

The statistical theory of linear selection indices from phenotypic to genomic selection

J. Jesus Cerón-Rojas | Jose Crossa 

Biometrics and Statistics Unit, International Maize and Wheat Improvement Center (CIMMYT), Km 45 Carretera Mexico-Veracruz, Edo. de Mexico, Mexico DF CP 52640, Mexico

Correspondence

Jose Crossa, Biometrics and Statistics Unit, International Maize and Wheat Improvement Center (CIMMYT), Km 45 Carretera Mexico-Veracruz, Mexico DF, CP 52640, Edo. de Mexico, Mexico.
Email: j.crossa@cgiar.org

Assigned to Associate Editor Manjit Kang.

Funding information

Bill and Melinda Gates Foundation and USAID, Grant/Award Numbers: INV-003439 BMGF/FCDO Accelerating Genetic Gains in, [Amend. No. 9 MTO 069033, USAID-CIMMYT Wheat/AGGMW]

Abstract

A linear selection index (LSI) can be a linear combination of phenotypic values, marker scores, and genomic estimated breeding values (GEBVs); phenotypic values and marker scores; or phenotypic values and GEBVs jointly. The main objective of the LSI is to predict the net genetic merit (H), which is a linear combination of unobservable individual traits' breeding values, weighted by the trait economic values; thus, the target of LSI is not a parameter but rather the unobserved random H values. The LSI can be single-stage or multi-stage, where the latter are methods for selecting one or more individual traits available at different times or stages of development in both plants and animals. Likewise, LSIs can be either constrained or unconstrained. A constrained LSI imposes predetermined genetic gain on expected genetic gain per trait and includes the unconstrained LSI as particular cases. The main LSI parameters are the selection response, the expected genetic gain per trait, and its correlation with H . When the population mean is zero, the selection response and expected genetic gain per trait are, respectively, the conditional mean of H and the genotypic values, given the LSI values. The application of LSI theory is rapidly diversifying; however, because LSIs are based on the best linear predictor and on the canonical correlation theory, the LSI theory can be explained in a simple form. We provided a review of the statistical theory of the LSI from phenotypic to genomic selection showing their relationships, advantages, and limitations, which should allow breeders to use the LSI theory confidently in breeding programs.

Abbreviations: BLP, best linear predictor; BLUP, best linear unbiased predictor; BP, best predictor; CCA, canonical correlation analysis; CDMLPSI, constrained decorrelated multi-stage linear phenotypic selection index; CLGSI, constrained linear genomic selection index; CLPSI, constrained linear phenotypic selection index; COMLPSI, constrained optimum multi-stage linear phenotypic selection index; DMLPSI, decorrelated multi-stage linear phenotypic selection index; ESIM, eigen selection index method; GEBV, genomic estimated breeding value; GESIM, genomic eigen selection index method; LGSI, linear genomic selection index; LMSI, linear molecular selection index; LPSI, linear phenotypic selection index; LSI, linear selection index; MESIM, molecular eigen selection index method; MSPE, minimum mean square prediction error; MVN, multivariate normal; OMLPSI, optimum multi-stage linear phenotypic selection index; PPG-ESIM, predetermined proportional gain eigen selection index method; QTL, quantitative trait loci; REML, restricted maximum likelihood; RESIM, restricted eigen selection index method; RL, rate of lay; RLPSI, restricted linear phenotypic selection index; SM, age at sexual maturity.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. Crop Science © 2021 Crop Science Society of America

1 | INTRODUCTION

Smith (1936) published a paper called “A discriminant function for plant selection,” in which he described a statistical method to select parents for the next selection cycle based on a linear combination of several joint quantitative traits. That method makes it possible to improve several traits that differ in additive variability, heritability, economic importance, and in the correlation among their phenotypes and genotypes in the plant and animal breeding context (Hazel et al., 1994). Hazel and Lush (1942) called that selection method “total score,” whereas Hazel (1943) called it “selection index,” and Cerón-Rojas and Crossa (2018) called it “linear phenotypic selection index” (LPSI). Smith (1936) made a clear distinction between the LPSI (discriminant function) and the net genetic merit (H). The LPSI is a linear combination of observable phenotypic records, and H is an unobservable linear combination of trait breeding values (\mathbf{g}) weighted by the trait economic values (\mathbf{W}).

The main assumptions made by Smith (1936) to develop the LPSI theory are as follows: the vector of genotypic values \mathbf{g} (the average of the phenotypic values across a [large] population of environments, or the genetic constitution of an organism or cell) is composed entirely of additive effects of genes and thus is the plant or animal breeding value, and H is the total individual genotypic value (Hazel & Lush, 1942; Kempthorne & Nordskog, 1959). Finally, the vectors of adjusted means of phenotypic values (\mathbf{y}) and H have a joint multivariate normal distribution. By those assumptions, the joint distribution of H and LPSI is bivariate normal, whereas the conditional distribution of H given LPSI (and the conditional distribution of LPSI given H) is univariate normal. Under those assumptions, the regression of H on \mathbf{y} is linear (Kempthorne & Nordskog, 1959), and it is possible to write the H and \mathbf{y} values as a multiple linear regression model, where H and \mathbf{y} are the dependent and independent variables, respectively, and both are random. In this model, the LPSI is the conditional expectation of H given \mathbf{y} . Then, to predict H , one must find the values of the LPSI regression coefficients so that the LPSI values may best discriminate those individuals that have the highest H values (Smith, 1936).

Smith (1936) wrote his paper before the Eisenhart (1947) article, which highlighted that there are two fundamentally different explanatory variables, called “fixed” and “random” effects. One important result of this distinction is that whereas fixed effects influence only the mean of \mathbf{y} , random effects influence only the variance of \mathbf{y} and thus the LPSI and its parameters. Assuming that the population mean of H and \mathbf{y} is zero, the selection response is the conditional expectation of H given the LPSI values, whereas the expected genetic gain per trait (or multi-trait selection response) is the conditional expectation of \mathbf{g} given the linear selection index (LSI) val-

Core Ideas

- Linear selection indices (LSIs) are useful in plant and animal breeding.
- The main LSI objective is to predict the net genetic merit of individuals.
- We did a complete review of the statistical theory of the LSI.
- We described the bases of genomic LSI.

ues (Cochran, 1951; Kempthorne & Nordskog, 1959; Smith, 1936). The estimates of these last parameters and the correlation between H and the LPSI are the main criteria to validate and compare the efficiency of any LSI.

Some authors (Finney, 1962; Hazel & Lush, 1942; Young, 1961) compared the LPSI efficiency with the independent culling method, where breeders sequentially select for a series of independent traits in the same population, and the tandem selection method, which involves sequential selection on a series of traits where repeated cycles of selection for a single trait are practiced until a desirable expression of the trait is reached. Those authors found that the LPSI selection response was higher than the selection response for the other two selection methods. Thus, under certain conditions (e.g., normality), the LPSI is the best option for selection.

The LPSI theory has been extended, and now a LSI can be a linear combination of marker scores (Lande & Thompson, 1990), genomic estimated breeding values (GEBVs) (Cerón-Rojas & Crossa, 2019; Cerón-Rojas et al., 2015), phenotypic values and marker scores (Lande & Thompson, 1990), and phenotypic values and GEBVs jointly (Dekkers, 2007). The last two indices are called “combined LSIs” in this review. An additional distinction is between single-stage and multi-stage LSIs (Cerón-Rojas & Crossa, 2020a; Cerón-Rojas et al., 2019a, 2019b; Cochran, 1951; Young, 1964; Xu & Muir, 1992). Multi-stage LSIs are methods for selecting traits at different times or stages of the crop, such as the age at which breeders can measure and select the trait of economic interest and are used to save money and time because these indices do not need to have a large number of individuals in the selection process.

The main objectives of all LSIs are to predict H and select parents for the next generation; thus, the target of study of any LSI is not a parameter but rather the unobserved random H values. Additional objectives of LSIs are to maximize the selection response, the expected genetic gain per trait, and the correlation with H . By imposing restrictions equal to some predetermined values on the expected genetic gain per trait, several authors (Harville, 1975; Kempthorne &

Nordskog, 1959; Mallard, 1972; Tallis, 1985) developed the constrained LPSI (CLPSI) theory. In addition, by applying the LPSI and CLPSI theories in the context of genomic selection, Cerón-Rojas et al. (2015) and Cerón-Rojas and Crossa (2019) developed the unconstrained and constrained linear genomic selection indices (LGSI and CLGSI, respectively), which are linear combinations of GEBVs. Cerón-Rojas et al. (2016b) have shown that the CLPSI vector of coefficients is a linear combination of the LPSI vector of coefficients or a projection of the unconstrained LPSI vector of coefficients to a different space.

All the LSIs described above assume that the economic weights are known constants; however, for several reasons, this assumption is rarely fulfilled (Vandepitte & Hazel, 1977). Furthermore, several authors (Hazel, 1943; Lin, 1978; Yamada et al., 1975) have indicated that there is no one general method for assigning economic weights to the traits because those economic weights can change from time to time or vary from one location to another in breeding programs. For this reason, modified indices, such as the base index (Williams, 1962), modified base index (Smith, 1983), and non-weighted multiplicative index (Elston, 1963), have been proposed in plant and animal breeding. However, those authors did not show that the latter LSIs maximize the correlation with the net genetic merit. In addition, it can be difficult to obtain the selection response and the expected genetic gain per trait for those indices.

The preceding problems led Cerón-Rojas et al. (2008a, 2016a) to develop the eigen selection index method (ESIM), the restricted ESIM (RESIM), and the predetermined proportional gain ESIM (PPG-ESIM) in the canonical correlation context (Hotelling, 1936). The ESIM is an unrestricted index, but the RESIM and PPG-ESIM allow imposing null and predetermined restrictions, respectively, on the expected genetic gains of some traits, while the rest remain without restrictions. These indices do not require economic weights when predicting H and making selections because they use the first eigenvector of the matrix of multi-trait heritabilities (in the ESIM case) or the first eigen vector of a linear transformation of the matrix of multi-trait heritabilities (in the RESIM and PPG-ESIM case) in the prediction. Cerón-Rojas et al. (2008b) adapted the ESIM to the marker context and developed an index (molecular ESIM [MESIM]) similar to the linear molecular selection index (LMSI), whereas Cerón-Rojas and Crossa (2018) adapted ESIM to the genomic selection context and developed a combined genomic eigen selection index method (genomic ESIM [GESIM]), similar to the Dekkers (2007) index.

All indices associated with the LPSI theory (unconstrained and constrained; phenotypic, genomic, multi-stage, etc.) are particular cases of the best linear predictor (BLP) theory, which is based on the minimum mean square prediction error (MSPE), because this approach allows us to obtain the unique

best MSPE predictor (Bickel & Doksum, 2015). The MSPE is the expected squared deviation between the true value and the estimate, and it provides an effective measure of proximity to the true parameter of interest (Carlin & Louis, 2000). On the other hand, all indices associated with ESIM are based on the canonical correlation analysis (CCA) theory (Hotelling, 1936).

Searle et al. (2006) have indicated that the best predictor (BP), BLP, and the best linear unbiased predictor (BLUP) are associated as follows. When all the parameters of the joint distribution of H and LSI are known, BP is available, whereas for BLP only the first and second moments are assumed to be known, and for BLUP the second but not first moments are assumed to be known. Under multivariate normality, BP and BLP are the same (Bickel & Doksum, 2015; Christensen, 2011; Harville, 2018); however, in practice, those approaches need estimates of their parameters. In this last case, the LSI could not be BLP (Searle et al., 2006). Nevertheless, with good estimates of the phenotypic and genotypic covariance matrices (e.g., by restricted maximum likelihood [REML] estimate), the BLP (or LPSI) is close to the BLUP (Christensen, 2011; Henderson, 1990).

Several authors (Baker, 1986; Bramscap, 1984; Dekkers & Settar, 2004; Hazel et al., 1994; Henderson, 1963; Lin, 1978; Moreau et al., 2007; Nordskog, 1978; Phillipsson et al., 1994) have given partial reviews of the LSIs theory, such as the unconstrained and constrained LPSI and the LMSI, among others. In addition, using minimum mathematical rigor, Cerón-Rojas and Crossa (2018) wrote a user manual for breeders based on the LSI theory, and they gave many illustrative practical examples. In the current study, we reviewed the statistical theory of the LSIs from the unconstrained and constrained LPSI to the unconstrained and constrained LGSI. That is, based on the BLP theory, in the present work, we unified all the statistical theory of the LSIs associated with the LPSI (e.g., CLPSI, LGSI, CLGSI, etc.) from phenotypic to genomic selection. This requires greater mathematical rigor; as such, it is not possible to provide as many illustrative examples as Cerón-Rojas and Crossa (2018) did. In addition, based on the CCA theory, we present a unified theory of the ESIM, RESIM, and PPG-ESIM.

This review has four main parts. The first describes two earlier approaches to the unconstrained LSI theory and a description of the BLUP theory. The second describes the unconstrained LSI theory, including the LPSI, multi-stage LPSI, LGSI, and combined LSI. Similarly, the third part contains the constrained LSI that includes the CLPSI, constrained multi-stage LPSI, and CLGSI. Finally, in the fourth part, we describe ESIM, RESIM, PPG-ESIM, and their relationships with the LPSI and CLPSI. The MESIM and GESIM are only variants of ESIM and the combined LSIs, which can be found in Cerón-Rojas and Crossa (2018). The main objective of this review is to introduce researchers to the statistical theory of

the LSIs. Our intended audience for this study is mainly plant breeders and graduate-level students who are interested in LSI theory and who come from animal and plant breeding programs.

In subsection “Numerical example: LPSI vs ESIM,” we provide an illustrative example using the LPSI and ESIM theory. In subsection “RIndSel,” we introduce **RIndSel** (Alvarado et al., 2018; Pacheco et al., 2017; Perez-Elizalde et al., 2014), an R-software that is useful for estimating the LPSI, LGSI, CLPSI, ESIM, etc., parameters and for selecting individual candidates as parents for the next selection cycle. Furthermore, **RIndSel** is completely automated, which eliminates the need to learn a specialized syntax for anyone who is a novice user of R-software. This means that the users only need to learn how to introduce their data into the program and how to interpret the **RIndSel** results. This software, and a complete **RIndSel** user manual, can be downloaded from <https://data.cimmyt.org/dataset.xhtml?persistentId=hdl:11529/10854>.

2 | NOTATION

We use uppercase letters to denote random variables and lowercase letters to denote their particular realizations. For example, X is a random variable, whereas x ($X = x$) is a particular realization of X . Vectors are denoted with bold lowercase letters (e.g., \mathbf{g} and \mathbf{y}) and matrices with bold uppercase letters (\mathbf{X}). Sometimes it is necessary to denote vectors with bold capital letters; for example, \mathbf{E} denotes the vector of expected genetic gain per trait.

The expectation of the random variable $X[E(X)]$ is scalar, whereas the conditional expectation of X given $Y = y$ [$E(X|Y = y)$] is a function of y . Thus, before we observe Y , the value of $E(X|Y = y)$ is unknown, so it is a random variable that we denote as $E(X|Y)$.

3 | TWO EARLY APPROACHES TO THE INDEX SELECTION THEORY

Pearson (1903) and Pearl and Surface (1909) were the first to describe the index theory before the Smith (1936) LPSI theory. As we shall see, Pearson (1903) predicted breeding values based on a linear mixed model, whereas Pearl and Surface (1909) used a nonlinear approximation to the index theory. In addition, whereas the Pearson (1903) index selects genotypes, the Pearl and Surface (1909) index selects phenotypes. This last index selects phenotypes because it does not maximize its correlation with the net genetic merit.

3.1 | Pearson index

Suppose that \mathbf{X} and \mathbf{Z} are design matrices and that $\boldsymbol{\theta}$ and \mathbf{g} are vectors of fixed and random (e.g., breeding values) effects, respectively, associated to the trait of n individual candidates (\mathbf{y}) to selection, and assume that \mathbf{y} and \mathbf{g} have multivariate normal (MVN) distribution, also denoted as “ \sim ”, i.e., $\mathbf{y} \sim \text{MVN}$ and $\mathbf{g} \sim \text{MVN}$. Consider the following linear mixed model:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\theta} + \mathbf{Z}\mathbf{g} + \boldsymbol{\varepsilon} \quad (1)$$

where $\mathbf{y} \sim \text{MVN}(\mathbf{X}\boldsymbol{\theta}, \mathbf{V})$ is a $n \times 1$ vector of observations for one trait, $\mathbf{X}\boldsymbol{\theta}$ and $\mathbf{V} = \sigma_g^2 \mathbf{Z}\mathbf{A}\mathbf{Z}' + \mathbf{I}_n \sigma_\varepsilon^2$ are the unconditional expectation and variance of \mathbf{y} , respectively, and $\text{var}(\mathbf{g}) = \mathbf{A}\sigma_g^2$ is the unconditional variance of $\mathbf{g} \sim \text{MVN}(\mathbf{0}, \mathbf{A}\sigma_g^2)$; $\text{var}(\mathbf{y}|\mathbf{g}) = \mathbf{I}_n \sigma_\varepsilon^2$ is the conditional variance of the vector of residuals $\boldsymbol{\varepsilon} \sim \text{MVN}(\mathbf{0}, \mathbf{I}_n \sigma_\varepsilon^2)$, $\mathbf{0}$ denotes the null expectations of \mathbf{g} and $\boldsymbol{\varepsilon}$, \mathbf{A} is the numerator relationship matrix (Lynch & Walsh, 1998; Mrode, 2014), and \mathbf{I}_n is an identity matrix of size $n \times n$.

Based on the multivariate normal distribution theory, Pearson (1903) predicted \mathbf{g} as

$$\hat{\mathbf{g}} = E(\mathbf{g}|\mathbf{y}) = \sigma_g^2 \mathbf{A}\mathbf{Z}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\theta}})$$

which now is called “single trait linear selection index” and where \mathbf{V}^{-1} is the inverse matrix of the variance of \mathbf{y} , whereas $\hat{\boldsymbol{\theta}} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}$ (Xu, 2003). Henderson (1990) indicated that the above equation is one of the main Pearson (1903) results. Currently, $\hat{\mathbf{g}} = E(\mathbf{g}|\mathbf{y})$ is called the BLUP of \mathbf{g} and is one of the main tools to make genotypic selection in plant (Piepho et al., 2008) and animal (Mrode, 2014) breeding programs.

The BLUP properties have been developed by Henderson (1949, 1950, 1963, 1984, 1990). This author also showed that the BLUP and BLP (or LPSI) are similar; however, whereas the BLUP predicts vector \mathbf{g} , the LPSI predicts a linear combination of \mathbf{g} (i.e., the net genetic merit), as shown in Equations 2–4. In addition, Mrode (2014) indicated that BLUP reduces to LPSI when no adjustments for environmental factors are needed. This means that one of the main advantages of BLUP compared with BLP occurs when records have to be pre-adjusted for fixed or for environmental effects. When the environmental effects are random, Xu (2003) writes Equation 1 as

$$\mathbf{y} = \mathbf{1}_n \mu + \mathbf{X}\boldsymbol{\gamma} + \mathbf{Z}\mathbf{g} + \mathbf{W}\boldsymbol{\pi} + \boldsymbol{\varepsilon}$$

where $\boldsymbol{\pi}$ is a vector of interaction genotypic–environment effects, $\mathbf{W} = \mathbf{X} \otimes \mathbf{Z}$, and \otimes denotes the Kronecker product (Lynch & Walsh, 1998). In this case, $E(\mathbf{y}) = \mathbf{1}_n \mu$, where $\mathbf{1}_n$ is a vector of n ones, μ is the population mean, $\mathbf{V} = \mathbf{X}\text{var}(\boldsymbol{\gamma})\mathbf{X}' + \sigma_g^2 \mathbf{Z}\mathbf{A}\mathbf{Z}' + \mathbf{W}\text{var}(\boldsymbol{\pi})\mathbf{W}' + \mathbf{I}_n \sigma_\varepsilon^2$, and $\text{var}(\boldsymbol{\gamma})$ and $\text{var}(\boldsymbol{\pi})$ are

the variance of $\boldsymbol{\gamma}$ and $\boldsymbol{\pi}$, respectively. Note that in this last case, nongenetic effects (included in $\boldsymbol{\gamma}$), such as location and year effects, are treated as random effects, whereas in Equation 1 those effects are treated as fixed and were included in $\boldsymbol{\theta}$. For additional details, see Xu (2003).

3.2 | Pearl and Surface index

The Pearl and Surface (1909) approach to the index theory is based on the index numbers theory (Ralph et al., 2015). By this reasoning, those authors called to their index “selection index numbers.” The Pearl and Surface (1909) arguments to develop the selection index numbers are the same as those indicated by Smith (1936) in his paper. Pearl and Surface (1909) defined the selection index numbers as

$$I = \frac{\mathbf{v}'_1 \mathbf{y}}{\mathbf{v}'_2 \mathbf{y}^*}$$

where \mathbf{y} is a vector of traits of economic interest for the breeder that become more desirable as their values increase, and \mathbf{y}^* is a vector of traits of economic interest that become more desirable as their values decrease. In the above equation, ν_1 and ν_2 are vector of constants to be given arbitrary values in the proportions that the different variables are to be weighted. The Pearl and Surface (1909) index can be considered a non-linear selection index and it does not maximize its correlation with the net genetic merit (Equation 2). As we have indicated in the Introduction, in this work we describe the LSIs associated to the LPSI theory and the ESIM index theory.

