Indian J. Genet., (2023); 83(1): 88-94
doi: 10.31742/ISGPB.83.1.11
ISSN: 0975-6906
www.ijgpb.org

RESEARCH ARTICLE

Comparative study of multi-trait genomic and phenotypic selection indexes for selection of superior genotypes

Md Asif Khan, Anil Rai, D. C. Mishra*, Neeraj Budhlakoti, Subhrajit Satpathy and Sayanti Guha Majumdar

Abstract

Genomic selection (GS) emerged as an efficient and cost-effective breeding technique that selects individuals based on their genetic merit via the prediction of genomic estimated breeding values (GEBVs) using molecular markers distributed over the entire genome. Genomic selection index (GSI) is a linear combination of GEBVs, while the phenotypic selection index (PSI) is a linear combination of multiple observable phenotypic traits. In this study, we compared the predictive performance of five parametric GS models such as RR-BLUP, Bayesian LASSO, Bayes A, Bayes B, and Bayes C for estimating GSI. Further, the GSI and PSI efficiency of breeding candidates was evaluated by applying suitable evaluation measures such as correlations of each indices with the net genetic merit, selection response, and expected genetic gain per trait. The findings of this study were further validated by two real datasets, suggesting that the GSI was more efficient than the PSI per unit of time.

Keywords: GEBVs, genomic selection, net genetic merit, selection index, selection response, genetic gain, GSI, PSI

Introduction

The conventional breeding approach for the genetic improvement of plants and animals using economically important quantitative traits solely relies upon phenotypic characters and available pedigrees. The advancement in molecular genetic techniques led to the availability of genome-wide high-density markers and high throughput genotyping techniques, which have made it possible to predict the breeding value of individuals through marker-assisted selection (MAS) more efficiently and accurately. MAS is an indirect selection approach of a trait of interest using marker information and integrating with phenotypic information (Fernando and Grossman 1989). However, MAS is considered as effective and more efficient only for the traits associated with one or a few major genes with large effect (Bernardo 2008). MAS has limitations of long selection cycles and is less effective for polygenic traits (i.e. traits govern by a few hundred to thousands of genes) (Heffner et al. 2009; Goddard and Hayes 2007; Xu et al. 2012).

To overcome the limitations of MAS, Meuwissen et al. (2001) proposed Genomic Selection (GS), where whole genome marker information is used to estimate the genetic merit of individuals. GS is a promising tool for improving the genetic gain of individuals for plant and animal breeding research in the current scenario. A typical GS process starts with forming a training population consisting of both genotypic (marker data) and phenotypic information (a trait of interest); this information is further used to build the statistical model and estimate the marker effects accordingly. Further, the estimate of the marker effects utilized to obtain genomic estimated breeding values (GEBVs), for the individuals in the testing/breeding population (*i.e.,* for which only genotypic information is available) (Heffner et al. 2009).

Current single-trait (univariate) GS (STGS) models do not perform well in scenarios like pleiotropy, missing data, and a trait with low heritability (Jia and Jannink 2012; Budhlakoti et al. 2019). To solve all these constraints, the multi-trait GS(MTGS) models are used to estimate the GEBV more precisely and accurately (Jia and Jannink 2012; Guo et al. 2014; Budhlakoti et al. 2019; Mishra et al. 2021). MTGS based

ICAR-Indian Agricultural Statistics Research Institute, Pusa, New Delhi 110 012, India

***Corresponding Author:** D. C. Mishra, ICAR-Indian Agricultural Statistics Research Institute, Pusa, New Delhi 110 012, India: E-mail: dwij.mishra@gmail.com

How to cite this article: Khan M.A., Rai A., Mishra D.C., Budhlakoti N., Satpathy S. & Majumdar S.G. 2023. Comparative study of multi-trait genomic and phenotypic selection indexes for selection of superior genotypes. Indian J. Genet. Plant Breed., **83**(1): 88-94.

Source of support: Nil

Conflict of interest: None.

Received: May 2022 **Revised:** Dec. 2022 **Accepted:** Dec. 2022

models uses the information of genetically correlated traits phenotypic information and low heritability to improve genomic prediction (Jia and Jannink 2012; Guo et al. 2014). On the other hand, combining multi-trait information with efficient selection index structure is often a complex and challenging task. The selection index (SI) is a linear combination of optimally weighted multiple traits to obtain greater overall genetic gains than if traits with independent thresholds are selected separately or collectively (Hazel and Lush 1942; Hazel 1943). Breeders and biometricians have advocated the use of both phenotypic and genomic selection indexes (Ceron-Rojaset al. 2015). The phenotypic selection index (PSI) is an optimally weighted linear combination of various observable phenotypic trait values (Smith 1936; Hazel and Lush 1942; Hazel 1943), whereas the genomic selection index (GSI) represents a linear combination of GEBVs used to evaluate an individual's net genetic merit and accordingly select superior individuals from the breeding population.

