

## Evaluation of different selection indices combining Pilodyn penetration and growth performance in *Eucalyptus* clones

Andrei Caíque Pires Nunes<sup>1\*</sup>, Marcos Deon Vilela de Resende<sup>2</sup>,  
Glêison Augusto dos Santos<sup>1</sup>, Rodrigo Silva Alves<sup>3</sup>

Crop Breeding and Applied Biotechnology  
17: 206-213, 2017  
Brazilian Society of Plant Breeding.  
Printed in Brazil  
<http://dx.doi.org/10.1590/1984-70332017v17n3a32>

**Abstract:** *The present study aimed to evaluate the selection indices efficiency for Pilodyn penetration combined with growth traits in Eucalyptus clones. It was carried out experiments in a randomized block design, with single tree plots and 30 replications. Diameter at breast height (DBH), total height (TH), and Pilodyn penetration as an indicator of basic density (BD) were measured. The volume was estimated. Based on predicted genotypic values, three indices presented the highest accuracies:  $I_8$  (based on partial correlation),  $I_7$  (based on the concept of multivariate BLUP) and  $I_3$  (based on two variables as ratio, which uses a third heritability estimate associated to the ratio DBH/BD, besides the two heritabilities of DBH and BD). Thus, it is possible to optimize the selection by combining properly the variables using their genetic control, precision and the relationships between them. The best options came from using only two no redundant traits DBH and BD.*

**Key words:** *Selection criteria, accuracy, wood quality, partial correlation, path analysis.*

### INTRODUCTION

Part of the increase in forest production is attributed to breeding programs (Costa et al. 2015). In this scenario, evaluation, characterization and selection of superior genotypes are critical steps in a forest breeding program that aims at maximizing genetic gain (Resende 2002, Bhering et al. 2015). Therefore, it is essential that the breeder carefully sets the goal of the selection, as well as the criteria that will be used to properly manage the breeding population and to generate significant gains in relation to the final commercial product.

Wood quality and volumetric analyses are essential in *Eucalyptus* improvement directed for pulp production (Gomide et al. 2010, Protásio et al. 2014). Although tree volume measurement is commonplace, wood quality studies are costly and time-consuming (Raymond and Apiolaza 2004). Basic density has been considered to be a universal index for assessing the quality of the wood, providing indirect information about other technological traits (Gomide et al. 2010). In spite of basic density importance, its determination is difficult due to high cost and need for tree felling.

Considering these fact, the Pilodyn method has been successfully applied for indirect estimation of wood basic density without to fall the tree (Gouvêa et

**\*Corresponding author:**  
E-mail: andreicaiquep@gmail.com

**Received:** 20 February 2016  
**Accepted:** 03 November 2016

<sup>1</sup> Universidade Federal de Viçosa (UFV), Departamento de Engenharia Florestal, Avenida Peter Henry Rolfs, s/n, Campus Universitário, 36.570-900, Viçosa, MG, Brazil

<sup>2</sup> Empresa Brasileira de Pesquisa Agropecuária, Centro Nacional de Pesquisa de Florestas, Estrada da Ribeira, km 111, Bairro Guaraituba, 83.411-000, Colombo, PR, Brazil

<sup>3</sup> UFV, Departamento de Biologia Geral,

al. 2011, Couto et al. 2013, Neves et al. 2013). With indirect estimation of basic density and information of tree growth (volume), the genetic selection of superior materials is technically correct.

Thus, in order to adopt more complex criteria to select superior genotypes considering several traits simultaneously, selection index theory can be employed. This theory was first described by Smith (1936), and later by Hazel (1943). According to Freitas et al. (2012), indices allow using a single value to select genotypes, since analysis is carried out by linear combinations of phenotypic data of many characters in study, whose weighting coefficients are estimated in order to maximize the correlation between the index and the true breeding values.

In addition, selection index is more efficient than direct selection, since it enables the distribution of gains between the several traits, in a more homogeneous way and in accordance with the purposes of the breeding program, generating higher total gain (Reis et al. 2011, Freitas et al. 2012, Cruz et al. 2014). Thus, these actions may greatly contribute to maximize the cost/benefit of the breeding program.

Several selection indices have been used for multivariate evaluation of characters of interest in different cultures, particularly *Eucalyptus* (Martins et al. 2006, Reis et al. 2011, 2015). Despite their importance, these indices are based on phenotypic values and do not consider the genotypic correlation and the cause and effect relationships between variables. According to Resende et al. (2014), the use of multivariate mixed models with multiple traits and unstructured covariance matrix is theoretically the most efficient method, and allows considering heterogeneity of variances and covariances. However, in practice, this approach is not used, due to the difficulty in convergence of the iterative analysis and the super parameterization (Resende et al. 2014).

