Research Article

Selection Index in the Study of Adaptability and Stability in Maize

Rogério Lunezzo de Oliveira,1 Renzo Garcia Von Pinho,2 Daniel Furtado Ferreira,3 Luiz Paulo Miranda Pires,4 and Wagner Mateus Costa Melo5

1 Engenheiro Agrônomo, Mestre em Genética e Melhoramento de Plantas, DuPont do Brasil S.A. Divisão Pioneer Sementes, Rodovia GO 210 km 04, Fazenda Santa Maria de Baixo, Zona Rural, Caixa Postal 1014, 75503-970 Itumbiara, GO, Brazil
2 Engenheiro Agrônomo, Doutor em Genética e Melhoramento de Plantas, Departamento de Agricultura/DAG, Universidade Federal de Lavras (UFLA), Lavras, MG, Brazil
3 Engenheiro Agrônomo, Pós-Doutor em Estatística e Experimentação Agronômica, Departamento de Ciências Exatas (DEX), Universidade Federal de Lavras (UFLA), Lavras, MG, Brazil
4 Engenheiro Agrônomo, Mestre em Genética e Melhoramento de Plantas, Departamento de Biologia/DBI, Universidade Federal de Lavras (UFLA), Lavras, MG, Brazil
5 Engenheiro Agrônomo, Doutor Agronomia/Fitotecnia, Advanta Comércio de Sementes Ltda, Florianópolis, SC, Brazil

Correspondence should be addressed to Rogério Lunezzo de Oliveira; rogerio.oliveira@pioneer.com

Received 10 November 2013; Accepted 29 December 2013; Published 13 February 2014

Academic Editors: B. Chen and W. Ma

Copyright © 2014 Rogério Lunezzo de Oliveira et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper proposes an alternative method for evaluating the stability and adaptability of maize hybrids using a genotype-ideotype distance index (GIDI) for selection. Data from seven variables were used, obtained through evaluation of 25 maize hybrids at six sites in southern Brazil. The GIDI was estimated by means of the generalized Mahalanobis distance for each plot of the test. We then proceeded to GGE biplot analysis in order to compare the predictive accuracy of the GGE models and the grouping of environments and to select the best five hybrids. The G × E interaction was significant for both variables assessed. The GGE model with two principal components obtained a predictive accuracy (PRECORR) of 0.8913 for the GIDI and 0.8709 for yield (t ha⁻¹). Two groups of environments were obtained upon analyzing the GIDI, whereas all the environments remained in the same group upon analyzing yield. Coincidence occurred in only two hybrids considering evaluation of the two features. The GIDI assessment provided for selection of hybrids that combine adaptability and stability in most of the variables assessed, making its use more highly recommended than analyzing each variable separately. Not all the higher-yielding hybrids were the best in the other variables assessed.

1. Introduction

In breeding programs, it is often necessary to obtain measures of various traits to select them simultaneously. This need led to the development of selection indices, allowing selection to be performed using a single value. The application of selection indices was initially proposed by Smith [1] through the optimized index. This index was later modified, and most of these modifications were also based on linear combinations of the phenotypic values observed. The most common modifications are those of Brim et al. [2], Kempthorne and Nordskog [3], Pešek and Baker [4], and Tai [5] and Smith et al. [6].

Other authors presented nonparametric indices, applied for the purpose of classifying genotypes. Elston [7] proposed a multiplicative index which considers all the traits with the same economic weight. Wricke and Weber [8] suggested the use of the Euclidean and Mahalanobis distances to classify the genotypes for various traits simultaneously, based on their distance from an ideal genotype (ideotype), defined by the researcher. Mulamba and Mock [9] developed a quite simple index, which uses the sum of the number of order that the
genotype presents for each trait; the lower the value obtained in this sum, the better the classification.

