

Selection of maize hybrids: an approach with multi-trait, multi-environment, and ideotype-design

Marco Antônio Peixoto¹, Igor Ferreira Coelho¹, Jeniffer Santana Pinto Coelho Evangelista¹, Suellen Sales de Oliveira Santos¹, Rodrigo Silva Alves¹, Jefferson Fernando Naves Pinto², Edésio Fialho dos Reis² and Leonardo Lopes Bhering^{1*}

Abstract: *The present study aimed to evaluate the applicability and efficiency of the FAI-BLUP index in the genetic selection of maize hybrids, using 84 maize hybrids that were evaluated for cycle, morphology, and yield traits in four environments. Models accounting for homogeneous and heterogeneous residual variances were tested, and variance components were estimated using the residual maximum likelihood. Genotypic values were predicted by best linear unbiased prediction, and factor analysis was applied to group the traits. The FAI-BLUP index was used for the selection of maize hybrids based on ideotype design. Three factors explained more than 70% of genotypic variability, with selective accuracies varying from low (0.46) to high (0.99). Predicted genetic gains were positive for traits related to yield and negative for traits related to cycle and morphology, as is desirable in maize crop.*

Keywords: *FAI-BLUP index, factor analysis, residual variance structures, multivariate analysis, genetic selection.*

INTRODUCTION


The growing demand for superior genotypes has led maize breeders to seek auxiliary techniques in the selection process. In plant breeding, multi-environment trials (MET) are useful for evaluating genotypes, testing their performance in a range of environments, and selecting the most superior (Alves et al. 2020). In MET, the genotype-by-environment (G×E) interaction is a factor influencing the performance of genotypes under environmental variation (Resende 2015), resulting in a change in genotyping ranking over different environments, which makes genetic selection difficult (Sripathi et al. 2018). In general, various traits have been evaluated in maize breeding, with the aim of supporting the selection and recommendation of the ideotype (*i.e.*, genotypes with simultaneous superior performance in many traits).

Multivariate analysis, which allows genetic selection based on a set of traits, is an important procedure when dealing with MET data. This multi-trait selection is relevant, as superior varieties combine optimal attributes for several traits simultaneously. One possibility in multivariate analysis is to explore and reduce data dimensionality by grouping traits (Yan and Frégeau-Reid 2018). However,

Crop Breeding and Applied Biotechnology
21(2): e34582122, 2021
Brazilian Society of Plant Breeding.
Printed in Brazil
<http://dx.doi.org/10.1590/1984-70332021v21n2a31>

***Corresponding author:**

E-mail: leonardo.bhering@ufv.br

 ORCID: 0000-0002-6072-0996

Received: 17 September 2020

Accepted: 12 June 2021

Published: 20 July 2021

¹ Universidade Federal de Viçosa, Departamento de Biologia Geral, Viçosa, 36.570-900, MG, Brazil

² Universidade Federal de Jataí, Departamento de Agronomia, Jataí, 75.801-615, Goiás, Brazil

combining several traits in an efficient structure of selection indexes is often a complex and difficult task (Paiva et al. 2020) because there is a reduction in selection gains as the number of traits assessed increases (Cruz et al. 2014).

An outstanding method to deal with MET data is factor analysis (FA) (Nuvunga et al. 2015, Peixoto et al. 2016, Barbosa et al. 2019). This approach allows the selection of genotypes considering both multi-trait and multi-environment information, and overcomes the loss of biological meaning by identifying a smaller group of latent variables (Rocha et al. 2018). The basic aim is to group highly correlated traits in common factors (latent variables) that contain a large amount of information about trait interrelation (Barbosa et al. 2019, Woyann et al. 2020). In FA, the factors that best explain the set of traits analyzed are established. These factors are then isolated to summarize the information and simplify subsequent analyses.

The factor analysis and genotype-ideotype design (FAI-BLUP) index (Rocha et al. 2018) has been identified as an efficient method for genetic selection considering multi-trait, multi-environment, and ideotype design information. It has the advantage of including the correlation structure between traits and the direction of selection identified by the breeder to select genotypes closer to the ideotype (Rocha et al. 2018). This methodology combines the use of FA and mixed models. One of the main advantages of this method is the ability to incorporate the predicted genetic values (BLUPs) in the analysis (Oliveira et al. 2019, Rocha et al. 2019, Woyann et al. 2019). The REML/BLUP methodology is considered the standard procedure for genetic evaluation in crop breeding (Resende 2016, Coelho et al. 2020, Peixoto et al. 2020), even though in some specific cases, the methods based on least squares return the same results.