3.3 | BLUP in the multi-trait context

Equation 1 is only useful to predict the individual breeding values associated to one trait. However, it is easy to extend Equation 1 to the multi-trait selection context, and it is possible to include in Equation 1 non-additive genetic effects, such as dominant genetic effects (or intra-locus dominance allelic interaction) and epistasis effects (or inter-loci allelic interaction). Multi-trait evaluation is the optimum methodology to evaluate plants and animals because this method increases the accuracy of predictions when it accounts for the phenotypic and genetic correlation between individual traits (Mrode, 2014). For t traits, we can write the vector of breeding and phenotypic values as $\mathbf{g}' = [\mathbf{g}'_1 \mathbf{g}'_2 \dots \mathbf{g}'_t]$ and $\mathbf{y}' = [\mathbf{y}'_1 \mathbf{y}'_2 \dots \mathbf{y}'_t]$, respectively, from where, for two traits, Equation 1 can be written as

$$\mathbf{y}_1 = \mathbf{X}_1 \boldsymbol{\theta}_1 + \mathbf{Z}_1 \mathbf{g}_1 + \boldsymbol{\varepsilon}_1$$

for Trait 1 and as

$$\mathbf{y}_2 = \mathbf{X}_2 \boldsymbol{\theta}_2 + \mathbf{Z}_2 \mathbf{g}_2 + \boldsymbol{\varepsilon}_2$$

for Trait 2. Thus, in this context, Equation 1 is equivalent to the following model:

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{\theta}_1 \\ \boldsymbol{\theta}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{Z}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_2 \end{bmatrix} \begin{bmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \end{bmatrix} + \begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \end{bmatrix}$$

where $\mathbf{0}$ is a null matrix and the other parameters are defined in a similar manner as in Equation 1. Nevertheless, in this case $\text{var} \begin{bmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \end{bmatrix} = \mathbf{C} \otimes \mathbf{A}$, where $\mathbf{C} = \begin{bmatrix} \sigma_{g_1}^2 & \sigma_{g_{12}}^2 \\ \sigma_{g_{21}}^2 & \sigma_{g_2}^2 \end{bmatrix}$ is an additive genetic covariance matrix, and \otimes and \mathbf{A} are as defined in Equation 1. In addition, $\text{var}(\mathbf{y}_1, \mathbf{y}_2 | \mathbf{g}_1, \mathbf{g}_2) = \mathbf{R} \otimes \mathbf{I}_n$, where $\mathbf{R} = \begin{bmatrix} \sigma_{\varepsilon_1}^2 & \sigma_{\varepsilon_{12}}^2 \\ \sigma_{\varepsilon_{21}}^2 & \sigma_{\varepsilon_2}^2 \end{bmatrix}$ is a residual covariance matrix. The unconditional covariance matrix of $[\mathbf{y}'_1 \mathbf{y}'_2]$ is $\mathbf{V} = \text{var} \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} = \mathbf{Z}^* [\mathbf{C} \otimes \mathbf{A}] \mathbf{Z}^{*'} + \mathbf{R} \otimes \mathbf{I}_n$, where $\mathbf{Z}^* = \begin{bmatrix} \mathbf{Z}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_2 \end{bmatrix}$, and \mathbf{I}_n is as defined in Equation 1. It is possible to extend the above equation to any number of traits of economic interest (Mrode, 2014).

For one trait, \mathbf{g}_1 can be partitioned into additive genetic (\mathbf{a}_1) effects (or intra-locus additive allelic interaction), dominant genetic ($\boldsymbol{\delta}_1$) effects, and epistasis (\mathbf{t}_1) effects, such that $\mathbf{g}_1 = \mathbf{a}_1 + \boldsymbol{\delta}_1 + \mathbf{t}_1$. Epistasis refers to the interaction among additive and dominance genetic effects, for instance, additive by additive, dominance by dominance, additive by dominance, additive by additive by dominance, etc.; that is, $\mathbf{t}_1 = \mathbf{aa}_1 + \mathbf{dd}_1 + \mathbf{ad}_1 + \dots$

When we separate non-additive genetic effects from additive effects, we remove the confounding factors that would otherwise bias the results of the analysis (Cockerham, 1954; Fisher, 1918; Holland, 2001; Kempthorne, 1954; Mrode, 2014). In this case, the variance of \mathbf{g}_1 is $\text{var}(\mathbf{g}_1) = \mathbf{A}\sigma_a^2 + \boldsymbol{\Theta}\sigma_d^2 + (\mathbf{A}\circ\mathbf{A})\sigma_{aa}^2 + (\mathbf{A}\circ\boldsymbol{\Theta})\sigma_{ad}^2 + (\boldsymbol{\Theta}\circ\boldsymbol{\Theta})\sigma_{dd}^2 + \dots$ where $\boldsymbol{\Theta}$ is the dominance relationship matrix, “ \circ ” denotes the Hadamard product of two matrices or the element-by-element multiplication for unlinked loci (Gianola & de los Campos, 2009; Hill, 2010), σ_a^2 is the additive variance component, σ_d^2 is the deviation of dominance variance component, etc.

For t traits, assuming additive and dominance effects, and epistasis effects such as additive by additive, dominance by dominance, and additive by dominance, the variance of $\mathbf{g}' = [\mathbf{g}'_1 \mathbf{g}'_2 \dots \mathbf{g}'_t]$ can be written as

$$\mathbf{C} \otimes \mathbf{A} + \mathbf{C}_D \otimes \boldsymbol{\Theta} + \mathbf{C}_{AA} \otimes (\mathbf{A}\circ\mathbf{A}) + \mathbf{C}_{DD} \otimes (\boldsymbol{\Theta}\circ\boldsymbol{\Theta}) + \mathbf{C}_{AD} \otimes (\mathbf{A}\circ\boldsymbol{\Theta})$$

where \mathbf{C} is as defined earlier; \mathbf{C}_D is the covariance matrix of the dominant genetic interactions of effects between alleles within loci; \mathbf{C}_{AA} , \mathbf{C}_{DD} , and \mathbf{C}_{AD} are the covariance matrices of the epistasis effects among loci; and \mathbf{A} , $\boldsymbol{\Theta}$, \otimes , and “ \circ ” are as defined above. Lynch and Walsh (1998) and Mrode (2014)

have given details of how Equation 1 can be written for additive and non-additive genetic effects jointly.

Although the above results seem very interesting, non-additive effects are not transmitted to offspring in non-inbreeding populations (Lynch & Walsh, 1998); the breeding value (or additive genetic effects \mathbf{a}_1) is the only component that can be selected and is thus the main component of interest in breeding programs (Mrode, 2014). Furthermore, in the index selection context, Andersson et al. (1998) found that the introduction of the dominance variance (e.g., 25% of the total phenotypic variance) has only a small positive effect (<1.5%) on the selection response; thus, to describe the LSI theory, the assumption that only additive effects are transmitted from generation to generation seems acceptable. This last assumption implies that the combined effects of all factors that make $\mathbf{y}' = [\mathbf{y}'_1 \mathbf{y}'_2 \dots \mathbf{y}'_t]$ different from $\mathbf{g}' = [\mathbf{g}'_1 \mathbf{g}'_2 \dots \mathbf{g}'_t]$, such as non-additive effects, are in the residual vector (Hazel & Lush, 1942). Mrode (2014) indicated that non-additive effects tend to be confounded with others, such as common maternal environment; thus, breeders should take care when they estimate non-additive effects. Du and Hoeschele (2000) described methods to estimate additive and non-additive variance components in the Bayesian context. In a similar manner, using linear mixed models theory, in the phenotypic and genomic selection context, Lynch and Walsh (1998) and Su et al. (2012), respectively, described methods to estimate additive and nonadditive variance components.

As we will see later, to predict H , the LPSI uses only matrix \mathbf{C} and includes matrix \mathbf{A} only for the estimation of \mathbf{C} (Cerón-Rojas & Crossa, 2018). The LPSI is the BLP of H , not the BLUP of H (Bulmer, 1980; Cochran, 1951; Henderson, 1963). However, as we have indicated in Equation 1, the BLP (LPSI) and BLUP theory are related, but, as we have indicated earlier, whereas the BLUP predicts $\mathbf{g}' = [\mathbf{g}'_1 \mathbf{g}'_2 \dots \mathbf{g}'_t]$, the LPSI predicts linear combination of the elements of \mathbf{g} (see Equations 2–4).

3.4 | Matrix \mathbf{P} as a special case of matrix \mathbf{V}

We have indicated that in the LSI context the assumption that only additive effects are transmitted from generation to generation is acceptable. For t traits, matrix \mathbf{V} is equal to

$$\text{var} \begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_t \end{bmatrix} = \mathbf{Z}^*[\mathbf{C} \otimes \mathbf{A}]\mathbf{Z}'^* + \mathbf{R} \otimes \mathbf{I}_n = \mathbf{V}$$

where

$$\mathbf{Z}^* = \begin{bmatrix} \mathbf{Z}_1 & 0 & \dots & 0 \\ 0 & \mathbf{Z}_2 & \dots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \dots & 0 & \mathbf{Z}_t \end{bmatrix}$$

whereas matrices \mathbf{C} and \mathbf{R} are of size $t \times t$. Here, matrix \mathbf{V} is associated with the prediction of vector $\mathbf{g}' = [\mathbf{g}'_1 \mathbf{g}'_2 \dots \mathbf{g}'_t]$ (Equation 1). Now suppose that matrix $\mathbf{Z}^* = \mathbf{I}_{nt}$ (where \mathbf{I}_{nt} is an identity matrix of size $nt \times nt$) and that the level of inbreeding among individuals is null; then $\mathbf{A} = \mathbf{I}_n$, and the matrix \mathbf{V} can be written as

$$\mathbf{V} = \mathbf{C} + \mathbf{R}$$

As we shall see later, in this last case matrix \mathbf{V} is associated with the BLP (LPSI) theory. Hereafter we denote it as \mathbf{P} , which is the usual notation in the LPSI context.

By the foregoing results it is evident that, in the prediction of the net genetic merit (Equation 2), the LSI theory does not use matrix \mathbf{A} . However, matrix \mathbf{A} is used for the estimation of \mathbf{C} (Cerón-Rojas & Crossa, 2018). This seems to be a disadvantage of the LPSI theory with respect to the BLUP theory. As we have indicated earlier, BLUP predicts vector $\mathbf{g}' = [\mathbf{g}'_1 \mathbf{g}'_2 \dots \mathbf{g}'_t]$, not the net genetic merit; this means that BLUP is associated only with a part of the LPSI theory: that associated to the expected genetic gain per trait (see Equations 8 and 10), which also predict the mean values of $\mathbf{g}' = [\mathbf{g}'_1 \mathbf{g}'_2 \dots \mathbf{g}'_t]$. Jeyaruban et al. (1995, Abstract, p. 1) compared the efficiency of BLUP and the LPSI and concluded that: “The relative selection response with the selection indices compared to the BLUP estimates were from 94.5 to 99.4% of BLUP. Inbreeding was higher in the BLUP selected populations, which could offset any advantage of BLUP.”

4 | UNCONSTRAINED LINEAR SELECTION INDICES

4.1 | The LPSI theory

In this section, we describe the LPSI theory and all the LSIs associated to this theory.

4.2 | Basic conditions to construct a valid LPSI in the single-stage selection context

Conditions to construct a valid LPSI are as follows. (a) The phenotypic value is additively composed of two main parts: the genotypic value and the environmental contribution. (b) The genotypic value is the individual breeding value; thus, it is composed only of additive genes effects. (c) The net genetic merit of an individual is the genotypic economic value (Kempthorne & Nordskog, 1959). (d) Selection should be made at only one stage of the plant's or animal's life cycle. (e) The generations do not overlap. (f) All individuals below a certain level of desirability are culled without exception. (g) Selected individuals have an equal

opportunity to have offspring (Hazel & Lush, 1942). (h) The LPSI values in the c^{th} selection cycle and the LPSI values in the $(c + 1)^{\text{th}}$ selection cycle are not correlated. (i) The correlation between the LPSI and the net genetic merit should be at its maximum in each selection cycle. These conditions are valid for all unconstrained or constrained LPSIs in the single-stage selection context. Cerón-Rojas and Crossa (2018) have given conditions for the unconstrained and constrained LGSIs.

4.3 | The net genetic merit

Let $\mathbf{g}_q' = [G_{q1} G_{q2} \dots G_{qt}]$ be a vector of true unobservable genotypic random variables for t traits with multivariate normal distribution and null expectation. Then the q^{th} ($q = 1, 2, \dots, n; n = \text{number of individuals}$) individual net genetic merit is

$$H_q = \mathbf{w}'\mathbf{g}_q \quad (2)$$

where $\mathbf{w}' = [w_1 w_2 \dots w_t]$ is a $1 \times t$ ($t = \text{number of traits}$) vector of known economic weights. The variance of H_q is denoted as $\sigma_H^2 = \mathbf{w}'\mathbf{C}\mathbf{w}$. Smith (1936) called Equation 2 “the genetic value of a plant or line.”

In Equation 2, $\mathbf{g}_q' = [G_{q1} G_{q2} \dots G_{qt}]$ is a vector of random variables whereas in the BLUP equation described earlier, $\mathbf{g}' = [\mathbf{g}'_1 \mathbf{g}'_2 \dots \mathbf{g}'_t]$ is a vector of particular realization of breeding values. Note that using the BLUP of $\mathbf{g}' = [\mathbf{g}'_1 \mathbf{g}'_2 \dots \mathbf{g}'_t]$ (i.e., $\hat{\mathbf{g}}' = [\hat{\mathbf{g}}'_1 \hat{\mathbf{g}}'_2 \dots \hat{\mathbf{g}}'_t]$), a predictor of Equation 2 is:

$$\hat{\mathbf{h}} = \sum_{j=1}^t w_j \hat{\mathbf{g}}_j$$

where $\hat{\mathbf{h}}' = [\hat{H}_1 \hat{H}_2 \dots \hat{H}_n]$ is a vector of estimated net genetic merit values for n individuals, and w_j is the j^{th} element of the vector of economic weights \mathbf{w} (Dekkers, 1997; Portes et al., 2020). The predictor $\hat{\mathbf{h}}$ combines the BLUP and LPSI theory to predict the net genetic merit.

4.4 | The net genetic merit and its relationship with \mathbf{y}

For the q^{th} individual, the relationship between Equation 2 and the vector of random phenotypic variables $\mathbf{y}'_q = [Y_{q1} Y_{q2} \dots Y_{qt}]$ for t traits is

$$H_q = m_H + \boldsymbol{\beta}'(\mathbf{y}_q - \boldsymbol{\mu}) + e_q \quad (3)$$

where m_H is the mean of H_q , $\boldsymbol{\mu}' = [\mu_1 \mu_2 \dots \mu_t]$ is a vector $1 \times t$ of phenotypic means of \mathbf{y}_q , and e_q is the deviation representing the extent to which $\boldsymbol{\beta}'(\mathbf{y}_q - \boldsymbol{\mu})$ differs from the expected value

of H_q (Young, 1964). In Equation 3, $\mathbf{y}'_q = [Y_{q1} Y_{q2} \dots Y_{qt}]$ is a vector of random variables, whereas in the BLUP equation described earlier, $\mathbf{y}' = [\mathbf{y}'_1 \mathbf{y}'_2 \dots \mathbf{y}'_t]$ is a vector of particular realization of phenotypes values for t traits. In addition, the j^{th} ($j = 1, 2, \dots, t$) element of vector $\boldsymbol{\mu}$ is equivalent to $\mathbf{X}\boldsymbol{\theta}$ (Equation 1) when $\mathbf{X} = \mathbf{1}$ (a vector of 1s) and $\boldsymbol{\theta} = \boldsymbol{\mu}_j$. We have assumed that e_q has a normal distribution, null expectation, and variance σ_e^2 and that the covariance between any pairs of e_q and e_r ($q \neq r$) is zero. As we will see later, $\boldsymbol{\beta}'(\mathbf{y}_q - \boldsymbol{\mu}) = I_q$ is the LPSI. It is possible to show that the random error (e_q) and the LPSI are uncorrelated (Bickel & Doksum, 2015). The variance of H_q can be written as $\sigma_H^2 = \mathbf{w}'\mathbf{C}\mathbf{w} = \sigma_I^2 + \sigma_e^2$, where $\sigma_I^2 = \boldsymbol{\beta}'\mathbf{P}\boldsymbol{\beta}$ is the variance of I_q , and $\mathbf{P} = \mathbf{C} + \mathbf{R}$ is the phenotypic covariance matrix of $\mathbf{y}'_q = [Y_{q1} Y_{q2} \dots Y_{qt}]$; \mathbf{C} is as defined earlier, and \mathbf{R} is the residual covariance matrix. All these last three matrices are of size $t \times t$.

Equation 3 is a multiple linear regression model where the independent (\mathbf{y}_q) and dependent (H_q) variables are random and have a joint multivariable normal distribution (Rencher & Schaalje, 2008). Nevertheless, whereas in the standard regression model the values of H_q and \mathbf{y}_q are observables, in Equation 3 only the \mathbf{y}_q values are observable. For this reason, H_q should be predicted using, for example, the BLP.

4.5 | The BLP of H

Hereafter we will denote the vector of phenotypic trait random variables as \mathbf{Y} and its particular realization values as \mathbf{y} ; and, to simplify notation, we will denote H_q and I_q as H and I , respectively. Suppose that H and \mathbf{Y} have a joint multivariate normal distribution and that we are looking for a linear predictor of H assuming that \mathbf{Y} is the only information that we can use to predict the unobservable random variable H . We want to find a predictor function of \mathbf{Y} such that this predictor is “close” to H in the sense of the MSPE, which is a measure used in the mathematical theory of prediction (Bickel, 2015). Suppose that $E(H) = m_H$, $E(\mathbf{Y}) = \boldsymbol{\mu}$, and that $\text{var}(H)$, $\text{var}(\mathbf{Y}) = \mathbf{P}$, and the covariance between H and \mathbf{Y} [$\text{cov}(H, \mathbf{Y}) = \mathbf{w}'\mathbf{C}$] exist and are finite, then the BLP of H (or the LPSI) is

$$E(H|\mathbf{Y} = \mathbf{y}) = m_H + \mathbf{w}'\mathbf{C}\mathbf{P}^{-1}(\mathbf{y} - \boldsymbol{\mu}) = \beta_0 + \boldsymbol{\beta}'\mathbf{y} \quad (4)$$

where $\beta_0 = m_H - \boldsymbol{\beta}'\boldsymbol{\mu}$, $\mathbf{w}'\mathbf{C}$ and $\boldsymbol{\mu}$ are as defined above, \mathbf{P}^{-1} is the inverse of the matrix \mathbf{P} , and $\boldsymbol{\beta} = \mathbf{P}^{-1}\mathbf{C}\mathbf{w}$. Cochran (1951) was the first to indicate that the LPSI is the BLP of H . Equation 4 is unique and unbiased (Christensen, 2011; Searle et al., 2006) and represents a model that is linear in \mathbf{y} as well as in $\boldsymbol{\beta}$. In addition, it indicates that the problem of predicting H is simply estimating its conditional mean, whereas the LPSI variance is $\sigma_I^2 = \boldsymbol{\beta}'\mathbf{P}\boldsymbol{\beta}$. The assumption of multivariate normal distribution of H and \mathbf{Y} is a sufficient condition for

Equation 4 to be linear; moreover, any function that is expressible as Equation 4 is regarded as linear even if the vector \mathbf{g} depends on the joint distribution of H and \mathbf{Y} (Bickel & Doksum, 2015; Harville, 2018). Smith (1936) called Equation 4 “discriminant function.”

By Equations 1–4, the random vectors \mathbf{g} and \mathbf{Y} have joint multivariate normal distribution with mean $E(\mathbf{g}) = \mathbf{0}$ and $E(\mathbf{Y}) = \boldsymbol{\mu}$, and covariance matrix $\text{var} \begin{pmatrix} \mathbf{Y} \\ \mathbf{g} \end{pmatrix} = \begin{bmatrix} \mathbf{P} & \mathbf{C} \\ \mathbf{C} & \mathbf{C} \end{bmatrix}$, where \mathbf{P} and \mathbf{C} are as defined above. Moreover, the distribution of LPSI and $H = \mathbf{w}'\mathbf{g}$ values is a bivariate normal distribution, with mean $E(H) = m_H$ and $E(I) = m_I$ and covariance matrix $\text{var} \begin{bmatrix} H \\ I \end{bmatrix} = \begin{bmatrix} \mathbf{w}'\mathbf{C}\mathbf{w} & \mathbf{w}'\mathbf{C}\boldsymbol{\beta} \\ \mathbf{w}'\mathbf{C}\boldsymbol{\beta} & \boldsymbol{\beta}'\mathbf{P}\boldsymbol{\beta} \end{bmatrix} = \begin{bmatrix} \sigma_H^2 & \sigma_{HI} \\ \sigma_{HI} & \sigma_I^2 \end{bmatrix}$, where $\boldsymbol{\beta}$, \mathbf{w}' , σ_H^2 , and σ_I^2 were defined earlier and $\sigma_{HI} = \mathbf{w}'\mathbf{C}\boldsymbol{\beta}$ is the covariance between H and LPSI.

As we will see later, because we can assume that $\beta_0 = m_H - \boldsymbol{\beta}'\boldsymbol{\mu}$ is a fixed constant, the maximized LPSI expected genetic gain per trait (Equation 10), the maximized correlation between H and LPSI (Equation 11), and the maximized LPSI selection response (Equation 12), are not affected if in Equation 4 we assume that $E(H) = 0$ and $E(\mathbf{Y}) = \mathbf{0}$. Likewise, when we select parents for the next generation, they will be the same in both cases, that is, when $E(H) = 0$ and $E(\mathbf{Y}) = \mathbf{0}$ and when $E(H) = m_H$ and $E(\mathbf{Y}) = \boldsymbol{\mu}$. Thus, we think that there are not loss of generality or loss of accuracy to describe the LSI theory if we assume that $E(H) = 0$ and $E(\mathbf{Y}) = \mathbf{0}$. In practice, $E(\mathbf{Y}) = \mathbf{0}$ only if \mathbf{y} is centered with respect to $\boldsymbol{\mu}$.

4.6 | The LPSI prediction error

Because $\boldsymbol{\beta} = \mathbf{P}^{-1}\mathbf{C}\mathbf{w}$, then $\sigma_{HI} = \mathbf{w}'\mathbf{C}\boldsymbol{\beta}$ and $\sigma_I^2 = \boldsymbol{\beta}'\mathbf{P}\boldsymbol{\beta}$ are the same, from where the variance of the LPSI prediction error is

$$\text{var}(H|I) = \sigma_H^2 - \frac{\sigma_{HI}^2}{\sigma_I^2} = \sigma_H^2 - \sigma_e^2 = \sigma_e^2 \quad (5)$$

Equation 5 and equation $\sigma_e^2 = \sigma_H^2(1 - \rho_{HI}^2)$ are the same, where $\rho_{HI}^2 = (\sigma_{HI}^2)/(\sigma_H^2\sigma_I^2)$ and $\sigma_{HI}^2 = (\mathbf{w}'\mathbf{C}\boldsymbol{\beta})^2$ are, respectively, the square of correlation and the square of covariance between H and LPSI. We defined the other parameters earlier.

The conditional mean of H is dependent on \mathbf{y} (Equation 4); however, the conditional variance of H (Equation 5) is independent of \mathbf{y} . Thus, under the multivariate normal assumption, Equations 4 and 5 provide a linear model with constant variance (Rencher & Schaalje, 2008). Moreover, if the marginal density of LPSI is normal and if the conditional density of H given LPSI is normal for every LPSI value, the assumption of bivariate normality of LPSI and H

is guaranteed (Arnold et al., 1999; Kagan & Wesolowski, 1996).

