



Selection index using the graphical area applied to sugarcane breeding

L.A. Silva¹, R.T. Resende¹, R.A.D.C. Ferreira¹, G.N. Silva², V. Kist³,
M.H.P. Barbosa⁴, M. Nascimento² and L.L. Bhering¹

¹Departamento de Biologia Geral, Universidade Federal de Viçosa, Viçosa, MG, Brasil

²Departamento de Estatística, Universidade Federal de Viçosa, Viçosa, MG, Brasil

³Instituto Federal Catarinense, Concórdia, SC, Brasil

⁴Departamento de Fitotecnia, Universidade Federal de Viçosa, Viçosa, MG, Brasil

Corresponding author: L.A. Silva
E-mail: lidianeufv2007@gmail.com

Genet. Mol. Res. 15 (3): gmr.15038711

Received April 15, 2016

Accepted June 14, 2016

Published September 16, 2016

DOI <http://dx.doi.org/10.4238/gmr.15038711>

Copyright © 2016 The Authors. This is an open-access article distributed under the terms of the Creative Commons Attribution ShareAlike (CC BY-SA) 4.0 License.

ABSTRACT. This study aimed to develop a multivariate selection index based on the graphical area of a polygon formed by standardized values, also known as radar chart. This methodology may be used to assist selection of superior genotypes in sugarcane breeding programs. Seven technological traits in 37 sugarcane genotypes were evaluated. An area index (AI) was constructed and the resulting polygon areas were used to rank genotypes under selection. In this study, we propose the use of restricted maximum likelihood to estimate genetic parameters and mixed model equations to predict genotypic and breeding values. The area of each polygon was calculated for phenotypic, genotypic, and estimated breeding values. Thereby, the genotypes with larger

area can be selected based on a detailed *a posteriori* evaluation of the radar charts. The proposed AI can be adjusted based on the breeders' specific interests, it is perfectly useful in other crops, and may also be applied to studies on genotype-environment interactions. Moreover, AI is a powerful tool that can evaluate trait stability of genotypes based on slight differences in the area formed by each genotype. Hence, this method is easy to apply and shows great potential for use in sugarcane breeding programs as well as in other breeding programs.

Key words: Visual index; Radar chart; Mixed models; Plant breeding

INTRODUCTION

Sugarcane (*Saccharum* spp) is one of the main crops in Brazil, making the country the major producer and exporter of sugar and ethanol (Mapa, 2016). Aiming at meeting the market demand, sugarcane breeding programs have sought to improve new analytical methodologies to optimize the process of obtaining and selecting superior genotypes, in order to develop genetic materials with high yield and expressing agronomic traits of interest. For this purpose, at the initial stages of the genetic breeding process of sugarcane, indirect selection based on one or a few traits have been used. However, this method may prove to be inadequate, since negative correlations between traits of interest are observed (Pedrozo et al., 2009).

The use of selection indices is an alternative method recommended by breeders. This is a multivariate analysis method for obtaining additional gains in a set of attributes considered relevant (Cameron, 1997). Several selection index methods that can be used in plant breeding programs have been described in the literature (Cruz et al., 2012). In general, these methods use phenotypic means to rank the genotypes. However, it is well known that mixed model equations used to predict the genotypic and breeding values provide paths for more accurate breeding selection (Piepho et al., 2008; van der Westhuizen and van der Westhuizen, 2009). Although the applicability of mixed models has been demonstrated in several crops, there are few studies employing this strategy in sugarcane. This is due to inaccuracies in the process of the covariance matrix estimate, as well as changes in genetic parameters with selection when applied to sugarcane (Pedrozo et al., 2009; Espósito et al., 2012). In order to increase the accuracy of these processes, graph visualization techniques are widely used in many fields, such as economics and politics (Keefer and Loayza, 2008), product quality evaluation (Li et al., 2010), health (Saary, 2008), and agriculture (Soriano et al., 2005; Banerjee et al., 2015). For example, Laino et al. (2015) applied radar charts to evaluate qualitative traits of wheat to characterize a germplasm collection, and Liu et al. (2010) evaluated quantitative and qualitative traits in tobacco.

Nunes et al. (2005) published a methodology to select individuals considering a multiple trait framework based on radar charts that does not require prior knowledge of the genetic and phenotypic covariance. This is an interesting procedure, which provides a descriptive analytical tool and uses a graphic index based on the sum of the standardized trait values. This allows the researcher to visualize which progeny traits have favorable phenotypes (Reis et al., 2011). The graph visualization presents the results in terms of indirect variables, such as means and coefficient of variation. In addition, this method offers results in terms of phenotypic values, given by fixed models. However, it is well known that methodologies

implementing mixed models provide more reliable estimates (Piepho et al., 2008; Castro et al., 2016; Almeida Filho et al., 2016).

The aim of the present study was to develop a selection index approach that considers the information of several sugarcane traits to identify superior genotypes using the polygon area of traditional radar charts. In addition, we adopted restricted maximum likelihood (REML) to estimate genetic parameters and mixed model equations (BLUP) to predict genotypic and breeding values.