While the application of this method in maize is not well documented, it has been implemented in other crops, such as the common bean, soybean, and sorghum (Silva et al. 2018, Oliveira et al. 2019, Rocha et al. 2019, Woyann et al. 2020). In this study, several maize hybrid traits were evaluated in four environments to: i) explore the residual variance structures for the prediction of genotypic values in MET; ii) investigate the relationship between traits through factor analysis; and iii) use the factors (latent variables) for the selection of hybrids through the FAI-BLUP index.

MATERIAL AND METHODS

The experiment was carried out between January and July 2018 at four sites (considered as environments: ENV1, ENV2, ENV3, and ENV4), located in the southwest of the state of Goiás, Brazil (Supplementary material, Appendix A). The experimental network consisted of 78 interpopulation hybrids and six commercial hybrids most used in the region (AS 1633, P3646, 30F53, BM 709, P4285, and 30K75), totaling 84 hybrids assessed. In each environment, trials were conducted using a complete block design with three replications and 44 plants per plot. Plots consisted of four 4-m rows, with a spacing of 0.40 m between plants and 0.45 m between rows. To eliminate the competition effects of each plot with its neighbors, only the two central rows were evaluated.

The following cycle-related traits were assessed: time to tasseling (TA) and time to female flowering (FL). TA refers to the days from seeding to 50% of tasseling (when more than half of the tassel releases pollen), and FL refers to the days from seeding to 50% of flowering (when the ear starts silking, and it is possible to see the silk outside of the husk). The following morphology-related traits were assessed: plant height (PH) and ear height (EH), measured in meters from the ground to the flag leaf. Ear length (EL) and ear diameter (ED) were two of the yield-related traits evaluated, based on the mean length (EL) in centimeters, of a row of five unhusked ears, and mean diameter (ED), in centimeters at the center of each of the five, side by side. Ear yield (EY) and grain yield (GY) were also chosen as yield-related traits, using the weight of the plot, with 13% moisture content, converted to hectares, kg ha^{-1} .

The estimation of variance components and prediction of genotypic values for the traits assessed were made through the residual maximum likelihood/best linear unbiased prediction (REML/BLUP) procedure, according to (Patterson and Thompson 1971, Henderson 1975). The statistical model associated with the evaluation of hybrids in a randomized complete block design with one observation per plot is given by the following equation:

$$y = Xr + Zg + e$$

where y is the vector of phenotypes; r is the vector of replication effects (assumed as fixed), added to the overall mean; g is the vector genotypic effects [(assumed as random) ($g \sim N(0, \sigma_g^2)$, where σ_g^2 is the genotypic variance)]; and e is the vector of residuals [(random) ($e \sim N(0, \sigma_e^2)$, where σ_e^2 is the residual variance)]. Uppercase letters (X and Z) represent the incidence matrices for r and g , respectively.

The statistical model associated with the evaluation of hybrids in a complete block design with one observation per plot in several environments is given by the following equation:

$$y = Xr + Zg + Wi + e$$

where y is the vector of phenotypes; r is the vector of replication-environment combinations (assumed as fixed), which comprises the effects of environment and replication within environment, added to the overall mean; g is the vector of genotypic effects [(assumed as random) ($g \sim N(0, \sigma_g^2$)]; i is the vector of G×E interaction effects [(random) ($i \sim N(0, \sigma_{ge}^2$) where σ_{ge}^2 is the G×E interaction variance]; and e is the vector of residuals [(random) ($e \sim N(0, \sigma_e^2$)]. Uppercase letters (X , Z , and W) represent the incidence matrices for the r , g and i effects, respectively.

For the random effects, significance was tested by the likelihood ratio test (LRT) using the chi-square statistics with 1 degree of freedom and 5% probability of error type I. Models with different residual variance structures (homogeneous and heterogeneous) were tested using the Bayesian information criterion (BIC) (Schwarz 1978).

The selective accuracies (r_{gg}) were calculated using the following equation (Resende et al. 2014): $r_{gg} = \sqrt{1 - \frac{PEV}{\sigma_g^2}}$, where PEV is the prediction error variance extracted from the diagonal of the generalized inverse of the coefficient matrix of the mixed model equations.

The eight traits analyzed were summarized in factors by FA. The factor analysis model is expressed by the following equation:

$$X_j = I_{j1} F_1 + I_{j2} F_2 + \dots + I_{jm} F_m + \varepsilon_j$$

where X_j is the j th environment, with $j = 1, 2, \dots$; I_{jk} being the factorial load for the j th environment, associated with the k th factor, where $k = 1, 2, \dots, m$; F_k is the k th common factor; ε is the specific factor. The number of factors formed was established such that the average proportion of variance of each environment explained by the common factors, or average commonality, reached at least 70%. The main goal of factor analysis is to describe the original variability of genotypic observations (BLUPs) in terms of a smaller number of random variables, called factors (Cruz et al. 2014). Therefore, the analysis starts with many variables that are reduced to a smaller number of latent variables (factors) representing the original variability. The rotation of the loads from the FA's latent variables (*varimax* rotation) was used (Mardia et al. 1979), with the aim of facilitating the interpretation of these latent variables.