Thus, structural equation modeling is an alternative that allows efficiently representing the standard multi-trait model (Resende et al. 2014). With a functional network of studied traits, it is possible to establish cause and effect relationships between the variables of interest and to compose optimum selection indices using the genotypic values predicted by univariate analysis by the REML/BLUP procedure (Maximum Restricted Likelihood/Best Linear Unbiased Prediction) (Resende et al. 2014, Viana and Resende 2014).

Therefore, the present study investigated the effectiveness of new classes of selection criteria based on partial correlations, direct effect of path analysis, ordinary correlations and heritability for *Eucalyptus*, using genotypic values predicted through the mixed models methodology.

## MATERIAL AND METHODS

### Experimental network

Experiments were carried out in the areas of CMPC Celulose Riograndense Company, in the municipalities of Minas do Leão (lat 30° 11' S, long 52° 00' W, alt 141 m asl, average temperature 17.5 °C and annual precipitation 1,422 mm), Encruzilhadas do Sul (lat 30° 27' S, long 52° 39' W, alt 250 m asl, average temperature 17 °C and annual precipitation 1,368 mm), Dom Feliciano (lat 30° 29' S, long 52° 19' W, alt 378 m asl, average temperature 16 °C and annual precipitation 1,564 mm) and Vila Nova do Sul (lat 30° 14' S, long 53° 49' W, alt 301 m asl, average temperature 16.8 °C and annual precipitation 1,133 mm), which are located in the state of Rio Grande do Sul (climate Cfa, according to the climatic classification of Koppen) Brazil. A network of clonal trials with 864 *Eucalyptus* clones was set in 2007. Trees were planted at a spacing of 3.5 x 2.6 m. At each site, it was established an experiment in a randomized block design, with single tree plots and 30 replications.

### Data collection

For the indirect estimate of basic density, Pilodyn's method (Greaves et al. 1996) was used when trees were three years old. The measurement with Pilodyn was carried out twice, on each north and south cardinal aspects of the tree. For analysis, the considered number was the mean of the two measurements. It was considered the inverse of Pilodyn penetration depth (mm) as the estimated basic density (BD).

Growth data of the trees were collected at three years of age, as well as the estimated basic density. It was measured the diameter at breast height (DBH), in centimeters (cm), and the total height (TH) of trees, in meters (m). DBH was

measured with the aid of a diameter tape, and TH was obtained using a relascope.

To calculate the volume (m<sup>3</sup>) without bark (Vol), the model of Leite et al. (1995) was used, as shown below:

$$Vol = 0.000048 \times DBH^{1.720493} \times TH^{1.180736} \times e^{(-3.00555) \times (tx/dbh) \times [1 - (d/dbh)^{1+0.228531 \times d}]} + E$$

In which DBH: diameter at 1.3 meters height; TH: total height; tx = 0, for the volume with bark and 1 for volume without bark; d: superior commercial diameter; E = experimental error.

### Statistical analysis

The statistical model for analysis of this experimental network in several environments (Resende 2002), with single tree plots is given by  $y = Xr + Zg + Hb + Wge + e$ , where:  $y$ ,  $r$ ,  $g$ ,  $h$ ,  $ge$  and  $e$  are vectors of data, replication effects (fixed), genotypic effects (random), block effects (random), effects of genotype x environment interaction (G x E) (random), and random errors, respectively. In addition,  $X$ ,  $Z$ ,  $H$  and  $W$  are the incidence matrices for  $r$ ,  $g$ ,  $h$  and  $ge$ , respectively. Predicted genotypic values free of interaction, considering all the environments were given by  $u + g$ , in which  $u$  is the mean of all sites. These values predicted for each variable, using univariate analysis, will be used in the selection indices. In addition, it was also obtained the genetic correlations between the analyzed variables. All analyses were carried out using the Selegen-REML/BLUP software (Resende 2016).

### Selection Indices

It follows a description of the index used in this paper (Table 1). Details of them and their accuracies are presented by Viana and Resende (2014) and Resende et al. (2014).