MATERIAL AND METHODS

Plant material

In this experiment, we used 37 sugarcane genotypes from the Active Germplasm Bank of the Sugarcane Breeding Program of Universidade Federal de Viçosa (PMGCA/UFV), which is part of the Inter-University Network for the Development of Sugarcane Industry. The PMGCA/UFV is located at Fundão Farm, in the municipality of Viçosa, in Minas Gerais State (20°45'14"S, 42°52'53"W, at 650 m above sea level).

The genotypes used in this study were chosen based on agronomic importance. We used genotypes that occurred or still occur in significant areas of commercial cultivation in Brazil, together with clones that are used in hybridization (Table 1).

Variable determination

At 18 months after planting, 10 stalks from each genotype were randomly collected by hand and without straw burning. A 500-g stalk subsample of each genotype was minced, homogenized, and weighed, to obtain the fresh sample weight. After drying this same material at 105°C for 24 h, the bagasse dry weight with juice was obtained. Dry matter (DM) content was obtained from the ratio between the bagasse dry weight with juice and the fresh sample weight.

Subsequently, another 500-g subsample was subjected to hydraulic press, at 250 kgf/cm², for 1 min. From this procedure, the juice was obtained and was subsequently used to determine the total soluble solids (BRIX), using an automatic digital refractometer (model RDA 9000 - Acatec, São Paulo, SP, Brazil), and to determine the apparent sucrose content in the juice (POL), using a polarimeter (model SDA 2500-Acatec). Moist cake weight was obtained from weighing the subsample bagasse that was removed from the press. Dry cake weight was obtained after drying the bagasse at 105°C for 24 h. From these traits and from the equations described in the literature (Consecana, 2006), we calculated estimates for fiber content (FIB), apparent sucrose content in sugarcane (POLS), purity (PUR), and total recoverable sugars (TRS).

Calculation of genetic parameters

The following seven traits were evaluated in this study: BRIX, TRS, PUR, POLS, FIB, DM, and POL. The phenotypic values (PV) were obtained from the phenotypic mean of each trait evaluated for each genotype. For each trait, the narrow-sense heritability coefficients (h^2) were calculated (described below). Genotypic values (GV) and estimated breeding values (EBV) were estimated using the mixed model methodology, REML/BLUP (Resende, 2007):

Table 1. Thirty-seven sugarcane genotypes from the Active Germplasm Bank of the Sugarcane Breeding Program of Universidade Federal de Viçosa (PMGCA/UFV), which is part of the Inter-University Network for the Development of Sugarcane Industry (RIDESA) and their respective parents.

| N | Genotype | Female parent | Male parent |
|----|-----------|---------------|-------------|
| 1 | IAC862210 | CP5248 | CO798 |
| 2 | IAC862480 | US71399 | ? |
| 3 | IAC873396 | CO740 | SP701143 |
| 4 | RB721012 | CO331 | ? |
| 5 | RB72454 | CP5376 | ? |
| 6 | RB739359 | IANE5534 | ? |
| 7 | RB739735 | CB52179 | ? |
| 8 | RB75126 | C278 | ? |
| 9 | RB83102 | NA5679 | SP701143 |
| 10 | RB835054 | RB72454 | NA5679 |
| 11 | RB835089 | RB72454 | NA5679 |
| 12 | RB835486 | L6014 | ? |
| 13 | RB845197 | RB72454 | SP701143 |
| 14 | RB845210 | RB72454 | SP701143 |
| 15 | RB845257 | RB72454 | SP701143 |
| 16 | RB855035 | L6014 | SP701284 |
| 17 | RB855036 | RB72454 | SP701143 |
| 18 | RB855113 | SP701143 | RB72454 |
| 19 | RB855156 | RB72454 | TUC717 |
| 20 | RB855453 | TUC717 | ? |
| 21 | RB855511 | SP711406 | ? |
| 22 | RB855536 | SP701143 | RB72454 |
| 23 | RB863129 | RB763411 | ? |
| 24 | RB867515 | RB72454 | ? |
| 25 | RB925211 | RB855206 | ? |
| 26 | RB925345 | H591966 | ? |
| 27 | RB92579 | RB75126 | RB72199 |
| 28 | RB928064 | SP701143 | ? |
| 29 | RB93509 | RB72454 | ? |
| 30 | RB935744 | RB835089 | RB765418 |
| 31 | SP775181 | HJ5741 | ? |
| 32 | SP791011 | NA5679 | CO775 |
| 33 | SP801816 | SP711088 | H575028 |
| 34 | SP801842 | SP711088 | H575028 |
| 35 | SP803280 | SP711088 | H575028 |
| 36 | SP813250 | CP701547 | SP711279 |
| 37 | SP832847 | HJ5741 | SP701143 |

?: unknown male parent.

$$y = X\beta + Zg + e \quad (\text{Equation 1})$$

in which y is the vector of observations; β is the vector of fixed effects (or the overall mean); g is the vector of random genotype effects; e is the random effect of residues; and X and Z are the incidence matrices of β and g , respectively. The matrix form of the Equation 1 is given by:

$$\begin{bmatrix} \hat{\beta} \\ \hat{g} \end{bmatrix} = \begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + V^{-1}k \end{bmatrix}^{-1} \begin{bmatrix} X'y \\ Z'y \end{bmatrix} \quad (\text{Equation 2})$$

in which $\hat{\beta}$ and \hat{g} are the estimated vectors of the fixed and genetic effects, respectively; X and Z are the known incidence matrices; k is a constant that contains the h^2 of each trait, being

