



## Economic weights for restriction of selection index as optimal strategy for combining multiple traits

E. Mancin,\*<sup>ORCID</sup> R. Mantovani,<sup>ORCID</sup> B. Tuliozi, and C. Sartori<sup>ORCID</sup>

Department of Agronomy, Food, Natural Resources, Animals and Environment, University of Padova, Viale dell'Università, 16, 35020 Legnaro, PD, Italy

### ABSTRACT

Presently, selection indices combine EBV belonging to many phenotypes accounting for production and functional traits. In this situation, ensuring positive genetic trends becomes difficult, especially when an antagonistic genetic relationship between traits exists. For this reason, guaranteeing zero genetic progress for some traits could be advantageous, as it would allow maximization of the productive characteristics while maintaining steady other antagonistic traits. This approach can also be useful for traits with an intermediate optimum, such as type traits, or for phenotypes for which it is difficult to quantify a specific economic weight. On this point, deriving the economic weight for a restricted selected index can be an optimal strategy. In this article, we provide a brief demonstration of how to derive economic weights by the restriction of some antagonistic traits. This procedure is then validated in a real-life situation using a dual-purpose cattle breed, due to the presence of more than one antagonistic trait with respect to milk yield. The R code is also provided to apply this procedure. The practical application to an actual situation demonstrates that the method produces feasible solutions in situations where more antagonistic traits are considered for selection, and prevents the detriment of some functional traits. Finally, we demonstrated how to obtain the desired restricted genetic progress when heterogeneous sources of information are present and its influence on selection response.

**Key words:** selection index, antagonistic traits, breeding goals

### INTRODUCTION

The aggregate economic selection index is commonly used to classify animals based on a combination of the EBV for traits of interest weighted by their respective

economic values (Hazel, 1943). Weights of combined traits are derived by considering the profit produced or the expectation of some functional aspects in a breed. However, the main difficulty in including many traits in the selection index is to ensure positive genetic progress for all traits, as many of them could be negatively correlated. For example, beef conformation in dual-purpose breeds is considered a productive trait alongside milk production (Cunningham and McClintock, 1974), but an antagonistic genetic relationship among the 2 traits exists (Fuerst-Waltl et al., 2016; Mazza et al., 2016; Sartori et al., 2018). Additionally, not all traits require positive genetic progress, and for some traits it is necessary to guarantee their maintenance (i.e., no genetic progress) over time (Miglior et al., 2005). For example, functional traits in dairy cattle, such as SCS, fertility, longevity, and some morphological traits, do not necessarily need improvement but just stability over time (Mancin et al., 2021). Furthermore, non-worsening genetic progress over time can be particularly suitable for local breeds, where the addition to the selection index of some non-tangible economical traits is sometimes required, as with breed-typical traits (Mancin et al., 2021).

In the literature, several methods to restrict the genetic gain of target traits to zero have been proposed (e.g., Kempthorne and Nordskog, 1959; Xie and Xu, 1997). Limiting the genetic progress to zero for some traits can be beneficial in all situations and allows the maximization of productive traits while maintaining functional and morphological traits intact.

This study aims to analytically demonstrate how to derive economic weights to be used in selection indexes to ensure zero genetic progress for certain traits. With this in mind, we analyzed the theoretical background behind both univariate and multivariate genetic progress, to demonstrate our procedure to restrict selection for some traits (i.e., ensure genetic progress equal to zero). Furthermore, we also provide a small example using genetic and phenotypic correlations in a dual-purpose cattle population, accounting for restriction on 2 traits antagonistic to milk yield, while maximizing

Received March 15, 2022.

Accepted August 1, 2022.

\*Corresponding author: [enrico.mancin@phd.unipd.it](mailto:enrico.mancin@phd.unipd.it)

the genetic gain for all considered traits. The R code to run all the requested steps is also provided in the article Appendix. We also showed the influence of heterogeneous information sources on the restricted selection indexes and how to consider this information during the selection process. A new approach for dealing with unequal accuracies across individuals is also proposed.

## METHODS

This study did not require any specific ethics permit. The cattle sampled belonged to commercial private herds and were not experimentally manipulated. Samples were collected by technicians from the Breeders Association of Alpine Grey breed.

### Theory Behind Univariate and Multivariate Selection Index and Economic Weights

The theory of deriving the economic weights in the restriction selection index has already been proposed in many other studies (Brascamp, 1984; Weller, 1994; Dekkers, 2003). For instance, Weller (1994) proposed a similar derivation to ours, which was generalized for traits that are under a specific constraint, not only under restriction, and Brascamp (1984) obtained the restricted selection weights by solving the equivalence  $\mathbf{a} = -\boldsymbol{\lambda}$ , which was, in turn, proposed in (Cunningham et al., 1970); see later. A completely different approach was proposed by (Dekkers, 2003), using an iterative procedure for obtaining economic weights in a nonlinear profit function.

However, in this article, economic restricted economic weights are derived in 2 different ways: (1) by back-solving the economic weights from the Hazel (1943) selection index and (2) by adapting the genetic selection index equation proposed in (Lin, 1990) to efficiently account for the different source of information. In addition, the theory behind restricted economic weights was demonstrated by starting from the classic univariate response to selection and then passing through the multivariate one; finally, the equations to obtain restricted or nonrestricted economic weights were demonstrated. At last, the influence of heterogeneous information in the restricted selection index is also presented and discussed. On this basis, a naive methodology that permits the user to account for different sources of information in the restricted genetic selection index was demonstrated.

Indeed, we can consider the response to selection ( $R$ ) as the difference between the mean phenotypic value of the progeny of selected parents and the mean phe-

notypic value of the whole parental generation before selection (Falconer and Mackay, 1996). The selection response for one trait of interest has been represented in many formulas. For consistency with what will be reported later, the formula presented in Harris (1964) is as follows:

$$R = \beta_{\mathbf{g}\mathbf{y}} \Delta_I, \quad [1]$$

where  $\beta_{\mathbf{g}\mathbf{y}}$  is the linear regression coefficient of the offspring genetic value  $\mathbf{g}$  on selected parents' phenotype ( $\mathbf{y}$ ) and  $\Delta_I$  stands for the selection differential, representing the deviation of selected parents' mean from the population mean. Replacing  $\beta_{\mathbf{g}\mathbf{y}}$  with  $\frac{\text{Cov}(\mathbf{g},\mathbf{y})}{\sigma_{\mathbf{y}}^2}$  and

then grouping  $\Delta_I/\sigma_{\mathbf{y}}$  into  $i$  (i.e., the intensity of selection in standardized units), we obtain (Falconer and Mackay, 1996)

$$R = \frac{\text{Cov}(\mathbf{g},\mathbf{y})}{\sigma_{\mathbf{y}}} i = r \sigma_{\mathbf{g}} i, \quad [2]$$

where, under the assumption of a normal distribution of the phenotype  $\mathbf{y}$ , the integration of  $i$  or  $\Delta_I/\sigma_{\mathbf{y}}$  provides the proportion of selected individuals (the parents) and  $r$  represents the accuracy  $\left[ \frac{\text{Cov}(\mathbf{g},\mathbf{y})}{\sigma_{\mathbf{y}}\sigma_{\mathbf{g}}} \right]$ . In a