**Table 1.** Description of the indexes and their accuracies used

Concept	Index <sup>1</sup>	Description	Accuracy <sup>1</sup>
Phenotypic index	$I_1$	$I_1 = \left( \frac{DBH}{S_{DBH}} \right) \left( \frac{DB}{S_{DB}} \right)$	$r_{gg} = \sqrt{\frac{n \times h^2}{1 + (n-1) \times h^2}}$
	$I_2$	$I_2 = \left( \frac{Vol}{S_{Vol}} \right) \left( \frac{DB}{S_{DB}} \right)$	
Optimum index using a ratio between two variables	$I_3$	$I_3 = \text{Log}DBH - \text{Log}(1/BD)$	$r_{gg} = \sqrt{\frac{n \times h_{yr}^2}{1 + (n-1) \times h_{yr}^2}}$
	$I_4$	$I_4 = \text{Log}Vol - \text{Log}(1/BD)$	
Genotypic index	$I_5$	$I_5 = \left( \frac{VG_{DBH}}{S_{VG_{DBH}}} \right) \left( \frac{VG_{BD}}{S_{VG_{BD}}} \right)$	$Ac_{I_5} = \frac{sd(I_5) \times Ac(I_3)}{sd(I_3)}$
	$I_6$	$I_6 = \left( \frac{VG_{Vol}}{S_{VG_{Vol}}} \right) \left( \frac{VG_{BD}}{S_{VG_{BD}}} \right)$	$Ac_{I_6} = \frac{sd(I_6) \times Ac(I_4)}{sd(I_4)}$
Multivariate BLUP index	$I_7$	$I_7 = b_1g_0 + b_2g_{o1} + b_3g_{o2}$	$r_{gg} = \sqrt{\text{Var}(Index)/\sigma_g^2}$
Partial correlation index	$I_8$	$I_8 = b_1g_0 + b_2g_{o1} + b_3g_{o2}$	
Ordinary correlation index	$I_9$	$I_9 = b_1g_0 + b_2g_{o1} + b_3g_{o2}$	$r_{gg} = \frac{sd(score I_x)}{sd(score I_{greater})}$
Path analysis index	$I_{10}$	$I_{10} = b_1g_0 + b_2g_{o1} + b_3g_{o2}$	

<sup>1</sup> Description of the indexes components and accuracies: *BD*: basic density indirectly estimated by the Pilodyn's method; *DBH*: diameter at breast height;  $S_{DBH}$ : standard deviation of diameter at breast height;  $S_{BD}$ : standard deviation of basic density; *Vol*: volume of wood without bark;  $S_{Vol}$ : standard deviation of the volume without bark;  $\text{Log}(1/BD)$ : inverse of the basic density indirectly estimated by the Pilodyn's method on the logarithmic scale;  $\text{Log}DBH$ : diameter at breast height on the logarithmic scale;  $\text{Log}Vol$ : volume of wood without bark on the logarithmic scale;  $VG_{BD}$ : genotypic value of the indirectly estimated basic density;  $VG_{DBH}$ : genotypic value of the diameter at breast height;  $S_{VG_{DBH}}$ : standard deviation of the genotypic value of the diameter at breast height;  $S_{VG_{BD}}$ : standard deviation of the genotypic value of the basic density;  $VG_{Vol}$ : genotypic value of volume without bark;  $S_{VG_{Vol}}$ : standard deviation of the genotypic value of *Vol*;  $g_i$ : standard genotypic value of the objective character (*DBH* x *BD*);  $g_{oi}$  is the standard genotypic value of the auxiliary characters (*DBH* and *BD*). The weighting coefficients ( $b_i$ ) of the index are given by (Viana and Resende 2014);  $r_{gg}$ : accuracy of the index;  $n$ : number of individuals per clone;  $h_{yr}^2$ : heritability of the ratio between two variables given by Resende et al. (2014);  $h^2$ : heritability of the phenotypic index;  $Ac_{I_5}$ : accuracy of the  $I_5$  genotypic index;  $sd(I_5)$ : standard deviation of the  $I_5$  genotypic index score;  $Ac(I_3)$ : accuracy of the  $I_3$  optimum index;  $sd(I_3)$ : standard deviation of the  $I_3$  optimum index score;  $Ac_{I_6}$ : accuracy of the  $I_6$  genotypic index;  $sd(I_6)$ : standard deviation of the  $I_6$  genotypic index score;  $Ac(I_4)$ : accuracy of the  $I_4$  optimum index;  $sd(I_4)$ : standard deviation of the  $I_4$  optimum index score;  $\text{Var}(Index)/\sigma_g^2$ : ratio between the index and additive variances of the objective trait;  $sd(score I_x)$ : standard deviation of the  $I_7$ ,  $I_9$  or  $I_{10}$  indexes score;  $sd(score I_{greater})$ : standard deviation of the score of the index with greater variance among  $I_7$  and  $I_{10}$ .

### Internal consistency of the indices

As mentioned by Resende et al. (2014), a comparison between alternative selection indices may be carried out by varying the degree of covariance of the variables between each other. Thus, the Cronbach's alpha coefficient (1951) (modified by Resende et al. (2014)) works as an indicator of internal consistency of an index involving  $n$  variables. Its formula is given by:

$$\alpha = \frac{n-1}{n} \left( 1 - \frac{\sum v_i^2}{v_t^2} \right)$$

In which  $\sum v_i^2$  = sum of the variances of the  $n$  variables;  $v_t^2$  = total variance of the scores of the selection index;  $n$  = number of variables.

## RESULTS AND DISCUSSION

### Genetic parameters and genetic correlations

Genetic parameters of the analyzed traits were estimated (Table 2). Since the present study only aims to evaluate the different selection criteria, genetic parameters related to  $G \times E$  interactions are not reported here, but the full publication on that can be found in Nunes et al. (2016).