4.7 | The intensity of selection in the single-stage context

Let I be transformed into a random variable T , as $T = (I - m_I)/\sigma_I$, where $E(I) = m_I$ and σ_I is the standard deviation of I . Assume that all I values higher than I^* value corresponding to any proportion of selected individuals (p) will be selected; then the variables $T^* = (I^* - m_I)/\sigma_I$ and T are related to selection intensity (k) as

$$k = \frac{1}{p}E(T) = \frac{1}{p} \int_{\tau^*}^{\infty} \frac{\tau}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\tau^2\right\} d\tau = \frac{z(\tau^*)}{p} \quad (6)$$

where $z(\tau^*) = \frac{\exp\{-0.5\tau^{*2}\}}{\sqrt{2\pi}}$ is the height of the ordinate of the normal curve at the lowest value of τ^* retained, and p is the proportion of the population of animal or plant lines that will be selected (Smith, 1936). In the selected population, Equation 6 is the unconditional expectation of the random variable T [$E(T)$] weighted by p and was obtained by Smith (1936). It is possible to show that the variance of T (weighted by p), over the interval (τ^*, ∞) , is equal to $1 + k(\tau^* - k)$. Several authors (Hazel & Lush, 1942; Kempthorne & Nordskog, 1959; Xu & Muir, 1992; Young, 1964) have given additional details of Equation 6.

4.8 | The conditional expectation of H and \mathbf{g} given I

Suppose that the expectations of H and \mathbf{g} are null, whereas the expectation of I is $E(I) = m_I$; then the conditional expectation of H and \mathbf{g} given I is, respectively,

$$E(H|I) = \frac{\text{Cov}(H, I)}{\sigma_I^2}(I - m_I) = \frac{\boldsymbol{\beta}'\mathbf{C}\mathbf{w}}{\sigma_I} \frac{I - m_I}{\sigma_I} = \frac{\boldsymbol{\beta}'\mathbf{C}\mathbf{w}}{\sigma_I} T \quad (7)$$

and

$$E(\mathbf{g}|I) = \frac{\text{Cov}(\mathbf{g}, I)}{\sigma_I^2}(I - m_I) = \frac{\mathbf{C}\boldsymbol{\beta}}{\sigma_I} \frac{I - m_I}{\sigma_I} = \frac{\mathbf{C}\boldsymbol{\beta}}{\sigma_I} T \quad (8)$$

where $\text{cov}(H, I) = \boldsymbol{\beta}'\mathbf{C}\mathbf{w}$ is the covariance between H and I , $\text{cov}(\mathbf{g}, I) = \mathbf{C}\boldsymbol{\beta}$ is the covariance between \mathbf{g} and I , $T = (I - m_I)/\sigma_I$, and $\sigma_I^2 = \boldsymbol{\beta}'\mathbf{P}\boldsymbol{\beta}$ and σ_I are the variance and standard deviations of I , respectively. In Equation 7, $\text{cov}(H, I)/\sigma_I^2$ is the relationship between $E(H|I)$ and I (or the regression of H on I), and $\text{cov}(\mathbf{g}, I)/\sigma_I^2$ is the relationship between $E(\mathbf{g}|I)$ and I in Equation 8.

4.9 | The LPSI selection response, expected genetic gain per trait, and correlation

By Equations 6–8, the LPSI selection response (R) and expected genetic gain per trait (\mathbf{E}) are, respectively,

$$R = k \frac{\boldsymbol{\beta}'\mathbf{C}\mathbf{w}}{\sigma_I} = k\sigma_H\rho_{HI} \quad (9)$$

and

$$\mathbf{E} = k \frac{\mathbf{C}\boldsymbol{\beta}}{\sigma_I} \quad (10)$$

where k was defined in Equation 6, $\sigma_H = \sqrt{\mathbf{w}'\mathbf{C}\mathbf{w}}$ is the standard deviation of H , and $\rho_{HI} = \sigma_{HI}/\sigma_H\sigma_I$ is the correlation between H and LPSI. All the other parameters of Equations 9 and 10 were defined earlier. Smith (1936) obtained Equation 9, and Hanson and Johnson (1957) were the first to describe Equation 10. In Equation 9, R is a scalar; in Equation 10, \mathbf{E} is a vector $t \times 1$ (t = number of traits) of the expected genetic gain value of each trait. In addition, R is the expectation of H , and \mathbf{E} is the expectation of \mathbf{g} .

To maximize the selection response (Equation 9), Smith (1936) used differential calculus to derive Equation 9 with respect to $\boldsymbol{\beta}$. The result reported by this author was the vector of regression coefficients $\boldsymbol{\beta} = \mathbf{P}^{-1}\mathbf{C}\mathbf{w}$. He indicated that $\boldsymbol{\beta} = \mathbf{P}^{-1}\mathbf{C}\mathbf{w}$ should maximize the selection response. Thus, Smith (1936) did not use the BLP theory to develop the LPSI theory because using differential calculus is not necessary to obtain $\boldsymbol{\beta} = \mathbf{P}^{-1}\mathbf{C}\mathbf{w}$ in the BLP context (Equation 4).

4.10 | The maximized correlation between H and I

Because $\boldsymbol{\beta} = \mathbf{P}^{-1}\mathbf{C}\mathbf{w}$, the maximized correlation between H and I is

$$\rho_{\max} = \frac{\sqrt{\boldsymbol{\beta}'\mathbf{P}\boldsymbol{\beta}}}{\sqrt{\mathbf{w}'\mathbf{C}\mathbf{w}}} = \frac{\sqrt{\mathbf{w}'\mathbf{C}\mathbf{P}^{-1}\mathbf{C}\mathbf{w}}}{\sqrt{\mathbf{w}'\mathbf{C}\mathbf{w}}} \quad (11)$$

where all the terms were defined earlier. Equation 11 represents the proportion of the variance of H that can be attributed to the regression relationship with I . The covariance (σ_{HI}) and correlation are good measures of relationship only for variables with linear trends and are generally unsuitable for non-normal random variables with a curvilinear relationship (Rencher, 2002). That is, Equation 11 is a measure of linear dependency between H and I and is called the “multiple correlation coefficient,” whereas $\rho_{HI_{\max}}^2 = \frac{\sigma_I^2}{\sigma_H^2}$ is called “the coefficient of determination or squared multiple correlation” (Rencher & Schaalje, 2008). Note that while $\rho_{HI} = \sigma_{HI}/\sigma_H\sigma_I$

can take any value in the interval -1.0 and 1.0 , Equation 11 gives the maximum value of ρ_{HI} , and when $\rho_{HI} = 0$, H and I are independent.

4.11 | The maximized selection response

Once again, because $\boldsymbol{\beta} = \mathbf{P}^{-1}\mathbf{C}\mathbf{w}$, the maximized LPSI selection response is

$$R_{\max} = k\sqrt{\boldsymbol{\beta}'\mathbf{P}\boldsymbol{\beta}} = k\sqrt{\mathbf{w}'\mathbf{C}\mathbf{P}^{-1}\mathbf{C}\mathbf{w}} \quad (12)$$

Equation 12 predicts the mean improvement in H attributable to indirect selection on I only when $\boldsymbol{\beta} = \mathbf{P}^{-1}\mathbf{C}\mathbf{w}$ and is proportional to $\sqrt{\boldsymbol{\beta}'\mathbf{P}\boldsymbol{\beta}}$ and k . Thus, whereas in Equation 9 the selection response can take any value, Equation 12 gives the maximum value of Equation 9. This is the main difference between the two equations.

Cerón-Rojas et al. (2006) showed that Equation 12 is related to the Cauchy–Schwarz inequality, which corroborates that $\boldsymbol{\beta} = \mathbf{P}^{-1}\mathbf{C}\mathbf{w}$ is a global minimum when the mean squared difference between I and H , $E[(H - I)^2]$, is minimized, and a global maximum when the ρ_{HI} correlation between I and H is maximized because Equation 12 is true only when $\boldsymbol{\beta} = \mathbf{P}^{-1}\mathbf{C}\mathbf{w}$.

4.12 | The maximized correlation, selection response, and expected genetic gains have the same form for all LSIs

Equations 9–12 have the same form for all LSIs described in this review. Thus, as we will see later, the only difference between those equations and the selection response, expected genetic gain per trait, and correlation, associated with the other LSIs, is the type of information that the LSI vector of coefficients and the matrices contain.

4.13 | Summary of the LPSI statistical properties

When H and I have a joint bivariate normal distribution; $\boldsymbol{\beta} = \mathbf{P}^{-1}\mathbf{C}\mathbf{w}$; and \mathbf{P} , \mathbf{C} , and \mathbf{w} are known, the statistical LPSI properties are the following:

1. Equations 10–12 are maximum.
2. The variance of I and the covariance between H and I are the same.
3. The maximized correlation between H and I is equal to Equation 11.
4. The LPSI is unbiased in the sense that $E(I) = E(H)$ (Searle et al., 2006).

5. The variance of the predicted error (Equation 5) is minimal.
6. The total variance of H explained by I is $\sigma_I^2 = \rho_{HI}^2 \sigma_H^2$.
7. Selecting any upper fraction of the population based on values of I (Equation 4) ensures that for that selected proportion, $E(H)$ is maximized (Equation 9). That is, with truncation selection, Equation 4 maximizes the expected mean of the selected individuals (Cochran, 1951; Henderson, 1990).
8. The I values are appropriate to make selection, i.e., selecting the plants or animals with the highest predictions to be parents of the next generation (Searle et al., 2006).
9. The LPSI is the BLP of H in the sense that it (a) is a linear function of \mathbf{y} ; (b) is “best” in the sense that it has minimum mean squared error of prediction; and (c) is optimal and unique among the class of linear predictors based on \mathbf{y} (Robinson, 1991; Searle et al., 2006).

Cerón-Rojas and Crossa (2018) showed that all the LSIs associated with the LPSI have similar properties as those described in Points 1–9, as exemplified by CLPSI and CLGSI. This occurs because the CLPSI and CLGSI vector of coefficients are linear combinations of the LPSI and LGSI vector of coefficients or projections of these vectors to a different space.

4.14 | Multi-stage LPSIs

One stage is the animal or plant age at which breeders can measure and select the trait of economic interest. Cochran (1951) and Young (1964) combined the LPSI theory with the independent culling method and developed the optimum multi-stage LPSI (OMLPSI), which can include several stages in each selection cycle. The OMLPSI takes into consideration the correlation among indices at different stages when making selections and requires multiple integration techniques to derive selection intensities. In addition, after the first selection stage, the OMLPSI values could be non-normally distributed, and there are problems of convergence when the traits and the index values are highly correlated (Börner & Reinsch, 2012).

Xu and Muir (1992) developed the decorrelated multi-stage LPSI (DMLPSI) as one possible solution to the OMLPSI problems. After the first selection stage, the DMLPSI values could be normally distributed, and this index does not require multiple integration techniques to derive selection intensities because it maximizes the correlation between the index and the net genetic merit at each stage under the restriction that the covariance between the DMLPSI values at different stages be zero. To obtain truncation points and selection intensities, Xu and Muir (1992) derived a set of nonlinear equations that can be solved iteratively using a multidimensional New-

ton method. One problem associated with DMLPSI is that the selection responses and correlation with H are lower than the OMLPSI selection response and correlation after the first selection stage (Cerón-Rojas et al., 2019a, 2019b).

4.15 | The OMLPSI

As we have indicated earlier, there is no loss of generality if we assume that $E(H) = 0$ and $E(\mathbf{y}) = \mathbf{0}$ to describe the LSI theory. Thus, to simplify notation, hereafter we will assume that $E(H) = 0$ and $E(\mathbf{y}) = \mathbf{0}$. These last two assumptions are common in the multi-stage LPSI context (Xie & Xu, 1997, 1998; Xu & Muir, 1992; Young, 1964). In practice, $E(\mathbf{y}) = \mathbf{0}$ only if \mathbf{y} is centered with respect to $\boldsymbol{\mu} = E(\mathbf{y})$.

Let $\mathbf{y}' = [Y_1 \ Y_2 \ \dots \ Y_t]$ be a vector of random variables for t traits. We can partition this vector into S ($t_i < S < t$, where t_i is the number of traits at Stage i) subvectors as $\mathbf{y}' = [\boldsymbol{\omega}'_1 \ \boldsymbol{\omega}'_2 \ \dots \ \boldsymbol{\omega}'_S]$, where $\boldsymbol{\omega}'_i = [Y_1 \ Y_2 \ \dots \ Y_{t_i}]$ is a subvector of \mathbf{y} at Stage i . Thus, at Stage S , $t = \sum_{i=1}^S t_i$ is the total number of selected traits from vector \mathbf{y} and, for each stage, the BLP (or OMLPSI) of H is

$$E(H|\boldsymbol{\omega}_1) = \boldsymbol{\beta}'_1 \boldsymbol{\omega}_1, \quad E(H|\boldsymbol{\omega}_1, \boldsymbol{\omega}_2) = \boldsymbol{\beta}'_2 \boldsymbol{\omega}_2^*, \quad \dots, \\ E(H|\boldsymbol{\omega}_1, \boldsymbol{\omega}_2, \dots, \boldsymbol{\omega}_S) = \boldsymbol{\beta}'_S \mathbf{y} \quad (13)$$

where $\boldsymbol{\omega}'_2 = [\boldsymbol{\omega}'_1 \ \boldsymbol{\omega}'_2]$. In Equation 13, the BLP changes at each stage; however, the form of this predictor is the same as that of the LPSI.

At Stage i , let $I_i = \boldsymbol{\beta}'_i \boldsymbol{\omega}_i$ be the index, where $\boldsymbol{\beta}'_i = [\beta_{i1} \ \beta_{i2} \ \dots \ \beta_{it_i}]$ is the OMLPSI vector of coefficients, and $\boldsymbol{\omega}_i$ and t_i are as defined earlier; then we can construct the OMLPSIs as follows:

$$\begin{bmatrix} I_1 \\ I_2 \\ \vdots \\ I_S \end{bmatrix} = \begin{bmatrix} \boldsymbol{\beta}'_1 & 0 & \dots & 0 \\ \boldsymbol{\beta}'_1 & \boldsymbol{\beta}'_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\beta}'_1 & \boldsymbol{\beta}'_2 & \dots & \boldsymbol{\beta}'_S \end{bmatrix} \begin{bmatrix} \boldsymbol{\omega}_1 \\ \boldsymbol{\omega}_2 \\ \vdots \\ \boldsymbol{\omega}_S \end{bmatrix}$$

This last result indicates that until Stage $S - 1$, each index is partial, but at Stage S , $I_S = \boldsymbol{\beta}'_1 \boldsymbol{\omega}_1 + \boldsymbol{\beta}'_2 \boldsymbol{\omega}_2 + \dots + \boldsymbol{\beta}'_S \boldsymbol{\omega}_S$ is a whole index. Young (1964) called to this procedure “the part and whole index selection,” whereas Xu and Muir (1992) called it “selection index updating” because as traits become available, each subsequent index contains all traits available up to that current stage. The above equation is valid for any unconstrained or constrained multi-stage index.

4.16 | The OMLPSI vector of coefficients

Let $\mathbf{P} = \{\mathbf{P}_{ij}\}$ be the covariance matrix of vector $\mathbf{y}' = [\boldsymbol{\omega}'_1 \ \boldsymbol{\omega}'_2 \ \dots \ \boldsymbol{\omega}'_S]$, where $\mathbf{P}_{ij} = \text{cov}(\boldsymbol{\omega}_i, \boldsymbol{\omega}_j)$ is the ij^{th} ($i, j = 1, 2, \dots, S$)

submatrix of \mathbf{P} , and let $\mathbf{\Lambda} = \{\text{cov}(\boldsymbol{\omega}_i, \mathbf{g})\}' = [\mathbf{C}_1 \ \mathbf{C}_2 \ \dots \ \mathbf{C}_S]$ be the covariance matrix between $\boldsymbol{\omega}_i$ ($i = 1, 2, \dots, S$) and \mathbf{g} , where $\text{cov}(\boldsymbol{\omega}_i, \mathbf{g}) = \mathbf{C}_i$ is a submatrix of matrix $\mathbf{\Lambda}$ at Stage i . Moreover, let \mathbf{Q}_{ii} and $\mathbf{\Lambda}'_i = [\mathbf{C}_1 \ \mathbf{C}_2 \ \dots \ \mathbf{C}_i]$ be submatrices of \mathbf{P} and $\mathbf{\Lambda}$ until Stage i (Cerón-Rojas et al. 2019a), respectively, then the OMLPSI vector of coefficients that maximizes Equations 9 and 10 (and the correlation between OMLPSI and H) at Stage i is

$$\boldsymbol{\beta}_i = \mathbf{Q}_{ii}^{-1} \mathbf{\Lambda}'_i \mathbf{w} \quad (14)$$

where \mathbf{Q}_{ii}^{-1} is the inverse of matrix \mathbf{Q}_{ii} , and $\mathbf{w}' = [w_1 \ w_2 \ \dots \ w_t]$ is as defined earlier.

4.17 | Maximized selection response, expected genetic gain per trait, and correlation at Stage i

By Equation 14, the maximized selection response (R_{\max_i}), expected genetic per trait (\mathbf{E}_{\max_i}), and correlation (ρ_{\max_i}) at Stage i , respectively, are

$$\begin{aligned} R_{\max_i} &= k_i \sqrt{\boldsymbol{\beta}'_i \mathbf{Q}_{ii} \boldsymbol{\beta}_i}, \mathbf{E}_{\max_i} = k_i \frac{\mathbf{\Lambda}_i \boldsymbol{\beta}_i}{\sqrt{\boldsymbol{\beta}'_i \mathbf{Q}_{ii} \boldsymbol{\beta}_i}}, \text{ and } \rho_{\max_i} \\ &= \frac{\sqrt{\boldsymbol{\beta}'_i \mathbf{Q}_{ii} \boldsymbol{\beta}_i}}{\sqrt{\mathbf{w}' \mathbf{C} \mathbf{w}}} \end{aligned} \quad (15)$$

where k_i is the selection intensity. For all stages, the total selection response and expected genetic gain per trait are $R_T = \sum_{i=1}^S R_{\max_i}$ and $\mathbf{E}_T = \sum_{i=1}^S \mathbf{E}_{\max_i}$, respectively.

4.18 | DMLPSI

4.18.1 | Covariance matrix to make the DMLPSI values among stages null

Let $\boldsymbol{\omega}'_{(i-1)} = [\boldsymbol{\omega}'_1 \ \boldsymbol{\omega}'_2 \ \dots \ \boldsymbol{\omega}'_{i-1}]$ and $\boldsymbol{\omega}'_{(i)} = [\boldsymbol{\omega}'_1 \ \boldsymbol{\omega}'_2 \ \dots \ \boldsymbol{\omega}'_i]$ (where $\boldsymbol{\omega}'_i = [Y_1 \ Y_2 \ \dots \ Y_{t_i}]$ is a subvector of $\mathbf{y}' = [Y_1 \ Y_2 \ \dots \ Y_t]$ at Stage i) be subvectors of \mathbf{y} until Stages $i - 1$ and i , and let t_{i-1} and $t_{i-1} + t_i$ be the size of vectors $\boldsymbol{\omega}'_{(i-1)}$ and $\boldsymbol{\omega}'_{(i)}$, respectively; then the covariance matrix between $\boldsymbol{\omega}'_{(i-1)}$ and $\boldsymbol{\omega}'_{(i)}$ ($\mathbf{Q}_{(i-1)i}$) will be of size $t_{i-1} \times (t_{i-1} + t_i)$ and can be written as

$$\mathbf{Q}_{(i-1)i} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1(n_{i-1})} & \sigma_{1(n_{i-1}+n_i)} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2(n_{i-1})} & \sigma_{2(n_{i-1}+n_i)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \sigma_{(n_{i-1})1} & \sigma_{(n_{i-1})2} & \dots & \sigma_{(n_{i-1})(n_{i-1})} & \sigma_{(n_{i-1})(n_{i-1}+n_i)} \end{bmatrix}$$

Matrix $\mathbf{Q}_{(i-1)i}$ is a nonsquare and nonsymmetric phenotypic variance matrix and is useful by making the DMLPSI

and constrained decorrelated multi-stage LPSI (CDMLPSI) values independent among stages. By definition, a covariance matrix is symmetric; however, in this case, matrix $\mathbf{Q}_{(i-1)i}$ is of size $t_{i-1} \times (t_{i-1} + t_i)$ because until Stages $i - 1$ there are t_{i-1} traits, whereas at Stage i there are $t_{i-1} + t_i$ traits. This means that at Stage $i - 1$ we have an index with t_{i-1} traits, whereas at Stage i we have an index with $t_{i-1} + t_i$; that is, the trait selected until Stage $i - 1$ and the traits selected until Stage i . By this reasoning, as we indicated earlier, Young (1964) called this procedure “the part and whole index selection,” whereas Xu and Muir (1992) called it “selection index updating” because as traits become available, each subsequent index contains all traits available up to that current stage.

4.18.2 | The DMLPSI vector of coefficients at Stage i

Let $I_{i-1} = \mathbf{b}'_{i-1} \boldsymbol{\omega}_{i-1}$ and $I_i = \mathbf{b}'_i \boldsymbol{\omega}_i$ be the DMLPSIs at Stages $i - 1$ and i , respectively, and let $\mathbf{J}'_{i-1} = [I_1 \ I_2 \ \dots \ I_{i-1}]$ be a vector of DMLPSIs values until Stage $i - 1$. We need to obtain the DMLPSI vector of coefficients at Stage i such that the covariance between I_i and \mathbf{J}_{i-1} will be null [i.e., $\text{cov}(I_i, \mathbf{J}_{i-1}) = \mathbf{0}$]. By this restriction, I_i and I_{i-1} are not correlated; hence the name “decorrelated multi-stage linear phenotypic selection index.” Xu and Muir (1992) showed that $\text{cov}(I_i, \mathbf{J}_{i-1}) = \mathbf{B}'_{i-1} \mathbf{Q}_{(i-1)i} \mathbf{b}_i$, where, \mathbf{B}_{i-1} is an upper triangular matrix (Cerón-Rojas et al. 2019a) where each column contain the vector of coefficients of the DMLPSI, $\mathbf{b}'_{i-1} = [b_{(i-1)1} \ b_{(i-1)2} \ \dots \ b_{(i-1)n_{i-1}}]$ is the DMLPSI vector of coefficients at Stage $i - 1$, $\mathbf{Q}_{(i-1)i}$ is as defined above, and \mathbf{b}_i is the DMLPSI vector of coefficients at Stage i .