$k = (1 - h^2) / h^2$ and $h^2 = \sigma_g^2 / (\sigma_g^2 + \sigma_e^2)$, where σ_g^2 is the genetic variance component and σ_e^2 is the residual variance component; V is the identity matrix (I) used to obtain GV, or a kinship matrix (A) described by Wright (1922) to obtain EBV. For the EBV model, σ_g^2 is the additive genetic variance, and thus the h^2 assumes narrow-sense heritability values (Lynch and Walsh, 1998).

Prediction of the GV and EBV was carried out using the R packages lme4 (Bates et al., 2007) and pedigreemm (Vazquez et al., 2010).

Area index (AI)

The AI was constructed based on the methodology that Nunes et al. (2005) used to study phenotypic stability and adaptability, which considers each trait measured as an environment in the seminal methodology. For the construction of the radar charts, all traits were standardized by assigning a mean of zero and a variance of one. To ensure better graphical visualization, a constant (the lowest value of each trait) was added to the standardized values, making them all positive.

For the graphical presentation, a polygon based on the genotypes' performance was obtained. Thus, in contrast to the initial proposal of Nunes et al. (2005), in this study, we calculated the area of each polygon using the coordinates of the polygon in the plan x and y (Equation 3), which was implemented in the R package splancs (Rowlingson and Diggle, 2014).

$$area = \frac{(x_1y_2 - y_1x_2) + (x_2y_3 - y_2x_3) + \dots + (x_ny_1 - y_nx_1)}{2} \quad (\text{Equation 3})$$

where x_1 is the x coordinate of vertex 1, and x_n is the x of the n -th vertex; y_1 is the y coordinate of vertex 1, and y_n is the y coordinate of the n -th vertex.

The values of these areas were used as AI for the selection of promising genotypes. Thus, the genotypes with the largest areas should be selected for *a posteriori* evaluation of their radar charts. When it was desired to select genotypes with lower values for a given trait of interest, it was necessary to invert the value order. For example, to evaluate the susceptibility to a disease in terms of damage, it must be converted into resistance scale, considering the highest score as the most resistant. The R codes used to generate the graphics and the AIs are attached in the [Figures S1](#) and [S2](#).

In order to compare the viability of the proposed methodology, three different strategies for creating the indices were used. An AI was constructed for each analysis approach described above (PV, GV, and EBV). These indices were denoted AI for phenotypic values (AI_{PV}), AI for genotypic values (AI_{GV}), and AI for estimated breeding values (AI_{EBV}).

Genotypes were ranked for each proposed AI, according to the polygon area formed by the original trait order in the data sheet (BRIX, TRS, PUR, POLS, FIB, DM, and POL). However, since the trait order can influence the polygon shape, and consequently its area, we calculated, for each genotype, the areas of all possible polygons formed, by order permutation among the seven traits. In addition, we calculated the mean of all the areas and the difference between the smallest and largest polygon area that changes in the trait order produced. To check whether the areas formed after permutation were statistically different between them, we perform the Scott Knott test adopting 5% significant level.

To check if one order was enough to account for all possible areas, we estimated the

correlation between the polygon area formed by the original trait order and the mean of all possible areas. Moreover, in order to make the values of the areas more reliable and simplify the data analysis, we estimated the correlation between the mean of the total number of permutation possibilities and the mean value of only 100 permutations. Given the difference in area obtained from the permutation, the smallest and the largest areas of each genotype were used as the confidence intervals of the mean. Thus, to investigate the potential of using only 100 permutations, we also carried out correlations between the intervals of the total number of permutations with the intervals of 100 permutations.

Finally, to test the reliability of the proposed methodology, the results obtained by selection from the AI presented in this study were compared with those obtained using the classical indices of Elston (1963) and Mulamba and Mock (1978). In addition, the rank of the five most cultivated genotypes in Brazil were also used for comparison (Chapola et al., 2013).

RESULTS

The h^2 estimated for the seven traits presented intermediate to low magnitude, ranging from 0.1263 (PUR) to 0.4095 (FIB) (Table 2).

Table 2. Narrow-sense heritability coefficients (h^2) for seven traits evaluated in 37 sugarcane genotypes.

| Variable | BRIX | POL | POLS | PUR | TRS | DM | FIB |
|----------|--------|--------|--------|--------|--------|--------|--------|
| h^2 | 0.2294 | 0.2091 | 0.2759 | 0.1263 | 0.2962 | 0.4022 | 0.4095 |

BRIX: total soluble solids; POL: apparent sucrose content in the juice; POLS: apparent sucrose content in sugarcane; PUR: purity; TRS: total recoverable sugars; DM: dry matter; FIB: fiber content.

Figure 1 shows the radar charts for the two genotypes of largest and smallest areas from AI_{PV} , AI_{GV} , and AI_{EBV} based on the original order of the seven traits (BRIX, TRS, PUR, POLS, FIB, DM, and POL). In this case, the larger the area of each genotype (black line), the better the performance of the genotype regarding the considered traits, compared with the mean of all genotypes.

