_1) * \mathrm{freq}(A_2) * \mathrm{freq}(B_1) * \mathrm{freq}(B_2)$$

freq (A1) is the frequency of A1 allele in the population likewise for other alleles in the population. D is the deviation of parental genotypes from the recombinant genotypes and estimated using the haplotype frequencies as follows:

$D=\mathrm{freq}(A_1\text{-}B_1)*\mathrm{freq}(A_2\text{-}B_2)-\mathrm{freq}(A_1\text{-}B_2)*\mathrm{freq}(A_2\text{-}B_1)$

In the generation 51, the size of the population increased to 1000 individual and labeled as the reference population for which the genotypic and phenotypic information were available. Subsequently, the breeding population was generated from the reference population. Animals in the breeding population only had genotypic information by which their genomic breeding values had to be estimated.

The simulated genome included ten chromosomes on which 10,000 single nucleotide polymorphisms (SNP) were evenly distributed. Minor allele frequency was set to 0.05 (Table 1). QTL effects were sampled from a gamma distribution with shape and scale parameters of 0.4 and 1.66, respectively, and assigned to 10% of the polymorphic SNPs (Meuwissen et al., 2001). Using prediction equations, the effects of all SNPs were estimated in the reference population by combining the genotypic and phenotypic information of all individuals. Then, the genomic breeding values of the selection candidates were estimated by summing the effects of all SNPs they carry according to SNPs effects previously estimated in the reference population.

Table 1. Parameters used for simulation program

Genome size	1,000 cM
Number of chromosomes	10
Number of SNP marker	10,000
Distribution of additive QTL effects	Gamma
Number of QTL	1,000 (10% of SNP)
Effective population size (*Ne*)	100
Heritability	0.10, 0.30, 0.50
Minor allele frequency (MAF)	0.05

Methods of genomic prediction

Regression tree

When the trait follows a continuous distribution (e.g., daily milk production), regression tree (RT) is used. Let $\boldsymbol{y}$ (n×1) be the vector observations, and $\boldsymbol{X} = \{\boldsymbol{x_i}\}$, where $\boldsymbol{x_i}$ is a (p×1) vector representing the genotype of each animal for p SNP. The RT model can be represented as follows:

$$\Psi(\boldsymbol{y}, \boldsymbol{X})$$

The RT is constructed as follows: 1) different samples from the training data set, i.e., {(x1, y1), . . . , (xn, yn)}, are drawn with replacement, 2) a small group of SNP are randomly selected from the p SNP marker and the SNP j which minimizes the lose function is selected, 3) according to the genotype of SNP j, the node is split in two child nodes and individuals go to one of the child nodes according to the SNP alleles they carry, 4) steps 2-3 are repeated until a minimum node size is reached and all the terminal nodes become maximally homogeneous. The predicted value of the genotype $\boldsymbol{x_i}$

for regression problems is the average phenotype of the individuals in the node (Gonzalez-Recio and Forni, 2011). The package *rpart* (Therneau et al., 2019) was used to run RT.

Bagging and Random Forest

Both Bagging and RF regression use an ensemble of regression trees, grown on bootstrap samples of the observations using a random subset of predictors to define the best split at each node. The prediction for a new observation $x(\hat{f}_{rf}^{B}(x))$, is computed by averaging the predictions over B trees, $\{T(x,\Psi_b)\}_1^B$, in which Ψ_b characterizes the b_{th} tree in terms of split variables, cut points at each node, and terminal node values. Bagging and RF were fitted using the following model (Ogutu et al., 2011):

$$\hat{f}_{rf}^{B}(x) = \frac{1}{B}\sum_{b=1}^{B} T(x,\Psi_b),$$

The fundamental difference between Bagging and RF is that in RF, only a subset of features is selected at random out of the total and the best split feature from the subset is used to split each node in a tree, but in Bagging all features are considered for splitting a node. RF has tree important parameters: the number of trees to grow (*ntree*), the number of SNPs randomly selected at each tree node (*mtry*), and the minimum size of terminal nodes of trees, below which no split is attempted (*nodesize*). An optimum combination of these parameters should be used while running RF. The combination that provided the highest accuracy [*ntree* =1500, *mtry* = 500 and *nodesize* =25] was used to analyze the data in each scenario. The package *randomForest* (Liaw and Wiener, 2018) was used to run RF. Important parameters in Bagging are *nbag* which is used to control the number of decision trees voting in the ensemble and complexity parameter (*cp*) which is used to control the size of the decision tree and to select the optimal tree size. Bagging was run with different combinations of these parameters, and the combination which provided the highest accuracy was: *nbag* =1000 and *cp* =500. This combination was used to analyze the data in each scenario. Bagging was fitted to the data using the *ipred* package (Peters et al., 2023).