Let $\mathbf{S}_{i(i-1)} = \mathbf{Q}_{i(i-1)} \mathbf{B}_{i-1}$ and $\mathbf{S}'_{i(i-1)} = \mathbf{S}_{(i-1)i} = \mathbf{B}'_{i-1} \mathbf{Q}_{(i-1)i}$ be the transpose of matrix $\mathbf{S}_{i(i-1)}$, and assume that matrices \mathbf{Q}_{ii} , $\mathbf{Q}_{i(i-1)}$, and $\mathbf{\Lambda}_i$ and vector \mathbf{w} are known. To minimize the mean squared difference between $H = \mathbf{w}' \mathbf{g}$ and $I_i = \mathbf{b}'_i \boldsymbol{\omega}_i$ (i.e., $E[(H - I_i)^2]$), under the restriction $\mathbf{S}_{(i-1)i} \mathbf{b}_i = \mathbf{B}'_{i-1} \mathbf{Q}_{(i-1)i} \mathbf{b}_i = \mathbf{0}$, it is necessary to minimize the function

$$f(\mathbf{b}_i, \boldsymbol{\xi}) = \mathbf{b}'_i \mathbf{Q}_{ii} \mathbf{b}_i - 2\mathbf{w}' \mathbf{\Lambda}_i \mathbf{b}_i + 2\boldsymbol{\xi}' \mathbf{S}_{(i-1)i} \mathbf{b}_i$$

with respect to vectors \mathbf{b}_i and $\boldsymbol{\xi}' = [\xi_1 \ \xi_2 \ \dots \ \xi_{i-1}]$, where $\boldsymbol{\xi}$ is a vector of Lagrange multipliers. In matrix notation, the derivative results of \mathbf{b}_i and $\boldsymbol{\xi}$ are

$$\begin{bmatrix} \mathbf{b}_i \\ \boldsymbol{\xi} \end{bmatrix} = \begin{bmatrix} \mathbf{Q}_{ii} & \mathbf{S}_{i(i-1)} \\ \mathbf{S}_{(i-1)i} & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{\Lambda}_i \mathbf{w} \\ \mathbf{0} \end{bmatrix}$$

from where the DMLPSI vector of coefficients that minimizes $E[(H - I_i)^2]$ at Stage i can be written as

$$\mathbf{b}_i = \mathbf{K}_i \boldsymbol{\beta}_i \quad (16)$$

where $\mathbf{K}_i = [\mathbf{I}_i - \boldsymbol{\Phi}_i]$, $\boldsymbol{\Phi}_i = \mathbf{Q}_{ii}^{-1} \mathbf{S}_{i(i-1)} [\mathbf{S}_{(i-1)i} \mathbf{Q}_{ii}^{-1} \mathbf{S}_{i(i-1)}]^{-1} \mathbf{S}_{(i-1)i}$, $\boldsymbol{\beta}_i = \mathbf{Q}_{ii}^{-1} \boldsymbol{\Lambda}_i \mathbf{w}$ (Equation 14), \mathbf{Q}_{ii}^{-1} is the inverse of matrix \mathbf{Q}_{ii} , and \mathbf{I}_i is an identity matrix of the same size as \mathbf{Q}_{ii} . When the restriction $\text{cov}(I_i, \mathbf{J}_{i-1}) = \mathbf{0}$ is not imposed, $\boldsymbol{\beta}_i = \mathbf{b}_i$, and at Stage 1 ($i = 1$), $\mathbf{S}_{i(i-1)} = \mathbf{S}_{1(0)} = \mathbf{Q}_{1(0)} \mathbf{B}_0 = \mathbf{0}$, $\boldsymbol{\Phi}_1 = \mathbf{Q}_{11}^{-1} \mathbf{S}_{1(0)} (\mathbf{S}_{(0)1} \mathbf{Q}_{11}^{-1} \mathbf{S}_{1(0)})^{-1} \mathbf{S}_{1(0)} = \mathbf{0}$, and $\mathbf{K}_1 = \mathbf{K}_1 = [\mathbf{I} - \mathbf{0}] = \mathbf{I}$, where \mathbf{I} is an identity matrix of the same size as $\boldsymbol{\Phi}_1$. Thus, $\mathbf{b}_1 = \mathbf{K}_1 \boldsymbol{\beta}_1 = \mathbf{Q}_{11}^{-1} \boldsymbol{\Lambda}_1 \mathbf{w} = \boldsymbol{\beta}_1$, where \mathbf{Q}_{11}^{-1} is the inverse of matrix \mathbf{Q}_{11} , $\boldsymbol{\Lambda}_1 = \mathbf{C}_1$, and $\text{cov}(\boldsymbol{\omega}_1, \mathbf{g}) = \mathbf{C}_1$. The method to obtain Equation 16 is valid for all selection indices associated with DMLPSI.

The DMLPSI expected genetic gain per trait, selection response, and correlation at Stage i are similar to those of Equation 15 changing $\boldsymbol{\beta}_i$ for $\mathbf{b}_i = \mathbf{K}_i \boldsymbol{\beta}_i$. The only difference between the DMLPSI and OMLPSI is the way the selection intensities are obtained. Xu and Muir (1992) described a general method to obtain the DMLPSI selection intensity, whereas Cerón-Rojas et al. (2019a, 2019b) described a method to obtain the OMLPSI selection intensity in the two-stage context. Methods to obtain the DMLPSI and OMLPSI selection intensities are valid for both the unconstrained and constrained multi-stage indices.

4.18.3 | OMLPSI selection intensity

As we have seen in Equation 6, the selection intensity (k), in the single-stage context, is related to the height of the ordinate of the normal curve $z(\tau^*)$ and the proportion selected (p) as $k = [z(\tau^*)]/p$. Kempthorne and Nordskog (1959) have indicated that

$$p = \int_{\tau^*}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \tau^2 \right\} d\tau$$

from where it is evident that $k = [z(\tau^*)]/p$ is a left truncated normal distribution (Arismendi, 2013; Wilhelm & Manjunath, 2010). A truncated distribution is a conditional distribution resulting when the domain of the parent distribution is restricted to a smaller region (Hattaway, 2010). The equation $k = [z(\tau^*)]/p$ is also called the “hazard function,” the “hazard rate,” or the “inverse Mill’s ratio” of the normal distribution (Rausand & Høyland, 2004).

Young (1964) showed that the intensity of selection for the two-stage case could be obtained in a similar manner as $k = [z(\tau^*)]/p$ (Equation 6). Thus, suppose that indices I_1 and I_2 have a joint normal distribution and can be transformed into the standardized random normal variables $T_1 = (I_1 - m_1)/\sigma_1$

and $T_2 = (I_2 - m_2)/\sigma$ with mean zero and variance 1, where m_1 and m_2 are the means, and σ_1 and σ_2 are the standard deviations of I_1 and I_2 , respectively. In this case, the method of selection is to retain animals or plants with $T_1 \geq c_1$ at Stage 1 and $T_1 + T_2 \geq c_2$ at Stage 2, where c_1 and $c_2 - T_1$ are truncation points for I_1 and I_2 , respectively. Thus, according to Young (1964), the selected population has the bivariate left truncated normal distribution with probability density function given by

$$h_{T_1, T_2}(\tau_1, \tau_2) = \frac{f_{T_1, T_2}(\tau_1, \tau_2)}{p}$$

where

$$f_{T_1, T_2}(\tau_1, \tau_2) = \frac{1}{2\pi \sqrt{1-\rho_{12}^2}} \times \exp \left\{ -\frac{1}{2(1-\rho_{12}^2)} [\tau_1^2 + \tau_2^2 - 2\rho_{12}^2 \tau_1 \tau_2] \right\}$$

and ρ_{12} is the correlation between T_1 and T_2 . The total proportion selected ($p = q_1 q_2$) can now be written as

$$p = q_1 q_2 = \int_{c_1}^{\infty} \int_{c_2 - \tau_1}^{\infty} f_{T_1, T_2}(\tau_1, \tau_2) d\tau_2 d\tau_1$$

where c_1 and $c_2 - \tau_1$ are truncation points for I_1 and I_2 , respectively. It is evident that when the number of stages increases, the number of variables in $f_{T_1, T_2}(\tau_1, \tau_2)$, and the number of integrals in p also increases.

By the foregoing results, in the multi-stage selection context it is usual to fix a proportion (p) before selection is carried out and then, based on the fixed p value, to determine the proportion q_i ($i = 1, 2, \dots, S$) for each stage under the restriction

$$p = \prod_{i=1}^S q_i$$

where S is the number of stages. In a two-stage selection context, $p = q_1 q_2$. It is assumed that p is known but q_i is unknown. One way of solving the equation $p = q_1 q_2$ is by trial and error, depending on the importance of each trait (Hazel & Lush, 1942) or each index at each stage. Cerón-Rojas et al. (2019a, 2019b) have given a general method to estimate the OMLPSI selection intensities in the two-stage context, whereas Xu and Muir (1992) have provided a general method to obtain the selection intensities for the DMLPSI. The methods to obtain the selection intensities described above are valid for any multi-stage index. Until now, there has been no general procedure to obtain the OMLPSI selection intensity for any number of stages.

4.19 | LGSi

Togashi et al. (2011) originally proposed the LGSi, which is a linear combination of GEBVs; however, Cerón-Rojas et al. (2015) and Cerón-Rojas and Crossa (2019) developed the complete unconstrained and constrained LGSi theory. These authors designed the LGSi for selecting genotypes in inbred populations; that is, LGSi exploits the linkage disequilibrium between markers and quantitative trait loci (QTL) produced when inbred lines are crossed to select parents for the next selection cycle. This index uses all available marker information in the prediction because it obtains the GEBV by multiplying the genomic best linear unbiased predictor of all estimated marker effects in the training population by the marker code values of the testing population (VanRaden, 2008).

4.19.1 | The genomic breeding values

Let \mathbf{u}_j be an $m \times 1$ vector of QTL additive effects associated with markers that affect the j th trait, and let \mathbf{L} be a matrix $n \times m$ (n = number of individual and m = number of markers in the population) of coded marker values (e.g., $2 - 2q$, $1 - 2q$, and $-2q$ for the marker genotypes AA , Aa , and aa , respectively); then

$$\boldsymbol{\alpha}_j = \mathbf{L}\mathbf{u}_j \quad (17)$$

is the vector of individual genomic breeding values associated with the j th characteristic ($j = 1, 2, \dots, t$; t = number of traits) of the candidates for selection. We assumed that $\boldsymbol{\alpha}_j$ has multivariate normal distribution with mean $\mathbf{0}$ and covariance matrix $\mathbf{G}\sigma_\alpha^2$, where σ_α^2 is the additive genomic variance of $\boldsymbol{\alpha}_j$; $\mathbf{G} = \mathbf{L}\mathbf{L}'\kappa$ is the matrix of additive genomic relationship, $\kappa = \sum_{l=1}^m 2q_l(1 - q_l)$ in an F_2 population; q_l is the frequency of allele A ; and $1 - q_l$ is the frequency of allele a for the l th marker ($l = 1, 2, \dots, m$).

4.19.2 | The net genetic merit and its relationship with the genomic breeding values

The net genetic merit is related to the vector of genomic breeding values (Equation 17) as

$$H = \boldsymbol{\beta}'\boldsymbol{\alpha} + \mathbf{e} \quad (18)$$

where $\boldsymbol{\beta}' = [\beta_1 \ \beta_2 \ \dots \ \beta_t]$ is the vector of regression coefficients, $\boldsymbol{\alpha}' = [\boldsymbol{\alpha}_1 \ \boldsymbol{\alpha}_2 \ \dots \ \boldsymbol{\alpha}_t]$, and \mathbf{e} is the residual with normal distribution, null expectation, and variance σ_e^2 . It is possible to show that the covariance between $\boldsymbol{\beta}'\boldsymbol{\alpha}$ and \mathbf{e} equals zero.

4.19.3 | The LGSi

The BLP of H is the conditional expectation of H given $\boldsymbol{\alpha}$, that is

$$E(H|\boldsymbol{\alpha}) = \mathbf{w}'\boldsymbol{\Gamma}\boldsymbol{\Gamma}^{-1}\boldsymbol{\alpha} = \mathbf{w}'\boldsymbol{\alpha} \quad (19)$$

where $\mathbf{w}'\boldsymbol{\Gamma} = \text{cov}(H, \boldsymbol{\alpha})'$ is the covariance between H and $\boldsymbol{\alpha}$, $\boldsymbol{\Gamma}^{-1}$ is the inverse matrix of the genomic covariance matrix $\boldsymbol{\Gamma} = \text{var}(\boldsymbol{\alpha}) = \text{cov}(\boldsymbol{\alpha}, \mathbf{g})$, \mathbf{w} is as defined in Equation 2, and $I_G = \mathbf{w}'\boldsymbol{\alpha}$ is the LGSi.

4.19.4 | The estimator of the LGSi

Let $\hat{\mathbf{u}}_j$ be a predictor of the \mathbf{u}_j vector of QTL effects associated with the markers that affect the j th trait (Cerón-Rojas & Crossa, 2019; Cerón-Rojas et al., 2015); then the expectation of \mathbf{u}_j given $\hat{\mathbf{u}}_j$ is $E(\mathbf{u}_j|\hat{\mathbf{u}}_j) = \hat{\mathbf{u}}_j$. This implies that the expectation of H , given the estimator of the LGSi,

$$\hat{I}_G = w_1\hat{\boldsymbol{\alpha}}_1 + w_2\hat{\boldsymbol{\alpha}}_2 + \dots + w_t\hat{\boldsymbol{\alpha}}_t$$

is equal to

$$E(H|\hat{I}_G) = \hat{I}_G$$

where $\hat{\boldsymbol{\alpha}}_j = \mathbf{L}\hat{\mathbf{u}}_j$ is the vector of GEBVs associated with trait j th.

4.19.5 | Selection response and expected genetic gain per trait

By Equations 8–10, the LGSi selection response and expected genetic gain per trait are, respectively,

$$R_G = k\sqrt{\mathbf{w}'\boldsymbol{\Gamma}\mathbf{w}} \quad \text{and} \quad E_G = k\frac{\boldsymbol{\Gamma}\mathbf{w}}{\sqrt{\mathbf{w}'\boldsymbol{\Gamma}\mathbf{w}}} \quad (20)$$

We have previously defined all parameters of Equation 20. To simplify notation, in Equation 20 we omitted the intervals between selection cycles, which denote the time required for LGSi to complete one selection cycle (Cerón-Rojas et al., 2015). Using the Cauchy–Schwarz inequality, Cerón-Roja and Crossa (2020b) showed that in the asymptotic context, when the number of markers and genotypes tend to be infinite, $R_G = k\sqrt{\mathbf{w}'\boldsymbol{\Gamma}\mathbf{w}}$ is the upper boundary for the unconstrained LPSi selection response. It is possible to show that, in the context of the constrained LPSi selection response, there is a similar result (Cerón-Roja & Crossa, 2019).

4.20 | Relationships between the LPSI and LGSI selection response

In the asymptotic context, when the number of markers and genotypes tend to infinite, matrix Γ approaches matrix \mathbf{C} , and, at the limit, we would expect that $\Gamma = \mathbf{C}$ (Cerón-Rojas & Sahagún-Castellanos, 2016). This means that matrix $\mathbf{P} = \mathbf{C} + \mathbf{R}$ approaches matrix $\mathbf{P} = \Gamma + \mathbf{R}$, from where $\mathbf{P}^{-1} = (\Gamma + \mathbf{R})^{-1} = \Gamma^{-1} - \Gamma^{-1}(\Gamma^{-1} + \mathbf{R}^{-1})^{-1}\Gamma^{-1}$ is the inverse of matrix \mathbf{P} (Searle et al., 2006), whereas Γ^{-1} and \mathbf{R}^{-1} are the inverses of matrices Γ and \mathbf{R} , respectively. In the asymptotic context, let $\mathbf{b} = \mathbf{P}^{-1}\Gamma\mathbf{w}$ be the LPSI vector of coefficients; then, the LPSI response can be written as

$$R = k\sqrt{\mathbf{b}'\mathbf{P}\mathbf{b}} = k\sqrt{\mathbf{w}'\Gamma\mathbf{P}^{-1}\Gamma\mathbf{w}} = k\sqrt{\mathbf{w}'\Gamma\mathbf{w} - \mathbf{w}'(\Gamma^{-1} + \mathbf{R}^{-1})^{-1}\mathbf{w}}$$

Assuming that $\mathbf{b}'\mathbf{P}\mathbf{b}$ and $\mathbf{w}'\Gamma\mathbf{w}$ are positive semi-definite (i.e., $\mathbf{b}'\mathbf{P}\mathbf{b} \geq 0$ and $\mathbf{w}'\Gamma\mathbf{w} \geq 0$), then $\mathbf{w}'\Gamma\mathbf{w} \geq \mathbf{w}'(\Gamma^{-1} + \mathbf{R}^{-1})^{-1}\mathbf{w} \geq 0$, from where in the asymptotic context, $R_G \geq R$.

In the non-asymptotic context, when the number of markers and genotypes is low, matrices \mathbf{C} and \mathbf{P} can be written as $\mathbf{C} = \Gamma + \Xi$ and $\mathbf{P} = (\Gamma + \Xi) + \mathbf{R}$, respectively, where $\Xi = \mathbf{C} - \Gamma$. Thus, the inverse of matrix \mathbf{P} is

$$\mathbf{P}^{-1} = [(\Gamma + \Xi) + \mathbf{R}]^{-1} = (\Gamma + \Xi)^{-1} - (\Gamma + \Xi)^{-1}[(\Gamma + \Xi)^{-1} + \mathbf{R}^{-1}]^{-1}(\Gamma + \Xi)^{-1}$$

from where $\mathbf{b} = \mathbf{P}^{-1}(\Gamma + \Xi)\mathbf{w}$, and the LPSI selection response (R) is

$$\begin{aligned} R &= k\sqrt{\mathbf{b}'\mathbf{P}\mathbf{b}} = k\sqrt{\mathbf{w}'(\Gamma + \Xi)\mathbf{P}^{-1}(\Gamma + \Xi)\mathbf{w}} \\ &= k\sqrt{\mathbf{w}'\Gamma\mathbf{w} + \mathbf{w}'\Xi\mathbf{w} - \mathbf{w}'[(\Gamma + \Xi)^{-1} + \mathbf{R}^{-1}]^{-1}\mathbf{w}} \end{aligned}$$

This last equation indicates that in the non-asymptotic context, R_G and R are related in three possible ways:

1. $R > R_G$ if $\mathbf{w}'\Xi\mathbf{w} > \mathbf{w}'[(\Gamma + \Xi)^{-1} + \mathbf{R}^{-1}]^{-1}\mathbf{w}$
2. $R = R_G$ if $\mathbf{w}'\Xi\mathbf{w} = \mathbf{w}'[(\Gamma + \Xi)^{-1} + \mathbf{R}^{-1}]^{-1}\mathbf{w}$
3. $R_G > R$ if $\mathbf{w}'\Xi\mathbf{w} < \mathbf{w}'[(\Gamma + \Xi)^{-1} + \mathbf{R}^{-1}]^{-1}\mathbf{w}$

By Points 2 and 3, in the non-asymptotic context, R_G may be equal to or larger than R depending on the size of $\mathbf{w}'\Xi\mathbf{w}$ and $\mathbf{w}'[(\Gamma + \Xi)^{-1} + \mathbf{R}^{-1}]^{-1}\mathbf{w}$.

5 | COMBINED LINEAR SELECTION INDICES

A combined LSI to predict H (Equation 2) is a linear combination of phenotypic values and marker scores (LMSI) or

phenotypic values and GEBVs (called combined LGSI). Both indices combine information on markers linked to QTL and phenotypic values of the traits in the prediction of H because it is impossible to identify all QTL affecting the economically important traits (Dekkers, 2007; Dekkers & Settar, 2004; Lande & Thompson, 1990; Li, 1998). The LMSI assumes that favorable alleles and their average effects on phenotype are known; however, this assumption is only valid for traits conditioned by major genes and not for quantitative traits that are influenced by the environment. This is important because many QTL with small effects could interact with the environment and among themselves (Heffner et al., 2009; Hospital et al., 1997). Linear molecular selection index efficiency depends on various factors, such as number and density of markers associated with QTL, population size, trait heritability, additive genetic variances explained by markers, and the precision of the estimated effect of gene substitution (Dekkers & Dentine, 1991). Several authors (Gimelfarb & Lande, 1994, 1995; Zhang & Smith, 1992, 1993) have pointed out the effectiveness of the LMSI in inbred populations with large population sizes and traits with low heritability values when only one trait and its associated molecular score are considered. Moreau et al. (2000, 2007) found that the LMSI was more effective than LPSI only in early-generation testing and that LMSI had the additional disadvantage of increased costs due to molecular marker evaluation.

In the genomic selection context, Dekkers (2007) developed the “combined LGSI” as a possible solution to LMSI problems. This index uses GEBVs (no marker scores) and phenotypic values jointly to predict H and is very similar to the LMSI. We describe the LMSI, the combined LGSI, and the combined optimum and decorrelated multi-stage LGSI (OMLGSI and DMLGSI, respectively) developed by Cerón-Rojas and Crossa (2020a).

5.1 | The LMSI molecular score

Suppose that the QTL effects combine additively both within and between loci; then the h^{th} unobservable genetic value G_h can be written as

$$G_h = \sum_{k=1}^N \vartheta_k v_k$$

where ϑ_k is the effect of the QTL k^{th} , v_k is the number of favorable alleles at the QTL k^{th} (2, 1, or 0), and N is the number of QTL affecting the j^{th} trait of interest. Because the QTL effect values are not observable, the G_h values are also not observable; however, we can use a linear combination of the markers linked to the QTL (s_h) that affect the j^{th} trait to predict G_h as

$$s_h = \sum_{k=1}^M b_k l_k$$

where s_h is a predictor of G_h , b_k is the regression coefficient of the linear regression model, l_k is the coded value of the k^{th} markers (e.g., 1, 0, and -1 for marker genotypes AA , Aa , and aa , respectively), and M is the number of selected markers linked to the QTL that affect the j^{th} trait. The above equation is called the “marker score” (Lande & Thompson, 1990). The number of selected markers is only a subset of potential markers linked to QTL in the population under selection; thus, the the variance of the s_h values should be lower than or equal to the variance of the G_h values.