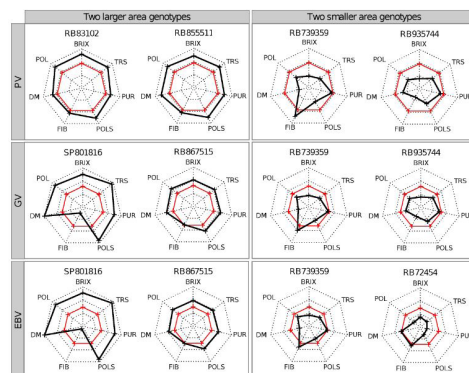


Figure 1. Radar chart for the two genotypes with largest and smallest polygon areas considering the mean phenotypic values (PV), the genotypic values (GV), and the estimated breeding values (EBV). Red lines: overall mean of the 37 genotypes for each trait; black lines: standardized mean of each trait for a given genotype. BRIX: total soluble solids; POL: apparent sucrose content in the juice; POLS: apparent sucrose content in sugarcane; PUR: purity; TRS: total recoverable sugars; DM: dry matter; FIB: fiber content.

We observed that AI_{PV} differed from the other indices in which genotypes were selected as the largest and smallest areas (Figure 1). On the other hand, AI_{GV} and AI_{EBV} selected the same two genotypes for the larger areas, RB867515 and SP801816. Both of these genotypes showed that only the FIB content was lower than or equal to the trait mean.

For the seven traits, we evaluated 5040 combinations (all possible permutations between the seven traits used) and formed different polygons for each genotype. Figure 2 shows the rank of all evaluated genotypes for each proposed AI formed by the different trait combinations. This graph highlights the difference between the lowest and the highest polygon areas that a change in trait order may result in for each genotype, as well as the mean of the 5040 possible areas.

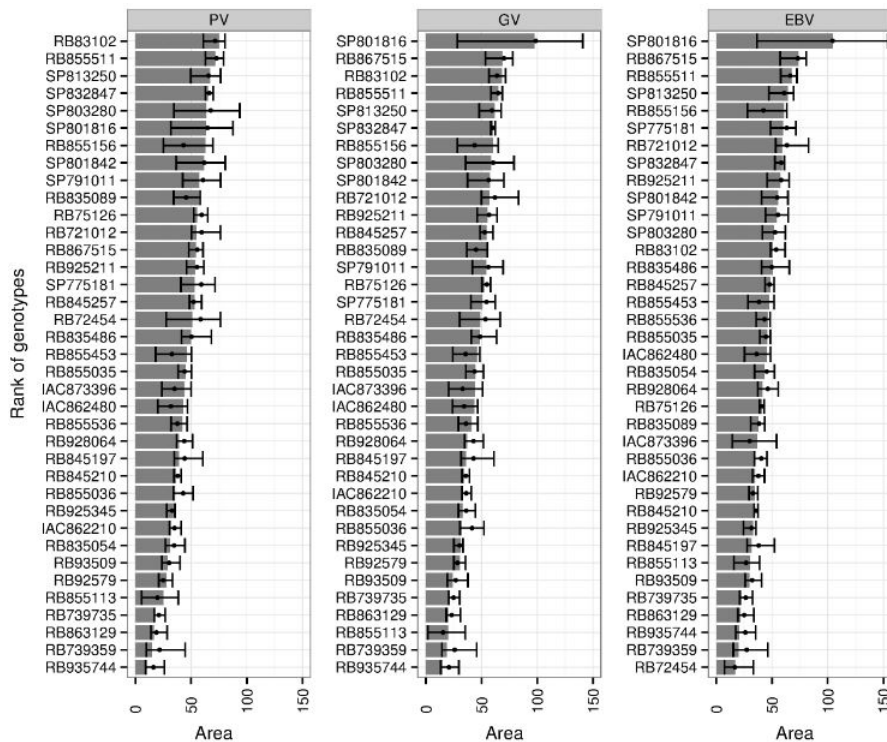


Figure 2. Rank of the 37 genotypes, from the largest to the smallest area (standardized value²), considering the original order of the seven traits, for phenotypic value (PV), genotypic value (GV), and estimated breeding values (EBV). Shaded area: area of each genotype for the original order of the seven traits; interval bar: interval between the smallest and largest possible area for each genotype; dot: mean value of all the possible areas for each genotype.

There were strong correlations between the areas of the original order and the mean of the 5040 possible areas for each genotype (Table 3). The highest correlation was observed for AI_{EBV} , followed by AI_{GV} and AI_{PV} (Table 3). Furthermore, we found that using only the mean of 100 combinations (around 2% of the total) was enough to represent the mean of all 5040 polygons. When comparing the difference between the superior range and lower range (mean interval), that is, for 5040 and 100 permutations, high correlations were also observed, especially for EBV, followed by GV and PV (Table 3).