Estimates of individual heritability of the studied characters may be considered low ( $h_g^2 = 0.07$  for TH), moderate (from 0.15 for Vol, to 0.23 for  $I_1$ ), and high (0.59 for BD), Table 2, according to the classification reported by Resende (2002). Elevated heritability value for basic density (0.64) was found by Wei and Borralho (1997) in *Eucalyptus urophylla* S.T. Blake. Muneri and Raymond (2000) and Kube et al. (2001) also reported high values of heritability for Pilodyn penetration and basic density, ranging from 0.60 to 0.70. In spite of high genetic control of basic density, the heritability for growth traits have been reported in literature ranging from 0.10 to 0.22 (Kube et al. 2001), which corroborates with the present work.

Notwithstanding the estimate of broad sense individual heritability, it is observed that the value of this parameter for  $I_2$  (0.16) was similar to the heritability of Vol (0.15), Table 2. Thus, there is the need to develop an index which enables the estimate of a balanced heritability, i.e., that not only resembles to only one of the traits of the index. The same reasoning can be applied to  $I_1$ .

**Table 2.** Estimates of genetic parameters (individual REML) and genotypic correlations (below the genetic parameters) for basic density (BD in  $\text{kg m}^{-3}$ ), diameter at breast height (DBH in cm), total height (Th in m), volume (Vol in  $\text{m}^3 \text{ha}^{-1} \text{year}$ ), phenotypic index DBH $\times$ BD ( $I_1$ ), and phenotypic index Vol $\times$ BD ( $I_2$ ) for *Eucalyptus* clones evaluated in the joint analysis between environments, at three years of age

Parameters <sup>1</sup>	BD	DBH	Th	Vol	$I_1$	$I_2$
$h_g^2$	0.59	0.18	0.07	0.15	0.23	0.16
$h_{mg}^2$	0.95	0.70	0.54	0.65	0.78	0.67
Acgen	0.97	0.84	0.73	0.80	0.88	0.82
$c_{bloc}^2$	0.24	0.03	0.08	0.05	0.15	0.07
Overall mean	382.64	13.26	14.72	0.08	6.03	44.92
CV $_{gi}$ (%)	12.39	8.16	5.09	17.89	12.04	19.20
CV $_e$ (%)	9.70	15.71	16.50	37.80	20.20	40.08
CV $_r$	1.27	0.51	0.30	0.47	0.59	0.47
	BD	DBH	Th	Vol	$I_1$	$I_2$
BD	-	-0.27	-0.01	-0.22	0.67	0.26
DBH		-	0.66	0.97	0.49	0.81
Th			-	0.77	0.48	0.71
Vol				-	0.52	0.86
$I_1$					-	0.86
$I_2$						-

<sup>1</sup> Description of genetic parameters:  $h_g^2$ : coefficient of individual heritability in the broad-sense (corrected to variance of block), free from interaction;  $h_{mg}^2$ : heritability of clone mean; Acgen: genetic accuracy in clone selection;  $c_{bloc}^2$ : coefficient of determination of block; Overall mean: overall mean of characters between different environments; CV $_{gi}$  (%): coefficient of genotypic variation; CV $_e$  (%): coefficient of experimental variation; CV $_r$ : coefficient of relative variation.

Prediction accuracy of genetic values of the clones was high (Table 2). According to Resende and Duarte (2007), accuracies above 0.70 are sufficient for evaluations in a breeding population, and when the goal is the evaluation of the Value of Cultivation and Use, accuracies must be greater than 0.90. These high accuracy levels justify the great experimental quality, the caution and the technical precision in the establishment and evaluation of experiments. Moreover, the high number of replications (30) enabled obtaining reliable results of clones ranking by their predicted genetic values.

The value of coefficient of environmental variation (CVe) for  $I_2$  (40.08) was twice as higher as the CVe of  $I_1$  (20.20) (Table 2). The prediction accuracy of the breeding values of  $I_2$  (0.82) was relatively lower than the prediction accuracy of  $I_1$  (0.88). These results show that  $I_1$  is more accurate than  $I_2$ . This fact corroborates with the highest value of CVe of Vol in relation to the CVe of DBH (Table 2). Considering that the growth traits Vol and DBH make up the indices  $I_2$  and  $I_1$ , respectively, by multiplying by BD, the difference in accuracy and in CVe between these two indices is related to greater uncertainty in the estimate of Vol. Thus, since the estimate equation of Vol is composed of DBH and TH, the inclusion of the latter in the estimate of Vol led to higher value of CVe and lower accuracy of  $I_2$ , in relation to  $I_1$ .