5.2 | The combined LSI

The combined LSI is

$$I_M = \beta'_y y + \beta'_C \boldsymbol{\varphi} = [\beta'_y \ \beta'_C] \begin{bmatrix} y \\ \boldsymbol{\varphi} \end{bmatrix} = \boldsymbol{\beta}' \mathbf{t} \quad (22)$$

where $\boldsymbol{\varphi}$ can be the vector of marker scores, $\mathbf{s}' = [s_1 \ s_2 \ \dots \ s_t]$, or the vector of genomic breeding values $\boldsymbol{\alpha}$ (Equation 18). When $\boldsymbol{\varphi} = \mathbf{s}$, Equation 22 is the LMSI, whereas when $\boldsymbol{\varphi} = \boldsymbol{\alpha}$, Equation 22 is the combined LGSI. Vector \mathbf{y} was defined earlier, whereas $\boldsymbol{\beta}' = [\beta'_y \ \beta'_C]$, and $\mathbf{t}' = [y' \ \boldsymbol{\varphi}']$.

5.3 | Vector of coefficients, maximized selection response, and correlation

Let $\text{var}(\mathbf{s}) = \mathbf{S}$ be the covariance matrix of marker score and $\boldsymbol{\Gamma}$ as defined in Equation 19; then, according to the LPSI theory, the vector of coefficients that maximizes the combined LSI selection response, expected genetic gain per trait, and correlation is

$$\boldsymbol{\beta} = \mathbf{O}^{-1} \boldsymbol{\Pi} \mathbf{v} \quad (23)$$

where when $\mathbf{O} = \text{var} \begin{bmatrix} y \\ \mathbf{s} \end{bmatrix} = \begin{bmatrix} \mathbf{P} & \mathbf{S} \\ \mathbf{S} & \mathbf{S} \end{bmatrix}$ and $\boldsymbol{\Pi} = \text{var} \begin{bmatrix} \mathbf{g} \\ \mathbf{s} \end{bmatrix} = \begin{bmatrix} \mathbf{C} & \mathbf{S} \\ \mathbf{S} & \mathbf{S} \end{bmatrix}$, Equation 23 is the LMSI vector of coefficients, whereas when $\mathbf{O} = \text{var} \begin{bmatrix} y \\ \mathbf{z} \end{bmatrix}$ and $\boldsymbol{\Pi} = \text{var} \begin{bmatrix} \mathbf{g} \\ \mathbf{z} \end{bmatrix}$, Equation 23 is the combined LGSI vector of coefficients. Vector $\mathbf{v}' = [\mathbf{w}' \ \mathbf{0}']$ is the vector of economic weights, where $\mathbf{0}$ is a null vector of size $t \times 1$. The other parameters were previously defined.

5.4 | The maximized combined LSI selection response

By Equation 23, the maximized LMSI and combined LGSI selection response of both indices is

$$R_C = k \sqrt{\boldsymbol{\beta}' \mathbf{O} \boldsymbol{\beta}} \quad (24)$$

where k is the selection intensity. When \mathbf{S} and $\boldsymbol{\Gamma}$ are null matrices, $\boldsymbol{\beta}_y = \mathbf{b} = \mathbf{P}^{-1} \mathbf{C} \mathbf{w}$ (the LPSI vector of coefficients), whereas $R_C = R$ (the LPSI selection response). In addition, when the number of markers and genotypes increases, matrix $\boldsymbol{\Gamma}$ tends to matrix \mathbf{C} and at the limit $\boldsymbol{\Gamma} = \mathbf{C}$, from where $\boldsymbol{\beta}_y = \mathbf{0}$, $\boldsymbol{\beta}_C = \mathbf{w}$, and Equation 24 is equal to $R_G = k \sqrt{\mathbf{w}' \boldsymbol{\Gamma} \mathbf{w}}$ (Equation 20). That is, the weights and selection response of the LGSI and the combined LGSI will be equal.

5.5 | The combined LGSI expected genetic gain per trait and correlation

The maximized combined LGSI expected genetic gain per trait (\mathbf{E}_C) and correlation (ρ_{HI_C}) are, respectively,

$$\mathbf{E}_C = k \frac{\boldsymbol{\Pi} \boldsymbol{\beta}}{\sqrt{\boldsymbol{\beta}' \mathbf{O} \boldsymbol{\beta}}} \text{ and } \rho_{HI_C} = \frac{\sqrt{\boldsymbol{\beta}' \mathbf{O} \boldsymbol{\beta}}}{\sqrt{\mathbf{w}' \mathbf{C} \mathbf{w}}} \quad (25)$$

where k is the selection intensity. Note that the denominator of ρ_{HI_C} (Equation 25) and the denominator of the maximized LPSI correlation (Equation 11) are very similar; however, the numerator of both equations is different. We would expect that $\sqrt{\boldsymbol{\beta}' \mathbf{O} \boldsymbol{\beta}} \geq \sqrt{\mathbf{b}' \mathbf{P} \mathbf{b}}$, and then $\rho_{HI_C} \geq \rho_{HI}$.

5.6 | Asymptotic relationship between combined LGSI and LGSI parameters

The asymptotic relationship between \mathbf{E}_C (Equation 25) and the LGSI expected genetic gain per trait \mathbf{E}_G (Equation 20) is as follows. When $\boldsymbol{\Gamma} = \mathbf{C}$, $\boldsymbol{\Pi} = \begin{bmatrix} \boldsymbol{\Gamma} & \boldsymbol{\Gamma} \\ \boldsymbol{\Gamma} & \boldsymbol{\Gamma} \end{bmatrix}$ and $\boldsymbol{\beta}' = [\mathbf{0}' \ \mathbf{w}']$, from where

$$\mathbf{E}_C = k \frac{2 \boldsymbol{\Gamma} \mathbf{w}}{\sqrt{\mathbf{w}' \boldsymbol{\Gamma} \mathbf{w}}} = 2 \mathbf{E}_G \quad (26)$$

This means that in the asymptotic context, the CLGSI expected genetic gain per trait is twice the LGSI expected genetic gain per trait. Of course, 2 is only a proportionality constant; thus, in reality, $\mathbf{E}_C = \mathbf{E}_G$. Similarly, when

$\Gamma = \mathbf{C}$ and $\boldsymbol{\beta}' = [\mathbf{0}' \ \mathbf{w}']$, $\boldsymbol{\beta}'\mathbf{O}\boldsymbol{\beta} = \mathbf{w}'\Gamma\mathbf{w}$ and $\mathbf{w}'\mathbf{C}\mathbf{w} = \mathbf{w}'\Gamma\mathbf{w}$ in ρ_{HI_C} (Equation 25); thus, the maximum correlation between H and I_C in the asymptotic context will be equal to 1.0, as we would expect.

5.7 | Combined multi-stage LGSIs

Based on the LMSI, Xie and Xu (1998) developed a decorrelated multi-stage LMSI. They assumed that their index would increase the efficiency of molecular assisted selection in a two-stage context by first (Stage 1) selecting immature seedlings (embryos) based on molecular score information only, and later, at Stage 2, by selecting individual traits based on molecular score and phenotypic value information jointly. The Xie and Xu (1998) index has the same LMSI problems as previously indicated. For this reason, Cerón-Rojas and Crossa (2020a) adapted the Dekkers (2007) index to the multi-stage context using the OMLPSI and DMLPSI approaches. Both these indices use GEBVs and phenotypic information jointly to predict the net genetic merit; then they are free of the problems of the Xie and Xu (1998) index. In the multi-stage context, the Cerón-Rojas and Crossa (2020a) indices are as follows: at Stage 1, we can select immature (seedlings, embryos) based on GEBVs information only and, in Stage 2, we select individual traits based on GEBVs and phenotypic values jointly. This is the Xie and Xu (1998) idea but in the genomic selection context.

The approach based on the OMLPSI optimum index will be called “optimum combined multi-stage LGSI” (OCMLGSI), whereas the index based on the DMLPSI will be called “decorrelated combined multi-stage LGSI” (DCMLGSI) because, at stage two, both indices use GEBVs and phenotypic information jointly to predict the net genetic merit.

5.8 | The OCMLGSI and DCMLGSI parameters at Stage 1

The LGSI and LPSI are part of the OCMLGSI and DCMLGSI. At stage-one the estimator of the LGSI ($\hat{I}_G = w_1\hat{\boldsymbol{\alpha}}_1 + w_2\hat{\boldsymbol{\alpha}}_2 + \dots + w_t\hat{\boldsymbol{\alpha}}_t$, where $\hat{\boldsymbol{\alpha}}_j = \mathbf{L}\hat{\mathbf{u}}_j$ is the vector of GEBV of trait j^{th}) is used to predict H because at this stage, both indices select immature seedlings (or embryos) based only on GEBV information. The OCMLGSI and DCMLGSI maximized selection responses are, respectively,

$$R_{O_1} = k_{O_1} \sqrt{\mathbf{w}'\Gamma\mathbf{w}} \text{ and } R_{D_1} = k_{D_1} \sqrt{\mathbf{w}'\Gamma\mathbf{w}} \quad (27)$$

where k_{O_1} and k_{D_1} (the selection intensities for each index) are the only difference between R_{O_1} and R_{D_1} . The maximized

correlation between $H = \mathbf{w}'\mathbf{g}$ and $I_G = \mathbf{w}'\boldsymbol{\alpha}$ is $\rho_{HI} = \frac{\sqrt{\mathbf{w}'\Gamma\mathbf{w}}}{\sqrt{\mathbf{w}'\mathbf{C}\mathbf{w}}}$, where $\sqrt{\mathbf{w}'\Gamma\mathbf{w}}$ is the standard deviation of the variance of $I_G = \mathbf{w}'\boldsymbol{\alpha}$, $\sigma_H = \sqrt{\mathbf{w}'\mathbf{C}\mathbf{w}}$ is the standard deviation of $H = \mathbf{w}'\mathbf{g}$, and \mathbf{C} is the covariance matrix of \mathbf{g} .

5.9 | Covariance matrices for OCMLGSI and DCMLGSI

The main difference between the multi-stage LPSIs described earlier (OMLPSI and DMLPSI) and OCMLGSI and DCMLGSI is that, in addition to $\mathbf{y}' = [Y_1 \ Y_2 \ \dots \ Y_t]$, the latter two indices incorporate $\boldsymbol{\alpha}$ (Equation 18) in the prediction of H . Let $\boldsymbol{\eta}' = [\boldsymbol{\alpha}' \ \mathbf{y}']$; then vector $\boldsymbol{\eta}$ can be partitioned into S subvectors as $\boldsymbol{\eta}' = [\boldsymbol{\alpha}' \ \mathbf{y}'] = [\boldsymbol{\omega}'_1 \ \boldsymbol{\omega}'_2 \ \dots \ \boldsymbol{\omega}'_S]$, where vector \mathbf{y} has been partitioned into $S - 1$ subvector as $\mathbf{y}' = [\boldsymbol{\omega}'_2 \ \boldsymbol{\omega}'_3 \ \dots \ \boldsymbol{\omega}'_S]$. This means that at Stage 1, $\boldsymbol{\omega}_1 = \boldsymbol{\alpha}$ (Equation 18), and $I_i = \boldsymbol{\beta}'_i \boldsymbol{\omega}_i$ is the index at Stage i ($i = 2, 3, \dots, S$), where $\boldsymbol{\beta}'_i = [\beta_{i1} \ \beta_{i2} \ \dots \ \beta_{it_i}]$.

Let $\mathbf{g}' = [G_1 \ G_2 \ \dots \ G_t]$ be a vector of true unobservable random breeding values for t traits, and let $\boldsymbol{\alpha}$, \mathbf{y} , and $\boldsymbol{\omega}_1$ be as defined earlier. The covariance between $\boldsymbol{\alpha}_j$ and G_j ($j = 1, 2, \dots, t$) is equal to the variance of $\boldsymbol{\alpha}_j$ [i.e., $\text{cov}(\boldsymbol{\alpha}_j, G_j) = \sigma_{\alpha_j}^2$] (Dekkers, 2007); the covariance matrix of $\boldsymbol{\alpha}$ is $\text{var}(\boldsymbol{\alpha}) = \Gamma$, and the covariance matrices between $\boldsymbol{\omega}_i$ ($i = 1, 2, \dots, S$) and \mathbf{g} for S stages is

$$\begin{bmatrix} \text{cov}(\boldsymbol{\omega}_1, \mathbf{g}) \\ \text{cov}(\boldsymbol{\omega}_2, \mathbf{g}) \\ \vdots \\ \text{cov}(\boldsymbol{\omega}_S, \mathbf{g}) \end{bmatrix} = \begin{bmatrix} \Gamma \\ \mathbf{C}_2 \\ \vdots \\ \mathbf{C}_S \end{bmatrix} = \boldsymbol{\Omega}_\Gamma$$

where at Stage 1, $\boldsymbol{\omega}_1 = \boldsymbol{\alpha}$, $\text{cov}(\boldsymbol{\omega}_1, \mathbf{g}) = \Gamma$ is the genomic covariance matrix, and from Stage 2 to S , $\text{cov}(\boldsymbol{\omega}_i, \mathbf{g}) = \mathbf{C}_i$ is the i^{th} genotypic covariance submatrix of $\boldsymbol{\Omega}_\Gamma$. Note that while the size of matrix Γ is $t \times t$, the size of matrix \mathbf{C}_i will be $n_i \times t$, where n_i is the size of vector $\boldsymbol{\omega}_i$. For example, if $\mathbf{g}' = [g_1 \ g_2 \ g_3 \ g_4]$, $\boldsymbol{\omega}'_1 = [\boldsymbol{\alpha}_1 \ \boldsymbol{\alpha}_2 \ \boldsymbol{\alpha}_3 \ \boldsymbol{\alpha}_4]$, and $\boldsymbol{\omega}'_1 = [Y_1 \ Y_2]$, then Γ will be of size 4×4 , and \mathbf{C}_2 will be of size 2×4 , where 2 is the number of rows and 4 is the number of columns in matrix \mathbf{C}_2 . By the foregoing results, the covariance matrix of $\boldsymbol{\eta}' = [\boldsymbol{\alpha}' \ \mathbf{y}'] = [\boldsymbol{\omega}'_1 \ \boldsymbol{\omega}'_2 \ \dots \ \boldsymbol{\omega}'_S]$ is

$$\boldsymbol{\Sigma} = \begin{bmatrix} \Gamma & \Gamma_{12} \cdots \Gamma_{1S} \\ \Gamma_{21} & \mathbf{P}_{22} \cdots \mathbf{P}_{2S} \\ \vdots & \vdots \ \ddots \ \vdots \\ \Gamma_{S1} & \mathbf{P}_{S2} \cdots \mathbf{P}_{SS} \end{bmatrix}$$

where all the gamma matrices are genomic matrices, while the others are phenotypic covariance matrices. Matrices \mathbf{Q}_{ii}

and Λ_i are constructed as

$$\mathbf{Q}_{ii} = \begin{bmatrix} \Gamma & \Gamma_{12} & \dots & \Gamma_{1i} \\ \Gamma_{21} & \mathbf{P}_{22} & \dots & \mathbf{P}_{2i} \\ \vdots & \vdots & \ddots & \vdots \\ \Gamma_{i1} & \mathbf{P}_{i2} & \dots & \mathbf{P}_{ii} \end{bmatrix} \text{ and } \Lambda_i = \begin{bmatrix} \Gamma \\ \mathbf{C}_2 \\ \vdots \\ \mathbf{C}_i \end{bmatrix}$$

which are submatrices of Σ and Ω_{Γ} , respectively, until Stage i .

5.10 | The OCMLGSI and DCMLGSI at Stage 2

At Stage 2, OCMLGSI and DCMLGSI is

$$I_C = \beta'_G \alpha + \beta'_y y = [\beta'_G \ \beta'_y] \begin{bmatrix} \alpha \\ y \end{bmatrix} = \beta'_2 t^* \quad (28)$$

where $y' = [\omega'_2 \ \omega'_3 \ \dots \ \omega'_S]$ and α is as defined earlier; β_G and β_y are vectors of coefficients of genomic and phenotypic weight values, respectively, $\beta'_2 = [\beta'_G \ \beta'_y]$, and $t^* = [\alpha' \ y']$. The only difference between OCMLGSI and DCMLGSI is the way vector $\beta'_2 = [\beta'_G \ \beta'_y]$ is obtained.

5.11 | The OCMLGSI parameters at Stage 2

The OCMLGSI vector of coefficients (β_2) that maximizes the response, correlation, and expected genetic gain per trait (Equation 25) at Stage 2 is

$$\beta_2 = \mathbf{Q}_{22}^{-1} \Lambda_2 \nu^* \quad (29)$$

where \mathbf{Q}_{22}^{-1} is the inverse of matrix $\mathbf{Q}_{22} = \begin{bmatrix} \Gamma & \Gamma \\ \Gamma & \mathbf{P}_{22} \end{bmatrix}$, where $\omega_2^* = [\omega'_1 \ \omega'_2]$ and $\mathbf{P}_{22} = \text{var}(\omega_2)$; $\Lambda_2 = \begin{bmatrix} \Gamma \\ \mathbf{C}_2 \end{bmatrix}$, $\mathbf{C}_2 = \text{cov}(\mathbf{g}, \omega_2)$, $\nu^* = [\mathbf{0}' \ \mathbf{w}']$ is a vector where $\mathbf{0}$ is a null vector of size $t \times 1$, and \mathbf{w} is the vector of economic weights. The maximized selection response, and the maximized correlation between H and OCMLGSI, respectively, are

$$R_{O_2} = k_{O_2} \sqrt{\beta_2' \mathbf{Q}_{22} \beta_2} \text{ and } \rho_{HI_C} = \frac{\sqrt{\beta_2' \mathbf{Q}_{22} \beta_2}}{\sqrt{\mathbf{w}' \mathbf{C} \mathbf{w}}} \quad (30)$$

where k_{O_2} is the OCMLGSI intensity at stage two. The total selection response for Stages 1 and 2 is $R_{O_i} = R_{O_1} + R_{O_2}$.

5.12 | The DCMLGSI parameters at Stage 2

The process to obtain the DCMLGSI vector of coefficients is similar to that described in subsection “The DMLPSI vector of coefficients at Stage i ,” where we described how to obtain the DMLPSI vector of coefficients (Equation 16). Subsequently, in this subsection we will only provide the main results of the DCMLGSI for Stage 2. Additional details associated with DCMLGSI are in Cerón-Rojas and Crossa (2020a).

The DCMLGSI vector of coefficients at Stage 2 is

$$\mathbf{b}_2 = \mathbf{K}_2 \beta_2 \quad (31)$$

where $\mathbf{K}_2 = [\mathbf{I}_2 - \Phi_2]$, $\Phi_2 = \mathbf{Q}_{22}^{-1} \mathbf{S}_{21} [\mathbf{S}_{12} \mathbf{Q}_{22}^{-1} \mathbf{S}_{21}]^{-1} \mathbf{S}_{12}$, and $\beta_2 = \mathbf{Q}_{22}^{-1} \Lambda_2 \alpha$ (Equation 29); and \mathbf{I}_2 is an identity matrix of the same size as matrix \mathbf{Q}_{22} , $\mathbf{S}_{21} = \mathbf{Q}_{21} \mathbf{B}_1$; and $\mathbf{S}'_{21} = \mathbf{S}_{12} = \mathbf{B}'_1 \mathbf{Q}_{12}$ is the matrix of constraints similar to the matrix defined in the DMLPSI context. Matrix \mathbf{K}_2 transforms the OCMLGSI vector of coefficients (β_2) into Equation 31 and is the only difference with respect to Equation 29. The maximized DCMLGSI selection response and correlation for two stages can be written as in Equation 30 by changing $\beta_2 = \mathbf{Q}_{22}^{-1} \Lambda_2 \nu^*$ by Equation 31 and the OCMLGSI selection intensity (k_{O_2}) by the DCMLGSI selection intensity (k_{D_2}). Finally, note that at Stage 2 $\mathbf{B}'_1 = [\mathbf{w}' \ \mathbf{0}'_{S-1}]$, where $\mathbf{w} = \mathbf{b}_1$ is the vector of economic weights and $\mathbf{0}'_{S-1}$ is a null vector of size $S - 1$. The size of matrix \mathbf{Q}_{12} depend on the number of traits selected at Stage 2.

6 | CONSTRAINED LINEAR SELECTION INDICES

By imposing null restrictions on Equations 10, Kempthorne and Nordskog (1959) developed the null restricted LPSI (denoted RLPSI) in the single-stage context, which imposes restrictions equal to zero on the trait expected genetic gain. Later, Mallard (1972) generalized the RLPSI and developed an optimum constrained LPSI (CLPSI), which imposes restrictions different to zero and includes the RLPSI as a particular case. Both indices mainly affect the expected genetic gain per trait (Equation 10). Whereas in the single-stage context the CLPSI vector of coefficients is a linear combination of the LPSI vector of coefficients ($\beta = \mathbf{P}^{-1} \mathbf{C} \mathbf{w}$), in the multi-stage context, the constrained optimum multi-stage LPSI (COMLPSI) vector of coefficients is a linear combination of the OMLPSI vector of coefficients (Equation 14). In turn, the CDMLPSI vector of coefficients is also a linear combination of the OMLPSI vector of coefficients.

6.1 | Single-stage CLPSI

6.1.1 | CLPSI main objective and Mallard matrix

Assume that μ_j is the population mean of the j^{th} trait before selection. The main objective of the CLPSI is to change μ_j to $\mu_j + d_j$, where d_j is a predetermined change in μ_j imposed by the breeder on Equation 10. Let

$$\mathbf{N}' = \begin{bmatrix} d_r & 0 & \dots & 0 & -d_1 \\ 0 & d_r & \dots & 0 & -d_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & d_r & -d_{r-1} \end{bmatrix}$$

a Mallard (1972) matrix $(r - 1) \times r$ of predetermined proportional gains, where d_j ($j = 1, 2, \dots, r$) is the j^{th} element of vector $\mathbf{d}' = [d_1 \ d_2 \ \dots \ d_r]$, and let \mathbf{U}' be a matrix $(t - r) \times t$ ($t =$ number of traits; $r =$ number of restricted traits) of 1s and 0s, where 1 indicates that the traits are restricted, and 0 indicates that the traits are unrestricted. The size of matrix \mathbf{U}' depends on the number of restricted traits.