Table 3. Correlation coefficients (r^2) of areas formed by the original order of traits, as well as after 100 permutations with the total number of possible permutations (5040), for each proposed area index (AI ; AI_{pv} , AI_{GV} and AI_{EBV}).

| AI | Area of original order ¹ | Mean of 100 permutations | |
|------------|-------------------------------------|--------------------------|-----------------------|
| | | Mean ² | Interval ³ |
| AI_{pv} | 0.93 | 0.94 | 0.84 |
| AI_{GV} | 0.95 | 0.97 | 0.89 |
| AI_{EBV} | 0.97 | 0.99 | 0.93 |

¹Correlation between the areas formed by the original order of the traits and mean areas of the 5040 permutations; ²correlation between mean areas formed by 100 permutations and mean areas of the 5040 permutations; ³correlation between intervals (the largest area minus the smallest area) of 100 and 5040 permutations.

The two genotypes ranked with the largest area for GV and EBV (SP801816 and RB867515, respectively) show statistically significant differences in area values 5% by the Scott Knott test (Table S1). However, considering the proposed interval, RB867515 showed low variation compared to SP801816, which had the largest range among all genotypes for both indices (Figure 2).

Comparing the rank from EBV (Figure 2) and its respective radar charts (Figure 3), SP832847 presented a high mean and small interval. As shown in Figure 3, it can be noted that all seven traits were above the mean for this genotype. Furthermore, RB721012 presented higher means and larger interval than the SP832847 genotype, and it was noted that all traits were above the mean for RB721012 too. However, RB721012 presented higher potential for FIB. In contrast, IAC873396 had a low mean and a large interval, and the radar chart suggested that only one variable (FIB) was higher than the grand mean.

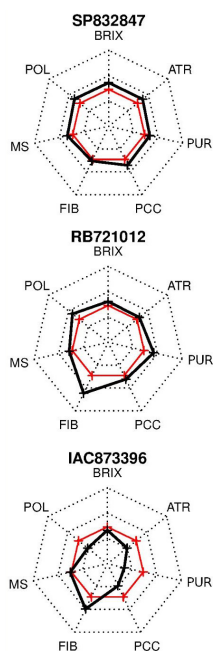


Figure 3. Radar chart of three random genotypes, considering the estimated breeding values (EBV). Red lines: overall mean of the 37 genotypes for each trait; black lines: standardized mean of each trait for a given genotype.

Table 4 ranked the top five genotypes using the classical indices of Mulamba and Mock (1978) and Elston (1963), the indices proposed in this study (AI_{PV} , AI_{GV} , and AI_{EBV}), and the five most cultivated genotypes in Brazil (Chapola et al., 2013). Among the different indices, only the AI_{GV} and AI_{EBV} included the most cultivated genotypes in Brazil (RB867515).

Table 4. Rank of the top five genotypes obtained using different selection indices from sugarcane technological variables evaluated in this study.

| Rank | Elston ¹ | Mulamba ² | AI_{PV} ³ | AI_{GV} ⁴ | AI_{EBV} ⁵ | Varietal census ⁶ |
|------|---------------------|----------------------|------------------------|------------------------|-------------------------|------------------------------|
| 1 | SP803280 | SP803280 | RB83102 | SP801816 | SP801816 | RB867515 |
| 2 | SP801816 | SP801816 | RB855511 | RB867515 | RB867515 | SP813250 |
| 3 | RB72454 | SP801842 | SP813250 | RB83102 | RB855511 | RB855453 |
| 4 | SP813250 | RB72454 | SP832847 | RB855511 | SP813250 | RB92579 |
| 5 | RB925211 | SP813250 | SP803280 | SP813250 | RB855156 | RB855536 |

¹Elston index (1963); ²Mulamba and Mock index; ³area index (AI) for phenotypic values (AI_{PV}); ⁴AI for genotypic values (AI_{GV}); ⁵AI for estimated breeding values (AI_{EBV}); ⁶rank of the most cultivated genotypes in Brazil (Chapola et al., 2013).

DISCUSSION

Genetic evaluation of traits used for the AI

Narrow-sense heritability is an important parameter in crops with vegetative propagation, such as sugarcane. For these types of reproduction, the genetic variability is fixed after crossing, and there is no segregation at later stages (Resende and Barbosa, 2005; Zhou and Joshi, 2012). In this study, the heritability estimates ranged from intermediate to low (Table 2) (Resende, 2002). A possible cause for the low heritability estimates may be the restricted genetic base of the population used in the study (Farias Neto and Resende, 2001; Almeida et al., 2014), since most of the genotypes originated from full- or half-sib families (Table 1).

Resende (2002) highlighted the importance of using information of progenies and parents, for which there is much information available, which allows increasing the selective accuracy of the trait from low to moderate. In situations in which h^2 exceeds 0.50, there would be practically no advantage in using family information. Thus, selection based solely on the individual's information would provide high h^2 (>0.70). However, in the present study, h^2 values were <0.50 for all traits. This highlights the importance of using the additive genetic values in order to make the selection process more effective.

AI

The selection index proposed in this study based on the radar chart areas turned out to be an efficient method. It presented great potential for application in sugarcane breeding programs, as well as in other breeding programs. Nunes et al. (2005) used the radar graphic to study the adaptability and stability of cultivars. Since then, the methodology has been applied to multi-trait analyses, including selection indices (Lima et al., 2012, 2015). Thus, the methodology developed in the present study can be used in multi-trait scenarios, but it could also be applied to a multi-environment scenario.