Boosting

Boosting belongs to a family of machine learning algorithms that convert the weak learners to strong learner. Here, regression trees were the weak learners. The Boosting adds regression trees to the residual (misclassified inputs) of the previous regression tree in such a way that by adding the new trees, the error function is decreased (Hastie et al., 2009; Oguto et al., 2011):

$$f(x) = \sum_{m=1}^{M} \beta_m b(x;\gamma_m)$$

where, β_m, m=1,2,.., M are basis expansion coefficient, and $b(x;\gamma_m)$ are simple functions of the multivariate argument, with a set of parameters $\gamma = (\gamma_1,\gamma_2,\ldots,\gamma_M)$.

Prediction is accomplished by weighting the ensemble outputs of all regression trees. Boosting was carried out using the package *gbm* (Greenwell et al., 2019). The tuning parameters in Boosting were the number of trees (*ntree*), tree depth or tree complexity (*tc*) and shrinkage rate or learning rate (*lr*). A series of values for each parameter was specified and the performance of the model with each combination of the tuning parameters was evaluated. In the model with highest predictive performance, these tuning parameters were: *ntree* = 1500, *tc* = 7 and *lr* = 0.02.

Comparison of methods

Prediction accuracy: This criterion was calculated as the Pearson's correlation between the predicted and true (simulated) breeding values (Legarra and Reverter, 2018):

$$\text{Pearson's correlation} = \frac{cov(true\ values, predicted\ values)}{var(predicted\ values)var(true\ values)}$$

Bias: This criterion was calculated as the difference between the average predicted breeding values and true breeding values (Legarra and Reverter, 2018):

$$Bias = \overline{predicted\ values - true\ values}$$

Reliability: Is the regression of predicted breeding values on true breeding values (Legarra and Reverter, 2018):

$$Reliability = \frac{cov(predicted\ values, true\ values)}{var(true\ values)}$$

The prediction model was analyzed ten times and the average computing time, memory requirement, accuracy, bias and reliability for each scenario was presented using a combination of heritability values (0.10, 0.30, 0.50).

Memory requirement and computing time

The package *pryr* (Wickham, 2018) was used to record the memory usage for each method. It records the amount of memory occupied by the objects created by executing a function in R. The computing time in each scenario was monitored and recorded with an R function. The computing time was measured as the time consumed for executing the codes of the methods studied, and did not include the time consumed for the simulation of population and genome.

Results

Figure 1 shows the prediction accuracy of the RT, Bagging, RF and Boosting at different levels of heritability (0.10, 0.30 and 0.50). By increasing the heritability from 0.10 to 0.50, prediction accuracy increased between 57% (Bagging) to 158% (RT); the lowest accuracy of prediction was attributed to RT. RF was superior to all other methods especially at lower levels of heritability and, therefore, ranked as the first. The accuracy of Bagging was equal to RF. At heritability 0.10, Bagging was superior to Boosting with a significant difference. At the heritability of 0.50, the differences

between Boosting with RF and Bagging was not significant (P>0.05).

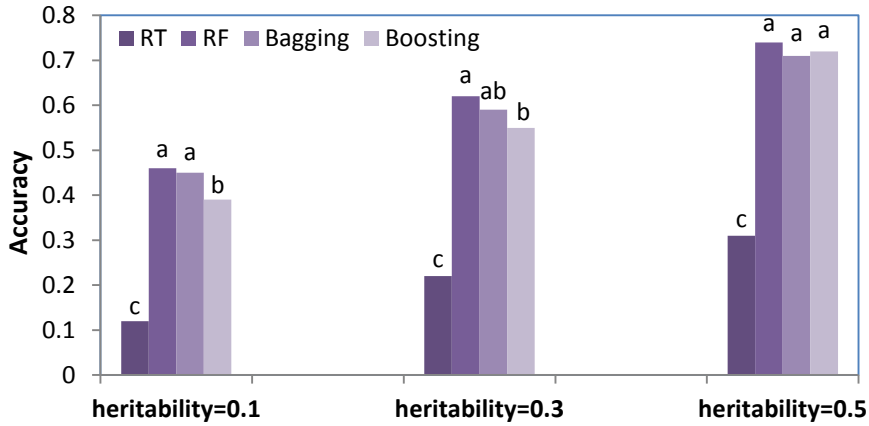

Figure 1. Accuracy of the studied methods in different scenarios of heritability

The bias of predicted GEBVs for different methods is shown in Figure 2. With increase in heritability from 0.10 to 0.50, the prediction bias decreased in a range from 32% (RF) to 35% (RT). RT and RF provided GBVs with maximum and minimum bias, respectively. The differences were more obvious at heritability=0.10, in which, the bias of RF was 3%, 20% and 73% smaller than Bagging, Boosting and RT, respectively. The difference between Bagging and Boosting was significant at all heritability levels, and accordingly, Bagging was ranked as the second.