multiple trait framework, the parents' phenotype ( $\mathbf{y}$ ) is represented in multivariate form by the phenotypic (co)variances  $\mathbf{P} = (\mathbf{y}'\mathbf{y})$ , and  $\mathbf{g}$  by the genetic (co)variances  $\mathbf{G} = (\mathbf{g}'\mathbf{g})$ . The maximization of the genetic progress for multiple traits is based on the indirect selection for an unobservable variable ( $\mathbf{H}$ ), by the truncated selection of an observable variable ( $\mathbf{I}$ ; Harris, 1964). The observable variable  $\mathbf{I}$  is the linear combination of ( $\mathbf{I} = \mathbf{b}'\mathbf{y}$ ), where vector  $\mathbf{b}$  represents the relative emphasis of each phenotype needed to achieve the desired goal. Therefore, the variance of selection index  $\mathbf{I}$  ( $\sigma_j$ ) can be expressed as  $\text{Var}(\mathbf{b}'\mathbf{y})$  or, in matrix form, as  $\mathbf{b}'\mathbf{P}\mathbf{b}$ . The role of vector  $\mathbf{b}$  is to maximize the genetic progress through maximization of the covariance between  $\mathbf{H}$  and  $\mathbf{I}$ ,  $\text{Cov}(\mathbf{H},\mathbf{I})$  (Harris, 1964). Regression of  $\mathbf{H}$  on  $\mathbf{I}$  is linear for any set of  $\mathbf{b}$  values. When no economic weight is applied,  $\mathbf{H}$  is equal to the vector of the true breeding values  $\mathbf{H} = \mathbf{g}'$ . Given that  $\text{Cov}(\mathbf{G},\mathbf{P}) = \mathbf{G}$ , the covariance between  $\mathbf{H}$  and  $\mathbf{I}$ ,  $\text{Cov}(\mathbf{y}'\mathbf{b},\mathbf{g}'\mathbf{y})$ , becomes  $\mathbf{G}\mathbf{b}'$ . Note that proof of  $\text{Cov}(\mathbf{G},\mathbf{P}) = \mathbf{G}$  can be demonstrated by setting the 2 different selection response formulas in Falconer and Mackay (1996) equal.

Considering  $i$ ,  $\mathbf{P}$ , and  $\mathbf{G}$  as constants, maximizing genetic progress is equal to maximizing  $\text{Cov}(\mathbf{H},\mathbf{I})$ , which

means minimizing the squared differences between  $\mathbf{H}$  and  $\mathbf{I}$ ; that is,  $(E[(\mathbf{H} - \mathbf{I})^2])$  (Céron-Rojas and Crossa, 2018), where  $E$  is the expected value. Setting the partial derivative ( $\partial$ ) of  $\mathbf{b}$  equal to zero,

$$\frac{\partial}{\partial \mathbf{b}} E[(\mathbf{H} - \mathbf{I})^2] = 0, \quad [3]$$

by changing  $\mathbf{H}$  and  $\mathbf{I}$  in [3], the formula becomes

$$\frac{\partial}{\partial \mathbf{b}} (\mathbf{G} + \mathbf{b}'\mathbf{P}\mathbf{b} - 2\mathbf{G}\mathbf{b}') = 0. \quad [4]$$

Deriving the formula above, we have

$$2\mathbf{b}\mathbf{P} - 2\mathbf{G} = 0; \quad [4a]$$

$$\mathbf{b} = \mathbf{P}^{-1}\mathbf{G}, \quad [4b]$$

with  $\mathbf{b}$  being the vector that maximized the genetic progress. Replacing  $\text{Cov}(\mathbf{G}, \mathbf{I})$  with  $\mathbf{b}'\mathbf{G}$  and  $\sigma_i$  with  $(\mathbf{b}'\mathbf{P}\mathbf{b})^{1/2}$  in formula [2], as in (Weller, 1994), we obtain

$$\mathbf{R} = \mathbf{b}'\mathbf{G}(\mathbf{b}'\mathbf{P}\mathbf{b})^{-1/2}\mathbf{i}. \quad [5]$$

Then, if we consider the different economic importance of traits, it becomes necessary to introduce a new element in equation [3]; that is, a vector ( $\mathbf{a}$ ) representing the weights of each trait. In accordance with Wolfová et al. (2001), vector  $\mathbf{a}$  can be standardized by dividing the economic weight of the traits ( $t$ ) by the respective genetic standard deviation ( $std$ ) as  $\mathbf{a}_{t\_std} = \mathbf{a}_t/\sigma_{at}$ .

From here when we refer to  $\mathbf{a}$ , we refer to the vector  $\mathbf{a}_{t\_std}$  (already standardized by trait's genetic standard deviations). Thus  $\mathbf{H}$  can be described as a function  $\mathbf{H} = \mathbf{a}'\mathbf{g}$ , called the linear aggregate genotype (Hazel, 1943), which is the function of the additive genetic values of the traits of interest, each having a specific economic weight. Therefore, in this form, the function  $\mathbf{H}$  includes the expected genetic progress for each trait. When economic weights are used, the variance of  $\mathbf{H}$  becomes  $\text{Cov}(\mathbf{a}'\mathbf{g}, \mathbf{a}'\mathbf{g})$ , which is equal to  $\mathbf{a}'\mathbf{G}\mathbf{a}$ , and correlation between  $\mathbf{H}$  and  $\mathbf{I}$   $\text{Cov}(\mathbf{b}'\mathbf{y}, \mathbf{a}'\mathbf{g})$  becomes  $\mathbf{a}'\mathbf{G}\mathbf{b}$ . Similarly, aiming to minimize the error of  $E[(\mathbf{H} - \mathbf{I})^2]$ , which is equal to

$$\frac{\partial}{\partial \mathbf{b}} E[(\mathbf{H} - \mathbf{I})^2] = 0,$$

equation [4a] becomes

$$\mathbf{a}'\mathbf{G}\mathbf{a} + \mathbf{b}'\mathbf{P}\mathbf{b} - 2\mathbf{a}'\mathbf{G}\mathbf{b} = 0. \quad [6]$$

After some math, this becomes

$$\mathbf{b} = \mathbf{P}^{-1}\mathbf{G}\mathbf{a}. \quad [7]$$

Thus, response to selection ( $\mathbf{R}$ ), as in Weller (1994), becomes equal to

$$\mathbf{R} = \mathbf{a}'\mathbf{G}\mathbf{b}(\mathbf{b}'\mathbf{P}\mathbf{b})^{-\frac{1}{2}}. \quad [8]$$

### Restricted Selection Index

A restricted selection index is adopted to preserve the genetic value of specific phenotypes over time; that is, setting the genetic progress of these traits to 0. Three different approaches have been developed in literature (Kempthorne and Nordskog, 1959; Tallis, 1962; Cunningham et al., 1970), all leading to the same results. However, Tallis (1962) and further integration given by Céron-Rojas and Crossa (2018) are the most intuitive. The basic idea of these 2 studies is that to maintain the genetic progress at zero is equivalent to assuming a null correlation between selection index and response to selection for the traits considered [i.e.,  $\text{Cov}(\mathbf{H}, \mathbf{I}) = 0$ ]. Traits under constraint are collected in matrix  $\mathbf{C}$ , where  $\text{Cov}(\mathbf{C}, \mathbf{I}) = 0$ , which means that  $\mathbf{C}\mathbf{b}' = 0$ .  $\mathbf{C}$  is constructed as  $\mathbf{C}' = \mathbf{U}'\mathbf{G}$ , where  $\mathbf{U}$  is an incidence matrix including as many 1's as the number of traits that have to be restricted, and 0's for the other traits. In this situation, the objectives of the selection index are two-fold: to minimize the error  $E[(\mathbf{H} - \mathbf{I})^2]$ , and to ensure null genetic progress for the traits of interest; that is,  $\text{Cov}(\mathbf{C}, \mathbf{I}) = 0$ . To solve both objectives, it becomes necessary to apply the Lagrange multipliers that allow maximizing (or minimizing) the value of a given function  $f(x, y, \dots)$  under another function of restriction,  $g(x, y, \dots) = c$ , as follows:

$$\nabla f(x_0, y_0) = \lambda \nabla g(x_0, y_0). \quad [9a]$$

It can be rearranged thus:

$$\mathcal{L}(x, y, \lambda) = f(x, y) - \lambda(g(x, y) - c), \quad [9b]$$

where  $\lambda$  is the Lagrange multiplier. Thus, considering  $E[(\mathbf{H} - \mathbf{I})^2]$  as  $f(x, y)$  and  $\text{Cov}(\mathbf{C}, \mathbf{I}) - 0$  as  $g(x, y) - c$ , equation [9a, b] becomes

$$\mathcal{L}(\mathbf{b}, \lambda) = E[(\mathbf{H} - \mathbf{I})^2] - 2\lambda'(\text{Cov}(\mathbf{C}, \mathbf{I}) - 0). \quad [10]$$