From the visualization of the proposed radar chart (Figure 1), the breeder is able to identify for which traits a particular genotype is below or above the desired mean, enabling

selection or disposal of materials. For this purpose, the area of the central polygon is the mean of all 37 genotypes evaluated for the seven traits. However, in practice, the breeder may replace this central area according to his/her interests. For example, a comparison can instead be made with a given superior genotype available in the market, or potential improvement goals that a company wants to achieve.

For this selection index, the shape and the polygon area are modified based on the order that the traits are arranged in. Thus, it would be more appropriate to use the mean of all possible areas formed for each genotype. Nunes et al. (2005) did not comment on this fact since the areas were evaluated only visually and were based on indirect measurements. Hence, the methodology based on the graphical presentation, together with the permutation test, provides the breeder a realistic measure of what is found in each graphic, regardless of the traits. It is important to highlight that in this study, only seven traits were investigated. Thus, for each genotype, 5040 areas were possible considering that the number of permutations increases with increasing number of evaluated traits. This procedure usually has high computational costs when the number of traits is above 10. In order to facilitate the breeders' work when using this tool, we investigated the correlation between the original trait combination and the mean of the 5040 possible areas. High correlations were observed in all cases (Table 3), which suggests that the method is accurate even when using a random trait order. Furthermore, we also found that the correlation between 100 and 5040 permutations was high, both for the means and the intervals. This indicates that it is not necessary to use all the permutations for determining the range between the largest and smallest areas of each genotype. Based on these results, it should be highlighted that AI_{GV} and AI_{EBV} were more accurate in discriminating genotypes than AI_{PV} , since the highest correlations were found when using EBV. It is important to use GV and EBV, since the traits studied here presented a relatively low h^2 (Habier et al., 2010).

Most selection indices used in breeding programs are constructed based on estimates of genetic parameters and phenotypic means obtained using analyses of variance (Cameron, 1997). However, the use of mixed model equations together with variance components estimated by REML/BLUP is an alternative that may be used in the construction of indices that can result in a more accurate selection process (Resende, 2002; Castro et al., 2016). These alternatives were not mentioned by Nunes et al. (2005). As presented in Table 4, unlike AI_{PV} , the AI_{GV} and AI_{EBV} indices agreed in which genotypes were identified as having the largest areas (SP801816 and RB867515). However, the EBV resulted in a change in the rank of the top genotypes, suggesting that the kinship coefficients of the matrix A, no matter how small, changed the GV results somehow (Munoz et al., 2014).

The SP801816 genotype presented the largest area for both AI_{GV} and AI_{EBV} . However, it also showed the largest variation in the largest and the smallest possible areas among all evaluated genotypes. In this case, depending on the order of the traits, this genotype may present a very large or very small area, which may be indicative of a disturbing factor for the breeder (i.e., some traits present discrepant values that make the interval larger). The visualization of this genotype suggests that this disturbance was due to a very low FIB content compared to the other traits [Leite et al. (2008) also found low FIB for this genotype]. This caused greater variation in the polygon area based on trait order, and consequently a larger interval. However, through the selection index methodology proposed here, this disturbing factor can be easily visually identified by the breeder, and can be considered a problem or not, depending on the trait in consideration.

Although radar charts can provide a good visual interpretation when a small number of

genotypes is evaluated, this interpretation gets increasingly complicated with increasing number of genotypes. Liu et al. (2010) evaluated only seven tobacco genotypes, and consequently, the genotypes were easily discriminated using the radar chart. However, breeders usually evaluate several genotypes in breeding programs, making it impossible to select superior genotypes by radar chart. For instance, Soriano et al. (2005) evaluated 461 pig progenies, and Laino et al. (2015) evaluated more than 4000 wheat accessions. Therefore, the application of AI in breeding programs that evaluate a very large number of genotypes is useful, since genotypes can be ranked and selected based on the area estimated by the AI combined with REML/BLUP.

Using the AI rank (Figure 2), showing the proposed interval for each genotype, together with their respective radar charts, the breeder can also easily identify genotypes of interest, as well as any kind of disturbance. Assuming that changing the trait order influences the polygon shape and area, it is noted that the genotype presents larger area and smaller interval, suggesting little influenced by the order. For instance, considering the genotypes SP832847, RB721012, and IAC873396, based on the EBV rank of Figure 2, and on their radar charts (Figure 3), the former has a very small interval and high mean, and its radar chart is very stable (i.e., only slight area changes after permutations), with all traits above the mean. In contrast, RB721012, which has the largest interval, has a radar chart with all traits above the mean. Nevertheless, it presents different behavior compared with FIB, which is a trait that causes greater variation in the polygon area. Finally, IAC873396 presents a low mean and large interval, and consequently its radar chart is very unstable and may undergo large variations in polygon shape.