Figure 3 shows the reliability of predicted GEBVs for different methods. There were 32% (RF) to 61% (RT) increases in reliability with increases in heritability values from 0.10 to 0.50. Regarding the reliability, the predictions of RF and Bagging showed significantly (P<0.05) higher reliability than Boosting and RT and, therefore both of which were ranked as the first. RT provided GEBVs with the lowest reliability and, therefore, was ranked as the last one.

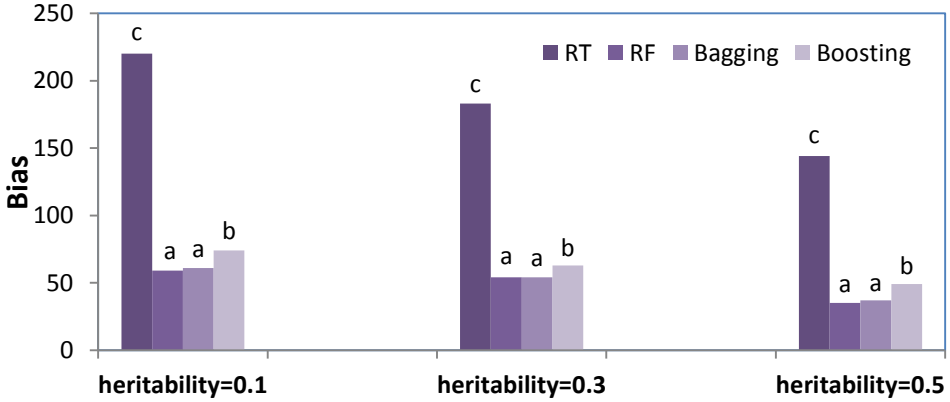

Figure 2. Bias of the studied methods in different scenarios of heritability

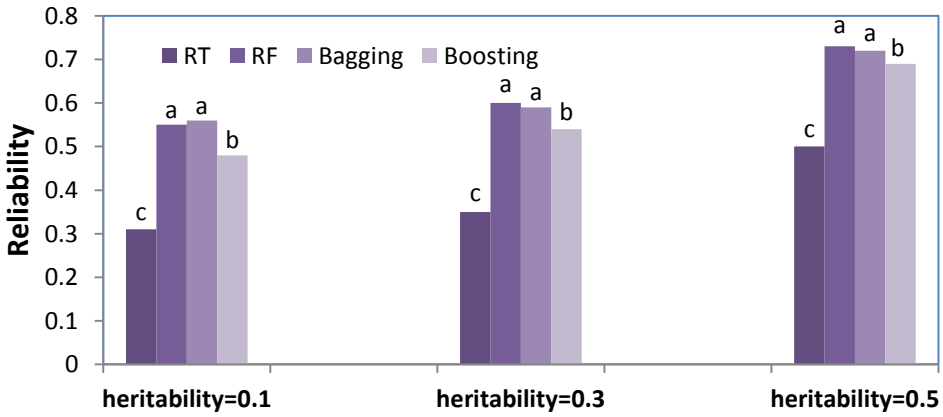

Figure 3. Reliability of the studied methods in different scenarios of heritability

The computing time for different methods is shown in Figure 4. The RT (0.81 minute) was the fastest method and Boosting (15.12 minutes) was the slowest one. Concerning the memory requirement, RT, RF, Bagging and Boosting ranked as the first, second, third and last fastest method. The memory requirement of the studied traits (Figure 5) ranged from 0.09 GB (RT) to 3.24 GB (Boosting) (Figure 5).