To select superior hybrids containing a set of traits that approach the ideotype, the FAI-BLUP index was applied using the predicted genotypic values (BLUPs) (Rocha et al. 2018). The number of ideotypes was determined based on desirable and undesirable factors for the traits under selection, and the distance between the genotype and the proposed ideotype was obtained. This distance was converted into a spatial probability, enabling hybrid ranking. The ideotype design used for the FAI-BLUP analysis was assigned as “maximum” for the traits related with productivity (EL, ED, EY, and GY), and “minimum” for cycle- and morphology-related traits (TA, FL, PH, and EH). To compare the efficiency of selection in the FAI-BLUP index, direct selection based on the main trait in maize crop (GY) was performed for each environment, and the harmonic means of relative performance of genotypic values (HMRPGV) (Resende et al. 2014) were assessed over the four environments considering the GY trait (selection based on adaptability, stability, and productivity). In addition, indirect selection through HMRPGV (based on the GY trait) was conducted, and selection gains were calculated for all traits.

The selection gain (SG), considering 20% as selective intensity (17 hybrids), was obtained as follows (Resende et al. 2014): $SG = \frac{\sum_{i=1}^n GV}{p}$, where GV is the genotypic value and p is the number of selected genotypes. The Kappa coefficient (K) (Cohen 1960) was applied to calculate the agreements among selected hybrids by all strategies used, as follows: $K = \frac{A - C}{D - C} * 100$, where A is the number of selected hybrids, between pairs of environments; C is the number of selected hybrids due to chance ($C = bD$, where b is the selection intensity = 0.2) and D is the number of selected hybrids (17).

All analyses were performed using the software Selegen REML/BLUP (Resende 2016), ASReml (Gilmour et al. 2015), RBio (Bhering 2017), and R (R Development Core Team 2020).

RESULTS AND DISCUSSION

The LRT indicated the existence of genotypic variability among hybrids, in both individual and joint analyses, for most traits (Supplementary material, Appendix B). In the joint analysis, only two traits (EY and GY) showed significant G×E interaction effects. The G×E interaction effects did not exhaust all genotypic variability of both traits; hence, the genotypic effects were significant. Regarding the residual variance structure, for EL and FL, the model with heterogeneous residual variance presented the best goodness-of-fit, according to BIC, whereas the remaining traits (TA, PH, EH, ED, EY, and GY) were best represented by the model with homogeneous residual variance (Table 1). Therefore, the superior model that accounted for a specific residual variance structure for each trait, *i.e.*, the model that presented the lowest BIC value, was applied for the estimation of variance components and for the prediction of genotypic values.

The mean selective accuracy (\bar{r}_{gp}) of the hybrids ranged from 0.46 (FL in ENV2) to 0.99 (FL in the joint analysis) (Figure 1). In addition, the joint analysis presented higher values of selective accuracy for all traits analyzed. Similar patterns have been reported in the literature for maize hybrids (Coelho et al. 2020). Selective accuracy measures the closeness of the genetic values estimated to the real genetic value, as a correlation between these values (Resende 2015). It depends on the number of replications, residual variance, and proportion of residual and genetic variation (Resende and Duarte 2007). In this context, increasing the number of environments by considering the joint analysis implies the maximization of selective accuracy.

In addition, the values of mean selective accuracy increased in the joint analysis compared with the individual analysis for FL and EL traits (Figure 1), where the model with the best fit was related to heterogeneous residual variance. In this case, the selective accuracy was affected by the residual

Table 1. Bayesian information criterion (BIC) for the models with homogeneous and heterogeneous residual variance structure for the traits: time to tasseling (TA); time to flowering (FL); plant height (PH); ear height (EH); ear length (EL); ear diameter (ED); ear yield (EY); grain yield (GY). Bold values indicate the model that presented the best fit for each trait

Traits	Residual variance structure	BIC
TA	Homogeneous	1682.34
	Heterogeneous	1697.41
FL	Homogeneous	1865.04
	Heterogeneous	1825.57
PH	Homogeneous	-3574.45
	Heterogeneous	-3573.04
EH	Homogeneous	-3895.83
	Heterogeneous	-3883.18
EL	Homogeneous	4566.10
	Heterogeneous	4496.58
ED	Homogeneous	1065.92
	Heterogeneous	1084.00
EY	Homogeneous	15245.94
	Heterogeneous	15252.97
GY	Homogeneous	14766.72
	Heterogeneous	14773.14