Estimation of marker effects to obtain GEBVs by a simple linear regression statistical model encounters the problem of over-parameterization and multi-colinearity in whole-genome regression. To overcome this limitation, penalized and Bayesian regression models were suggested to estimate the marker effects (Meuwissen et al. 2001; de los Campos et al. 2011; Gianola 2013). Ridge regression (RR) uses penalized least squares techniques by shrinking the effect of the markers equally toward zero. However, RR considers that all the predictors (i.e. markers in our case) contribute to equal variance, which is not true for all the traits (Piepho 2009; Meuwissen et al. 2001). Several Bayesian models are available to solve this issue, which assume some prior distribution for marker effects. The Bayesian approach, based on the marker effects' posterior distributions, provides inference about model parameters (Meuwissen et al. 2001; Jia and Jannink 2012; Wang et al.2018). In this study, MTGSbased models such as standard penalized regression (RR-BLUP) and Bayesian regression like Bayes A, Bayes B, Bayes C, and Bayesian LASSO (Meuwissen et al. 2001; Habier et al. 2011; VanRaden 2008) are used to construct the GSI.

Lande and Thompson (1990) introduced a marker-assisted selection (MAS) based selection index that incorporates marker and phenotypic information and evaluated the same using simulated data, further concluded that the estimated selection response was higher than when only the phenotypic data were used. Togashi et al. (2011) suggested four selection indices, based on the BLUP theory, in the context of GS; however, these indices are not validated with any real or simulated dataset. Ceron-Rojaset al. (2015) proposed PSI and GSI theory to estimate the net genetic merit, selection response and expected genetic gains per trait that enable the breeders with an objective criteria for evaluating and selecting parents based on single and multiple traits and also evaluated these indices using real and simulated data. The main purpose of both indices (PSI and GSI) is to provide objective criteria to breeders to select the candidates as parents for a breeding programme based on their net genetic merit and maximize selection response.

Therefore, this study aimed at applying PSI methodology and multi-trait based GS models:standard penalized regression method (RR-BLUP) and Bayesian regression methods (Bayesian LASSO, Bayes A, Bayes B, and Bayes C) to estimate GSIs from breeding population and to evaluate the performance of PSI and GSIs on real datasets using suitable evaluation measure.

Materials and methods

This study, MTGS based models, such as standard penalized regression (RR-BLUP) and Bayesian regression (i.e., Bayes A, Bayes B, Bayes C, and Bayesian LASSO) to construct the GSI. One such basic model used in MTGS is the multivariate linear mixed model; where the marker effects are treated as random. Details of same is as follows (Jia and Jannink 2012):

$$\mathbf{y} = \mathbf{u} + \sum_{j=1}^{p} X_j \boldsymbol{a}_j \delta_j + \boldsymbol{e} \qquad (1)$$

where $\mathbf{y}$ is a matrix $(n \times q)$ of q phenotypic traits of n individuals: $\mathbf{u}$ is the overall population mean; X is a design matrix $(n \times p)$ corresponding to the p marker genotypes of n individuals; $\boldsymbol{a}_j$is a vector $(1 \times q)$ for the j marker effects on all q phenotypic traits and assumed normally distributed $\boldsymbol{a}_j \sim N(0, \Sigma_{a_j})$; Σ_{a_j} *is* the variance-covariance matrix $(q \times q)$ for marker; j, δ_j is an indicator variable with a value of 1 in the presence of a marker j otherwise, its value is 0; $\boldsymbol{e}$ is a residual matrix $(n \times q)$ with row variance $\boldsymbol{\Sigma}_{\boldsymbol{e}}(q \times q)$ (Jia and Jannink 2012; Wang et al. 2018).

Hereby, MTGS models are used in combining optimal attributes for multiple traits simultaneously to construct an efficient selection index. A selection index or total score is constructed by combining all the traits' values corresponding to their weight components for all the traits simultaneously. The weight allocated to each trait is determined by its relative economic value, heritability, and genetic and phenotypic correlations among the traits (Smith 1936). It may be based on phenotypic (i.e., Phenotypic Selection Index) or genomic (Genomic Selection Index) information.