6.1.2 | Matrix \mathbf{U}'

We can construct matrix \mathbf{U}' as follows. Suppose that we restrict one of the t traits; then, we can restrict the first of them as $\mathbf{U}' = [1 \ 0 \ 0 \ \dots \ 0]$, the second one as $\mathbf{U}' = [0 \ 1 \ 0 \ \dots \ 0]$, etc. In addition, we can restrict the first and second traits jointly as $\mathbf{U}' = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \end{bmatrix}$, the first and third traits as $\mathbf{U}' = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \end{bmatrix}$, the second and third traits as $\mathbf{U}' = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \end{bmatrix}$, etc. The procedure used to construct matrix \mathbf{U}' is valid for any number of restricted traits (Cerón-Rojas & Crossa, 2018).

6.1.3 | The CLPSI vector of coefficients

Let $\mathbf{M}' = \mathbf{N}'\Psi'$ be a Mallard (1972) matrix of restrictions, where $\Psi' = \mathbf{U}'\mathbf{C}$ is a matrix of null restrictions, and \mathbf{N}' , \mathbf{U}' , and \mathbf{C} are as defined earlier. To obtain the CLPSI vector of coefficients we need to minimize the mean squared difference $E[(I - H)^2]$ under the restriction $\mathbf{M}'\boldsymbol{\beta} = \mathbf{0}$ assuming that \mathbf{P} , \mathbf{C} , \mathbf{U}' , \mathbf{N}' , and \mathbf{w} are known; that is, we need to minimize the function

$$f(\boldsymbol{\beta}, \boldsymbol{\xi}) = \boldsymbol{\beta}'\mathbf{P}\boldsymbol{\beta} + \mathbf{w}'\mathbf{C}\mathbf{w} - 2\mathbf{w}'\mathbf{C}\boldsymbol{\beta} + 2\xi'\mathbf{M}'\boldsymbol{\beta}$$

with respect to vectors $\boldsymbol{\beta}$ and $\boldsymbol{\xi}' = [\xi_1 \ \xi_2 \ \dots \ \xi_{r-1}]$, where $\boldsymbol{\xi}$ is a vector of Lagrange multipliers. The derivative results from $\boldsymbol{\beta}$ and $\boldsymbol{\xi}$ are

$$\begin{bmatrix} \mathbf{P} & \mathbf{M}' \\ \mathbf{M}' & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta} \\ \boldsymbol{\xi} \end{bmatrix} = \begin{bmatrix} \mathbf{C}\mathbf{w} \\ \mathbf{0} \end{bmatrix}$$

from where the CLPSI vector of coefficients that minimizes $E[(I - H)^2]$ under the restriction $\mathbf{M}'\boldsymbol{\beta} = \mathbf{0}$ is

$$\boldsymbol{\beta}_C = \mathbf{K}\boldsymbol{\beta} \quad (32)$$

where $\mathbf{K} = [\mathbf{I}_t - \mathbf{D}]$, $\mathbf{D} = \mathbf{P}^{-1}\mathbf{M}(\mathbf{M}'\mathbf{P}^{-1}\mathbf{M})\mathbf{M}'$, and \mathbf{I}_t is an identity matrix of size $t \times t$. It is possible to show that when $\mathbf{N} = \mathbf{U}$, the restriction $\mathbf{M}'\boldsymbol{\beta} = \mathbf{0}$ changes to $\Psi'\boldsymbol{\beta} = \mathbf{0}$, from where Equation 32 changes to the RLPSI vector of coefficients, that is,

$$\mathbf{b}_R = \mathbf{K}_R\boldsymbol{\beta} \quad (33)$$

where $\mathbf{K}_R = [\mathbf{I} - \mathbf{D}]$, $\mathbf{D} = \mathbf{P}^{-1}\Psi(\Psi'\mathbf{P}^{-1}\Psi)^{-1}\Psi'$, \mathbf{P}^{-1} is the inverse of matrix \mathbf{P} , and \mathbf{I} is an identity matrix $t \times t$. If $\mathbf{N} = \mathbf{U}$ and \mathbf{U} is a null matrix, then $\mathbf{b}_R = \boldsymbol{\beta}$ (the LPSI vector of coefficients). This means that the CLPSI is the most general index and it includes the RLPSI and LPSI as particular cases.

Using restriction $\mathbf{U}'\mathbf{C}'\boldsymbol{\beta} = \theta^*\mathbf{d}$, Tallis (1985) showed that Equation 32 can be written as

$$\mathbf{b}_T = \mathbf{b}_R + \theta^*\boldsymbol{\delta}^* \quad (34)$$

where $\mathbf{b}_R = \mathbf{K}_R\boldsymbol{\beta}$ (Equation 34), $\boldsymbol{\delta}^* = \mathbf{P}^{-1}\Psi(\Psi'\mathbf{P}^{-1}\Psi)^{-1}\mathbf{d}$, $\mathbf{d}' = [d_1 \ d_2 \ \dots \ d_r]$, and $\theta^* = [\boldsymbol{\beta}'\Psi(\Psi'\mathbf{P}^{-1}\Psi)^{-1}\mathbf{d}]/[\mathbf{d}'(\Psi'\mathbf{P}^{-1}\Psi)^{-1}\mathbf{d}]$. When $\theta^* = 0$, $\mathbf{b}_T = \mathbf{b}_R$, and if $\theta^* = 0$ and \mathbf{U}' is a null matrix, $\mathbf{b}_T = \boldsymbol{\beta}$ (Equation 14). In addition, when $\theta^* = 1.0$, Equation 34 is equal to $\mathbf{b}_T = \mathbf{b}_R + \boldsymbol{\delta}^*$, which is not an optimum CLPSI (Mallard, 1972; Tallis, 1962). Itoh and Yamada (1987) have shown that Equation 34 can be written as Equation 32, and vice versa. Thus, Equations 32 and 34 express the same result in a different mathematical way.

There are some problems associated with θ^* . For example, the θ^* values should be higher than zero and $\theta^* \neq 1.0$ because when the value of θ^* is negative, the CLPSI moves the population trait means in the opposite direction to the desired direction (Itoh & Yamada, 1987). Problems associated with the RLPSI and CLPSI are (a) due to the constraints, the correlation between RLPSI and CLPSI with the net genetic merit decreases as the number of constraints increases and (b) the CLPSI may change the means in the opposite direction to the desired direction due to the opposite sign values between \mathbf{w} and \mathbf{d} .

According to Itoh and Yamada (1987), one possible solution to the above problems could be to use the desired gains LPSI theory (Brascamp, 1984; Itoh & Yamada, 1986) to make a selection. The most important aspect of the desired LPSI gains is that it does not require economic weights, which

means that the covariance between I and H [$\text{Cov}(H, I) = \mathbf{w}'\mathbf{C}\boldsymbol{\beta}$] is not defined. In addition, this index maximizes neither the correlation between I and H (ρ_{IH}) nor the selection response (Equation 9). Another problem with this index is that it is not a predictor of $H = \mathbf{w}'\mathbf{g}$. Thus, we think that the CLPSI is a better option to select than the desired gains LPSI.

6.2 | The maximized CLPSI expected genetic gain per trait, correlation, and response

By changing vector $\boldsymbol{\beta} = \mathbf{P}^{-1}\mathbf{C}\mathbf{w}$ in Equations 32, 33, or 34 in Equations 10–12, it is possible to obtain the maximized CLPSI expected genetic gain per trait, correlation, and selection response. That is, the LPSI and CLPSI parameters are equivalent, and the only difference is how the vector of coefficients is obtained.

6.3 | COMLPSI

According to the single-stage CLPSI (Equation 32) and OMLPSI theory described above, to obtain the COMLPSI vector of coefficients at Stage i we need to minimize the mean squared difference between the net genetic merit $H = \mathbf{w}'\mathbf{g}$ and the index $I_i = \boldsymbol{\beta}_i' \boldsymbol{\omega}_i$, i.e., $E[(H - I_i)^2]$, under the restrictions $\mathbf{M}_i' \boldsymbol{\beta}_i = \mathbf{0}$, where $\mathbf{M}_i' = \mathbf{N}'\mathbf{U}'\boldsymbol{\Lambda}_i'$ and $\boldsymbol{\Lambda}_i'$ is as defined earlier. Supposing that matrices \mathbf{Q}_{ii} , \mathbf{U} , and $\boldsymbol{\Lambda}_i'$ and vectors \mathbf{d} and \mathbf{w} are known, then at Stage i it is necessary to minimize the function

$$f_O(\boldsymbol{\beta}_i, \mathbf{v}) = \boldsymbol{\beta}_i' \mathbf{Q}_{ii} \boldsymbol{\beta}_i - 2\mathbf{w}\boldsymbol{\Lambda}_i' \boldsymbol{\beta}_i + 2\xi' \mathbf{M}_i' \boldsymbol{\beta}_i \quad (30)$$

concerning the vector of coefficients $\boldsymbol{\beta}_i$ and the vector of Lagrange multipliers $\boldsymbol{\xi}' = [\xi_1 \ \xi_2 \ \dots \ \xi_{r-1}]$. The COMLPSI vector of coefficients at Stage i is

$$\boldsymbol{\beta}_{C_i} = \mathbf{K}_{O_i} \boldsymbol{\beta}_i \quad (35)$$

where $\boldsymbol{\beta}_i$ is as defined in Equation 14, $\mathbf{K}_{O_i} = [\mathbf{I}_i - \mathbf{F}_{O_i}]$, $\mathbf{F}_{O_i} = \mathbf{Q}_{ii}^{-1} \mathbf{M}_i' (\mathbf{M}_i' \mathbf{Q}_{ii}^{-1} \mathbf{M}_i')^{-1} \mathbf{M}_i'$, and \mathbf{I}_i is an identity matrix of the same size as \mathbf{Q}_{ii} . When $\mathbf{N} = \mathbf{U}$, the vector of coefficients of Equation 35 is similar to Equation 33, and when $\mathbf{N} = \mathbf{U}$ is a null matrix, Equation 35 is equal to Equation 14. Thus, the COMLPSI is more general than the OMLPSI and includes the OMLPSI and the multi-stage null phenotypic restricted index as particular cases.

6.4 | CDMLPSI

Xie and Xu (1997) extended the DMLPSI to a constrained DMLPSI (CDMLPSI); however, their approach is based on the single-stage Tallis (1962) constrained index, which is not optimal. Based on the Mallard (1972) index (Equation 32), which is an optimal index, Cerón-Rojas et al. (2019b) developed an optimal CDMLPSI, which we will describe in this subsection. The main difference between the CDMLPSI and the COMLPSI is that the COMLPSI imposes only one restriction when it solves the LPSI equations to obtain its vector of coefficients, whereas the CDMLPSI imposes the additional restriction that the covariance among the CDMLPSI values between stages be zero. As we shall see, the CDMLPSI vector of coefficients is a linear combination of the OMLPSI vector of coefficients (Equation 14).

Let $I_{i-1} = \boldsymbol{\delta}'_{i-1} \boldsymbol{\omega}_{i-1}$ and $I_i = \boldsymbol{\delta}'_i \boldsymbol{\omega}_i$ be the CDMLPSI at Stages $i-1$ and i , respectively, and let $\mathbf{J}'_{i-1} = [I_1 \ I_2 \ \dots \ I_{i-1}]$ be a vector of CDMLPSIs values until Stage $i-1$ so that the covariance between $I_i = \boldsymbol{\delta}'_i \boldsymbol{\omega}_i$ and \mathbf{J}'_{i-1} is null. To obtain $\boldsymbol{\delta}_i$, we need to minimize the mean squared difference between $H = \mathbf{w}'\mathbf{g}$ and $I_i = \boldsymbol{\delta}'_i \boldsymbol{\omega}_i$ $\{E[(H - I_i)^2]\}$ under the joint restrictions $\mathbf{M}_i' \boldsymbol{\delta}_i = \mathbf{0}$ and $\text{cov}(I_i, \mathbf{J}_{i-1}) = \mathbf{B}'_{(i-1)} \mathbf{Q}_{(i-1)} \boldsymbol{\delta}_i = \mathbf{0}$, where $\mathbf{M}_i' = \mathbf{N}'\mathbf{U}'\boldsymbol{\Lambda}_i'$, and $\boldsymbol{\Lambda}_i'$ is as defined earlier. In this case, \mathbf{B}_{i-1} is an upper triangular matrix (Cerón-Rojas et al., 2019a, 2019b) where each column contains the vectors of coefficients of the CDMLPSI until Stage $i-1$.

In a manner similar to the DMLPSI, let $\mathbf{S}_{i(i-1)} = \mathbf{Q}_{i(i-1)} \mathbf{B}_{(i-1)}$ and $\mathbf{S}'_{i(i-1)} = \mathbf{S}_{(i-1)i}$; then at Stage i , the CDMLPSI vector of coefficients that maximizes Equations 9 and 10 and the correlation is

$$\boldsymbol{\delta}_i = \mathbf{K}_{D_i} \boldsymbol{\beta}_i \quad (36)$$

where $\boldsymbol{\beta}_i$ is as defined in Equation 14, $\mathbf{K}_{D_i} = [\mathbf{I}_i - \mathbf{F}_{D_i}]$, $\mathbf{F}_{D_i} = \mathbf{Q}_{ii}^{-1} \mathbf{V}_i (\mathbf{V}_i' \mathbf{Q}_{ii}^{-1} \mathbf{V}_i)^{-1} \mathbf{V}_i'$, $\mathbf{V}_i = [\mathbf{M}_i \ \mathbf{S}_{i(i-1)}]$, $\mathbf{V}_i' = \begin{bmatrix} \mathbf{M}_i' \\ \mathbf{S}_{(i-1)i} \end{bmatrix}$, and \mathbf{I}_i is an identity matrix of the same size as \mathbf{Q}_{ii} . When matrix $\mathbf{S}_{i(i-1)}$ is null, Equations 36 and 35 are the same. When $\mathbf{N} = \mathbf{U}$, Equation 36 imposes null restrictions, giving us a null restricted DMLPSI index similar to Equation 33.

According to Equations 35 and 36, matrices $\mathbf{K}_{O_i} = [\mathbf{I}_i - \mathbf{F}_{O_i}]$ and $\mathbf{K}_{D_i} = [\mathbf{I}_i - \mathbf{F}_{D_i}]$ transform the OMLPSI vector of coefficients ($\boldsymbol{\beta}_i$) into the COMLPSI and the CDMLPSI vectors of coefficients, respectively. In addition, the CDMLPSI expected genetic gain per trait, selection response, and the correlation at Stage i are the same as those of the OMLPSI when replacing $\boldsymbol{\beta}_i$ with $\boldsymbol{\delta}_i = \mathbf{K}_{D_i} \boldsymbol{\beta}_i$.

6.5 | The single-stage constrained LGSI

In a manner similar to the single-stage CLPSI context (Equation 32), it can be shown (Cerón-Rojas & Crossa, 2019) that the difference between the unconstrained LGSI described above and the constrained LGSI (CLGSI) is the projection matrix $\mathbf{K}_G = [\mathbf{I}_t - \mathbf{D}_G]$, where $\mathbf{D}_G = \mathbf{UN}(\mathbf{N}'\mathbf{U}'\mathbf{TUN})^{-1}\mathbf{N}'\mathbf{U}'\mathbf{T}$. Matrices \mathbf{N} and \mathbf{U} were defined in Equation 32, whereas \mathbf{I}_t is an identity matrix $t \times t$. Matrix \mathbf{K}_G is idempotent and projects the LGSI vector of coefficients ($\mathbf{w} = \boldsymbol{\beta}$) into a space that is smaller than the original space of \mathbf{w} . Thus, while \mathbf{w} is the vector of coefficients of the unconstrained LGSI,

$$\mathbf{b}_G = \mathbf{K}_G \mathbf{w} \quad (37)$$

is the CLGSI vector of coefficients. In the CLGSI context, the selection response and expected genetic gain per trait are the same as those described in Equation 20 when changing \mathbf{w} for $\mathbf{b}_G = \mathbf{K}_G \mathbf{w}$.

7 | ESIM, RESIM, and PPG-ESIM

In the current review, we describe ESIM, RESIM, and PPG-ESIM, which were developed in the canonical correlation context, and do not use economic weights when predicting $H = \mathbf{w}'\mathbf{g}$ and making selections. The other indices associated with ESIM (e.g., MESIM and GESIM) can be seen in Cerón-Rojas and Crossa (2018).

7.1 | ESIM

The fundamental difference between ESIM and the LPSI is the interpretation of the vector \mathbf{w} . Whereas \mathbf{w} is a vector of economic weights (known and fixed) in the LPSI context, in ESIM \mathbf{w} is fixed but unknown and can be estimated in each selection cycle as a linear combination of the ESIM vector of coefficients. Moreover, in ESIM, the breeding values (\mathbf{g}) and the phenotypic records (\mathbf{y}) are two sets of variables that should have maximum correlation (λ_j), from where the CCA theory is the base of ESIM. The ESIM index can be written as $I = \mathbf{b}'_{E_j}\mathbf{y}$, where $\mathbf{b}'_{E_j} = [b_1 \ b_2 \ \dots \ b_t]$ is the unknown index vector of coefficients, and t is the number of traits. In the CCA context, the vector of ESIM coefficients (\mathbf{b}'_{E_j}) is called the “canonical vector,” whereas λ_j is called the “canonical correlation” (Wilms & Croux, 2016).

For the j th correlation (λ_j), the CCA theory allows us to write ESIM and $H = \mathbf{w}'\mathbf{g}$ as $I_E = \mathbf{b}'_{E_j}\mathbf{y}$ and $H_E = \mathbf{w}'_{E_j}\mathbf{g}$, respectively, where $\mathbf{w}_{E_j} = \mathbf{C}^{-1}\mathbf{P}\mathbf{b}_{E_j}$, and \mathbf{b}_{E_j} and λ_j

are obtained from

$$(\mathbf{P}^{-1}\mathbf{C} - \lambda_j^2\mathbf{I})\mathbf{b}_{E_j} = \mathbf{0} \quad (38)$$

Note that \mathbf{b}_{E_j} is the j th vector ($j = 1, 2, \dots, t$) of the multi-trait heritability matrix $\mathbf{P}^{-1}\mathbf{C}$ and maximizes the correlation $\rho_{HI} = \sigma_{HI}/\sigma_H\sigma_I$ described in Equation 9 (Cerón-Rojas & Crossa, 2018). Equation 38 is the ESIM fundamental results obtained by Cerón-Rojas, Sahagún-Castellanos, et al. (2008) and is the base for constructing RESIM, PPG-ESIM, MESIM, and GESIM. To select with ESIM, breeders should use the first eigenvector (\mathbf{b}_{E_1}) of matrix $\mathbf{P}^{-1}\mathbf{C}$ in $I_E = \mathbf{b}'_{E_1}\mathbf{y}$ and the first λ_1 in the ESIM selection response.

7.2 | Changing the signs and proportions of the ESIM vector of coefficients

Equation 38 can be written as $(\mathbf{P}^{-1}\mathbf{C})\mathbf{I}\mathbf{b}_{E_j} = \lambda_j^2\mathbf{I}\mathbf{b}_{E_j}$, where $\mathbf{I} = \mathbf{F}^{-1}\mathbf{F}$ is an identity matrix of size $t \times t$ ($t =$ number of traits), and $\mathbf{F} = \text{diag}\{f_1, f_1 \dots f_t\}$ is a diagonal matrix with values equal to any real number, except zero values. Thus, another way of writing Equation 38 is

$$[\mathbf{F}(\mathbf{P}^{-1}\mathbf{C})\mathbf{F}^{-1} - \lambda_j^2\mathbf{I}]\boldsymbol{\beta}_{E_j} = \mathbf{0}$$

where $\boldsymbol{\beta}_{E_j} = \mathbf{F}\mathbf{b}_{E_j}$. Matrices $\mathbf{P}^{-1}\mathbf{C}$ and $\mathbf{F}(\mathbf{P}^{-1}\mathbf{C})\mathbf{F}^{-1}$ are similar, and both have the same eigenvalues but different eigenvectors (Harville, 2008). When the \mathbf{F} values are only 1s, vector \mathbf{b}_{E_j} is not affected; when the \mathbf{F} values are only -1 s, vector \mathbf{b}_{E_j} will change its direction, and if the \mathbf{F} values are different from 1 and -1 , matrix \mathbf{F} will change the proportional values of \mathbf{b}_{E_j} . In practice, \mathbf{b}_{E_j} is first obtained from Equation 38 and then multiplied by matrix \mathbf{F} to obtain $\boldsymbol{\beta}_{E_j} = \mathbf{F}\mathbf{b}_{E_j}$; that is, $\boldsymbol{\beta}_{E_j}$ is a linear transformation of \mathbf{b}_{E_j} . Matrix $\mathbf{F}(\mathbf{P}^{-1}\mathbf{C})\mathbf{F}^{-1}$ is called “similarity transformation,” and matrix \mathbf{F} is called the “transforming matrix” (Cerón-Rojas & Crossa, 2018). Vector $\boldsymbol{\beta}_{E_j} = \mathbf{F}\mathbf{b}_{E_j}$ can substitute \mathbf{b}_{E_j} , and in this case, the optimized ESIM index should be written as $I_{E_j} = \boldsymbol{\beta}'_{E_j}\mathbf{y}$.

7.3 | The maximized ESIM selection response, expected genetic gain per trait, and correlation

Changing $\boldsymbol{\beta} = \mathbf{P}^{-1}\mathbf{C}\mathbf{w}$ (Equation 4) for \mathbf{b}_{E_1} (Equation 38) in Equation 12, the ESIM selection response can be written similarly as the LPSI response. The same is true for the maximized expected genetic gain per trait, while the ESIM correlation is λ_1 , the square root of λ_1^2 . Cerón-Rojas and Crossa (2018) showed the ESIM correlation could also be written as Equation 11.

There are some problems associated with ESIM. For example, when matrix \mathbf{P} is not positive definite and matrix \mathbf{C} is not positive semidefinite, the eigenvalues of Equation 38 could be negative, and, in such cases, ESIM should not be used to make selections. Nevertheless, Okamoto (1973) showed that when the number of observations was higher than the number of traits, \mathbf{P} was symmetric and positive definite, and its eigenvalues were different with probability one. In turn, Gentle (2007) showed that when \mathbf{P} and \mathbf{C} had the same size, all eigenvalues of $\mathbf{P}^{-1}\mathbf{C}$ existed and were finite. Thus, under the Okamoto (1973) and Gentle (2007) results, ESIM is a good option for making phenotypic selection in breeding programs when it is not possible to obtain vector \mathbf{w} for the LPSI.