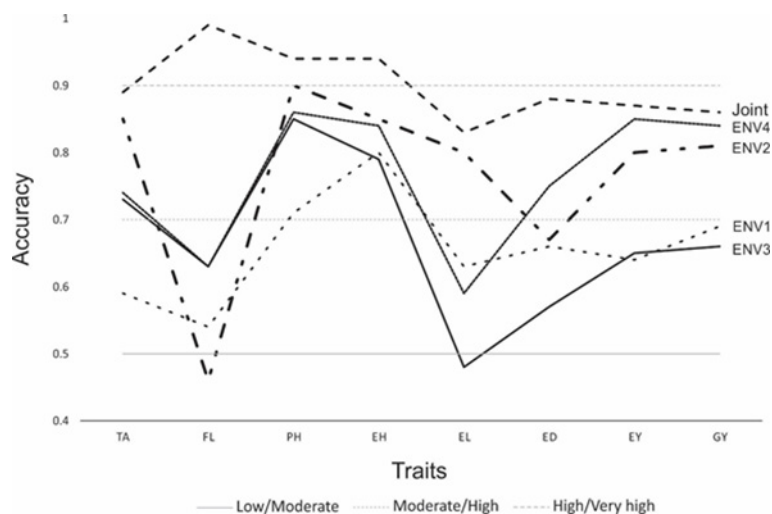


Figure 1. Mean accuracy values for the analyzed traits. ENV1: environment 1; ENV2: environment 2; ENV3 environment 3; ENV4: environment 4; and, Joint: Joint analysis. TA: tasseling; FL: flowering; PH: plant height; EH: ear height; EL: ear length; ED: ear diameter; EY: ear yield; GY: grain yield. The grey lines represent the intervals for selective accuracy' classification according to Resende and Duarte (2007).

variance (Resende and Duarte 2007), as this model represents one residual variance for each environment (De Faveri et al. 2015). For MET analysis, selecting the best-fit model by modeling the residual variance structure allows the selective accuracy to be maximized. MET data are generally related to heterogeneous residual variance structures in many annual crops, such as maize, cotton, and the common bean (So and Edwards 2011, Rocha et al. 2019, Melo et al. 2020, Peixoto et al. 2020). Therefore, the modeling of the residual variance structure is a straightforward and reliable in MET analysis.

It was verified by FA that the average commonality was superior to 75% when three factors (latent variables) were considered (considering those latent variables related to eigenvalues superior to 1) for the individual analyses (Table 2). According to Cruz et al. (2014), each latent variable consisted of a group of traits that were strongly associated with each other, but weakly correlated with other traits. In this sense, the first latent variables from the individual analyses were more influenced by TA and FL traits, being interpreted as cycle-related latent variables, as they were more influenced by cycle traits. Similarly, the second latent variable was more associated with PH and EH, allowing the association between the second factor and morphology-related traits; hence, it was directly influenced by traits in the morphology group. The third latent variable was influenced by EY and GY and related to traits from the yield group; therefore, these were termed yield-related traits.

EL and ED presented lower values of commonalities (< 0.60), which did not allow for their association with the latent variables formed. Four latent variables were formed for joint analyses (Table 2). The first and second were formed by the same traits from the individual analyses, whereas the third was configured by EY, GY, and EL (all yield-related traits). The fourth latent variable was constituted by ED, which is considered a yield-related factor.

According to Murakami and Cruz (2004) and Oliveira et al. (2005), the factor analysis clusters high genetic correlation traits into the same factor and low genetic correlated traits into different ones. Therefore, each latent variable has a biological meaning based on the genetic correlation between pairs of traits. However, the absence of association with other traits (EL and ED) in the individual analyses and the association of these traits with others in the joint analysis was a result of the lower commonality and an indication of the presence of G×E interaction. Even though there is a correlation between these two traits and the others from the yield group, the lower proportion of the variation included

Table 2. Loads for each latent variable in relation to the first four factors, commonalities, and variance accumulation in percentage (VA%) considering each environment and the joint analysis. TA: tasseling; FL: flowering; PH: plant height; EH: ear height; EL: ear length; ED: ear diameter; EY: ear yield; GY: grain yield