Genotypic correlations between characters was estimated (Table 2) and it was found high correlation value between DBH and Vol (0.97). Nunes et al (2016) reported that is advantageous to perform the indirect selection of *Eucalyptus* clones aiming at gains in Vol through DBH. Negative values of genetic correlation between BD and DBH, and BD and Vol evidence the need for the study of selection indices involving wood quality and growth characters, simultaneously. Negative values (Kube et al. 2001, Bison et al. 2006) and positive values (Paula et al. 2002, Reis et al. 2011) of genetic correlation between basic density in *Eucalyptus* and growth characters were reported by different authors. Thus, discrepancies in genetic correlations in each cited work are caused by the genetic variation that exists in the evaluated population and by the different genes that are segregating in relation to the control of growth and wood quality characters (Reis et al. 2011).

### Heritability and correlations of optimal indices based on a ratio between two variables

The genetic analyses of a trait using variables as a ratio seems to be unused in forest tree breeding so far. This paper is the first one to evaluate its effectiveness. The results have shown that it is a very promising technique. It was ranked among the three best out of the ten selection indices evaluated.

When comparing the heritability of the phenotypic indices (Table 2),  $I_1$  (0.23) and  $I_2$  (0.16), with the heritability of  $I_3$  and  $I_4$  (Table 3), an increase is observed in the genetic control, and therefore, greater efficiency of these last indices is also observed. For  $I_3$ , there was a 37% increase in relation to the phenotypic heritability index, while for  $I_4$ , this increase was 18%. These results show that  $I_3$  and  $I_4$  were more efficient than the phenotypic indices in weighing the genotypic values for each variable under study, and thus they were more efficient in weighting the effects of each variable on the index as a whole.

It is observed that the calculated heritability (optimum) of  $I_4$  was lower than the heritability of  $I_3$  (Table 3). This fact can be explained since the elasticity coefficient ( $K^2$ ) of  $I_4$  (0.53) was lower than the  $K^2$  of  $I_3$  (1.05) (Table 3). Thus, the  $I_4$  had its calculated heritability penalized by the higher variance of volume (Vol), when compared with the variance of the diameter at breast height (DBH). Therefore,  $I_3$  index is more accurate and ideal for selection of superior genotypes in relation to  $I_4$ , due to greater accuracy in the measurement of DBH, when compared with the estimate of Vol.

Correlations between indices based on a ratio between two traits and their constituent variables were calculated (Table 3). Genotype correlations of the constituent variables

**Table 3.** Heritabilities and correlations of a ratio between two variables

Coefficients <sup>1</sup>	$I_3$	$I_4$
$h_{y^*}^2$	0.60	0.34
$r_{y^*W^*}^g$	0.58	0.65
$r_{y^*X^*}^g$	-0.63	-0.61
$r_{y^*W^*}^p$	0.30	0.70
$r_{y^*X^*}^p$	-0.37	-0.18
$k^2$	1.05	0.53
$k_h^2$	1.02	1.77
$k$	1.03	0.72
$k_h$	1.01	1.33

<sup>1</sup> Description of the coefficients:  $h_{y^*}^2$ : heritability of the ratio between two variables (optimum index);  $r_{y^*W^*}^g$ : genotypic correlation of the index with the variable  $W^*$ , which for  $I_3$  is DBH, and for  $I_4$  is Vol;  $r_{y^*X^*}^g$ : genotypic correlation of the index with variable  $X^*$ , which for both indices is 1/BD;  $r_{y^*W^*}^p$ : phenotypic correlation of the index with variable  $W^*$ , which for  $I_3$  is DBH, and for  $I_4$  is Vol;  $r_{y^*X^*}^p$ : phenotypic correlation of the index with the variable  $X^*$ , which for both indices is 1/BD;  $k^2$ : elasticity or relationship between variances, being the phenotypic variance of  $X$  (always BD) in the numerator, and the phenotypic variance of  $W$  (DBH or Vol) in the denominator;  $k_h^2$ : ratio between heritability of variable  $X$  (BD) estimated in the original scale and the calculated heritability of the ratio between two variables;  $k$ : square root of  $k^2$ ;  $k_h$ : square root of  $k_h^2$ . As mentioned in the material and methods, coefficient details can be found in Resende et al. (2014).

of each phenotypic index with these same indices (Table 2) were different from these obtained for  $I_3$  and  $I_4$ . The negative correlation between BD and  $I_3/I_4$  is justified, since for calculating the ratio between two variables, it was necessary to carry out analyses considering  $1/BD$ . Therefore, for comparison, it should be noted the magnitude of the correlation, not the negative signal. Thus, it is evident that indices based on the ratio between two variables do not present high correlation with only one of the component character of this index, as found for  $I_1$  and  $I_2$ .  $I_3$  and  $I_4$  indices provide a better balance between the two variables that compose it.  $I_3$  presented genotypic correlation with DBH in the order of 0.58 and with  $1/BD$  of -0.63 (Table 3), while  $I_4$  genetic correlation with DBH and BD is 0.49 and 0.67, respectively (Table 2).