7.4 | Summary of the ESIM statistical properties

1. The variance of I_E and the covariance between H_E and I_E are the same.
2. The ESIM variance prediction error is $(1 - \lambda_1^2)\sigma_{H_E}^2$, where $\sigma_{H_E}^2$ is the variance of H_E .
3. The relative effectiveness of I_E in predicting H_E is the ratio of $(1 - \lambda_1^2)\sigma_{H_E}^2$ over $\sigma_{H_E}^2$ (i.e., $1 - \lambda_1^2$); thus, the greater λ_1^2 is, the more effective I_E is in predicting H_E .
4. The mean squared effect of I_E on H_E or the total variance of H_E explained by I_E is $\sigma_{I_E}^2 = \lambda_1^2\sigma_{H_E}^2$, and the relative mean squared effect can be measured by λ_1^2 .

Cerón-Rojas and Crossa (2018) showed that all the above results are valid for RESIM, PPG-ESIM, MESIM, and GESIM. Note that the above ESIM properties are very similar to the LPSI described earlier. This is so because the only difference among the LPSI and ESIM is the method to obtain the vector of coefficients.

7.5 | The constrained eigen selection index method

Similar to the CLPSI (Equation 32), the objective of the predetermined proportional gain ESIM (PPG-ESIM) is to fix r of t ($r < t$) traits by predicting only the genetic gains of $(t - r)$ of them. In Equation 32, we gave details associated with the CLPSI that will be used in the subsequent subsections. The statistical bases and objectives of the PPG-ESIM are the same as those of the CLPSI described in Equation 32. That is, PPG-ESIM tries to change μ_j to $\mu_j + d_j$, where d_j is a predetermined change in μ_j . In the RESIM, $d_j = 0$ ($j = 1, 2,$

\dots, r), where r is the number of predetermined proportional gains.

7.6 | The PPG-ESIM vector of coefficients

According to Cerón-Rojas et al. (2016a), the PPG-ESIM vector of coefficients that maximizes the PPG-ESIM selection response and expected genetic gain per trait is

$$(\mathbf{K}_E\mathbf{P}^{-1}\mathbf{C} - \lambda_P^2\mathbf{I}_t)\mathbf{b}_P = \mathbf{0} \quad (39)$$

where matrices $\mathbf{K}_E = [\mathbf{I}_t - \mathbf{Q}_E]$ and $\mathbf{Q}_E = \mathbf{P}^{-1}\mathbf{M}(\mathbf{M}'\mathbf{P}^{-1}\mathbf{M})\mathbf{M}'$ are the same as those obtained in Equation 32, \mathbf{I}_t is an identity matrix $t \times t$, and λ_P^2 and \mathbf{b}_P are the eigenvalue and the eigenvector of matrix $\mathbf{K}_E\mathbf{P}^{-1}\mathbf{C}$, respectively. When $\mathbf{N} = \mathbf{U}$, $\mathbf{b}_P = \mathbf{b}_R$ (the vector of coefficients of RESIM), and when \mathbf{U}' is a null matrix, $\mathbf{b}_P = \mathbf{b}_E$ (the vector of coefficients of ESIM). That is, PPG-ESIM is more general than RESIM and ESIM and includes the latter two indices as particular cases.

7.7 | Relationships among ESIM, PPG-ESIM, LPSI, and CLPSI

Cerón-Rojas et al. (2008a, 2016b) described the relationships between the ESIM and PPG-ESIM vector of coefficients and the LPSI and CLPSI vector of coefficients as follows. When the LPSI and CLPSI vector of coefficients ($\boldsymbol{\beta}$ and $\mathbf{b}_C = \mathbf{K}\boldsymbol{\beta}$, respectively) are in the space generated by the eigenvectors of Equations 38 and 39, respectively, $\boldsymbol{\beta}$ and \mathbf{b}_C can be written as

$$\boldsymbol{\beta} = \sum_{j=1}^t \alpha_j^* \mathbf{b}_{E_j} \text{ and } \mathbf{b}_C = \sum_{j=1}^t a_j^* \mathbf{b}_{P_j} \quad (35)$$

where α_j^* and a_j^* are constants that can be estimated by least squares, and \mathbf{b}_{E_j} and \mathbf{b}_{P_j} are the eigenvectors of Equations 38 and 39. Based on the Bessel inequality, Cerón-Rojas et al. (2008a) described a method to determine conditions under which $\boldsymbol{\beta}$ and \mathbf{b}_C are in the space generated by the eigenvectors of Equations 38 and 39. The above results allowed Cerón-Rojas et al. (2016b) to find, in the asymptotic context, approximate expectations and variances of the estimators of $\boldsymbol{\beta}$ and \mathbf{b}_C .

7.8 | Estimators of the LPSI parameters

Cerón-Rojas and Crossa (2018) have described the REML method to estimate the LSI parameters. Here, we consider

the estimators of the LPSI vector of coefficients, selection response, correlation, and expected genetic gain per trait, as

$$\hat{\boldsymbol{\beta}} = \hat{\mathbf{P}}^{-1} \hat{\mathbf{C}} \mathbf{w}, \hat{R} = k \sqrt{\hat{\boldsymbol{\beta}}' \hat{\mathbf{P}} \hat{\boldsymbol{\beta}}}, \hat{\rho}_{HI} = \frac{\sqrt{\hat{\boldsymbol{\beta}}' \hat{\mathbf{P}} \hat{\boldsymbol{\beta}}}}{\sqrt{\mathbf{w}' \hat{\mathbf{C}} \mathbf{w}}}, \text{ and } \hat{\mathbf{E}} = k \frac{\hat{\mathbf{C}} \hat{\boldsymbol{\beta}}}{\hat{\sigma}_I} \quad (36)$$

respectively, and the estimated LPSI value will be expressed as $\hat{I} = \hat{\boldsymbol{\beta}}' \mathbf{y}$. In the above equations, $\hat{\mathbf{P}}$ ($\hat{\mathbf{P}}^{-1}$) and $\hat{\mathbf{C}}$ are REML estimators of the phenotypic (\mathbf{P}) and genotypic (\mathbf{C}) covariance matrices, whereas $\hat{\sigma}_I = \sqrt{\hat{\boldsymbol{\beta}}' \hat{\mathbf{P}} \hat{\boldsymbol{\beta}}}$ is the estimator of the standard deviation of \hat{I} , and $\sqrt{\mathbf{w}' \hat{\mathbf{C}} \mathbf{w}}$ is the estimator of the standard deviation of H .

The REML method is based on projecting the data into a subspace free of fixed effects and maximizing the likelihood function in this subspace (Blasco, 2001). Thus, because $\hat{\mathbf{P}}$ ($\hat{\mathbf{P}}^{-1}$) and $\hat{\mathbf{C}}$ are REML estimators, by the invariance property of maximum likelihood estimators (Rencher & Schaalje, 2008), $\hat{\boldsymbol{\beta}}$, \hat{R} , $\hat{\rho}_{HI}$, and $\hat{\mathbf{E}}$ are maximum likelihood estimators and they are minimum variance and asymptotic unbiased estimators for $\boldsymbol{\beta}$, R , ρ_{HI} , and \mathbf{E} (Cerón-Rojas & Crossa, 2020b; Muirhead, 2005). Cerón-Rojas and Crossa (2020b) showed that, in effect, \hat{R} and $\hat{\rho}_{HI}$ are asymptotic unbiased estimators and these authors described methods to obtain confidence intervals for the expectations of \hat{R} and $\hat{\rho}_{HI}$. The REML estimators can also be used in the CLPSI, ESIM, combined LGSI, multi-stage LSI, etc.

7.9 | Numerical example: LPSI vs ESIM

We compare the ESIM efficiency versus LPSI efficiency using a real data set from commercial egg poultry lines obtained from Akbar et al. (1984). The estimated phenotypic ($\hat{\mathbf{P}}$) and genetic ($\hat{\mathbf{C}}$) covariance matrices among the rate of lay (RL, number of eggs), age at sexual maturity (SM, days),

and egg weight (kg), were $\hat{\mathbf{P}} = \begin{bmatrix} 240.57 & -95.62 & 2.07 \\ -95.62 & 167.20 & 4.58 \\ 2.07 & 4.58 & 22.80 \end{bmatrix}$ and

$\hat{\mathbf{C}} = \begin{bmatrix} 29.86 & -17.90 & -4.13 \\ -17.90 & 18.56 & 1.49 \\ -4.13 & 1.49 & 9.24 \end{bmatrix}$, respectively. Furthermore, the

number of genotypes and the vector of economic weights were $n = 3,330$ and $\mathbf{w}' = [19.54 \ -3.56 \ 17.01]$, respectively, and the selection intensity was 10% ($k = 1.755$) for both indices.

The estimated LPSI vector of coefficients was $\hat{\boldsymbol{\beta}} = \mathbf{w}' \hat{\mathbf{P}}^{-1} \hat{\mathbf{C}} = [1.82 \ -1.38 \ 3.25]$, and the LPSI estimated selection response, expected genetic gain per trait, and correlation was $\hat{R} = k \sqrt{\hat{\boldsymbol{\beta}}' \hat{\mathbf{P}} \hat{\boldsymbol{\beta}}} = 74.91$, $\hat{\mathbf{E}}' = k \frac{\hat{\boldsymbol{\beta}}' \hat{\mathbf{C}}}{\sqrt{\hat{\boldsymbol{\beta}}' \hat{\mathbf{P}} \hat{\boldsymbol{\beta}}}} =$

$[2.70 \ -2.20 \ 0.84]$, and $\hat{\rho}_{HI} = \frac{\sqrt{\hat{\boldsymbol{\beta}}' \hat{\mathbf{P}} \hat{\boldsymbol{\beta}}}}{\sqrt{\mathbf{w}' \hat{\mathbf{C}} \mathbf{w}}} = 0.362$, respectively.

Because in the ESIM context $\hat{\mathbf{b}}'_{E_1} \hat{\mathbf{b}}_{E_1} = 1$, the best way of comparing ESIM results versus LPSI results is when the LPSI coefficient vector is normalized (i.e., when the LPSI coefficient vector is equal to $\hat{\boldsymbol{\beta}}^* = \hat{\boldsymbol{\beta}} / \hat{\boldsymbol{\beta}}' \hat{\boldsymbol{\beta}}$ and then $\hat{\boldsymbol{\beta}}^* \hat{\boldsymbol{\beta}}^* = 1$); however, it can be shown that the normalization process only affects the estimated LPSI selection response because in that case, $\hat{R} = 74.91$ is divided by $\hat{\boldsymbol{\beta}}' \hat{\boldsymbol{\beta}}$. For example, for this data set result, $\hat{\boldsymbol{\beta}}' \hat{\boldsymbol{\beta}} = 15.76$; then, the estimated LPSI selection response using $\hat{\boldsymbol{\beta}}^* = \hat{\boldsymbol{\beta}} / \hat{\boldsymbol{\beta}}' \hat{\boldsymbol{\beta}}$ will be $\hat{R} = \frac{74.91}{15.74} = 4.75$, while the rest of the estimated LPSI parameters will be the same. When $0 < \hat{\boldsymbol{\beta}}' \hat{\boldsymbol{\beta}} < 1$ and $1 < \hat{R}$, the values of \hat{R} will increase, but when $1 < \hat{\boldsymbol{\beta}}' \hat{\boldsymbol{\beta}}$, the values of \hat{R} will decrease, as in this example. The product $\hat{\boldsymbol{\beta}}' \hat{\boldsymbol{\beta}}$ does not affect $\hat{\rho}_{HI}$ because it is invariant to scale change. Also, $\hat{\boldsymbol{\beta}}' \hat{\boldsymbol{\beta}}$ does not affect $\hat{\mathbf{E}}$ because $\hat{\boldsymbol{\beta}}' \hat{\boldsymbol{\beta}}$ appears in the numerator and denominator of both estimated parameters.

In ESIM, the sign and proportion of the expected genetic gain values for traits RL, SM, and egg weight should be according to the breeder's interest. For example, if the breeder's interest is that the expected genetic gain per trait for RL be positive and negative for SM, the sign and proportion of the values of the first eigenvector should be modified using a linear combination of $\hat{\mathbf{b}}_{E_1}$ as described earlier (i.e., $\hat{\mathbf{b}}_{E_1} = \mathbf{F} \hat{\mathbf{b}}_{E_1}$) in order to achieve expected genetic gain per trait values in RL and SM.

The information needed to obtain the estimated ESIM parameters is matrix $\hat{\mathbf{P}}^{-1} \hat{\mathbf{C}}$. We need to find the eigenvalues and eigenvectors of the equation $[(\hat{\mathbf{P}}^{-1} \hat{\mathbf{C}})(\hat{\mathbf{P}}^{-1} \hat{\mathbf{C}})' - \hat{\mu}_j \mathbf{I}] \hat{\mathbf{b}}_{E_j} = \mathbf{0}$, where $\hat{\mu}_j = \hat{\lambda}_{E_j}^4$ (Cerón-Rojas & Crossa, 2018). In this case, $\hat{\mu}_1 = \hat{\lambda}_{E_1}^4 = 0.2115$, $\hat{\lambda}_{E_1}^2 = 0.4599$ and the estimated ESIM correlation was $\hat{\lambda}_{E_1} = 0.6782$. The estimated ESIM eigenvector of coefficients is $\hat{\mathbf{b}}'_{E_1} = [-0.1701 \ 0.0259 \ 0.9851]$, and the estimated ESIM index can be constructed as

$$\hat{I}_E = -0.170RL + 0.026SM + 0.985EW$$

The estimated ESIM selection response and expected genetic gain per trait were $\hat{R}_E = 1.755 \sqrt{\hat{\mathbf{b}}'_{E_1} \hat{\mathbf{P}} \hat{\mathbf{b}}_{E_1}} = 9.54$ and $\hat{\mathbf{E}}'_E = 1.755 \frac{\hat{\mathbf{b}}'_{E_1} \hat{\mathbf{C}}}{\sqrt{\hat{\mathbf{b}}'_{E_1} \hat{\mathbf{P}} \hat{\mathbf{b}}_{E_1}}} = [-3.10 \ 1.61 \ 3.18]$, respectively. Because the estimated LPSI selection response was $\hat{R} = (74.91/15.74) = 4.75$, the estimated ESIM selection response was higher than the estimated LPSI response for this data set.

Now suppose that the breeder's interest is to increase RL and decrease SM; then $\hat{\mathbf{E}}'_E$ is a good result but $\hat{\mathbf{E}}'_E$ is wrong. We can change the sign and proportion of $\hat{\mathbf{E}}'_E$ by transforming $\hat{\mathbf{b}}_{E_1}$ as $\hat{\boldsymbol{\beta}}_{E_1} = \mathbf{F} \hat{\mathbf{b}}_{E_1}$ using a convenient matrix \mathbf{F} such as

$$\mathbf{F} = \begin{bmatrix} -9 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

In such a case, $\hat{\beta}'_{E_1} = \hat{\mathbf{b}}'_{E_1} \mathbf{F} = [1.531 \ 0.026 \ 0.981]$, $\hat{R}_E = 1.755\sqrt{\hat{\beta}'\hat{\mathbf{P}}\hat{\beta}} = 42.44$, and $\hat{\mathbf{E}}'_E = 1.755(\hat{\beta}'\hat{\mathbf{C}})/(\sqrt{\hat{\beta}'\hat{\mathbf{P}}\hat{\beta}}) = [2.990 \ -1.85 \ 0.205]$. However, vector $\hat{\beta}'_{E_1}$ has not been normalized. To normalize $\hat{\beta}'_{E_1}$, we need to divide it by $\hat{\beta}'_{E_1} \hat{\beta}_{E_1} = 3.314$, but $\hat{\beta}'_{E_1} \hat{\beta}_{E_1}$ should only affect $\hat{R}_E = 42.44$, which should be divided by 3.314; that is, $\hat{R}_E = 42.44/3.314 = 12.806$. According to the theory of similar matrices (Harville, 2008), the estimated maximized ESIM correlation ($\hat{\lambda}_{E_1} = 0.6782$) should not be affected by matrix \mathbf{F} .

7.10 | RIndSel

RIndSel is a powerful freely available Java software useful for estimating the LPSI, LGSI, CLPSI, ESIM, etc., parameters and selecting individual candidates as parents for the next selection cycle. It is available in the CIMMYT (for its Spanish acronym, or International Maize and Wheat Improvement Center) **RIndSel Selection indices for plant breeding** repository at <https://data.cimmyt.org/dataverse/root?q=rindsel>. **RIndSel** is compatible with Windows XP, 7, 8, and 10 and can be installed on 32-bit and 64-bit computers. To input the data information, the user must create an Excel file (saved as a comma-delimited “.csv”) with the data information to be analyzed, from where the software reads that information and estimates the index parameters.

In the linear mixed model context (Searle et al., 2006), **RIndSel** uses REML and two experimental designs (a randomized complete block design and a lattice or alpha lattice design) to estimate the phenotypic and genotypic covariance matrices. This software presents the results of the estimated index parameters in two Excel files (e.g., *alloutSmith.csv* and *outSmith.csv*, where, in this case, “Smith” denotes the LPSI used for the analysis) and one text file (*outtextSmith.txt*). File “*alloutSmith.csv*” contains all the trait adjusted means for all genotypes, and the estimated index values, whereas file “*outSmith.csv*” contains the genotypes and the trait values selected with the index values after these have been ranked. File “*outtextSmith.txt*” contains all information associated with the estimated index parameters such as the estimated phenotypic and genotypic covariance matrices, the estimated selection response, etc. A complete user manual with instructions on how to use **RIndSel** in plant breeding is available at <https://data.cimmyt.org/dataset.xhtml?persistentId=hdl:11529/10854>.

7.11 | Summary of the main results of this review

In this manuscript, we have summarized and expanded upon the results of the following authors:

1. Smith (1936) established the LPSI theory on which all LSIs described in this work are based.
2. Cochran (1951) was the first to indicate that the LPSI is the BLP of the net genetic merit and was the first to develop the two-stage optimum LPSI.
3. Kempthorne and Nordskog (1959) were the first to develop the RLPSI.
4. Young (1964) combined the LPSI theory with the culling selection method and developed the OMLPSI in the multi-stage context, from where Cerón-Rojas et al. (2019b) developed the COMLPSI and Cerón-Rojas and Crossa (2020a) developed a COMLPSI.
5. Mallard (1972) developed the first optimum constrained LPSI, which allows imposing constraints different to zero on the expected genetic gain per trait and includes the RLPSI as a particular case.
6. Xu and Muir (1992) developed the DMLPSI, from where Cerón-Rojas et al. (2019b) developed a constrained optimum DMLPSI and Cerón-Rojas and Crossa (2020a) developed a DCMLPSI.
7. Cerón-Rojas et al. (2008a) and Cerón-Rojas et al. (2016b) developed the ESIM, RESIM, and PPG-ESIM in the CCA context.
8. Cerón-Rojas et al. (2015) and Cerón-Rojas and Crossa (2019) expanded the unconstrained and constrained LPSI to the genomic selection context and developed the unconstrained and constrained LGSI.

This work complements the classic characterizations of the LSI made by several authors (Bulmer, 1980; Cochran, 1951; Henderson, 1963) and indicates that the LSI theory is a general mathematical-statistical framework that should allow breeders to confidently use it in plant and animal breeding programs.

8 | CONCLUSIONS

Populations of plants or animals are sets of individuals that, in turn, are sets of phenotypic characteristics, such as color, weight, height, etc. Thus, when breeders select parents in a breeding program, they select individuals, not single characteristics, and the selection of any set of traits on the individuals changes not only the mean of the individual's unselected traits but also the correlation between selected and unselected

traits (Pearson, 1903). Then, in breeding programs, breeders should select parents using LSI because the prediction procedures with univariate models do not contemplate the genetic correlations between traits. In practice, evaluating plants or animals requires several simultaneous traits. For example, breeders studying yield and grain quality register phenotypic data that include yield components (e.g., grain weight or biomass), grain quality (e.g., flavor, shape, color, nutrient content), and resistance to biotic and abiotic stresses (Jia & Jannink, 2012).

We have described the LSI statistical theory assuming that the net genetic merit and the vector of phenotypic values have multivariate normal distribution. Under this assumption, and when the phenotypic and genotypic covariance matrices are known, the LSIs are particular cases of the BLP theory. Furthermore, we found that regardless of the LSI characteristics (single-stage or multi-stage, unconstrained or constrained, etc.), the LSI parameters (selection response, expected genetic gain per trait, and correlation) are the same.

The BLP theory provides the mathematical-statistical framework to develop the LSI theory. Thus, the LSI is a particular case or an application of the BLP to plant and animal breeding selection. This result allows breeders to locate the LSI theory in a general mathematical framework and then confidently use it when making selections. Nevertheless, some problems of the LSI not discussed in this work are associated with the estimation of the LSI parameters. For example, when the phenotypic and genotypic covariance matrices are estimated, the LPSI could not be BLP; indeed, it is not linear in \mathbf{y} (Searle et al., 2006). Nevertheless, Cerón-Rojas and Crossa (2020b) found that the LPSI is a good predictor of the net genetic merit including when REML estimates the phenotypic and genotypic covariance matrices. Because of this, breeders should use the LSI theory in plant breeding with confidence.

ACKNOWLEDGMENTS

The authors are grateful to Dr. Roberto Fritsche-Neto, Senior Scientist-IRRI, for reading and revising the original version of the manuscript. We very much appreciate the contribution of one anonymous reviewer, who carefully read and corrected the original manuscript work several times during the past months. We introduced many important suggestions made by the anonymous reviewer that significantly improved the quality of the manuscript and its presentation and readability. We also thank Associate Editor Manjit Kang, who invested a significant time handling the revision of this study. Open Access fees are received from the Bill and Melinda Gates Foundation. We acknowledge the financial support provided by the Bill and Melinda Gates Foundation [INV-003439 BMGF/FCDO Accelerating Genetic Gains in Maize and Wheat for Improved Livelihoods (AG2MW)] as well as USAID projects [Amend. No. 9 MTO 069033, USAID-CIMMYT Wheat/AGGMW,

AGG-Maize Supplementary Project, AGG (Stress Tolerant Maize for Africa)]. We are also thankful for the financial support provided by the Foundations for Research Levy on Agricultural Products (FFL) and the Agricultural Agreement Research Fund (JA) in Norway through NFR grant 267806, as well as the CIMMYT CRP.