Phenotypic Selection Index (PSI)

PSI is a linear combination of several observable phenotypic traits with optimal economic weighted values and same can be expressed in the following form (Smith 1936):

$$I_p = \boldsymbol{b}'\boldsymbol{p}(2)$$

$$\boldsymbol{b} = \boldsymbol{P}^{-1}\boldsymbol{C}\boldsymbol{w}(3)$$

where, $\boldsymbol{p}' = [\,p_1 \quad p_2 \quad \dots \quad p_q\,]$ be a vector of q phenotypic trait values and, $\boldsymbol{b}' = [b_1 \quad b_2 \quad \dots \quad b_q]$ is the $I_p$vector of coefficients. $\boldsymbol{P}^{-1}$ is the inverse of the phenotypic covariance matrix ($\boldsymbol{P}$), and $\boldsymbol{C}$ is the covariance matrix of true breeding values $\boldsymbol{g}$.

Genomic Selection Index (GSI)

The GSI (I_G) can be defined as (Ceron-Rojas et al. 2015):

$$I_G = \boldsymbol{w}'\boldsymbol{\gamma}(4)$$

where, $\boldsymbol{\gamma}' = [\gamma_1\gamma_2 \cdots \gamma_q]$ is a $1 \times q$ vector of genomic breeding values and $\boldsymbol{w}' = [w_1 \quad w_2 \quad \dots \quad w_q]$ is a vector of economic weights for individuals.

The GEBV ($\boldsymbol{\gamma}$) can be calculated as (Ceron-Rojas et al. 2015):

$$\hat{\boldsymbol{\gamma}} = \boldsymbol{X}\hat{\boldsymbol{\alpha}}(5)$$

where, $\hat{\boldsymbol{\alpha}}$ will be used as the estimator of the vector of marker effects $\boldsymbol{\alpha}' = [\boldsymbol{\alpha}'_1 \quad \boldsymbol{\alpha}'_2 \dots \quad \boldsymbol{\alpha}'_q]$ for q traits in the base population and $\boldsymbol{X}$ is anincidence matrix of size $n \times p$ of the coded values of marker genotypes.

The genetic merit (H) can be written as:

$$H = \boldsymbol{w}'\boldsymbol{g}(6)$$

where, $\boldsymbol{g}' = [g_1 \quad g_2 \quad \dots \quad g_q]$ be a vector of true breeding values and $\boldsymbol{w}' = [w_1 \quad w_2 \quad \dots \quad w_q]$ is a vector of economic weights for individuals.

Performance evaluation measures

- Response to selection

PSI (R_P) and GSI (R_G) selection response Ceron-Rojas et al. (2015) can be explained as (Ceron-Rojas et al. (2015)):

$$R_P = \frac{k}{L_P}\sqrt{\boldsymbol{b}'\boldsymbol{P}\boldsymbol{b}}(7)$$

$$R_G = \frac{k}{L_G}\sqrt{\boldsymbol{w}'\boldsymbol{\Gamma}\boldsymbol{w}}(8)$$

where, L_P and L_G denotes the time required to complete one selection cycle and $\boldsymbol{\Gamma}$ is a covariance matrix of additive genomic breeding values $\boldsymbol{\gamma}$.

- Expected genetic gain per trait

The expected genetic gain per traitper selection cycle for GSI ($\boldsymbol{E}_G$) and PSI ($\boldsymbol{E}_P$) (Ceron-Rojas et al. 2015):

$$\boldsymbol{E}_P = \frac{k}{L_P}\frac{\boldsymbol{C}\boldsymbol{b}}{\sqrt{\boldsymbol{b}'\boldsymbol{P}\boldsymbol{b}}}(9)$$

$$\boldsymbol{E}_G = \frac{k}{L_G}\frac{\boldsymbol{\Gamma}\boldsymbol{w}}{\sqrt{\boldsymbol{w}'\boldsymbol{\Gamma}\boldsymbol{w}}}(10)$$

where, $\boldsymbol{E}_P$ and $\boldsymbol{E}_G$ are the expected genetic gain per selection cycle for each trait in the PSI and the GSI, respectively (Togashi et al. 2011).

- Technow inequality

To compare GSI and PSI efficiency the ratio of estimated GSIs and the PSI selection response given by Technow et al. (2013) was used:

$$\lambda = \frac{\hat{R}_G}{\hat{R}_P} = \frac{L_P}{L_G}\sqrt{\frac{\mathbf{w}'\hat{\mathbf{\Gamma}}_l\mathbf{w}}{\hat{\mathbf{b}}'\hat{\mathbf{P}}\hat{\mathbf{b}}}} = \frac{L_P}{L_G}\frac{\hat{\rho}_{H,GSI}}{\hat{\rho}_{H,PSI}}$$

Using this criterion, (i)if $\lambda > 1$, GSI efficiency will be greater than PSI efficiency, (ii)if $\lambda = 1$, the efficiency of both selection indices will be equal, (iii)and if $\lambda < 1$, PSI will be more efficient than GSI.