ENV		TA	FL	PH	EH	EL	ED	EY	GY	VA (%)	Mean
1	Factor 1	-0.259	0.147	-0.251	-0.042	-0.609	-0.596	-0.937	-0.945	39.831	
	Factor 2	0.839	0.880	0.045	0.011	-0.091	0.158	0.009	0.002	58.812	
	Factor 3	-0.013	-0.029	-0.875	-0.953	-0.364	0.103	-0.231	-0.189	75.908	
	Communalities	0.771	0.798	0.831	0.909	0.512	0.391	0.931	0.929		0.759
2	Factor 1	-0.097	0.069	-0.183	0.013	-0.491	-0.511	-0.905	-0.906	31.679	
	Factor 2	0.944	0.951	0.105	0.070	0.220	0.154	-0.230	-0.238	58.394	
	Factor 3	-0.044	-0.136	-0.915	-0.953	-0.054	0.128	-0.205	-0.202	75.608	
	Communalities	0.902	0.927	0.882	0.913	0.293	0.301	0.914	0.916		0.756
3	Factor 1	-0.078	0.014	-0.171	-0.091	-0.678	-0.546	-0.963	-0.955	35.829	
	Factor 2	-0.973	-0.974	0.148	-0.029	-0.153	-0.150	0.122	0.144	61.280	
	Factor 3	0.064	0.046	-0.920	-0.948	-0.162	-0.149	-0.014	-0.031	80.533	
	Communalities	0.957	0.950	0.898	0.908	0.509	0.343	0.943	0.934		0.805
4	Factor 1	-0.052	0.025	-0.152	-0.069	-0.683	-0.541	-0.958	-0.946	35.021	
	Factor 2	-0.974	-0.963	0.135	-0.036	-0.136	-0.268	0.139	0.164	60.589	
	Factor 3	-0.036	-0.065	0.934	0.946	0.142	0.192	0.007	0.012	80.843	
	Communalities	0.953	0.933	0.914	0.902	0.505	0.401	0.937	0.922		0.808
Joint	Factor 1	-0.126	0.100	-0.240	-0.116	-0.832	-0.165	-0.906	-0.912	39.712	
	Factor 2	0.961	0.952	0.045	0.066	0.121	-0.006	-0.036	-0.042	63.539	
	Factor 3	0.022	-0.133	-0.912	-0.963	-0.187	-0.097	-0.141	-0.146	80.148	
	Factor 4	-0.080	0.102	-0.209	0.053	0.300	-0.939	-0.331	-0.312	92.857	
	Communalities	0.947	0.945	0.935	0.948	0.832	0.919	0.952	0.951		0.929

in the analyses explained by the factors (communality) and the presence of G×E interaction as a disturbance factor, result in these unexpected patterns.

The FAI-BLUP index ranked the hybrids toward the maize genotype-ideotype (Supplementary material, Appendix C). The hybrid ranking changed across the environments, indicating the existence of a G×E interaction. However, hybrids 51, 2, and 26 presented the best performance, including all traits analyzed, across all environments. Factor analysis and its strengths (mentioned previously) were used in the FAI-BLUP index to assess the performance of genotypes based on BLUPs for each environment, and accounting for all traits simultaneously (Rocha et al. 2018).

The Kappa index values (Figure 2) highlight the dissimilarities between environments in relation to hybrid performance through the FAI-BLUP index, apart from ENV3-ENV4, where higher similarities between the best genotypes were found (agreement of 0.85). However, these results demonstrated that the G×E interaction from yield-related traits (EY and GY) caused some disturbance in the analysis. In addition, this would be expected in the secondary traits in the analyses, where the oscillations caused by the G×E interaction also displayed an important factor that ultimately implies lower values of the coincidence index. Barbosa et al. (2019) has found similar patterns in the study of coffee genotypes and Nardino et al. (2020) revealed that the G×E interaction is a factor of disturbance after FA in maize diallel hybrids.

The results indicate that the G×E interaction was significant in only two traits, and the latent variables presented the same pattern in each environment. However, the lower values of coincidence demonstrate that the G×E interaction changes the ranking and reduces the success of the hybrid’s indirect selection. This is a direct response to FA, where the traits are considered together in the selection process, resulting in low values of agreement by the Kappa index. According to van Eeuwijk et al. (2016), the presence of a G×E interaction affects the genotype performance over the environment, as gene expression is different at each site.

The three selection methodologies explored here are conceptually different. The FAI-BLUP index ranked the genotypes based on the genotypic aggregate of several traits combined, whereas direct selection ($\mu + g$) ranked the genotypes based on the specific trait (GY in our case), and HMRPGV ranked the genotypes for each location, considering the effects of the G×E interaction ($\mu + g + ge$). These dissimilarities among the selection strategies caused the values of Cohen’s Kappa between each strategy to be lower (Figure 3), thereby indicating that the selection conducted by the breeding companies, considering only the GY trait, does not favor the selection of the maize genotype-ideotype.