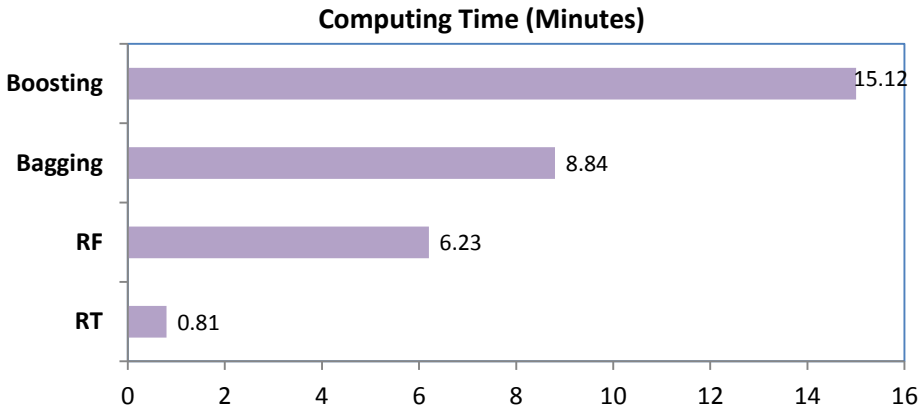

Figure 4. Computing time of the studied methods

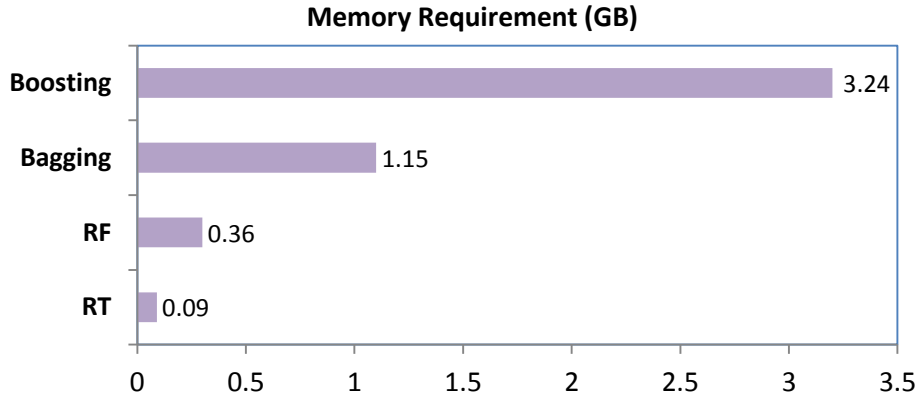

Figure 5. Memory requirement of the studied methods

Discussion

The effect of heritability on the accuracy of genomic evaluation has been studied by several authors including Hayes et al. (2010), Mohammadi Chamachar et al. (2015), Zhang et al. (2017), Ahmadi et al. (2021) and Ashoori-Banaei et al. (2021). A common finding of these studies is that by increasing the heritability, the accuracy and reliability of GEBVs increased and the bias of GEBVs decreased. For example, Hayes et al. (2010) reported that at heritability levels of 0.10, 0.30, 0.50, 0.70 and 0.90, the accuracy of genomic evaluation was 0.35, 0.50, 0.60, 0.65 and 0.72, respectively. By increasing the heritability, the effect of environmental noises on the phenotypic variation decreased (Ahmadi et al., 2021). Therefore, most of the phenotypic variation is caused by the genetic variation. In such a situation, the power of models to extract SNPs effects increases leading to increased accuracy (Ahmadi et al., 2021; Ashoori-Banaei et al., 2021).

The accuracy of prediction, the speed of calculations and the amount of RAM required are the factors that affect the overall performance of genomic prediction method (Ashoori-Banaei et al., 2021). The current findings showed that fitting a regression tree on the data did not provide GEBVs with acceptable accuracy, bias and reliability. Therefore, it could not be recommended for genomic selection. Complexity, instability and unwieldy are disadvantages of RT. While decision trees are simple and interpretable models for regression and classification**,** they suffer from high bias and high variance which makes them less useful for most practical applications (Ashoori et al., 2021). Hastie et al. (2009) emphasized that the decision trees' low predictive accuracy can be improved by the use of their refinements methods such as Bagging, RF and Boosting. These strategies combine multiple RT to reduce the variability and build more accurate prediction models; for example, Bagging reduces the variance observed in decision trees, RF improves accuracy by avoiding high tree correlation, and Boosting reduces the