- Correlations or accuracy ofselection indexes (GSI and PSI) withthe net genetic merit (H)

The estimated correlation (or accuracy) between H and GSI ($\hat{\rho}_{H,GSI}$), and between H and PSI ($\hat{\rho}_{H,PSI}$) are given below (Ceron-Rojas et al. 2015):

$$\hat{\rho}_{GSI,H} = \frac{\sqrt{\boldsymbol{w}'\boldsymbol{\Gamma}\boldsymbol{w}}}{\sqrt{\boldsymbol{w}'\boldsymbol{C}\boldsymbol{w}}}(12)$$

$$\hat{\rho}_{PSI,H} = \frac{\sqrt{\boldsymbol{b}'\boldsymbol{P}\boldsymbol{b}}}{\sqrt{\boldsymbol{w}'\boldsymbol{C}\boldsymbol{w}}}(13)$$

Estimation of the phenotypic variance-covariance matrix $\hat{\boldsymbol{P}} = \hat{\boldsymbol{C}} + \hat{\boldsymbol{R}}$, the covariance matrix of true breeding values ($\hat{\boldsymbol{C}}$), $\boldsymbol{\Gamma}$ the covariance matrix of additive genomic breeding values $\boldsymbol{\gamma}$ and residuals matrix ($\hat{\boldsymbol{R}}$), respectively can be obtained by using analysis of variance (ANOVA) (Lynch and Walsh 1998), maximum likelihood, or restricted maximum likelihood (REML) (Patterson and Thompson 1971).

Experimental dataset

In order to evaluate the performance measures of different GS models, two F_2 biparental maize populations from Beyene et al. (2015) were used in the study, which were denoted as JMpop1 DTMA Mexico optimum environment (dataset 1) and JMpop3 DTMA Mexico optimum environment (dataset 2). Each dataset contains genotypic data and four phenotypic traits: grain yield (GY, t/ha), plant height (PHT, cm), ear height (EHT, cm), and thesis days (AD, days). Dataset 1 contained 247 lines (individuals) which were genotyped for 195 markers, while dataset 2 had 396 lines genotyped for 190 markers. As the breeding objective was to increase GY while decreasing PHT, EHT, and AD, so accordingly economic weights were provided (Beyene et al. 2015) denoted as, w = (2 – 0.3 – 0.3 – I), respectively. The top 10% of individuals w.r.t. PSI and GSI (i.e., k = 1.75) were selected from the two data sets, respectively. Here, the time required to complete one GSI selection cycle (generation interval) was L_G = 1.5 years, whereas for one PSI selection cycle, it was L_P = 4 years (Beyene et al. 2015).

In the current study, two real datasets were analysed considering all the four traits (*i.e.,* GY, PHT, EHT, and AD). In each selection cycle, four different evaluation measures were used in the analysis. The estimated GSI and PSI selection responses, the estimated expected genetic gain for each trait in the PSI and the GSI, the estimated comparison of PSI and GSI efficiency and the correlations between the GSI and PSI with the true net genetic merit were computed. To rank the overall performance of GSI and PSI models used in the current study, the Technique for order of preference by similarity to ideal solution (TOPSIS) was used. TOPSIS is based on multi-criteria decision-making/analysis, which was developed by Hwang and Yoon (1981) and implemented in R package 'topsis'. All the analyses were performed using the R programming platform (R Development Core Team 2022).

Results and discussion

Correlations/accuracy of GSI and PSI with the true net genetic merit (H)

As expected, the correlation between GSI and H tends to decrease more than the correlation between PSI and H in both datasets (Fig. 1). The likely reason behind this is that the estimated value of PSI was calculated using only the phenotypic information, whereas the estimated value of GSI was calculated using five different GS models (i.e., RR-BLUP, Bayesian LASSO, BayesA, BayesB and BayesC) using only the newly generated SNP genotypic information from the breeding population. It was observed that the correlation between PSI and H for dataset 1 and 2 was 0.95 and 0.42, respectively, whereas the correlation between GSI and H using five different GS models ranged between 0.45–0.57 for dataset1 and 0.19–0.29 for dataset 2, respectively (Fig. 1). More specifically, Bayesian-based GSI (GSI_BLASSO, GSI_Bayes A, GSI_Bayes B and GSI_Bayes C) methods have higher correlation with net genetic merit than penalized regression based GSI (GSI_RR-BLUP) method in both datasets. The maximized correlation of the indexes with the net genetic merit signifies the accuracy of the indexes, which is different from the Technow inequality (i.e., efficiency comparison among indexes in terms of generation interval). We used the Fisher z-test to test the significance of the correlation coefficient and observed that the correlation coefficient is significant (p-value<0.0001) for all possible comparisons.