It was also observed that neither the classical indices (Elston, 1963; Mulamba and Mock, 1978) nor AI_{pv} ranked RB867515 among the top five ranked genotypes based on the seven evaluated traits. AI_{GV} and AI_{EBV} both classified this genotype in the second position in the rank (Table 4). This result was confirmed by the rank of the most cultivated genotypes in Brazil, in which RB867515 ranks first. The best results were obtained using genotypic values and EBV, compared with the results of the analysis based only on the phenotypic means. This was probably due to the more accurate estimates and predictions obtained in this study.

To conclude, the proposed index methodology combined with REML/BLUP is efficient and can be used as an auxiliary tool for breeders. The radar chart allows a clear visualization of genotypes with larger and smaller areas and is suitable both for sugarcane breeding and others crops. For this selection index, different traits can be evaluated, and the central mean value can be adjusted according to the breeders' interest. For instance, the mean of the company's genotypes, or the mean of the genotypes that have already been implemented in the breeding program may be considered. Although the obtained area differs depending on the trait order, a permutation test provides a general area. We found high correlations with all possible trait combinations also with fewer permutations. Furthermore, the interval between the largest and the smallest area after permutation can be used as a parameter for evaluating the stability of the traits in comparison to the central mean.

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

We thank Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq),

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), and Fundação Arthur Bernardes (FUNARBE) for granting financial support for the implementation of the research project.

REFERENCES

- Almeida LM, Viana AP, do Amaral Júnior AT and Júnior JBC (2014). Breeding full-sib families of sugar cane using selection index. *Cienc. Rural* 44: 605-611. <http://dx.doi.org/10.1590/S0103-84782014000400005>
- Almeida Filho JE, Tardin FD, Guimarães JF, Resende MD, et al. (2016). Multi-trait BLUP model indicates sorghum hybrids with genetic potential for agronomic and nutritional traits. *Genet. Mol. Res.* 15: <http://dx.doi.org/10.4238/gmr.15017071>.
- Banerjee V, Krishnan P, Das B, Verma APS, et al. (2015). Crop Status Index as an indicator of wheat crop growth condition under abiotic stress situations. *Field Crops Res.* 181: 16-31. <http://dx.doi.org/10.1016/j.fcr.2015.06.009>
- Bates D, Sarkar D, Bates MD and Matrix LT (2007). The lme4 package. R package version 2.
- Cameron ND (1997). Selection indices and prediction of genetic merit in animal breeding. CAB international, Wallingford.
- Castro RD, Peternelli LA, Resende MD, Marinho CD, et al. (2016). Selection between and within full-sib sugarcane families using the modified BLUPIS method (BLUPISM). *Genet. Mol. Res.* 15: <http://dx.doi.org/10.4238/gmr.15017334>.
- Chapola RC, Cruz JA, Nunes IK and Fernandes Júnior AR (2013). Censo varietal 2012. Editora CCA-UFSCar, Araras.
- Consecana (2006). Conselho dos produtores de cana-de-açúcar, açúcar e álcool do Estado de São Paulo. Manual de instruções, Piracicaba.
- Cruz CD, Regazzi AJ and Carneiro PCS (2012). Modelos biométricos aplicados ao melhoramento genético. Universidade Federal de Viçosa, Viçosa.
- Elston RC (1963). A weight-free index for the purpose of ranking or selection with respect to several traits at a time. *Biometrics* 19: 85-97. <http://dx.doi.org/10.2307/2527573>
- Espósito DP, Peternelli LA, de Paula TOM and Barbosa MHP (2012). Análise de trilha usando valores fenotípicos e genotípicos para componentes do rendimento na seleção de famílias de cana-de-açúcar. *Cienc. Rural* 42: 38-44. <http://dx.doi.org/10.1590/S0103-84782011005000152>
- Farias Neto JT and Resende MDV (2001). Application of the mixed model methodology (reml/blup) in variance components estimation and prediction of genetic values in peach palm (*Bactris gasipaes*). *Rev. Bras. Frutic.* 23: 320-324. <http://dx.doi.org/10.1590/S0100-29452001000200024>
- Habier D, Tetens J, Seefried FR, Lichtner P, et al. (2010). The impact of genetic relationship information on genomic breeding values in German Holstein cattle. *Genet. Sel. Evol.* 42: 5. <http://dx.doi.org/10.1186/1297-9686-42-5>
- Keefer P and Loayza N (2008). Terrorism, economic development, and political openness. Cambridge University Press, Cambridge.
- Laino P, Limonta M, Gerna D and Vaccino P (2015). Morpho-physiological and qualitative traits of a bread wheat collection spanning a century of breeding in Italy. *Biodivers. Data J.* 3: e4760. <http://dx.doi.org/10.3897/BDJ.3.e4760>
- Leite GMV, Andrade LAB, Garcia JC and Anjos IA (2008). Effects of sources and dosages of calcium silicate on yield and quality of sugarcane, cultivar SP80-1816. *Cienc. Agrotec.* 32: 1120-1125. <http://dx.doi.org/10.1590/S1413-70542008000400013>
- Li G, Li G, Yang X and Zhou M (2010). A comprehensive power quality evaluation model based on radar chart method. *Dianli Xitong Zidonghua* 34: 70-74.
- Lima DC, Abreu ÁFB, Ferreira RADC and Ramalho MAP (2015). Breeding common bean populations for traits using selection index. *Sci. Agric.* 72: 132-137. <http://dx.doi.org/10.1590/0103-9016-2014-0130>
- Lima LK, Ramalho MAP and Abreu AF (2012). Implications of the progeny x environment interaction in selection index involving characteristics of the common bean. *Genet. Mol. Res.* 11: 4093-4099. <http://dx.doi.org/10.4238/2012.September.19.5>
- Liu K, Wang Y, Luo C, Chen Z, et al. (2010). Application of radar chart analytical method in tobacco variety trial. *Chin. Tobacco Sci.* 6: 016.
- Lynch M and Walsh B (1998). Genetics and analysis of quantitative traits. Sinauer Associates, Sunderland.
- Mapa (2016). Ministério da agricultura, pecuária e abastecimento. Available at [<http://www.agricultura.gov.br/vegetal/culturas/cana-de-acucar>]. Accessed June 4, 2016.
- Mulamba NN and Mock JJ (1978). Improvement of yield potential of the Eto Blanco maize (*Zea mays* L.) population by breeding for plant traits. *Egypt. J. Genet. Cytol.* 7: 40-51.