error by building trees sequentially using information from previously built trees (Valiati Barreto et al., 2024). Our findings showed that Bagging could provide GEBVs with accuracy, bias and reliability equal to RF and higher than Boosting and RT. For economic traits in livestock, Bagging has not been compared with other genomic selection methods in terms of predictive performance, computing time and memory requirement. However, in maize, Valiati Barreto et al. (2024) used RF, Bagging and Boosting for genomic evaluation of grain yield and flowering time and reported that Bagging and RF produced very similar predictions with is in agreement with our findings. Also, de Sousa et al. (2020) compared regression-tree based models (RT, RF, Bagging and Boosting) for genomic selection of leaf rust resistance to Arabica coffee and reported that Bagging presented higher accuracy and lower apparent error rate compared to other methods studied. According to James et al. (2013), the advantage of Bagging over RF depends on the number of correlated predictor variables (SNPs) in the dataset. If there are many correlated predictor variables, RF will be superior to Bagging. Otherwise, RF does not result in improvement compared to Bagging. In addition, de Sousa et al. (2020) reported that since rust resistance is an oligogenic trait, the framework of RF, whereby the division of the nodes in the RT is performed using a small random number of markers, it is possible that certain nodes could have only chosen the markers that were not associated with the trait, thus explaining the lower performance of RF compared with Bagging. Our simulated scenarios were purely additive. Therefore, the high percentage of agreement between the RF and Bagging could indicate that the simulated trait does not present complicating factors for modeling such as the dominance, epistasis and imprinting to reveal the latent abilities of these methods in complex modeling of different genetic effects. As a result, it is the genetic architecture of the traits which determines the winner of the competition.

The GBLUP equipped with Bagging, i.e., BagBLUP has been tested for genomic selection both with empirical and simulated data. Gianola et al. (2014) reported that bagging can ameliorate the predictive performance of GBLUP and make it more robust against over-fitting. Abdollahi-Arpanahi (2015) used BagBlup for genomic selection of broiler chicken traits and reported that Bagging did reduce the variability of GBLUP predictions and enhanced the predictive performance when the model was 'under-regularized'. Sahebalam et al. (2022) compared BagBlup, GBLUP and Bayesian LASSO and reported no significant difference between methods in terms of the accuracy of prediction.

The methods with higher memory requirement (Figures 4 and 5) also needed longer time to finish the same job. When comparing the computational cost of the techniques to Bagging, only the Boosting required more computational time and memory (being 1.70 times slower and needed 2.90 times more memory to finish the same analysis) inconsistence with de Sousa et al. (2020). It could be a limitation for Bagging especially where large genomic datasets are used for genomic selection.

Conclusion

Bagging showed a decent performance in terms of the prediction accuracy, computing time and memory requirement. Regarding the predictive performance, Bagging presented equal outcome compared with RF. However, computational costs of Bagging were noticeably higher than RF which can limit the overall performance of this method, especially when analysis includes big dataset. Because, according to previous reports, the superiority of RF over Bagging varies with the size and structure of the genomic data, it is reasonable to apply both models and use the results of the better model. Based on the predictive performance and computational costs, the RT and Boosting recorded poor performance compared to RF and Bagging and, therefore, may not be recommended for genomic selection.

References

Abdollahi-Arpanahi, R., Morota, G., Valente, B.D., Kranis, A., Rosa, G.J.M., Gianola, D., 2015. Assessment of bagging GBLUP for whole-genome prediction of broiler chicken traits. *Journal of Animal Breeding and Genetics* 132, 218-228.

Ahmadi, Z., Ghafouri-Kesbi, F., Zamani, P., 2021. Assessing the performance of a novel method for genomic selection: rrBLUP method6. *Journal of Genetics* 100, 24.

Ashoori-Banaei, S., Ghafouri-Kesbi, F., Ahmadi, A., 2021. Comparison of regression tree-based methods in genomic selection. *Journal of Genetics* 100, 85.

Budhlakoti, N., Kushwaha, A.K., Rai, A., Chaturvedi, K.K., Kumar, A., Pradhan, A.K., 2022. Genomic selection: a tool for accelerating the efficiency of molecular breeding for development of climate resilient crops. *Frontiers in Genetics* 13, 66.

Dekkers, J.C., 2004. Commercial application of marker- and gene-assisted selection in livestock: strategies and lessons. *Journal of Animal Science* 82, E313-E32

de Sousa, I.C.D., Nascimento, M., Silva, G.N., Nascimento, A.C.C., Cruz, C.D., Silva, F.F., Almeida, D.P.D., Pestana, K.N., Azevedo, C.F., Zambolim, L., 2021. Genomic prediction of leaf rust resistance to Arabica coffee using machine learning algorithms. *Scientia Agricola* 78, e20200021.

Fernando, R.L., Grossman, M., 1989. Marker-assisted selection using best linear unbiased prediction. *Genetic Selection Evolution* 2, 246-477.