Estimated PSI and GSI selection response

The selection response represents the mean value of the future unseen population generated from the selected parents. The observations in the present study, that most of the GSI selection response was more efficient than the PSI selection response in both datasets, where the generation interval are taken into consideration to elaborate. When the generation interval increases, the GSI selection response decreases because the additive genomic variance-covariance decreases after each generation. The estimated selection could be more specifically observed (Table 1), it was concluded that Bayesian-based GSI selection response (R_{GBL}, R_{GBA}, R_{GBB} and R_{GBC}) was higher in comparison to penalized regression based GSI response to selection (R_{GRR}) in both the datasets.

Estimated expected genetic gains for PSI and GSI

Expected genetic gain per trait or multi-trait selection response represents the mean of each trait of candidates under selection. It was observed that the estimation of expected genetic gain for each trait, it was believed that GY, PHT, and EHT traits were higher in GSI based (E_{GBL} ,E_{GBA} ,E_{GBB} and E_{GBC}) genetic gain in comparison to the PSI based expected genetic gains for dataset 1 (Table 2).

The results of dataset 2 were presented in Table 3, it can be observed that genetic gain per trait was higher in GSI-based (E_{GRR} , E_{GBL} ,E_{GBA} , E_{GBB} and E_{GBC}) genetic gain in comparison to the PSI-based expected genetic gains. In both the datasets, all GSI-based models led to increased rates of genetic gain per trait than PSI-based model, presumably, because PSI takes about years to complete each selection cycle, whereas GSI takes around 1.5 years (Beyene et al. 2015). To rank, the overall performance of the GSI-based and PSI model TOPSIS method was used (Hwang and Yoon 1981).

Technow Inequality f

Technow inequality is the selection response ratio for comparing the efficiency of selection indices. Technow inequality *i.e.,* selection response ratio, is a good criterion as it contains all the information on genetic gain of each selection indices considering generation interval. The results in Table 4, the estimated GSI (including five different GS models-based GSI selection responses) and the PSI selection response ratios considering varying generation intervals with same selection intensity. However, GSI efficiency was higher than PSI efficiency because the period between selection cycles in GSI is shorter in comparison to PSI, with respect to datasets under consideration in the current study.

Table 1. Estimated PSI and GSIs (for five GS models) selection response

Response to Selection (*R)	Dataset 1	Dataset 2
RPSI	0.93	1.94
$R_{GSI\_RR-BLUP}$	0.38	1.98
R_{GSI_BLASSO}	0.94	1.97
$R_{GSI\_Bayes A}$	0.95	2.05
$R_{GSI\_Bayes B}$	0.95	2.05
$R_{GSI\_Bayes C}$	1.16	1.99

*R_{PSI} = Response to selection of PSI, $R_{GSI\_RR-BLUP}$ = Response to selection of GSI using RR-BLUP, R_{GSI_BLASSO} = Response to selection of GSI using Bayesian LASSO, $R_{GSI\_Bayes A}$ = Response to selection of GSI using Bayes A, $R_{GSI\_Bayes B}$ = Response to selection of GSI usingBayes B and $R_{GSI\_Bayes C}$ = Response to selection of GSI using Bayes C

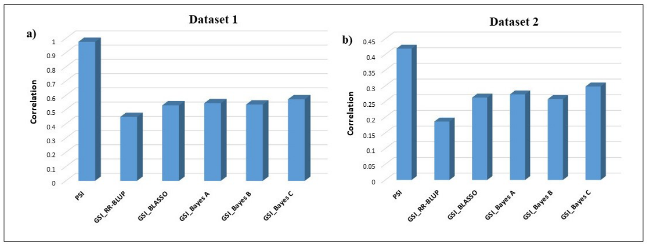

Fig. 1. Correlations of GSI and the PSI with the true net genetic merit (H)

Table 2. Expected genetic gains for the PSI and GSIs (for five GS models) using dataset 1