### Efficiency of selection indices

According to the accuracy values, it can be concluded that the most effective indices are  $I_8$ ,  $I_7$ ,  $I_3$ ,  $I_5$  and  $I_{10}$ , while  $I_6$  presents the lowest accuracy value, along with  $I_9$  (Table 4). In general, prediction accuracies were high. This result was obtained due to the high number of replications and consistency in setting up and running the experiment. According to Resende and Duarte (2007), accuracies above 0.90 are considered too high and ensure reliable selection of superior genetic materials, as obtained for  $I_8$ ,  $I_7$ ,  $I_3$ ,  $I_5$ ,  $I_{10}$  and  $I_4$ .

This result corroborates the fact that partial correlations and path analyses ( $I_8$  and  $I_{10}$ ) are more efficient procedures than the ordinary correlations of Pearson ( $I_9$ ), since they are conditional correlations, unlike the latter (Cruz et al. 2014, Resende et al. 2014). Thus, in the composition of  $I_8$  and  $I_{10}$  indices, genotypic values are optimally weighted, and the considered correlations are odd (Resende et al. 2014), and there is no overestimation or underestimation of the index score.

As a report by Resende et al. (2014), for the analysis of a multivariate vector of observations of several traits, the multivariate mixed model is theoretically the most efficient, since it allows considering the complete heterogeneity of variance and covariance. Also, according to these authors, in practice, the use of the multivariate mixed model does not apply, due to the problematic convergence of the iterative analysis and super parametrization. Thus, it is important to use optimum selection indices that incorporate the concept of multivariate BLUP, by global maximization; the use of genotypic correlations and heritabilities; as well as the indices based on the concept of structural equation (Resende et al. 2014). This new approach becomes crucial for the optimization of the selection process, since it generates the same results of the multivariate mixed model, with less effort and high accuracy. In this context, the indices developed in this work can be used for any species and in any situation, in order to optimize the process of selection of superior genetic materials.

According to Resende et al. (2014), the use of structural equations (path analysis) is similar to the use of partial correlation matrices, instead of total correlations. This reduces the complexity of the multivariate mixed model, since it works with clean correlations between each pair of variables, making full rank the covariance matrices (Resende et al. 2014). In the present study, the index of greater efficiency was  $I_8$ , which is based on partial correlations between  $DBH \times BD$  with DBH and BD. However,  $I_{10}$ , based on the direct effect of path analysis also showed high accuracy. Thus, it is verified the equivalence of the use of structural equation models and partial correlations, since the path analysis depends on the partial correlations, as reported by Resende et al. (2014).

Among the three best indices,  $I_3$  was slightly higher, since it has greater internal consistency measured by the alpha coefficient. Internal consistency of an index may be studied by the degree of covariance of the variables between each other. Cronbach's alpha coefficient (1951) can be used as an indication of consistency of an index involving these variables (Resende et al. 2014). Thus, the higher the value of the coefficient, the more reliable is the index.  $I_3$ ,  $I_5$  and  $I_1$  presented the highest internal consistencies (0.50). According to Resende et al. (2014), the lower the specific variance of each variable and the higher the total variance that they produce together, the higher is the alpha coefficient. Thus, when the sum of the variances of the individual variables is reduced, it increases the variance they have in common, that is, the one that ensures the coherence or internal consistency of the index (Resende et al. 2014). Therefore, it is verified that the constituent variables of these indices combine well, i.e., they covary in the index to which they belong (Resende et al. 2014).

The  $I_5$  index presented high accuracy, and was higher than  $I_1$  (Table 4). The index with the use of genotypic values proves to be more efficient, when compared to the phenotypic index. Resende (2002) reports that the use of genotypic

values will be advantageous when genetic correlations of the variables are close to zero, and when prediction accuracies of genetic values for each character, individually, are high. Such conditions for success in the use of genotypic index are found in the present study, in which the genetic correlation between DBH and BD is null (-0.27) (Table 2), and the prediction accuracies for these two characters are high (0.97 for BD, and 0.84 for DBH) (Table 2). If these conditions are not met,  $I_8$ ,  $I_7$  or indices based on a ratio between two variables should be used, since they take into account the accuracy of each trait, their correlation and the relationship between variances of the constituent variables of the index.

Genotypic index in function of DBH×BD ( $I_5$ ), which considers only the heritability of the characters, was as efficient as  $I_3$ , which optimally weighs the genotypic values by heritability and correlations between variables. However, since  $I_3$  is optimum by the above mentioned reasons, it was more efficient (Table 4), and should be used especially in experiments in which the constituent variables of the index have non-null genotype correlations. Contrary to what happened to  $I_5$  (0.95 accuracy),  $I_6$  (0.78 accuracy) is the least efficient index. This result can be explained given that the conditions for use of the genotypic indices cited by Resende et al. (2002) are not met in the case of  $I_6$ , due to the inherent imprecision of Vol.