Indeed, when selection is based only on one trait, indirect selection does not present persuasive results in secondary traits, depending on the magnitude of the genetic correlation between pairs of traits, thereby indicating a correlated

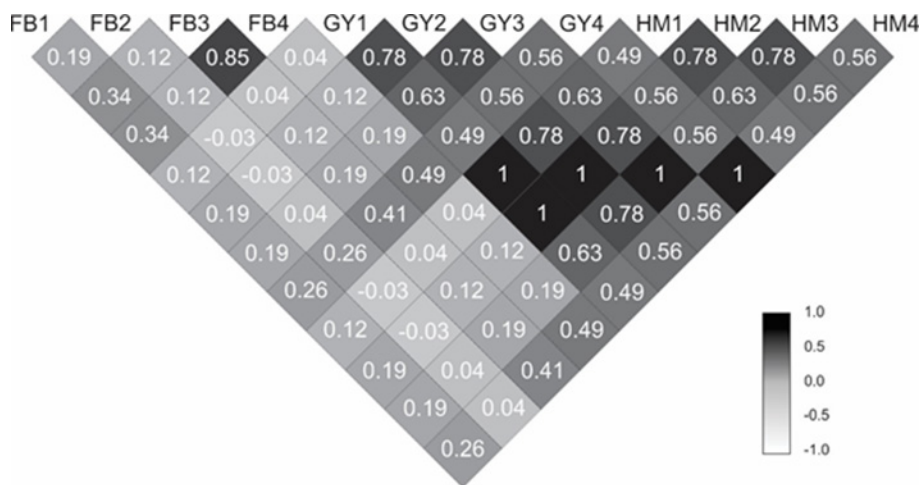


Figure 2. Cohen Kappa agreement index of the 17 selected hybrids (selection intensity of 20%) by FAI-BLUP, direct selection based on grain yield trait, and harmonic mean of relative performance of genotypic values (HMRPGV), between pairs of environments. FB1, FB2, FB3, and FB4 = FAI-BLUP index in environment 1, 2, 3, and 4, respectively; GY1, GY2, GY3, and GY4 = grain yield direct selection at environment 1, 2, 3, and 4, respectively; HM1, HM2, HM3, and HM4 = indirect selection ($\mu + g + ge$) based on HMRPGV for the environment 1, 2, 3, and 4, respectively.

response (Ertiro et al. 2020). For example, indirect selection through HMRPGV presented small gains for all traits except GY and EY, compared with FAI-BLUP (Table 3), whereas gains with values opposite to what was considered desired were found (*i.e.*, PH trait: negative gains were desirable, and only positive gains were found). However, for GY and EY, the gains overtook those gains from FAI-BLUP, as expected, because the HMRPGV was built based on the GY trait, and the EY trait was highly correlated with GY. These facts also highlight the outstanding performance of the FAI-BLUP index in the selection of hybrids searching for the maize ideotype.

The selection of several traits combined often result in reduced selection gain for all traits assessed (Cruz et al. 2014), compared with the direct selection for GY, indirect selection through HMRPGV, and selection for adaptability, stability, and productivity through HMRPGV (Table 3). For example, the top superior hybrids selected in all environments by the FAI-BLUP index (51, 2, and 26), were not presented as superior by direct (only considering FY trait) and HMRPGV selection. Indeed, the aims of each strategy are slightly different, and even though the direct and HMRPGV selections presented higher gains compared to the FAI-BLUP index, the latter should be preferred in the selection process of superior hybrids considering a group of traits. Furthermore, the gain with selection from the FAI-BLUP index presented reasonable values for GY traits (Table 3), as well as for the other traits. Thus, multi-trait selection should be considered as an alternative to reduce costs and to select genotypes close to the maize ideotype.

Among the top hybrids indicated by the FAI-BLUP index were some interpopulational hybrids that overtook commercial hybrids. In general, in crops such as maize, selection is carried out based on GY, which can lead to a selection toward production, neglecting other traits (Lima and Borém 2018, Coelho et al. 2020). However, the results indicated that selection of the ideotype was feasible. For example, in commercial fields, reducing the time to flowering (fitting the shorter cycle maize between other crop seasons), or the time for the selection process (speeding up the generation advancing, thereby creating more generations in a shorter time) is desirable for breeding programs, and the productivity traits are not harmed. The development of superior hybrids for ideotype design seldom involves simultaneous trait selection. A breeding program must identify hybrids that combine desirable traits from all groups (yield-, morphology-, and cycle-related traits).