- Munoz PR, Resende MFR, Huber DA, Quesada T, et al. (2014). Genomic relationship matrix for correcting pedigree errors in breeding populations: impact on genetic parameters and genomic selection accuracy. *Crop Sci.* 54: 1115-1123. <http://dx.doi.org/10.2135/cropsci2012.12.0673>
- Nunes JAR, Ramalho MAP and Abreu AFB (2005). Graphical method in studies of adaptability and stability of cultivars. *Annual Report.* 48: 182-183.
- Pedrozo CÂ, Benites FRG, Barbosa MHP, Resende MDV, et al. (2009). Eficiência de índices de seleção utilizando a metodologia REML/BLUP no melhoramento da cana-de-açúcar. *Sci. Agrar.* 10: 31-36.
- Piepho HP, Möhring J, Melchinger AE and Büchse A (2008). BLUP for phenotypic selection in plant breeding and variety testing. *Euphytica* 161: 209-228. <http://dx.doi.org/10.1007/s10681-007-9449-8>
- Reis CAF, Gonçalves FMA, Ramalho MAP and Rosado AM (2011). Seleção de progênies de eucalipto pelo índice Z por MQM e Blup. *Pesqui. Agropecu. Bras.* 46: 516-522. <http://dx.doi.org/10.1590/S0100-204X2011000500009>
- Resende MDV (2002). Genética biométrica e estatística no melhoramento de plantas perenes. Embrapa Informação Tecnológica, Brasília.
- Resende MDV (2007). Matemática e estatística na análise de experimentos e no melhoramento genético. Embrapa Florestas, Brasília.
- Resende MDV and Barbosa MHP (2005). Melhoramento genético de plantas de propagação assexuada. Embrapa Florestas, Brasília.
- Rowlingson B and Diggle P (2014). splancs: spatial and space-time point pattern analysis. R package version 2.01-36.
- Saary MJ (2008). Radar plots: a useful way for presenting multivariate health care data. *J. Clin. Epidemiol.* 61: 311-317. <http://dx.doi.org/10.1016/j.jclinepi.2007.04.021>
- Soriano A, Quiles R, Mariscal C and García Ruiz A (2005). Pig sire type and sex effects on carcass traits, meat quality and physicochemical and sensory characteristics of Serrano dry-cured ham. *J. Sci. Food Agric.* 85: 1914-1924. <http://dx.doi.org/10.1002/jsfa.2197>
- van der Westhuizen RR and van der Westhuizen J (2009). A proposed selection index for feedlot profitability based on estimated breeding values. *Genet. Mol. Res.* 8: 448-455. <http://dx.doi.org/10.4238/vol8-2gmr599>
- Vazquez AI, Bates DM, Rosa GJ, Gianola D, et al. (2010). Technical note: an R package for fitting generalized linear mixed models in animal breeding. *J. Anim. Sci.* 88: 497-504. <http://dx.doi.org/10.2527/jas.2009-1952>
- Wright S (1922). Coefficients of inbreeding and relationship. *Am. Nat.* 65: 330-338. <http://dx.doi.org/10.1086/279872>
- Zhou M and Joshi S (2012). Trends in broad sense heritability and implications for sugarcane breeding in South Africa. *Sugar Tech.* 14: 40-46. <http://dx.doi.org/10.1007/s12355-011-0128-7>

Supplementary material

Table S1. Statistical test of Scott Knott (SK) to check the differences between areas of polygons formed from the phenotypic values (PV), genotypic values (GV) and estimated breeding values (EBV) adopting 5% significance level. The statistical test was based on all areas formed by trait order permutation process.

Figure S1. The R codes to generate graphics and area index (AI).

Figure S2. Example application of the AI() function considering a hypothetical set of 20 genotypes and six traits. AI (dat=, dim=, perm=) arguments: dat: matrix with genotypes in rows and traits in columns; dim: number of charts per column and total number of columns; perm: logical argument, if TRUE: the function computes all permutations (combinations) of trait order; if FALSE: the function uses the original database.