AUTHOR CONTRIBUTIONS

J. Jesus Cerón-Rojas: Conceptualization; Investigation. Jose Crossa: Conceptualization; Investigation.

CONFLICT OF INTEREST

The authors do not have any conflict of interest.

ORCID

Jose Crossa  <https://orcid.org/0000-0001-9429-5855>

REFERENCES

- Akbar, M. K., Lin, C. Y., Gyles, N. R., Gavora, J. S., & Brown, C. J. (1984). Some aspects of selection indices with constraints. *Poultry Science*, 63, 1899–1905. <https://doi.org/10.3382/ps.0631899>
- Alvarado, G., Pacheco, A., Perez-Elizalde, S., Burgueño, J., & Rodriguez, F. M. (2018). RinSel: Selection indices with R. pp. 243–256. In *Linear selection indices in modern plant breeding (p. 10)*. Springer.
- Andersson, E., Spanos, K., Mullin, T., & Lindgren, D. (1998). Phenotypic selection compared to restricted combined index selection for many generations. *Silva Fennica*, 32(2), 111–120. <https://doi.org/10.14214/sf.689>
- Arimendi, J. C. (2013). Multivariate truncated moments. *Journal of Multivariate Analysis*, 117, 41–75. <https://doi.org/10.1016/j.jmva.2013.01.007>
- Arnold, B. C., Castillo, E., & Sarabia, M. (1999). *Conditional specification of statistical models*. Springer.
- Baker, R. J. (1986). *Selection indices in plant breeding*. CRC Press.
- Bickel, P. J., & Doksum, K. A. (2015). *Mathematical statistics: Basic ideas and selected topics* (Volume 1, 2nd ed.). CRC Press Taylor & Francis Group.
- Blasco, A. (2001). The Bayesian controversy in animal breeding. *Journal of Animal Science*, 79, 2023–2046. <https://doi.org/10.2527/2001.7982023x>
- Börner, V., & Reinsch, N. (2012). Optimising multi-stage dairy cattle breeding schemes including genomic selection using decorrelated or optimum selection indices. *Genetics Selection Evolution*, 44, 1–11. <https://doi.org/10.1186/1297-9686-44-1>
- Bramscap, E. W. (1984). Selection indices with constraints. *Animal Breeding Abstract*, 52(9), 645–654.
- Bulmer, M. G. (1980). *The mathematical theory of quantitative genetics. Lectures in biostatistics*. Clarendon Press.
- Carlin, B. P., & Louis, T. A. (2000). *Bayes and empirical Bayes methods for data analysis* (2nd ed.). Chapman & Hall/CRC.
- Cerón-Rojas, J. J., Crossa, J., Sahagún-Castellanos, J., Castillo-González, F., & Santacruz-Varela, A. (2006). A selection index method based on eigen analysis. *Crop Science*, 46, 1711–1721. <https://doi.org/10.2135/cropsci2005.11-0420>
- Cerón-Rojas, J. J., Sahagún-Castellanos, J., Castillo-González, F., Santacruz-Varela, A., & Crossa, J. (2008). A restricted selection index

- method based on eigenanalysis. *Journal of Agricultural Biological and Environmental Statistics*, 4, 421–438.
- Cerón-Rojas, J. J., Castillo-González, F., Sahagún-Castellanos, J., Santacruz-Varela, A., Benítez-Riquelme, I., & Crossa, J. (2008). A molecular selection index method based on eigenanalysis. *Genetics*, 180, 547–557. <https://doi.org/10.1534/genetics.108.087387>
- Cerón-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.
- Cerón-Rojas, J. J., & Sahagún-Castellanos, J. (2016). Multivariate empirical Bayes to predict the plant breeding values. *Agrociencia*, 50, 633–648.
- Cerón-Rojas, J. J., Crossa, J., Toledo, F. H., & Sahagún-Castellanos, J. (2016a). A predetermined proportional gains eigen selection index method. *Crop Science*, 56, 2436–2447. <https://doi.org/10.2135/cropsci2015.11.0718>
- Cerón-Rojas, J. J., Crossa, J., & Sahagún-Castellanos, J. (2016b). Statistical sampling properties of the coefficients of three phenotypic selection indices. *Crop Science*, 56, 51–58. <https://doi.org/10.2135/cropsci2015.03.0189>
- Cerón-Rojas, J. J., & Crossa, J. (2018). *Linear selection indices in modern plant breeding*. Springer. <https://doi.org/10.1007/978-3-319-91223-3>
- Cerón-Rojas, J. J., & Crossa, J. (2019). Efficiency of a constrained linear genomic selection index to predict the net genetic merit in plants. *Genes/Genomes/Genetics*, 9, 3981–3994. <https://doi.org/10.1534/g3.119.400677>
- Cerón-Rojas, J. J., Toledo, F. H., & Crossa, J. (2019a). The relative efficiency of two multi-stage linear phenotypic selection indices to predict the net genetic merit. *Crop Science*, 59, 1037–1051. <https://doi.org/10.2135/cropsci2018.11.0678>
- Cerón-Rojas, J. J., Toledo, F. H., & Crossa, J. (2019b). Optimum and decorrelated constrained multi-stage linear phenotypic selection indices theory. *Crop Science*, 59, 2585–2600. <https://doi.org/10.2135/cropsci2019.04.0241>
- Cerón-Rojas, J. J., & Crossa, J. (2020a). Combined multi-stage linear genomic selection indices to predict the net genetic merit in plant breeding. *G3: Genes/Genomes/Genetics*, 10, 2087–2101. <https://doi.org/10.1534/g3.120.401171>
- Cerón-Rojas, J. J., & Crossa, J. (2020b). Expectation and variance of the estimator of the maximized selection response of linear selection indices with normal distribution. *Theoretical and Applied Genetics*, 133(9), 2743–2758. <https://doi.org/10.1007/s00122-020-03629-6>
- Christensen, R. (2011). *Plane answers to complex questions: The theory of linear models* (4th ed.). Springer.
- Cochran, W. G. (1951). Improvement by means of selection. In J. Neyman (Ed.), *Proceedings of the second Berkeley symposium on mathematical statistics and probability* (pp. 449–470). University of California Press.
- Cockerham, C. C. (1954). An extension of the concept of partitioning hereditary variance for analysis of covariances among relatives when epistasis is present. *Genetics*, 39, 859–882. <https://doi.org/10.1093/genetics/39.6.859>
- Crawley, M. J. (2013). *The R book* (2nd ed.). John Wiley & Sons.
- Dekkers, J. C. M. (2007). Prediction of response to marker-assisted and genomic selection using selection index theory. *Journal of Animal Breeding and Genetics*, 124, 331–341. <https://doi.org/10.1111/j.1439-0388.2007.00701.x>
- Dekkers, J. C. M., & Settar, P. (2004). Long-term selection with known quantitative trait loci. *Plant Breeding Reviews*, 24, 311–335.
- Dekkers, J. C. M., & Gibson, J. P. (1997). Applying breeding objectives to dairy cattle improvement. *Journal of Dairy Science*, 81(2), 19–35.
- Dekkers, J. C. M., & Dentine, M. R. (1991). Quantitative genetic variation associated with chromosomal markers in segregating populations. *Theoretical and Applied Genetics*, 81, 212–220. <https://doi.org/10.1007/BF00215725>
- Du, F.-X., & Hoeschele, I. (2000). Estimation of additive, dominance and epistatic variance components using finite locus models implemented with a single-site Gibbs and a descent graph sampler. *Genetics Research*, 76, 187–198. <https://doi.org/10.1017/S0016672300004614>
- Eisenhart, C. (1947). The assumptions underlying the analysis of variance. *Biometrics*, 3(1), 1–21. <https://doi.org/10.2307/3001534>
- Elston, R. C. (1963). A weight-free index for the purpose of ranking or selection with respect to several traits at a time. *Biometrics*, 19, 85–97. <https://doi.org/10.2307/2527573>
- Finney, D. J. (1962). Genetics gains under three methods of selection. *Genetics Research*, 3, 417–423. <https://doi.org/10.1017/S0016672300003256>
- Fisher, R. A. (1918). The correlation between relatives on the supposition of Mendelian inheritance. *Earth and Environmental Science*, 52, 399–433. <https://doi.org/10.1017/S0080456800012163>
- Gentle, J. E. (2007). *Matrix algebra theory: Computations and applications in statistics*. Springer.
- Gianola, D., & de los Campos, G. (2009). Inferring genetic values for quantitative traits non-parametrically. *Genetics Research*, 90, 525–540. <https://doi.org/10.1017/S0016672308009890>
- Gianola, D., Perez-Enciso, M., & Toro, M. A. (2003). On marker-assisted prediction of genetic value: Beyond the ridge. *Genetics*, 163, 347–365. <https://doi.org/10.1093/genetics/163.1.347>
- Gimelfarb, A., & Lande, R. (1994). Simulation of marker-assisted selection in hybrid populations. *Genetic Research*, 63, 39–47. <https://doi.org/10.1017/S0016672300032067>
- Gimelfarb, A., & Lande, R. (1995). Marker-assisted selection and marker-QTL associations in hybrid populations. *Theoretical and Applied Genetics*, 91, 522–528. <https://doi.org/10.1007/BF00222983>
- Hanson, W. D., & Johnson, H. W. (1957). Methods for calculating and evaluating a general selection index obtained by pooling information from two or more experiments. *Genetics*, 42(4), 421–432. <https://doi.org/10.1093/genetics/42.4.421>
- Harville, D. A. (1975). Index selection with proportionality constraints. *Biometrics*, 31, 223–225. <https://doi.org/10.2307/2529722>
- Harville, D. A. (2008). *Matrix algebra from a statistician's perspective*. Springer.
- Harville, D. A. (2018). *Linear models and the relevant distributions and matrix algebra*. CRC Press Taylor & Francis Group.
- Hattaway, J. T. (2010). *Parameter estimation and hypothesis testing for the truncated normal distribution with applications to introductory statistics grades* [Master's thesis]. <https://scholarsarchive.byu.edu/etd/2053/>
- Hazel, L. N. (1943). The genetic basis for constructing selection indexes. *Genetics*, 8, 476–490. <https://doi.org/10.1093/genetics/28.6.476>
- Hazel, L. N., Dickerson, G. E., & Freeman, A. E. (1994). The selection index: Then, now, and for the future. *Journal of Dairy Science*, 77, 3236–3251. [https://doi.org/10.3168/jds.S0022-0302\(94\)77265-9](https://doi.org/10.3168/jds.S0022-0302(94)77265-9)

- Hazel, L. N., & Lush, J. L. (1942). The efficiency of three methods of selection. *Journal of Heredity*, 33, 393–399. <https://doi.org/10.1093/oxfordjournals.jhered.a105102>
- Henderson, C. R. (1990). Statistical methods in animal improvement: Historical overview. In D. Gianola & K. Hammond (Eds.), *Advances in statistical methods for genetic improvement of livestock* (pp. 2–14). Springer-Verlag.
- Henderson, C. R. (1984). *Applications of linear models in animal breeding*. University of Guelph.
- Henderson, C. R. (1963). Selection index and expected genetic advance. In *Statistical genetics and plant breeding* (pp. 141–163). National Academy of Science–National Research Council.
- Henderson, C. R. (1949). Estimation of changes in herd environment. *Journal of Dairy Science*, 32, 706–715.
- Henderson, C. R. (1950). Estimation of genetic parameters. *Annals of Mathematical Statistics*, 21, 309–310.
- Heffner, E. L., Sorrells, M. E., & Jannink, J.-L. (2009). Genomic selection for crop improvement. *Crop Science*, 49, 1–12. <https://doi.org/10.2135/cropsci2008.08.0512>
- Hill, W. G. (2010). Understanding and using quantitative genetic variation. *Philosophical Transactions of the Royal Society*, 365(B), 73–85. <https://doi.org/10.1098/rstb.2009.0203>
- Holland, J. B. (2001). Epistasis and plant breeding. *Plant Breeding Review*, 21, 27–112.
- Hospital, F., Moreau, L., Lacoudre, F., Charcosset, A., & Gallais, A. (1997). More on the efficiency of marker-assisted selection. *Theoretical and Applied Genetics*, 95, 1181–1189. <https://doi.org/10.1007/s001220050679>
- Hotelling, H. (1936). Relations between two sets of variables. *Biometrika*, 28, 321–377. <https://doi.org/10.1093/biomet/28.3-4.321>
- Itoh, Y., & Yamada, Y. (1986). Re-examination of selection index for desired gains. *Genetics, Selection, Evolution.*, 18(4), 499–504. <https://doi.org/10.1186/1297-9686-18-4-499>
- Itoh, Y., & Yamada, Y. (1987). Comparisons of selection indices achieving predetermined proportional gains. *Genetics, Selection, Evolution.*, 19(1), 69–82. <https://doi.org/10.1186/1297-9686-19-1-69>
- Jeyaruban, M. G., Gibson, J. P., & Gowe, R. S. (1995). Comparison of index selection and best linear unbiased prediction for simulated layer poultry data. *Poultry Science*, 74, 1566–1576. <https://doi.org/10.3382/ps.0741566>
- Jia, Y., & Jannink, J.-L. (2012). Multiple-trait genomic selection methods increase genetic value prediction accuracy. *Genetics*, 192, 1513–1522. <https://doi.org/10.1534/genetics.112.144246>
- Kempthorne, O., & Nordskog, A. W. (1959). Restricted selection indices. *Biometrics*, 15, 10–19. <https://doi.org/10.2307/2527598>
- Kempthorne, O. (1954). The correlation between relatives in a random mating population. *Proceedings of the Royal Society of London*, 143(B), 102–113.
- Kagan, A., & Wesolowski, J. (1996). Normality via conditional normality of linear forms. *Statistics & Probability Letters*, 29, 229–232.
- Lande, R., & Thompson, R. (1990). Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics*, 124, 743–756. <https://doi.org/10.1093/genetics/124.3.743>
- Li, Z. (1998). Molecular analysis of epistasis affecting complex traits. In A. H. Paterson (Ed.), *Molecular dissection of complex traits* (pp. 119–130). CRC Press.
- Lin, C. Y. (1978). Index selection for genetic improvement of quantitative characters. *Theoretical and Applied Genetics*, 52, 49–56. <https://doi.org/10.1007/BF00281316>
- Lynch, M., & Walsh, B. (1998). *Genetics and analysis of quantitative traits*. Sinauer.
- McCulloch, C. A., & Searle, S. R. (2001). *Generalized, linear, and mixed models*. John Wiley and Sons.
- Mallard, J. (1972). The theory and computation of selection indices with constraints: A critical synthesis. *Biometrics*, 28, 713–735. <https://doi.org/10.2307/2528758>
- Moreau, L., Hospital, F., & Whittaker, J. (2007). Marker-assisted selection and introgression. pp. 718–751. In D. J. Balding, M. Bishop, & C. Cannings (Eds.), *Handbook of statistical genetics* (Volume 1, 3rd ed.), Wiley.
- Moreau, L., Lemarié, S., Charcosset, A., & Gallais, A. (2000). Economic efficiency of one cycle of marker-assisted selection efficiency. *Crop Science*, 40, 329–337. <https://doi.org/10.2135/cropsci2000.402329x>
- Mrode, R. A. (2014). *Linear models for the prediction of animal breeding values* (3rd ed.), CABI.
- Muirhead, R. J. (2005). *Aspects of multivariate statistical theory*. John Wiley & Sons.
- Nordskog, A. W. (1978). Some statistical properties of an index of multiple traits. *Theoretical and Applied Genetics*, 52, 91–94. <https://doi.org/10.1007/BF00281322>
- Okamoto, M. (1973). Distinctness of the eigenvalues of a quadratic form in a multivariate sample. *The Annals of Statistics*, 1(4), 763–765. <https://doi.org/10.1214/aos/1176342472>
- Pacheco, A., Pérez, S., Alvarado, G., Ceron, J., Rodríguez, F., Crossa, J., & Burgueño, J. (2017). *RIndSel: Selection indices for plant breeding*. CIMMYT Research Data & Software Repository Network.
- Pearl, R., & Surface, F. M. (1909). Selection index numbers and their use in breeding. *The American Naturalist*, 43(511), 385–400. <https://doi.org/10.1086/279071>
- Pearson, K. (1903). Mathematical contributions to the theory of evolution. XI. On the influence of natural selection on the variability of organs. *Philosophical Transactions of the Royal Society*, 200(A), 1–66.
- Perez-Elizalde, S., Cerón-Rojas, J. J., Crossa, J., Fleury, D., & Alvarado, G. (2014). Rindsel: An R package for phenotypic and molecular selection indices used in plant breeding. In D. Fleury & R. Whitford (Eds.), *Crop breeding. Methods in molecular biology (methods and protocols)* (Volume 1145). Humana Press.
- Philipson, J., Banos, G., & Arnason, T. (1994). Present and future uses of methodology in dairy cattle selection index. *Journal of Dairy Science*, 77, 3252–3261. [https://doi.org/10.3168/jds.S0022-0302\(94\)77266-0](https://doi.org/10.3168/jds.S0022-0302(94)77266-0)
- Piepho, H. P., Möhring, J., Melchinger, A. E., & Büchse, A. (2008). BLUP for phenotypic selection in plant breeding and variety testing. *Euphytica*, 161, 209–228. <https://doi.org/10.1007/s10681-007-9449-8>
- Portes, J. V., Menezes, G. R. O., Macneil, M. D., Da Silva, L. O. C., Gondo, A., & Braccini Neto, J. (2020). Selection indices for Nelore production systems in the Brazilian Cerrado. *Livestock Science*, 242, 1–8. <https://doi.org/10.1016/j.livsci.2020.104309>
- Ralph, J., O'Neill, R., & Winton, J. (2015). *A practical introduction to index numbers*. John Wiley & Sons.
- Rencher, A. C. (2002). *Methods of multivariate analysis* (2nd ed.). Wiley.
- Rencher, A. C., & Schaalje, G. B. (2008). *Linear models in statistics* (2nd ed.). John Wiley and Sons.
- Robinson, G. K. (1991). The BLUP is a good thing. *Statistical Science*, 6(1), 15–51.
- Rausand, M., & Høyland, A. (2004). *System reliability theory: Models, statistical methods, and applications* (2nd ed.). Wiley.

- Searle, S. R., Casella, G., & McCulloch, C. E. (2006). *Variance components*. John Wiley and Sons.
- Smith, C. (1983). Effects of changes in economic weights on the efficiency of index selection. *Journal of Animal Science*, *56*(5), 1057–1064. <https://doi.org/10.2527/jas1983.5651057x>
- Smith, H. F. (1936). A discriminant function for plant selection. *Annals of Eugenics*, *7*, 240–250. <https://doi.org/10.1111/j.1469-1809.1936.tb02143.x>
- Su, G., Christensen, O. F., Ostensen, T., Henryon, M., & Lund, M. S. (2012). Estimating additive and non-additive genetic variances and predicting genetic merits using genome-wide dense single nucleotide polymorphism markers. *PLOS ONE*, *7*(9), e45293. <https://doi.org/10.1371/journal.pone.0045293>
- Tallis, G. M. (1962). A selection index for optimum genotype. *Biometrics*, *18*, 120–122. <https://doi.org/10.2307/2527716>
- Tallis, G. M. (1985). Constrained selection. *The Japanese Journals of Genetics*, *60*, 151–155. <https://doi.org/10.1266/jjg.60.151>
- Togashi, K., Lin, C. Y., & Yamazaki, T. (2011). The efficiency of genome-wide selection for genetic improvement of net merit. *Journal of Animal Science*, *89*, 2972–2980. <https://doi.org/10.2527/jas.2009-2606>
- Vandepitte, W. M., & Hazel, L. N. (1977). The effect of errors in the economic weights on the accuracy of selection indexes. *Annales de Genetique et de Selection Animale*, *9*, 87–103.
- Vanraden, P. M. (2008). Efficient methods to compute genomic predictions. *Journal of Dairy Science*, *91*, 4414–4423. <https://doi.org/10.3168/jds.2007-0980>
- Wilhelm, S., & Manjunath, B. G. (2010). tmvtnorm: A package for the truncated multivariate normal distribution. *The R Journal*, *2*(1), 24–28. <https://doi.org/10.32614/RJ-2010-005>
- Williams, J. S. (1962). The evaluation of a selection index. *Biometrics*, *18*, 375–393. <https://doi.org/10.2307/2527479>
- Wilms, I., & Croux, C. (2016). Robust sparse canonical correlation analysis. *BMC Systems Biology*, *10*, 72. <https://doi.org/10.1186/s12918-016-0317-9>
- Xie, C., & Xu, S. (1997). Restricted multi-stage selection indices. *Genetics Selection Evolution*, *29*(2), 193–203. <https://doi.org/10.1186/1297-9686-29-2-193>
- Xie, C., & Xu, S. (1998). Efficiency of multi-stage marker-assisted selection in the improvement of multiple quantitative traits. *Heredity*, *80*, 489–498. <https://doi.org/10.1046/j.1365-2540.1998.00308.x>
- Xu, S. (2003). Advanced statistical methods for estimating genetic variances in plants. *Plant Breeding Reviews*, *22*, 113–163.
- Xu, S., & Muir, W. M. (1992). Selection index updating. *Theoretical and Applied Genetics*, *83*, 451–458. <https://doi.org/10.1007/BF00226533>
- Yamada, Y., Yokouchi, K., & Nishida, A. (1975). Selection index when genetic gains of individual traits are of primary concern. *The Japanese Journal of Genetics*, *50*(1), 33–41. <https://doi.org/10.1266/jjg.50.33>
- Young, S. S. Y. (1964). Multi-stage selection for genetic gain. *Heredity*, *19*, 131–145. <https://doi.org/10.1038/hdy.1964.11>
- Young, S. S. Y. (1961). A further examination of the relative efficiency of three methods of selection for genetic gains under less-restricted conditions. *Genetics Research*, *2*, 106–121. <https://doi.org/10.1017/S0016672300000598>
- Zhang, W., & Smith, C. (1992). Computer simulation of marker-assisted selection utilizing linkage disequilibrium. *Theoretical and Applied Genetics*, *83*, 813–820. <https://doi.org/10.1007/BF00226702>
- Zhang, W., & Smith, C. (1993). Simulation of marker-assisted selection utilizing linkage disequilibrium: The effects of several additional factors. *Theoretical and Applied Genetics*, *86*, 492–496. <https://doi.org/10.1007/BF00838565>

How to cite this article: Cerón-Rojas, J. J., & Crossa, J. (2022). The statistical theory of linear selection indices from phenotypic to genomic selection. *Crop Science*, 1–27. <https://doi.org/10.1002/csc2.20676>