Table 3. Predicted genetic gain, in percentage, considering the 17 maize hybrids selected based on the FAI-BLUP index for each environment, direct selection (DS) for each environment, and indirect selection for all environments combined and for each environment (u +g +ge) based on harmonic mean of relative performance of genotypic values (HMRPGV). TA: tasseling; FL: lowering; PH: plant height; EH: ear height; EL: ear length; ED: ear diameter; EY: ear yield; GY: grain yield

ENV	Selection strategy								
	FAI-BLUP								DS
	TA	FL	PH	EH	EL	ED	EY	GY	GY
1	-0.83	-0.45	-0.93	-2.69	0.98	0.54	3.6	4.16	15.12
2	-0.55	-0.63	-2.14	-4.28	0.3	0.37	3.06	3.09	13.31
3	-0.23	-0.21	-0.95	-2.68	0.86	0.37	5.18	6.31	13.18
4	-0.12	-0.1	-0.43	-1.31	0.62	0.11	3.38	4.35	15.94
ENV	Indirect selection through HMRPGV								HMRPGV
	TA	FL	PH	EH	EL	ED	EY	GY	GY
	1	0.61	-0.01	1.64	0.75	2.57	0.43	7.84	9.10
2	0.40	0.22	1.76	1.02	0.44	0.19	4.44	4.50	11.02
3	-0.01	-0.11	1.06	1.02	1.49	0.67	3.89	4.45	11.01
4	-0.01	-0.06	0.64	0.66	0.91	0.39	2.09	2.36	11.19

CONCLUSION

Models accounting for different residual structures were indicated for each trait. Model selection is an important step in MET data, particularly for annual crops. The results of this study demonstrated that it is possible to reduce the number of traits in the maize genetic assessment using FA, as the groups formed were related to each trait group analyzed (yield-, morphology-, and cycle-related traits). The FAI-BLUP index is suitable for combined genetic selection and can be used to achieve selection gains in all traits analyzed simultaneously; it is superior when compared with direct selection for GY traits, indirect selection through HMRPGV, and selection for adaptability, stability, and productivity (HMRPGV).

REFERENCES

- Alves RS, Resende MDV, Azevedo CF, Silva FF, Rocha JRASC, Nunes ACP, Carneiro APS and Santos GA (2020) Optimization of Eucalyptus breeding through random regression models allowing for reaction norms in response to environmental gradients. **Tree Genetics & Genomes** **16**: 38.
- Barbosa IPP, Costa WG, Nascimento M, Cruz CD, Oliveira ACB, Costa WG, Nascimento M, Cruz CD and Oliveira ACB (2019) Recommendation of *Coffea arabica* genotypes by factor analysis. **Euphytica** **215**: 1-10.
- Bhering LL (2017) RBio: A tool for biometric and statistical analysis using the R platform. **Crop Breeding and Applied Biotechnology** **17**: 187-190.
- Coelho IF, Alves RS, Rocha JRASC, Peixoto MA, Teodoro LPR, Teodoro PE, Pinto JFN, Reis EF and Bhering LL (2020) Multi-trait multi-environment diallel analyses for maize breeding. **Euphytica** **216**: 1-17.
- Cohen J (1960) A coefficient of agreement for nominal scales. **Educational and Psychological Measurement** **20**: 37-46.
- Cruz CD, Carneiro PCS and Regazzi AJ (2014) **Modelos biométricos aplicados ao melhoramento genético**. 3rd edn, Editora UFV, Viçosa, 668p.
- De Faveri J, Verbyla AP, Pitchford WS, Venkatanagappa S and Cullis BR (2015) Statistical methods for analysis of multi-harvest data from perennial pasture variety selection trials. **Crop and Pasture Science** **66**: 947-962.
- Ertiro BT, Olsen M, Das B, Gowda M and Labuschagne M (2020) Efficiency of indirect selection for grain yield in maize (*Zea mays* L.) under low nitrogen conditions through secondary traits under low nitrogen and grain yield under optimum conditions. **Euphytica** **216**: 1-12.
- Gilmour AR, Gogel BJ, Cullis BR, Welham SJ and Thompson R (2015) **ASReml user guide**. Release 4.1. Structural specification. VSN international Ltd, Hemel hempstead, 375p.
- Henderson CR (1975) Best linear unbiased estimation and prediction under a selection model. **Biometrics** **31**: 423-447.
- Lima RO and Borém A (2018) **Melhoramento de milho**. Editora UFV, Viçosa, 396p.
- Mardia K, Kent J and Bibby J (1979) **Multivariate analysis**. Academic Press, New York, 521p.
- Melo VL, Marçal TS, Rocha JRASC, Anjos RSR, Carneiro PCS and Carneiro JES (2020) Modeling (co)variance structures for genetic and non-genetic effects in the selection of common bean progenies. **Euphytica** **216**: 77.
- Murakami D and Cruz CD (2004) Proposal of methodologies for environment stratification and analysis of genotype adaptability. **Crop Breeding and Applied Biotechnology** **4**: 7-11.
- Nardino M, Barros WS, Olivoto T, Cruz CD, Silva FF, Pelegrin AJ, Souza VQ, Carvalho IR, Szareski VJ, Oliveira AC, Maia LC and Konflanz VA (2020) Multivariate diallel analysis by factor analysis for establish mega-traits. **Anais da Academia Brasileira de Ciências** **92**: 1-19.
- Nuvunga JJ, Oliveira LA, Pamplona AKA, Silva CP, Lima RR and Balestre M (2015) Factor analysis using mixed models of multi-environment trials with different levels of unbalancing. **Genetics and Molecular Research** **14**: 14262-14278.
- Oliveira GV, Carneiro PCS, Dias LAS, Carneiro JES and Cruz CD (2005) Factor analysis in the environment stratification for the evaluation of common bean cultivars. **Crop Breeding and Applied Biotechnology** **5**: 166-173.
- Oliveira ICM, Marçal TDS, Bernardino KDC, Ribeiro PCDO, Parrella RADC, Carneiro PCS, Schaffert RE and Carneiro JEDS (2019) Combining ability of biomass sorghum lines for agroindustrial characters and multitrait selection of photosensitive hybrids for energy cogeneration. **Crop Science** **59**: 1554-1566.
- Paiva JT, Oliveira HR, Nascimento M, Nascimento ACC, Silva HT, Henriques RF, Lopes PS, Silva FF, Veroneze R, Ferraz JBS, Eler JP, Mattos EC and Gaya LG (2020) Genetic evaluation for latent variables derived from factor analysis in broilers. **British Poultry Science** **61**: 3-9.
- Patterson HD and Thompson R (1971) Recovery of inter-block information when block sizes are unequal. **Biometrika** **58**: 545-554.
- Peixoto MA, Coelho IF, Evangelista JSPC, Alves RS, Rocha JRASC, Farias FJC, Carvalho LP, Teodoro PE and Bhering LL (2020) Reaction norms-based approach applied to optimizing recommendations of cotton genotypes. **Agronomy Journal** **112**: 4613-4623.
- Peixoto LS, Nunes JAR and Furtado DF (2016) Factor analysis applied to the G+GE matrix via REML/BLUP for multi-environment data. **Crop Breeding and Applied Biotechnology** **16**: 1-6.
- R Development Core Team (2020) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.
- Resende MDV (2015) **Genética quantitativa e de populações**. Suprema, Visconde do Rio Branco, 463p.
- Resende MDV (2016) Software Selegen-REML/BLUP: A useful tool for plant breeding. **Crop Breeding and Applied Biotechnology** **16**: 330-339.
- 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-QWAS, estatística espacial e temporal, competição, sobrevivência**. UFV, Viçosa, 881p.
- Rocha JRASC, Machado JC and Carneiro PCS (2018) Multitrait index based on factor analysis and ideotype-design: proposal and application on elephant grass breeding for bioenergy. **GCB Bioenergy** **10**: 52-60.
- Rocha JRASC, Nunes KV, Carneiro ALN, Marçal TS, Salvador FV, Carneiro PCS and Carneiro JES (2019) Selection of superior inbred progenies toward the common bean ideotype. **Agronomy Journal** **111**: 1181-1189.