As demonstrated above,  $E[(\mathbf{H} - \mathbf{I})^2]$  is equivalent to  $\mathbf{a}'\mathbf{G}\mathbf{a} + \mathbf{b}'\mathbf{P}\mathbf{b} - 2\mathbf{a}'\mathbf{G}\mathbf{b}$ , and  $\text{Cov}(\mathbf{C}, \mathbf{I})$  is equal to  $\mathbf{C}\mathbf{b}'$ :

$$\mathcal{L}(\mathbf{b}, \lambda) = \mathbf{a}'\mathbf{G}\mathbf{a} + \mathbf{b}'\mathbf{P}\mathbf{b} - 2\mathbf{a}'\mathbf{G}\mathbf{b} - 2\lambda'(\mathbf{C}\mathbf{b}' - 0). \quad [11]$$

Setting the partial derivative of  $\mathbf{b}$  and  $\lambda$  equal to zero, we obtain

$$\frac{\partial}{\partial \mathbf{b}, \lambda} \mathbf{a}'\mathbf{G}\mathbf{a} + \mathbf{b}'\mathbf{P}\mathbf{b} - 2\mathbf{a}'\mathbf{G}\mathbf{b} + 2\lambda'\mathbf{C}\mathbf{b} - 0 = 0. \quad [12]$$

The derivative results from  $\mathbf{b}$  and  $\lambda$  of equation [12] are as follows:

$$\frac{\partial}{\partial \mathbf{b}} \mathbf{P}\mathbf{b} + \mathbf{C}\lambda - \mathbf{G}\mathbf{a} = 0, \quad [13]$$

$$\frac{\partial}{\partial \lambda} \mathbf{C}'\mathbf{b} = 0. \quad [13a]$$

In matrix notation, equations [13] and [13a] can be represented as

$$\begin{bmatrix} 0 & \mathbf{C}' \\ \mathbf{C} & \mathbf{P} \end{bmatrix} \begin{bmatrix} \lambda \\ \mathbf{b}' \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{G}\mathbf{a} \end{bmatrix}. \quad [14]$$

The solution of the matrix

$$\begin{bmatrix} 0 & \mathbf{C}' \\ \mathbf{C} & \mathbf{P} \end{bmatrix}$$

must be obtained by inversion (Céron-Rojas and Crosa, 2018); that is,

$$\begin{bmatrix} 0 & \mathbf{C}' \\ \mathbf{C} & \mathbf{P} \end{bmatrix}^{-1} = \begin{bmatrix} (-\mathbf{C}\mathbf{P}^{-1}\mathbf{C})^{-1} & (\mathbf{C}\mathbf{P}^{-1}\mathbf{C})^{-1}\mathbf{C}'\mathbf{P}^{-1} \\ \mathbf{P}^{-1}\mathbf{C}(-\mathbf{C}\mathbf{P}^{-1}\mathbf{C})^{-1} & -\mathbf{P}^{-1}\mathbf{C}(-\mathbf{C}\mathbf{P}^{-1}\mathbf{C})^{-1}\mathbf{C}'\mathbf{P}^{-1} + \mathbf{P}^{-1} \end{bmatrix}, \quad [15]$$

and therefore

$$\begin{bmatrix} \lambda \\ \mathbf{b} \end{bmatrix} = \begin{bmatrix} (-\mathbf{C}\mathbf{P}^{-1}\mathbf{C})^{-1} & (-\mathbf{C}\mathbf{P}^{-1}\mathbf{C})^{-1}\mathbf{C}'\mathbf{P}^{-1} \\ \mathbf{P}^{-1}\mathbf{C}(-\mathbf{C}\mathbf{P}^{-1}\mathbf{C})^{-1} & -\mathbf{P}^{-1}\mathbf{C}(-\mathbf{C}\mathbf{P}^{-1}\mathbf{C})^{-1}\mathbf{C}'\mathbf{P}^{-1} + \mathbf{P}^{-1} \end{bmatrix} \begin{bmatrix} 0 \\ \mathbf{G}\mathbf{a} \end{bmatrix}. \quad [15a]$$

Solving [15a] for  $\mathbf{b}$  could be considered as the equation that maximizes the function under the  $\mathbf{C}'\mathbf{b}$  restriction:

$$\mathbf{b} = 0 + \mathbf{G}\mathbf{a}(-\mathbf{P}^{-1}\mathbf{C}(-\mathbf{C}\mathbf{P}^{-1}\mathbf{C})^{-1}\mathbf{C}'\mathbf{P}^{-1} + \mathbf{P}^{-1}). \quad [16]$$

Gathering by  $\mathbf{P}^{-1}$ ,

$$\mathbf{b} = \mathbf{P}^{-1}\mathbf{G}\mathbf{a}(-\mathbf{P}^{-1}\mathbf{C}(-\mathbf{C}\mathbf{P}^{-1}\mathbf{C})^{-1}\mathbf{C}' + \mathbf{I}), \quad [16a]$$

and replacing  $(-\mathbf{P}^{-1}\mathbf{C}(-\mathbf{C}\mathbf{P}^{-1}\mathbf{C})^{-1}\mathbf{C}' + \mathbf{I})$  with matrix  $\mathbf{K}$ , we obtain

$$\mathbf{b} = \mathbf{K}\mathbf{P}^{-1}\mathbf{G}\mathbf{a}. \quad [17]$$

For clarity, the vector  $\mathbf{b}$  identified in equation [17] is called  $\mathbf{b}_r$  to distinguish it from vector  $\mathbf{b}$  (the vector that maximized  $E[(\mathbf{H} - \mathbf{I})^2]$ ); therefore,

$$\mathbf{b}_r = \mathbf{K}\mathbf{P}^{-1}\mathbf{G}\mathbf{a}. \quad [18]$$

Because  $\mathbf{b} = \mathbf{P}^{-1}\mathbf{G}\mathbf{a}$ , equation [16] can be written as  $\mathbf{b}_r = \mathbf{K}\mathbf{b}$ , where matrix  $\mathbf{K}$  is a projection matrix applying a linear transformation on  $\mathbf{b}$  to reduce the space in which this vector is projected. The genetic progress ( $\mathbf{R}$ ) is then obtained by replacing  $\mathbf{b}$  with  $\mathbf{b}_r$  in formula [8], that is  $\mathbf{b}_r = \mathbf{P}^{-1}\mathbf{G}\mathbf{a}$ , and then using the new weights in formula [5]:

$$\mathbf{R} = \mathbf{b}'_r\mathbf{G}(\mathbf{b}'_r\mathbf{P}\mathbf{b}_r)^{-\frac{1}{2}} \mathbf{i}. \quad [19]$$

As mentioned at the beginning of this section, the theory behind multivariate response to selection has been extensively reported in the literature (Brascamp, 1984; Gibson, 1992; Weller, 1994; Dekkers, 2003). In this study, we analytically demonstrate how it is possible to obtain the economic weights (reported in vector  $\mathbf{a}$ ) when applying the restriction to variation for some traits in the selection indices.

This passage is not described in the literature but has considerable practical importance.