For the first time in the literature, it was compared so many selection indices, using somehow different concepts such as ratio between two variables, multivariate BLUP, partial correlations and direct effects of path analysis as part of the calculation of the weights. Path analysis is of great importance for the identification of direct and indirect effects of given characters in an objective variable (Cruz et al. 2014). Coefficients of path analysis were estimated based on ordinary genetic correlations of the characters, based on predicted genotypic values calculated by the REML/BLUP approach. Silva et al. (2009) report that path analysis becomes more effective when it is based on predicted genotypic values than when it is applied on phenotypic values. Thus, efficiency of the breeding program increases. Brasileiro et al. (2013), studied the consistency of path analyses using phenotypic and genotypic correlation, and concluded that in unbalanced cases, the use of genetic correlation produces more consistent results. Thus, it is noteworthy the precision of the analysis carried out in this study by the use of genotypic correlations obtained via REML/BLUP.

The five most efficient indices were those based on partial correlation, on the concept of multivariate BLUP, on two variables as a ratio and direct effects from path analysis. Basically, this can be attributed to the use of the following basic quantities: the genetic control of the trait, reliability and precision of the predicted genotypic values, and partial correlations between traits and the breeding objective. For the two variable as a ratio an additional feature is taken into account, the heritability of a third variable which is the own ratio. Thus, the use of indices based on these cited concepts are efficient and effective alternatives in selecting superior *Eucalyptus* genotypes based on several characters, without the complex procedures of multivariate mixed models.

## ACKNOWLEDGMENTS

The authors thank the Universidade Federal de Viçosa, CNPq, Capes and the company CMPC Celulose Riograndense.

## REFERENCES

Bhering LL, Cruz CD, Peixoto LAP, Rosado AM, Laviola BG and Nascimento M (2015) Application of neural networks to predict volume in eucalyptus. *Crop Breeding and Applied Biotechnology* 15: 125-131.

**Table 4.** Accuracies and Cronbach's alpha coefficients (1951) for  $I_1$ ,  $I_2$ ,  $I_3$ ,  $I_4$ ,  $I_5$ ,  $I_6$ ,  $I_7$ ,  $I_8$ ,  $I_9$ ,  $I_{10}$  indices, which are based on the concepts of phenotypic index, ratio between two variables (Resende et al. 2014), genotypic index, multivariate blup (Viana and Resende 2014), partial correlation (Viana and Resende 2014), ordinary correlation and direct effects of path analysis (Resende et al. 2014), respectively

Concept	Index	Accuracy	Alfa <sup>4</sup>
Partial Correlation	$I_8$	1.00	0.18
Mult Blup <sup>2</sup>	$I_7$	0.97	0.20
RBTV <sup>1</sup>	$I_3$	0.96	0.50
Genotypic	$I_5$	0.95	0.50
Path <sup>3</sup>	$I_{10}$	0.92	0.20
RBTV <sup>1</sup>	$I_4$	0.91	-. <sup>5</sup>
Phenotypic	$I_1$	0.88	0.50
Phenotypic	$I_2$	0.82	-. <sup>5</sup>
Ordinary correlation	$I_9$	0.79	0.21
Genotypic	$I_6$	0.78	0.49

<sup>1</sup> RBTV: Ratio between two variables; <sup>2</sup> Multi Blup: multivariate Blup; <sup>3</sup> Path: direct effect of path analysis; <sup>4</sup> Cronbach's alpha coefficient (1951) modified by Resende et al. (2014); <sup>5</sup> Negative value.