Selection of maize hybrids: an approach with multi-trait, multi-environment, and ideotype-design

- Schwarz G (1978) Estimating the dimension of a model. **The Annals of statistic** **6**: 461-464.
- Silva MJ, Carneiro PCS, Carneiro JE, Damasceno CMB, Parrella NNLD, Pastina MM, Simeone MLF, Schaffert RE and Parrella RA (2018) Evaluation of the potential of lines and hybrids of biomass sorghum. **Industrial Crops and Products** **125**: 379-385.
- So YS and Edwards J (2011) Predictive ability assessment of linear mixed models in multi-environment trials in corn. **Crop Science** **51**: 542-552.
- Sripathi R, Conaghan P, Grogan D and Casler MD (2018) Modeling genotype \times environment correlation structures in long-term multilocation forage yield trials. **Crop Science** **58**: 1447-1457.
- van Eeuwijk FA, Bustos-Korts D V and Malosetti M (2016) What should students in plant breeding know about the statistical aspects of Genotype \times Environment interactions? **Crop Science** **56**: 2119-2140.
- Woyann LG, Meira D, Matei G, Zdziarski AD, Dallacorte LV, Madella LA and Benin G (2020) Selection indexes based on linear-bilinear models applied to soybean breeding. **Agronomy Journal** **112**: 175-182.
- Woyann LG, Meira D, Zdziarski AD, Matei G, Milioli AS, Rosa AC, Madella LA and Benin G (2019) Multiple-trait selection of soybean for biodiesel production in Brazil. **Industrial Crops and Products** **140**: 111721.
- Yan W and Frégeau-Reid J (2018) Genotype by Yield \times Trait (GYT) Biplot: A novel approach for genotype selection based on multiple traits. **Scientific Reports** **8**: 1-10.