The new economic weights can be obtained by considering equations [19] and [8] equal. Thus, we obtain  $\mathbf{b}_r = \mathbf{K}\mathbf{b}$ , and, inverting formula [9b], we get  $\mathbf{a} = \mathbf{G}^{-1}\mathbf{P}\mathbf{b}$ . Then the values of new economic weights, which we call  $\mathbf{a}_r$ , are obtained substituting  $\mathbf{b}$  with  $\mathbf{K}\mathbf{b}$  in the previous equations:

$$\mathbf{a}_r = \mathbf{K}\mathbf{G}^{-1}\mathbf{P}\mathbf{b}_r. \quad [20]$$

The term  $\mathbf{a}_r$  represents the vector of the new weights to provide the traits to realize a null genetic progress

for some target traits and a positive genetic gain for the others. It is possible to express the new economic weights on a scale from 0 to 1 by dividing each weight by the sum of the absolute values of the others.

Finally, the final expected genetic progress can be obtained by using the new vector of economic weights in the general formula for genetic progress [5], but dividing  $\mathbf{a}_r$  for the genetic standard deviation of each trait  $t$  (i.e.,  $\mathbf{a}_{rt}/\sigma_{at}$ ; see Wolfová et al., 2001). In this way, the genetic progress with or without restriction can be compared.

A similar equation for obtaining the “pseudo” economic weights when all traits have a specific constraint (e.g., maximizing fat and protein production under the constraint that genetic progress for these traits is equal) was proposed by Weller (1994) using a different approach. However, both procedures to calculate the restricted selection response and its economic weights (the one proposed here and the one in Weller, 1994), albeit approximate, can be helpful during the decision-making process of how to drive the genetic improvement in a breed under selection for traits of interest. For example, the weights found in equation [20] can provide a rough estimate of the emphasis to be placed on each trait, which would allow evaluation of how and whether to move to a new breeding plan or evaluation of the inclusion (or not) of new phenotypes in the selection index.

However, the procedure proposed up to this point is not suitable for properly ranking the animals in some routine evaluations, as the  $\mathbf{b}_r$  or  $\mathbf{a}_r$  proposed respectively in equations [18] and [20] do not account for unequal accuracies across animals—a common situation in dairy cattle breeding schemes. In the next section, we will demonstrate that it is necessary to apply some adjustments to the equation of  $\mathbf{b}_r$  or  $\mathbf{a}_r$  before using it to obtain a selection index restricted for some traits. Note that the  $\mathbf{b}_r$  value is used to rank animals in the restricted phenotypic selection index (i.e., a combination of  $\mathbf{b}$  obtained in equation [17] with  $\mathbf{y}$ ), whereas  $\mathbf{a}_r$  can be used in the restricted genetic selection index (i.e., a combination of  $\mathbf{a}$  obtained in equation [20] with the estimated genetic values,  $\hat{\mathbf{g}}$ ).

### Restricted Selection Index in the Presence of Heterogeneous Information

As mentioned before, the main drawback of equations [1] and [19] is that they assume that all individuals share the same amount of information (i.e., the observed phenotype) and the same selective pressure. However, this is an oversimplification of what happens in livestock farming systems, where phenotypic information differs between sexes, and, consequently,

accuracy may also change. Therefore, considering 2 pathways of selection (males and females), this discrepancy, equation [1] becomes [2], and considering unequal accuracy and selection intensity among sexes, the selection response becomes

$$R = r_{H,I_s} \sigma_H i_s + r_{H,I_d} \sigma_H i_d \quad [21]$$

The subscripts  $s$  and  $d$  refer to sire and dam, respectively; in this equation, for a matter of simplicity, we considered only 2 groups of candidates, but they might be further divided into “sire of sire”; “sire of dam,” and so on. The term  $r_{H,I}$  represents the correlation coefficient, which is equivalent to

$$r_{H,I} = \frac{\text{Cov}(I, H)}{\sigma_H \sigma_I} = \frac{\sigma_I^2}{\sigma_H \sigma_I} = \frac{\sigma_I}{\sigma_H}.$$

From the multivariate response to selection expressed in equation [6], we know that  $\sigma_H = (\mathbf{a}'\mathbf{G}\mathbf{b})^{\frac{1}{2}}$  and  $\sigma_I = (\mathbf{b}'\mathbf{P}\mathbf{b})^{\frac{1}{2}}$ ; thus, equation [21] becomes

$$R = (\mathbf{a}'_s \mathbf{G} \mathbf{b}_s)^{\frac{1}{2}} (\mathbf{b}'_s \mathbf{P}_s \mathbf{b}_s)^{-\frac{1}{2}} (\mathbf{a}'_s \mathbf{G} \mathbf{b}_s)^{\frac{1}{2}} i_s + (\mathbf{a}'_d \mathbf{G} \mathbf{b}_d)^{\frac{1}{2}} (\mathbf{b}'_d \mathbf{P}_d \mathbf{b}_d)^{-\frac{1}{2}} (\mathbf{a}'_d \mathbf{G} \mathbf{b}_d)^{\frac{1}{2}} i_d, \quad [22]$$

where subscripts  $s$  and  $d$  refer to sires and dams, respectively. This formulation can be further generalized, at the cost of higher complexity, considering multiple selection pathways (Rendel and Robertson, 1950; Hill, 1974).

Note that equation [22] can be rearranged in equation [8] by multiplying the  $2 \mathbf{a}'\mathbf{G}\mathbf{b}$ . Therefore, to ensure the desired zero genetic progress for the target traits, it becomes necessary to account for these heterogeneous sources of information ( $\mathbf{P}_s$ ,  $\mathbf{P}_d$ ) and then recalculate the restricted  $\mathbf{b}$  for each cohort ( $\mathbf{b}_d$ ,  $\mathbf{b}_s$ ). The different selection intensities ( $i_s$ ,  $i_d$ ) do not affect the estimates of  $\mathbf{a}_r$  and  $\mathbf{b}_r$ , but they do affect the predicted  $R$ . The effects of using homogeneous or heterogeneous  $\mathbf{b}_r$  in the restricted selection index are shown in a simulation reported in Supplemental File S1 ([https://github.com/enmancio/Restricted\\_Selection\\_Index](https://github.com/enmancio/Restricted_Selection_Index)). We simulated the breeding scheme for Gray Alpine cattle using the AlphaSimR program (Gaynor et al., 2021; see Supplemental File S1).

A more effective method to account for the different heterogeneous sources of information among individuals is to use a genetic index as an alternative to the phenotypic selection index. The genetic index calculates the overall genetic merit of each individual as the com-

bination of the individual estimated breeding values ( $\hat{\mathbf{g}}$ ) for each trait and the vector  $\mathbf{b}$ :

$$\mathbf{I} = \mathbf{b}'\hat{\mathbf{g}}, \tag{23}$$

where  $\hat{\mathbf{g}}$  is estimated as best linear unbiased prediction using mixed model equations (Henderson, 1975). In the genetic selection index, it is possible to obtain the desired genetic progress by simply combining the vector of restricted economic weights ( $\mathbf{a}$ ) with  $\hat{\mathbf{g}}$ , because (1)  $\mathbf{b} = \mathbf{a}$  (see further), and (2) heterogeneous sources of information are implicitly taken into account in  $\hat{\mathbf{g}}$ .