Expected genetic gain (*E)	GY	PHT	EHT	AD	Rank
E_{PSI}	0.16	-0.85	-0.48	0.04	4
$E_{GSI\_RR\text{-}BLUP}$	0.05	7.40E-07	6.51E-07	-0.12	6
E_{GSI_BLASSO}	0.16	-0.16	-0.13	-0.04	2
$E_{GSI\_Bayes\ A}$	0.16	-0.12	-0.18	0.04	5
$E_{GSI\_Bayes\ B}$	0.15	-0.16	-0.31	-0.05	3
$E_{GSI\_Bayes\ C}$	0.20	-0.18	-0.24	-0.03	1

*E_{PSI} = Expected genetic gain of PSI, $E_{GSI\_RR\text{-}BLUP}$ = Expected genetic gain of GSI using RR-BLUP, E_{GSI_BLASSO} = Expected genetic gain of GSI using Bayesian LASSO, $E_{GSI\_Bayes\ A}$ = Expected genetic gain of GSI using Bayes A, $E_{GSI\_Bayes\ B}$ = Expected genetic gain of GSI using Bayes B and $E_{GSI\_Bayes\ C}$ = Expected genetic gain of GSI using Bayes C

Table 3. Expected genetic gains for the PSI and GSIs using dataset 2

Expected genetic gain (*E)	GY	PHT	EHT	AD	Rank
E_{PSI}	0.002	-7.86	-7.65	1.32	5
$E_{GSI\_RR\text{-}BLUP}$	0.0006	-7.10	-6.63	2.14	6
E_{GSI_BLASSO}	0.003	-6.93	-6.45	2.10	1
$E_{GSI\_Bayes\ A}$	0.002	-7.24	-6.71	2.14	3
$E_{GSI\_Bayes\ B}$	0.002	-7.06	-6.82	2.13	4
$E_{GSI\_Bayes\ C}$	0.003	-6.91	-6.37	2.11	2

*As per footnote in Table 1

Table 4. Selection response ratio/Technow inequality for comparison of GSI and PSI efficiency

Selection efficiency ratio ($\lambda = \frac{\hat{R}_{GSI}}{\hat{R}_{PSI}}$)	Dataset 1	Dataset 2
$\hat{R}_{GSI\_RR-BLUP}/\hat{R}_{PSI}$	0.40	1.02
$\hat{R}_{GSI\_BLASSO}/\hat{R}_{PSI}$	1.01	1.01
$\hat{R}_{GSI\_Bayes\ A}/\hat{R}_{PSI}$	1.02	1.05
$\hat{R}_{GSI\_Bayes\ B}/\hat{R}_{PSI}$	1.02	1.05
$\hat{R}_{GSI\_Bayes\ C}/\hat{R}_{PSI}$	1.24	1.02

However, the success of GS based breeding programme is determined by the characteristics of the population, such as trait heritability and the genetic architecture under study. The GEBVs based superior individual genotypes have proven to be incredibly accurate in empirical research (Lande et al. 2011). Because GEBVs are themselves indices, the GSI is defined as a linear combination of them, and it is a genomic predictor of net genetic merit, whereas the PSI is a phenotypic predictor of net genetic merit for the selection of a candidate (Robinson 1991; Togashi et al. 2011). In general, the PSI is expected to have a greater selection response and be more accurate than the GSIs in cases, where the generation interval is ignored, otherwise vice-versa. When the generation interval is taken into account, the GSI requires one-third or less of the time required by PSI in terms of genetic gain per trait per unit of time (Lorenz et al. 2011). Many reports have been published on generation interval reduction in breeding methods, such as in beef cattle (Gutierrez et al. 2003; Mc Parland et al. 2007), sheep (Joakimsen 1969; Mokhtari et al. 2014), and goat (Rashidi et al. 2015) populations. The implementation of generation interval under the GS paradigm has significantly reduced breeding interval in chickens from 12 to 6 months (Dekkers 2007) and in dairy cattle from 5–6 years to 1.5 years (Pyrce and Daetwyler, 2012). Implementation of generation interval in GS programmes, GSI will be more efficient than PSI, as evident by the Technow inequality (Table 4). However, if the number of markers is limited, GSI selection responses will provide lower values than PSI selection responses, and PSI efficiency will outperform GSI efficiency. Moreover, the PSI and GSI selection responses will be comparable, if the number of markers is very high. In conclusion, as indicated by the Technow inequality in terms of a unit of time, GSI will be more efficient than PSI only if $L_P > L_G$ otherwise, PSI will be more efficient than GSI.