Greenwell, B., Bradley, B., Cunningham, J., 2019. *gbm*: Generalized Boosted Regression Models. Available at:

https://cran.r-project.org/web/packages/gbm/index.html

Gianola, D., Weigel, K.A., Kramer, N., Stella, A., Schon, C.C., 2014. Enhancing genome-enabled prediction by bagging genomic BLUP. *PlosOne*: e91693.

González-Recio, O., Forni, S., 2011. Genome-wide prediction of discrete traits using Bayesian regressions and machine learning. *Genetics Selection Evolution* 43, 7.

Hastie, T.J., Tibshirani, R., Friedman, J., 2009. The Elements of Statistical Learning. 2nd ed. Springer, New York, USA.

Hayes, B.J., Daetwyler, H.D., Bowman, P., Moser, G., Tier, B., Crump, R., Khatkar, M., Raadsma H.W., Goddard M.E., 2010 Accuracy of genomic selection: comparing theory and results. In: Proceedings of the 18th Conference of the Association for the Advancement of Animal Breeding and Genetics. Barossa Valley (Australia). 27 September–2 October 2009, (Barossa Valley, Australia).

Hill, W.G., Robertson, A., 1968. Linkage disequilibrium in finite populations. *Theoretical and Applied Genetics* 38, 226-231.

Howard, R., Carriquiry, A.L., Beavis, W.D., 2014. Parametric and nonparametric statistical methods for genomic selection of traits with additive and epistatic genetic architectures. *Genetics* 4, 1027-1046.

Jafarzadeh, H., Mahdianpari, M., Gill, E., Mohammadimanesh, F., Homayouni, S., 2021. Bagging and boosting ensemble classifiers for classification of multispectral, hyperspectral and PolSAR data: A comparative evaluation. *Remote Sensors* 13, 4405.

James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. An Introduction to Statistical Learning: with Applications in R. Springer, New York, USA.

Legarra, A., Reverter, A., 2018. Semi-parametric estimates of population accuracy and bias of predictions of breeding values and future phenotypes using the LR method. *Genetic Selection Evolution* 50, 53.

Liaw, A., Wiener, M., 2018. Breiman and Cutler's random forests for classification and regression. Available at: http://cran.r-project.org/web/packages/randomForest/index.html.

Mohammadi-Chamachar, N., Hafezian, S.H., Honarvar, M., Farhadi, A., 2015. Effects of heritability and number of quantitative trait loci (QTL) on accuracy of genomic estimated breeding value. *Journal of Ruminant Research* 3, 111-124 (in Farsi with English Abstract).

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

Oguto, J.O., Piepho, H.P., Schulz-Streeck, T., 2011. A comparison of random forests, Gradient Boosting and support vector machines for genomic selection. *BMC Proceedings* 5, 11.

R Development Core Team., 2023. R: A language and environment for statistical computing. Vienna, Austria. Available at: https://www.R-project.org/.

Spelman, R., Garrick, D., 1997. Utilization of marker assisted selection in a commercial dairy cow population. *Livestock Production Science* 47, 139-147.

Sahebalam, H., Gholizadeh, M., Hafezian, H., Ebrahimi, F., 2022. Evaluation of Bagging approach versus GBLUP and Bayesian LASSO in genomic prediction. *Journal of Genetics* 101, 19.

Technow, F., 2013. *hypred*: Simulation of genomic data in applied genetics. Available at: https://github.com/cran/hypred.

Therneau, T., Atkinsonm., B., Ripley, B., 2019. *rpart*: Recursive partitioning for classification, regression and survival trees. An implementation of most of the functionality of the 1984 book by Breiman, Friedman, Olshen and Stone. Available at: https://cran.r-project.org/web/packages/rpart/index.html.

Valiati Barreto, C.A., Dias, K., de Sousa, I.C., Azevedo, C.F., Nascimento, A.C.C., Guimarães, L.G.M., Guimarães, C.T., Pastina, M.M., Nascimento, M., 2024. Genomic prediction in multi‑environment trials in maize using statistical and machine learning methods. *Scientific Reports* 14, 1062.

Wickham, H., 2018. *pryr*: Useful tools to pry back the covers of R and understand the language at a deeper level. Available at: https://cran.r-project.org/web/packages/pryr/index.html.

Zhang, A., Wang, H., Beyene, Y., Semagn, K., 2017. Effect of trait heritability, training population size and marker density on genomic prediction accuracy estimation in 22 bi-parental tropical maize populations. *Frontiers in Plant Science* 8, 1916.