The first point can be explained by maximizing the relationship  $E[(\mathbf{H} - \mathbf{I})^2] = 0$ , where  $\mathbf{H}$  is equal to  $\text{var}(\mathbf{a}'\mathbf{g})$  and  $\mathbf{I}$  is equal to  $\text{var}(\mathbf{b}'\hat{\mathbf{g}})$ . Therefore, this maximization is

$$\text{var}(\mathbf{a}'\mathbf{g}) - 2\text{cov}(\mathbf{a}'\mathbf{g}, \mathbf{b}'\hat{\mathbf{g}}) + \text{var}(\mathbf{b}'\hat{\mathbf{g}}) = 0. \tag{24}$$

We know from (Henderson, 1975) that  $\text{cov}(\mathbf{g}, \hat{\mathbf{g}})$  is equal to  $\text{var}(\hat{\mathbf{g}})$ ; thus equation [24] becomes

$$\begin{aligned} \text{var}(\mathbf{a}'\mathbf{g}) - 2\text{var}(\mathbf{a}'\hat{\mathbf{g}}, \mathbf{b}'\hat{\mathbf{g}}) + \text{var}(\mathbf{b}'\hat{\mathbf{g}}) \\ = \mathbf{a}'\text{var}(\mathbf{g})\mathbf{a} - 2\mathbf{a}'\text{var}(\hat{\mathbf{g}})\mathbf{b} + \mathbf{b}'\text{var}(\hat{\mathbf{g}})\mathbf{b}. \end{aligned} \tag{25}$$

Setting the partial derivate for  $\mathbf{b}$ ,

$$\frac{\partial}{\partial \mathbf{b}} \mathbf{a}'\text{var}(\mathbf{g})\mathbf{a} - 2\mathbf{a}'\text{var}(\hat{\mathbf{g}})\mathbf{b} + \mathbf{b}'\text{var}(\hat{\mathbf{g}})\mathbf{b}, \tag{26a}$$

$$-2\mathbf{a}'\text{var}(\hat{\mathbf{g}}) + 2\mathbf{b}'\text{var}(\hat{\mathbf{g}}) = 0, \tag{26b}$$

$$\mathbf{b}' = \mathbf{a}'. \tag{26c}$$

Thus,

$$\mathbf{I} = \mathbf{b}'\hat{\mathbf{g}} = \mathbf{a}'\hat{\mathbf{g}}. \tag{27}$$

The other advantage of the genetic restricted selection index is that it is not necessary to calculate the weight for each cohort because different family structures and availability of information have already been accounted for in  $\hat{\mathbf{g}}$  (Schneeberger et al., 1992). For this reason, in the genetic index this is equivalent to considering each animal (1 to  $n$ ) as a different cohort, with some of them sharing the same selection intensity (i.e., sires and dams):

$$\begin{aligned} R = \left( r_{H,I s_1} \sigma_H + \dots + r_{H,I s_n} \sigma_H \right) i_s \\ + \dots \left( r_{H,I d_1} \sigma_H + \dots + r_{H,I d_n} \sigma_H \right) i_d, \end{aligned} \tag{28}$$

where  $s_1 \dots s_n$  are animals belonging to the sire's cohort and  $d_1 \dots d_n$  are animals belonging to the dam's cohort. In the genetic index, the term  $r_{H,I}$  is equal to

$$\begin{aligned} r_{H,I} &= \frac{\text{cov}(I, H)}{\sigma_H \sigma_I} = \frac{\text{cov}(\mathbf{a}'\mathbf{g}, \mathbf{a}'\hat{\mathbf{g}})}{\sqrt{\text{var}(\mathbf{a}'\mathbf{g})} \sqrt{\text{var}(\mathbf{b}'\hat{\mathbf{g}})}} \\ &= \frac{\text{cov}(\mathbf{a}'\mathbf{g}, \mathbf{a}'\hat{\mathbf{g}})}{\sqrt{\text{var}(\mathbf{a}'\mathbf{g})} \sqrt{\text{var}(\mathbf{a}'\hat{\mathbf{g}})}} = \frac{\sqrt{\text{var}(\mathbf{a}'\hat{\mathbf{g}})}}{\sqrt{\text{var}(\mathbf{a}'\mathbf{g})}}. \end{aligned} \tag{29}$$

We know from (Henderson, 1975) that  $\text{var}(\hat{\mathbf{g}}) = \text{var}(\mathbf{g}) - \text{PEV}$ , where PEV is the predicted error variance. Assuming that all animals of a population have a null inbreeding, PEV is equal to  $\mathbf{C}^{\text{aa}}$ ; thus, equation [29] becomes

$$r_{H,I} = \frac{\sqrt{\text{var}(\mathbf{a}\mathbf{g}) - \text{PEV}}}{\sqrt{\text{var}(\mathbf{a}\mathbf{g})}} = 1 - \frac{\sqrt{\text{PEV}}}{\sigma_H} = 1 - \frac{\sqrt{\mathbf{C}^{\text{aa}}}}{\sigma_H}, \tag{30}$$

where  $\mathbf{C}^{\text{aa}}$  is the diagonal of the inverse of the left-hand side corresponding to the genetic value. Note that  $\text{PEV} = \text{var}(\mathbf{a}'\hat{\mathbf{g}} - \mathbf{a}'\mathbf{g}) = \text{var}(\hat{\mathbf{g}} - \mathbf{g})$ .

Thus, equations [21] and then [28] become

$$\begin{aligned} R = \left( 1 - \frac{\sqrt{\mathbf{C}^{\text{aa}}_{s_1}}}{\sigma_H} \sigma_H + \dots + 1 - \frac{\sqrt{\mathbf{C}^{\text{aa}}_{s_n}}}{\sigma_H} \sigma_H \right) i_s \\ + \left( 1 - \frac{\sqrt{\mathbf{C}^{\text{aa}}_{d_1}}}{\sigma_H} \sigma_H + \dots + 1 - \frac{\sqrt{\mathbf{C}^{\text{aa}}_{d_n}}}{\sigma_H} \sigma_H \right) i_d. \end{aligned} \tag{31}$$

Interestingly, when there is no relationship between animals and each individual has its own observation, equation [30] becomes equal to equation [1], since

$$r_{H,I} = \frac{\sigma_I}{\sigma_H}.$$

Despite the genetic selection index accounting for differences in accuracy among candidates, the vector of the restricted economic weights in equation [20] cannot be directly integrated in [27], because  $\mathbf{a}_r$  is not a linear function (Dekkers, 2003). Therefore, considering a single matrix  $\mathbf{G}$  and  $\mathbf{P}$  on equation [11] may not guarantee the desired restricted genetic progress for the reasons mentioned above.

This situation can be addressed using 2 different approaches. The first approach, which is similar to that applied to obtain equation [22], is based on the

re-estimation of  $\mathbf{a}_r$  by using different  $\mathbf{P}$  and  $\mathbf{G}$  values in equation [20]; that is, as many as the groups of candidates to selection.

The second approach is based on adapting the equation proposed in Lin (1990), allowing a simple procedure to compute the restricted economic weights directly in the genetic selection index. That procedure was based on the idea of the minimization of the error between  $\mathbf{I}$  and  $\mathbf{I}^*$ , where  $\mathbf{I} = \mathbf{a}'\mathbf{g}$  and  $\mathbf{I}^* = \mathbf{a}_r'\mathbf{g}$ . Lin (1990) solved that minimization through a similar approach to that used in this manuscript, and obtained  $\mathbf{a}_r$  values that are equal to

$$\mathbf{a}_r = [\mathbf{I} - \mathbf{G}^{-1}\mathbf{C}'(\mathbf{C}\mathbf{G}^{-1}\mathbf{C}')^{-1}\mathbf{C}]\mathbf{a}, \quad [32]$$

where  $\mathbf{C} = (\mathbf{U}'\mathbf{G})$ . However, economic weights obtained in equation [32] cannot be applied in real-life scenarios, as Lin (1990) assumed that covariances among predicted genetic values ( $\hat{\mathbf{g}}$ ) are equal to the true ones ( $\mathbf{g}$ ). Indeed, Schneeberger et al. (1992) claim that the equation proposed in Lin (1990) “ignores the effect of differing accuracies of prediction of individual EBVs for each animal, particularly when animal models are fitted to large unbalanced field data sets” (p. 1). Therefore, that problem can be overcome by substituting the  $\mathbf{G}$  presented in equation [32] with its predicted value ( $\hat{\mathbf{G}}$ ) that represents the (co)variance of the  $\hat{\mathbf{g}}$ 's. The  $\hat{\mathbf{G}}$  can be obtained, again, by using Henderson's (1975) formulations:

$$\hat{\mathbf{G}} = \text{Var}(\mathbf{g}) - \mathbf{C}^{\text{aa}} = \mathbf{A} \otimes \mathbf{G} - \mathbf{C}^{\text{aa}}, \quad [33]$$

where  $\mathbf{A}$  is the pedigree relationship matrix (Henderson, 1975),  $\mathbf{G}$  is the genetic (co)variance matrix, and  $\otimes$  represents the Kronecker product. Note that in this case  $\mathbf{C}^{\text{aa}}$  represents the whole block of the inverse of the left-hand-side, not only its diagonal, representing the animal prediction error covariance.