In the present study, four independent criteria were used to evaluate the performance of GSI and PSI methodology for multiple-trait using two real datasets. Correlations of the PSI and GSI with the true net genetic merit, the PSI and GSI selection response, the PSI and GSI predicted genetic gain per trait per unit of time, and the Technow inequality, to examine the theory of PSI and GSI efficiency. The observation of our study shows that GSI efficiency was slightly higher than PSI efficiency per unit of time or generation interval except for GSI based on rrBLUP. Another important finding from this study was that the Bayesian-based GSI has slightly better performance than penalized regression-based GSI. It can be suggested, that GSI is a better option for selecting candidates or individuals under multi-trait genomic selection.

Authors' contribution

Conceptualization of research (MAK, AR, DCM); Designing of the experiments (MAK, AR, DCM, NB); Contribution of experimental materials (MAK, AR, DCM, NB); Execution of field/lab experiments and data collection (MAK, AR, DCM); Analysis of data and interpretation (MAK, AR, DCM, SS, SGM); Preparation of the manuscript (MAK, AR, DCM, NB).

References

Bernardo R. 2008. Molecular markers and selection for complex traits in plants: Learning from the last 20 years. Crop Sci., **48**: 1649-1664.

Bernardo R. and Yu J. 2007. Prospects for genome-wide selection for quantitative traits in maize. Crop Sci., **47**: 1082-1090.

Beyene Y., Semagn K., Mugo S., Tarekegne A., Babu R. et al. 2015. Genetic gains in grain yield through genomic selection in eight bi-parental maize populations under drought stress. Crop Sci., **55:** 154–163.

Budhlakoti N., Mishra, D.C., Rai A., Lal S.B., Chaturvedi K.K. and Kumar R.R. 2019. A comparative study of single-trait and multi-trait genomic selection. J. Comput. Biol., **26**(10): pp.1100-1112.

Ceron-Rojas J. J., Crossa J., Arief V. N., Basford K., Rutkoski J., Jarquín D., Alvarado G., Beyene Y., Semagn K., DeLacy I. 2015. A genomic selection index applied to simulated and real data. Genes/Genomes/Genetics, **5**: 2155–2164.

Cochran W. G. 1951. Improvement by means of selection. In: Neyman J. (ed) Proceedings of the second Berkeley symposium on mathematical statistics and probability. University of California Press, Berkeley, CA: 449–470.

Costa M. M., Di Mauro A.O., Unêda-Trevisoli S. H., Castro Arriel N. H., Bárbaro I. M., Dias da Silveira G., Silva Muniz F. R. 2008. Analysis of direct and indirect selection and indices in soybean segregating populations. Crop Breed. Appl. Biotechnol., **8**: 47–55.

de los Campos G., Hickey J. M., Pong-Wong R.,Daetwyler H. D. and Calus P. L. 2013. Whole-genome regression and prediction methods applied to plant and animal breeding. Genetics, **193**: 327–345.

de los Campos G., Naya H.,Gianola D., CrossaJ., Legarra A. et al.2009. Predicting quantitative traits with regression models for dense molecular markers and pedigrees. Genetics, **182**(1): 375–385.

Endelman J. B. 2011. Ridge regression and other kernels for genomic selection with R package rrBLUP. Plant Genome, **4**: 250-255.

Fernando R. and Grossman M. 1989. Marker assisted selection using best linear unbiased prediction. Genet. Sel. Evol.,**21**: 467.

Goddard M. E. and Hayes B. J. 2007. Genomic selection. J. Anim. Breed. Genet., **124**: 323-330.

Guo, Zhao Y., Wang Y., Zhang Y., Du L. and Su G. 2014. Comparison of single-trait and multiple-trait genomic prediction models. BMC Genetics, 15-30.

Gianola D. 2013. Priors in whole-genome regression: the Bayesian alphabet returns. Genetics, **194**: 573–596.

Gutiérrez J.P., Altarriba J., Díaz C., Quintanilla R., Cañón J., and Piedrafita J. 2003. Pedigree analysis of eight Spanish beef cattle breeds. Genet. Sel. Evol., **35**: 43–64.

Hwang C. L. and Yoon K. 1981. In Multiple attribute decision making.Springer, 58–191.

Hayes B., Visscher P. M.and Goddard M. E. 2009. Increased accuracy of artificial selection by using the realized relationship matrix. Genet. Res,, **91**: 47–60.

Hazel L. N. 1943. The genetic basis for constructing selection indices. Genetics, **38**: 476–490.

Hazel L. N. and Lush J. L. 1942. The efficiency of three methods of selection. J. Hered., **33**: 393.

Heffner E. L. et al. 2009. Genomic selection for crop improvement. Crop Sci., **49**: 1-12.