## Evaluation of different selection indices combining Pilodyn penetration and growth performance in *Eucalyptus* clones

- Brasileiro BP, Peternelli LA and Barbosa MHP (2013) Consistency of the results of path analysis among sugarcane experiments. **Crop Breeding and Applied Biotechnology** **13**: 113-119.
- Costa RB, Martinez DT, Silva JC and Almeida BC (2015) Variabilidade e ganhos genéticos com diferentes métodos de seleção em progênes de *Eucalyptus camaldulensis*. **Revista de Ciências Agrárias** **58**: 69-74.
- Couto AM, Trugilho PF, Neves TA, Protásio TP and Sá VA (2013) Modeling of basic density of wood from *Eucalyptus grandis* and *Eucalyptus urophylla* using nondestructive methods. **Cerne** **19**: 27-34.
- Cruz CD, Carneiro PCS and Regazzi AJ (2014) **Modelos biométricos aplicados ao melhoramento genético**. Vol 2, Editora UFV, Viçosa, 668p.
- Freitas JPX, Oliveira EJ, Jesus ON, Neto AJC and Santos LR (2012) Formação de população base para seleção recorrente em maracujazeiro amarelo com uso de índices de seleção. **Pesquisa Agropecuária Brasileira** **47**: 393-401.
- Gomide JL, Neto HF and Regazzi AJ (2010) Análise de critérios de qualidade da madeira de eucalipto para produção de celulose kraft. **Revista Árvore** **34**: 339-344.
- Gouvêa AFG, Trugilho PF, Gomide JL, Silva JRM, Andrade CR and Alves ICN (2011) Determinação da densidade básica da madeira de *Eucalyptus* por diferentes métodos não destrutivos. **Revista Árvore** **35**: 349-358.
- Greaves BL, Borralho NMG, Raymond CA and Farrington A (1996) Use of a pilodyn for the indirect selection of basic density in *Eucalyptus nitens*. **Canadian Journal of Forest Research** **26**: 1643-1650.
- Hazel LV (1943) The genetic basis for constructing selection indexes. **Genetics** **28**: 476-490.
- Kube PD, Raymond CA and Banham PW (2001) Genetic parameters for diameter, basic density, cellulose content and fibre properties for *Eucalyptus nitens*. **Forest Genetics** **8**: 285-294.
- Leite HG, Guimarães DP and Campos JCC (1995) Descrição e emprego de um modelo para estimar múltiplos volumes de árvores. **Revista Árvore** **19**: 75-79.
- Martins IS, Martins RCC and Pinho DS (2006) Alternativas de índices de seleção em uma população de *Eucalyptus grandis* Hill ex Maiden. **Cerne** **12**: 287-291.
- Muneri A and Raymond CA (2000) Genetic parameters and genotype-by-environment interactions for basic density, pilodyn penetration and stem diameter in *Eucalyptus globulus*. **Forest Genetics** **7**: 317-328.
- Neves AT, Protásio TP, Trugilho PF, Valle MLA, Sousa LC and Vieira CMM (2013) Qualidade da madeira de clones de *Eucalyptus* em diferentes idades para a produção de bioenergia. **Revista de Ciências Agrárias** **56**: 139-148.
- Nunes ACP, Santos GA, Resende MDV, Silva LD, Higa A and Assis TF (2016) Estabelecimento de zonas de melhoramento para clones de eucalipto no Rio Grande do Sul. **Scientia Forestalis** **44**: 00-00.
- Paula RC, Pires IE, Borges RCG and Cruz CD (2002) Predição de ganhos genéticos em melhoramento florestal. **Pesquisa Agropecuária Brasileira** **37**: 159-165.
- Protásio TP, Goulart SL, Neves TA, Trugilho PF, Ramalho FMG and Queiroz LMRSB (2014) Qualidade da madeira e do carvão vegetal oriundos de floresta plantada em Minas Gerais. **Pesquisa Florestal Brasileira** **34**: 111-123.
- Raymond CA and Apiolaza LA (2004) Incorporating wood quality and deployment traits in *Eucalyptus globulus* and *Eucalyptus nitens*. In Walter C and Carson M (org) **Plantation forest biotechnology for the 21st century**. Research Signpost, Kerala, India, p. 87-99.
- Reis CAF, Gonçalves FMA, Ramalho MAP and Rosado AM (2011) Seleção de progênes de eucalipto pelo índice Z por MQM e Blup. **Pesquisa Agropecuária Brasileira** **46**: 517-523.
- Reis CAF, Gonçalves FMA, Ramalho MAP and Rosado AM (2015) Estratégias na seleção simultânea de vários caracteres no melhoramento do *Eucalyptus*. **Ciência Florestal** **25**: 457-467.
- Resende MDV (2002) **Genética biométrica e estatística no melhoramento de plantas perenes**. Editora Embrapa, Brasília, 975p.
- Resende MDV and Duarte JB (2007) Precisão e controle de qualidade em experimentos de avaliação de cultivares. **Pesquisa Agropecuária Tropical** **37**: 182-194.
- Resende MDV, Silva FF and Azevedo CF (2014) **Estatística matemática, biométrica e computacional: modelos mistos, multivariados, categóricos e generalizados (reml/blup), inferência bayesiana, regressão aleatória, seleção genômica, qtl-gwas, estatística espacial e temporal, competição, sobrevivência**. Editora Suprema, Viçosa, 881p.
- Resende MDV (2016) Software Selegen-REML/BLUP: a useful tool for plant breeding. **Crop Breeding and Applied Biotechnology** **16**: 330-339.
- Silva FL, Pedrozo CA, Barbosa MHP, Resende MDV, Peternelli LA, Costa PMA and Vieira MS (2009) Análise de trilha para os componentes de produção de cana-de-açúcar via blup. **Ceres** **56**: 308-314.
- Smith HF (1936) A discriminant function for plant selection. **Annals of Eugenics** **7**: 240-250.
- Viana AP and Resende MDV (2014) **Genética quantitativa no melhoramento de fruteiras**. Interciência, Rio de Janeiro, 296p.
- Wei X and Borralho NMG (1997) genetic control of wood basic density and bark thickness and their relationships with growth traits of *eucalyptus urophylla* in south east china. **Silvae Genetica** **46**: 1997.