Therefore we can say that replacing  $\hat{\mathbf{G}}$  with  $\mathbf{G}$  in equation [32] is possible to obtain the desired genetic gain even when accuracies differ among candidates. In fact, as for predicted genetic values, the predicted (co)

**Table 1.** Traits included in the example, with their phenotypic means (mean), coefficients of variation (CV), heritability ( $h^2$ ), and genetic variances ( $\sigma_a^2$ ) of the traits

Trait	Mean	CV	$\sigma_a^2$	$h^2$
Milk yield (kg)	16.300	0.329	2.210	0.219
Fat yield (kg)	0.621	0.339	2.600 <sup>1</sup>	0.178
Protein yield (kg)	0.564	0.304	1.895 <sup>1</sup>	0.125
SCS (points) <sup>2</sup>	2.333	0.788	0.379	0.133
Muscularity (points) <sup>2</sup>	27.141	0.222	9.144	0.328

<sup>1</sup>Variances have been multiplied by  $10^3$ .

<sup>2</sup>Traits to which a restriction to the genetic progress was applied.

variances matrix  $\hat{\mathbf{G}}$  considers each animal implicitly as a group of candidates to selection due to different accuracies. Despite this, that approach is not computationally straightforward, especially when many traits are considered. However, some approximations of predicted error variance can be used (Misztal and Wiggans, 1988; Hickey et al., 2009). The changes of the response to selection when  $\mathbf{G}$  is substituted with  $\hat{\mathbf{G}}$  are also reported in the simulation provided in Supplemental File S1.

## RESULTS

### Application

This study considered actual genetic and phenotypic correlations estimated in a local dual-purpose breed (Mancin et al., 2021; Table 1, Table 2). Traits accounted for were milk yield (**MY**, kg), fat yield (**FY**, kg), protein yield (**PY**, kg), SCS (points), and rear muscularity (**MUSC**; 1 to 5 points).

Milk yield (kg/d), FY (kg/d), PY (kg/d), and SCS were calculated from test-day data. Test-day records were collected regularly from 1997 to 2018, following the official Italian milk registration system, where data collection is done with an average interval of 4 wk between test-day collections. The SCS is obtained by normalizing the number of SCC (no./mL) as  $\text{SCS} = 3 + \log_2(\text{SCC}/10^5)$ , following (Ali and Shook, 1980). The MUSC trait is one of the morphological measurements (a total of 20 traits, others not included in this

**Table 2.** Genetic correlations (above diagonal) and phenotypic correlations (below diagonal) of the traits considered in the study

Trait	MY	FY	PY	SCS	MUSC
Milk yield (MY, kg)		0.758	0.845	0.069	-0.458
Fat yield (FY, kg)	0.768		0.824	0.067	-0.413
Protein yield (PY, kg)	0.905	0.766		0.088	-0.397
SCS (points)	-0.149	-0.080	-0.111		-0.156
Muscularity (MUSC, points)	-0.134	-0.079	-0.086	0.001	

**Table 3.** Traits included in the example, with their initial economic weights ( $a$ ) and the final economic weights ( $a_r$ ) applied after a restriction for the genetic progress of 2 traits and the corresponding response to selection ( $\Delta G$  and  $\Delta G_r$ ); for the purposes of our example, we assume that the sum of weights is 1

Trait	$a$	$\Delta G$	$a_r$	$\Delta G_r$
Milk yield (kg)	0.000	0.844	0.000	0.513
Fat yield (kg)	0.300	0.030	0.518	0.025
Protein yield (kg)	0.400	0.051	0.465	0.041
SCS (points) <sup>1</sup>	0.000	0.079	-0.008	0.000
Muscularity (points) <sup>1</sup>	0.300	-0.878	0.010	0.000

<sup>1</sup>Traits to which a restriction to the genetic progress was applied.

study) routinely collected once in the lifetime of each individual, around the time of first calving, by expert classifiers; these traits are rated on a scale of 1 to 50 points with biological extremes. These measurements ranged from the year 2010 to 2020. First, multivariate response to selection was computed by using economic weights that were assigned to traits according to breeders' preference (see Step 1) and therefore aimed mostly to improve the dairy attitude of the breed (the largest income for this breed) through a direct selection on fat and protein yield. Then, a restriction to the genetic gain was applied to SCS and MUSC to prevent detriment to the traits. Constraining MUSC would allow safeguarding of the the beef attitude, which is part of the breed's uniqueness. Furthermore, restricting SCS allows prevention against detriment of udder health. The initial economic weights, the final economic weights obtained after restriction of SCS and MUSC to zero, and the corresponding responses to selection ( $\Delta G$ ) are reported in Table 3.

Note that, before applying restriction in the selection index, both SCS and MUSC reported negative genetic progress, due to their negative genetic correlations with milk yield traits. The main equations described previously have been used throughout the example and are recalled at each of the 5 steps.

The procedure is presented as R code (see Appendix for details; R Core Team, 2017).

### Step 1: Definition of Matrices and Vectors of Data

A fixed selection intensity of  $i = 1.755$  corresponding to a percentage  $P = 0.1$  of parents selected has been set together with the matrices of phenotypic and genetic (co)variances ( $\mathbf{P}$  and  $\mathbf{G}$ ) for the target traits. The initial economic weights of the traits have been reported in the vector  $\mathbf{a}$ , mimicking a supposed initial weight in the selection index of 0.3 and 0.4 for FY and PY, respectively, and 0.3 for MUSC. Both MY and SCS had an economic weight of zero at the initial step. These economic weights of traits were then standardized to calculate the expected initial genetic progress to be compared with the final one.

### Step 2. Multivariate Genetic Progress Calculated Without Restrictions for Any Traits

Equations [4], [5b], and [6] are applied.

### Step 3. Application of the Restriction

Milk yield traits (MY, FY, and PY) show a positive genetic progress, whereas SCS showed a positive increase (indicating detrimental effect on udder health), and MUSC results worsened by selection. Two restrictions to genetic progress were then applied to SCS and MUSC, to prevent their negative influence on udder health and reduction of muscularity. The procedure moved from the definition of matrix  $\mathbf{C}$ , including the genetic variances for SCS and MUSC and their covariances with the other traits, in columns 3 and 4. This matrix is included within the restriction matrix  $\mathbf{K}$  and used for finding the new coefficients  $\mathbf{b}_r$ . This step applied equations [12], [14], and [6], including the new coefficients.

### Step 4. Obtaining the New Economic Weights for Traits Arising from the Restrictions in the Selection Index

Equation [15] was used in this case. The new economic weights were then divided by the sum of their absolute value to obtain a zero sum. Looking at the results (Table 2), it is possible to observe that applying the restriction for SCS and MUSC is equivalent to providing a negative economic weight of  $-0.008$  to SCS, an economic weight of 0.010 to MUSC, and economic weights of 0.518 and 0.465 for FY and PY, respectively.