Heffner E. L., Jannink J.L., Iwata H.et al. 2011. Genomic selection accuracy for grain quality traits in biparental wheat populations. Crop Sci., **51**: 2597-2606.

Howard R., Carriquiry A. L. and Beavis W. D.2014. Parametric and Nonparametric Statistical Methods for Genomic Selection of Traits with Additive and Epistatic Genetic Architectures. Genes/Genomes/Genetics, **4**: 1027-1046.

Jia Y., Jannink J. L. 2012. Multiple-trait genomic selection methods increase genetic value prediction accuracy. Genetics, **192**(4): 1513–1522.

Joakimsen Ø. 1969. Generation interval in Norwegian sheep. Acta Agric. Scand., **19**: 175–177.

Kempthorne O.and Nordskog A. W. 1959. Restricted selection indices. Biometrics, **15**: 10–19.

Lande R. and Thompson R. 1990. Efficiency of marker-assisted selection in the improvement of quantitative traits. Genetics, **124**: 743–756.

Lorenz A. J., Chao S., Asoro F. G., Heffner E. L., Hayashi T. et al. 2011. Genomic selection in plant breeding: knowledge and prospects. Adv. Agron., **110**: 77–123.

Lynch M. and Walsh B. 1998. Genetics and Analysis of Quantitative Traits. Sinauer Accoc. Inc., Sunderland, MA.

Mc Parland S., Kearney J., Rath M., and Berry D. 2007. Inbreeding trends and pedigree analysis of Irish dairy and beef cattle populations. J. Anim. Sci., **85**: 322–331.

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

Mokhtari M., Shahrbabak M.M., Esmailizadeh A., Shahrbabak H.M., and Gutierrez J. 2014. Pedigree analysis of Iran-Black sheep and inbreeding effects on growth and reproduction traits. Small Rumin. Res., **116**: 14–20.

Mishra D. C., Budhlakoti N., Mujumdar S. G., Rai A. 2021. Innovations in Genomic Selection: Statistical Perspective, Special Proceedings: ISBN #: 978-81-950383-0-5. 23rd Annual Conference. Pp 101-111.

Park T. and Casella G. 2008. The Bayesian Lasso. J. Am. Stat. Assoc,**103**: 681–686.

Patterson H. D. and Thompson R. 1971. Recovery of inter-block information when block sizes are unequal. Biometrika, **58**: 545–554.

Perez rodriguez P. and de los Campos G. 2011. Bayesian Generalized Linear Regression (BGLR R pacakage)

Piepho H. P. 2009. Ridge regression and extensions for genome wide selection in maize. Crop Sci., **49**: 1165-1176.

Pryce J. E. and Daetwyler H. D. 2012. Designing dairy cattle breeding schemes under genomic selection: areview of international research. Anim. Prod. Sci., **52**: 107-114.

Rashidi A., Mokhtari M., and Gutiérrez J. 2015. Pedigree analysis and inbreeding effects on early growth traits and greasy fleece weight in Markhoz goat. Small Rumin. Res., **124**: 1–8.

Robinson G. K. 1991. That BLUP is a good thing: The estimation of random effects. Stat. Sci., **6**: 15–51.

Smith H. F. 1936. A discriminant function for plant selection, pp. 466–476 in Papers on Quantitative Genetics and Related Topics. Department of Genetics, North Carolina State

College, Raleigh, North Carolina.

Technow F., Bürger A. and Melchinger A. E. 2013. Genomic prediction of northern corn leaf blight resistance in maize with combined or separated training sets for heterotic groups. G3 Genes/Genomes/Genetics (Bethesda), **3**: 197–203.

Tibshirani R. 1996. Regression shrinkage and selection via the lasso. J. R. Stat. Soc. Series B (Methodological), Pp 267-288.

Togashi K., Lin C. Y.and Yamazaki T. 2011. The efficiency of genomewide selection for genetic improvement of net merit. J. Anim. Sci., **89**: 2972–2980.

VanRaden P., Van Tassell C., Wiggans G. et al. 2009. Invited review: Reliability of genomic predictions for North American Holstein bulls. J. Dairy Sci., **92**: 16-24.

Verbyla K. L., Bowman P. J., Hayes B. J. and Goddard M. E. 2010. Sensitivity of genomic selection to using different prior distributions. BMC Proc., **4**(Supp. 1): S5.

Xu S. Z. 2003. Estimating polygenic effects using markers of the entire genome, Genetics, **163**: 789–801.

Xu Y. et al. 2012. Whole-genome strategies for marker-assisted plant breeding. Mol. Breed., **29**: 833-854.