### Step 5. Check the Genetic Progress

Aiming to check the final genetic progress to compare with the initial one, the new economic weights obtained are used in the equation of genetic progress (equation [6]) after dividing the economic weight by the genetic standard deviation of each trait.

## DISCUSSION

Finding the economic weights that make genetic progression equal to zero for some traits is helpful in dairy cattle breeding (as well as in other livestock species and breeds), in which many different traits are included in the selection index.

Particularly, our approach could be convenient for livestock, as in dairy and beef cattle, where selection indexes also include many morphological characters with an intermediate optimum, such as leg traits or udder conformation (Jeyaruban et al., 2012). In these traits, genetic improvement aims to maintain the phenotype at intermediate values (i.e., the traits will develop neither positively nor negatively). Other traits that should be maintained at the existing value while selecting for production improvement are the functional traits that show negative genetic correlations with production, such as birth weight versus growth in beef cattle. In the study of Winder et al. (1990), in Red Angus cattle, the genetic gain of birth weight was restricted to zero, aiming to preserve this trait despite the negative genetic correlations (around  $-0.7$ ) with the growth rate. Other functional traits typically showing negative genetic correlations with production are fertility and longevity (Oltenuacu and Broom, 2010). Depending on the breeds' attitude and the breeders' preferences, these traits are aimed at maintenance (as in dual-purpose cattle), or, in the case of dairy cattle breeds, they are subject to genetic improvements (positive selections). In fact, in recent decades, selection aimed at milk to improve production (Miglior et al., 2005) has caused a drastic decline in the quality of the udder conformation of dairy cows. A restriction could also be implemented on traits that have a low or not calculable economic value but are related to characteristics that are typical of the breed or that make the breed adapted to the territory in which it is reared (Krupová et al., 2016). This scenario involves traits such as feed efficiency (Fuerst-Waltl et al., 2016) or important morphological traits (Mancin et al., 2021). Because these traits could be lost during the selection process for productive traits such as MY, a solution consists of guaranteeing at least a non-negative genetic progress by applying a restriction while increasing the production for the most economically important traits.

The genetic gain restriction could also be applied to traits whose positive increase corresponds to the detriment of some functional characteristics, such as traits like SCS, whose positive variation means a loss in udder health. A restriction to zero of this trait has been adopted to prevent positive mammary health issues, as in Alpine Grey (Mancin et al., 2021) and Rendena cattle (Sartori et al., 2018).

The restriction on the selection index could also be a solution for productive traits with lesser economic importance than others but showing negative genetic correlations with the main ones. This situation is likely to occur in some dual-purpose breeds, in which the 2 aptitudes, milk and beef, show negative genetic correlations, and a much greater economic weight is assigned to the first aptitude. Under these circumstances, the genetic trend of the other aptitude may be negative, and restriction to zero could be a valuable solution to prevent a negative variation. This situation is the case for muscularity in the practical example reported above, which shows a negative genetic correlation (around  $-0.4$ ) with milk yield traits, as observed in many studies on dual-purpose breeds (Sartori et al., 2018; Mancin et al., 2021). A restriction to the genetic gain of MUSC has been applied to obtain the greatest positive increase as possible for milk without a detriment in muscularity, as also proposed in Mancin et al. (2021).

According to Wolfová et al. (2001), it should be an uncorrected prior assumption to consider some traits, such as MUSC and SCS, as restricted traits because they have their own specific economic value. However, in local breeds, no previous studies have ever been done about the economic assessment of these target traits. According to Krupová et al. (2016), it could therefore be considered a preliminary process to establish a suitable selection index satisfying the needs of any breed associations.

In conclusion, after presenting the theory behind the univariate and multivariate selection indices and the use of restriction for some traits within a selection index, this study focused on the problems of accounting for unequal individual information when restricting target traits. A new approach for dealing with this issue was offered. A simulation was also provided as a supporting document (Supplemental File S1). Moreover, this study showed a practical application using real data of a local breed, the Alpine Grey cattle, with possible implementation in situations where negative correlations occur among target selected traits. Applying restriction allows simultaneous improvement and non-deterioration of all target traits, which is especially beneficial when many traits need to be accounted for selection, as is common today in many dairy, dual-purpose, and beef breeds. Therefore, this study offered a predictive method for computing economic weights accounting for a restriction to zero of the genetic gain for traits of interest, useful for reasoning around the consequences of selection when new traits should be included in a selection index, or when different weights should be applied to traits that are already included in it. Moreover, this study provided a new method to

account for unequal accuracies of individual data when restricting traits, which is useful for implementation in real-life scenarios.

## ACKNOWLEDGMENTS

The authors thank the National Breeders Association of Alpine Grey breed (ANAGA; Bolzano, Italy). This work received funding from the DUALBREEDING project (PSRN 2014–2020, Sottomisura 10.2). The authors have not stated any conflicts of interest.

## REFERENCES

- Ali, A. K. A., and G. E. Shook. 1980. An optimum transformation for somatic cell concentration in milk. *J. Dairy Sci.* 63:487–490. [https://doi.org/10.3168/jds.S0022-0302\(80\)82959-6](https://doi.org/10.3168/jds.S0022-0302(80)82959-6).
- Brascamp, E. W. 1984. Selection indices with constraints. *Anim. Breed. Abstr.* 52:645–654.
- Céron-Rojas, J. J., and J. Crossa. 2018. *Linear Selection Indices in Modern Plant Breeding*. Springer. <https://doi.org/10.1007/978-3-319-91223-3>.
- Cunningham, E. P., and A. E. McClintock. 1974. Selection in dual purpose cattle populations: Effect of beef crossing and cow replacement rates. *Genet. Sel. Evol.* 6:227–239. <https://doi.org/10.1186/1297-9686-6-2-227>.
- Cunningham, E. P., R. A. Moen, and T. Gjedrem. 1970. Restriction of selection indexes. *Biometrics* 26:67–74. <https://doi.org/10.2307/2529045>.
- Dekkers, J. C. M. 2003. *Design and Economics of Animal Breeding Strategies*. Notes for summer short course, Iowa State University, July 14–18, 2003.
- Falconer, D. S., and T. F. C. Mackay. 1996. *Introduction to Quantitative Genetics*. 4th ed. Longmans Green.
- Fuerst-Waltl, B., C. Fuerst, W. Obritzhauser, and C. Egger-Danner. 2016. Sustainable breeding objectives and possible selection response: Finding the balance between economics and breeders' preferences. *J. Dairy Sci.* 99:9796–9809. <https://doi.org/10.3168/jds.2016-11095>.
- Gaynor, R.C., G. Gorjanc, and J.M. Hickey. 2021. AlphaSimR: An R package for breeding program simulations. *G3* 11:jkaa017. <https://doi.org/10.1093/g3journal/jkaa017>.
- Gibson, J. P. 1992. *Economic breeding plans*. Nord. Grad. course, Greve, Denmark.
- Harris, D. L. 1964. Expected and predicted progress from index selection involving estimates of population parameters. *Biometrics* 20:46–72. <https://doi.org/10.2307/2527617>.
- Hazel, L. N. 1943. The genetic basis for constructing selection indexes. *Genetics* 28:476–490. <https://doi.org/10.1093/genetics/28.6.476>.
- Henderson, C. R. 1975. Best linear unbiased estimation and prediction under a selection model. *Biometrics* 31:423–447. <https://doi.org/10.2307/2529430>.
- Hickey, J. M., R. F. Veerkamp, M. P. Calus, H. A. Mulder, and R. Thompson. 2009. Estimation of prediction error variances via Monte Carlo sampling methods using different formulations of the prediction error variance. *Genet. Sel. Evol.* 41:23. <https://doi.org/10.1186/1297-9686-41-23>.
- Hill, W. G. 1974. Prediction and evaluation of response to selection with overlapping generations. *Anim. Sci.* 18:117–139. <https://doi.org/10.1017/S0003356100017372>.
- Jeyaruban, G., B. Tier, D. Johnston, and H. Graser. 2012. Genetic analysis of feet and leg traits of Australian Angus cattle using linear and threshold models. *Anim. Prod. Sci.* 52:1–10. <https://doi.org/10.1071/AN11153>.
- Kempthorne, O., and A. W. Nordskog. 1959. Restricted selection indices. *Biometrics* 15:10–19. <https://doi.org/10.2307/2527598>.
- Krupová, Z., E. Krupa, M. Michaličková, M. Wolfová, and R. Kasarda. 2016. Economic values for health and feed efficiency traits of dual-purpose cattle in marginal areas. *J. Dairy Sci.* 99:644–656. <https://doi.org/10.3168/jds.2015-9951>.
- Lin, C. Y. 1990. A unified procedure of computing restricted best linear unbiased prediction and restricted selection index. *J. Anim. Breed. Genet.* 107:311–315. <https://doi.org/10.1111/j.1439-0388.1990.tb00040.x>.
- Mancin, E., C. Sartori, N. Guzzo, B. Tuliozi, and R. Mantovani. 2021. Selection response due to different combination of antagonistic milk, beef, and morphological traits in the Alpine Grey cattle breed. *Animals (Basel)* 11:1340. <https://doi.org/10.3390/ani11051340>.
- Mazza, S., N. Guzzo, C. Sartori, and R. Mantovani. 2016. Genetic correlations between type and test-day milk yield in small dual-purpose cattle populations: The Aosta Red Pied breed as a case study. *J. Dairy Sci.* 99:8127–8136. <https://doi.org/10.3168/jds.2016-11116>.
- Miglior, F., B. L. Muir, and B. J. Van Doormaal. 2005. Selection indices in Holstein cattle of various countries. *J. Dairy Sci.* 88:1255–1263. [https://doi.org/10.3168/jds.S0022-0302\(05\)72792-2](https://doi.org/10.3168/jds.S0022-0302(05)72792-2).
- Misztal, I., and G. R. Wiggans. 1988. Approximation of prediction error variance in large-scale animal models. *J. Dairy Sci.* 71:27–32. [https://doi.org/10.1016/S0022-0302\(88\)79976-2](https://doi.org/10.1016/S0022-0302(88)79976-2).
- Oltenuacu, P. A., and D. M. Broom. 2010. The impact of genetic selection for increased milk yield on the welfare of dairy cows. *Anim. Welf.* 19:39–49.
- R Core Team. 2017. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing.
- Rendel, J. M., and A. Robertson. 1950. Estimation of genetic gain in milk yield by selection in a closed herd of dairy cattle. *J. Genet.* 50:1–8. <https://doi.org/10.1007/BF02986789>.
- Sartori, C., N. Guzzo, S. Mazza, and R. Mantovani. 2018. Genetic correlations among milk yield, morphology, performance test traits and somatic cells in dual-purpose Rendena breed. *Animal* 12:906–914. <https://doi.org/10.1017/S1751731117002543>.
- Schneeberger, M., S. A. Barwick, G. H. Crow, and K. Hammond. 1992. Economic indices using breeding values predicted by BLUP. *J. Anim. Breed. Genet.* 109:180–187. <https://doi.org/10.1111/j.1439-0388.1992.tb00395.x>.
- Tallis, G. M. 1962. Note: A selection index for optimum genotype. *Biometrics* 18:120–122. <https://doi.org/10.2307/2527716>.
- Weller, J. I. 1994. *Economic Aspects of Animal Breeding*. Springer Science and Business Media.
- Winder, J. A., J. S. Brinks, R. M. Bourdon, and B. L. Golden. 1990. Genetic analysis of absolute growth measurements, relative growth rate and restricted selection indices in red Angus cattle. *J. Anim. Sci.* 68:330–336. <https://doi.org/10.2527/1990.682330x>.
- Wolfová, M., G. Nitter, J. Wolf, and J. Fiedler. 2001. Impact of crossing system on relative economic weights of traits in purebred pig populations. *J. Anim. Breed. Genet.* 118:389–402. <https://doi.org/10.1046/j.1439-0388.2001.00304.x>.
- Xie, C., and S. Xu. 1997. Restricted multistage selection indices. *Genet. Sel. Evol.* 29:193–203. <https://doi.org/10.1051/gse:19970206>.

## ORCIDS

- E. Mancin  <https://orcid.org/0000-0002-4130-4639>  
 R. Mantovani  <https://orcid.org/0000-0002-1919-0559>  
 C. Sartori  <https://orcid.org/0000-0002-2091-2280>

## APPENDIX

Complete R code (R version 4.2.1) of multivariate selection with restricted economic weights.

```

# selection intensity
i = 1.755
# definition of traits to be accounted
traits = c("MY", "FY", "PY", "SCS", "MUSC")
# matrix of phenotypic (co)variances
P = matrix(c(10.1072, 0.2507, 0.4118, -0.7921, -1.4849,
             0.2507, 0.0105, 0.0112, -0.0137, -0.0017,
             0.4118, 0.0112, 0.0205, -0.0265, -0.0018,
             -0.7921, -0.0137, -0.0265, 2.7916, 0.0317,
             -1.4849, -0.0017, -0.0018, 0.0317, 27.3200),
           ncol=5)
# matrix of genetic (co)variances
G = matrix(c( 2.2353, 0.0494, 0.0815, 0.0630, -2.0559,
             0.0494, 0.0019, 0.0023, 0.0018, -0.0540,
             0.0815, 0.0023, 0.0042, 0.0035, -0.0769,
             0.0630, 0.0018, 0.0035, 0.3768, -0.2906,
             -2.0559, -0.0540, -0.0769, -0.2906, 9.0014),
           ncol=5)
colnames(G) = traits; row.names(G) = traits
colnames(P) = traits; row.names(P) = traits
# genetic standard deviations of traits
ds_G = sqrt(diag(G))
# initial economic weights of traits
a = c(0,0.3,0.4,0,0.3)
# standardized economic weights
a_std = a/(ds_G)
# Inverse of P matrix
Pinv = solve(P)
# Coefficient that maximizes equations Pb = Ga
b = Pinv %*% G %*% (a_std)
# calculate the standard deviation of the index
ds_I = sqrt(t(b) %*% P %*% b)
# calculate and print initial genetic progress
gp = (i/ds_I) %*% t(b) %*% G
print(round(gp,3))
## MY FY PY SCS MUSC
## [1,] 0.846 0.031 0.051 0.079 -0.878
# create a matrix of restriction for SCS and MUSC
C = G[c('SCS','MUSC'),]
C = t(C)
# create an identity matrix with the dimension of the matrix correlations:
I = diag(rep(1,ncol(G)))
# calculate matrix K (matrix of restrictions)
K = I - (Pinv %*% ((C) %*% solve(t(C) %*% Pinv %*% (C)) %*% t(C)))
# coefficient that maximizes restriction
br = K %*% Pinv %*% G %*% (a_std)
# calculate genetic progress
gp = ((i/(ds_I))%*%t(br)%*%G)
# calculate economic weight according with the restrictions
new_a = solve(G)%*%(P)%*%(br)

```

```
a = c(new_a)/(sum(abs(new_a)))
print(round(a,3))
## 0.000 0.518 0.464 -0.008 0.010
a_std = a/(ds_G)
b = Pinv %*% G %*% (a)
# calculate standard deviation of the index
ds_I = sqrt(t(b) %*% P %*% b)
# calculate and final print genetic progress
gp = (i/ds_I) %*% t(b) %*% G
print(round(gp,3))
## MY FY PY SCS MUSC
## [1,] 0.593 0.024 0.041 0 0
```