There are various methodologies for evaluation of the G × E interaction. Crossa [10] suggests that the application of multivariate methods may be useful to better exploit the information contained in the data. Recently, a modification of the conventional AMMI analysis proposed by Yan et al. [11], called the GGE biplot, has been used for study of the G × E interaction. GGE analysis groups the genotype effect, which is an additive effect in the AMMI analysis, with the G × E interaction, which is a multiplicative effect, subjecting these effects to a multiplicative model of regression for sites (SREG - Site Regression). The main advantage of this technique in relation to AMMI analysis resides in the fact that the GGE biplot method always explains an intermediate portion of the sum of squares of genotypes (G) + interaction (G × E) in relation to the AMMII (with one principal component) and AMMII2 (with two principal components) models, Yan et al. [12].

Garcia and Júnior [13] and Farshadfar [14] incorporated stability and adaptability parameters in selection indices to select individuals of greatest stability and best performance. Nevertheless, these studies that applied nonparametric selection indices as parameters of incorporated stability did it with the mean values of the cultivars in all the environments.

It is difficult to find studies in the literature that have analyzed these selection indices at plot level and have analyzed the use of these indices for adaptability and stability analyses by multivariate methods such as the GGE biplot so that the applicability of a nonparametric selection index may be better evaluated and utilized. For that reason, the aim of this study was to evaluate the adaptability and stability of the genotype-ideotype distance index and selection of the best hybrids by means of this alternative.

2. Materials and Methods

Assessment data were composed of seven traits from 25 maize hybrids subjected to performance evaluation trials in six locations of the south of Brazil (Vacaria, RS; Abela Lóz, SC; Candoi, PR; Canoinhas, SC; Castro, PR; and Ponta Grossa, PR) during the 2010/2011 crop year. The trials were prepared in a completely randomized block design with two replications. The plots were composed of four five-meter rows, with a spacing of 70 cm between rows. Fertilizations were carried out following the recommendations made for each location by means of soil analysis. Crop treatments necessary for control of army worm (Spodoptera frugiperda), corn earworm (Helicoverpa zea), and weeds were carried out.

The traits assessed were grain yield—kg ha⁻¹, percentage of damaged grains, percentage of lodging, percentage of breakage, percentage of fallen plants, common rust (Puccinia sorghi) score, and gray leaf spot (Cercospora zeae-maydis) score.

The percentage of damaged grains was determined through the proportion of grains with symptoms of rotting in a 1000 grain sample. The percentage of lodging was calculated through the proportion of plants at an inclination of more than 30° due to weakening of the roots in relation to the total number of plants in the plot. The percentage of fallen plants, for its part, corresponded to the proportion of plants at an inclination of more than 30° due to weakening of the base of the stalk in relation to the total number of plants in the plot. The disease severity score was based on the diagrammatic scale proposed by Agroceres [15], which ranges from 1 to 9, in which score 1 represents a leaf without symptoms and score 9 corresponds to the material with more than 80% of the leaf area affected.

The ideotype for yield (kg ha⁻¹) was determined seeking the value of the highest yielding plot of all the trails and using the next thousand value above that. As the highest yielding plot was 16959 kg ha⁻¹, the ideotype was 17000 kg ha⁻¹. For the traits of percentage of damaged grains, percentage of lodging, percentage of fallen plants, and percentage of breakage, the ideotype was 0%. For common rust and grey leaf spot, the ideotype was score 1.

To obtain the matrix of variances and covariances among the traits assessed, multivariate analysis of variance (MANOVA) was performed using the data from all the hybrids, traits, and locations. For that purpose, the model according to Ferreira [16] was used, expressed in the vectorial form in the following expression:

\[
Y_{ijk} = \mu + \alpha_i + \beta_j + \delta_{ij} + \varepsilon_{ijk},
\]

in which the vector \(Y_{ijk} = [Y_{ijk1}, Y_{ijk2}, \ldots, Y_{ijkp}]\) refers to the multivariate observations associated with the \(i\)th hybrid, \(j\)th location \((i = 1, 2, \ldots, 25), (j = 1, 2, \ldots, 6)\) in the \(k\)th replication of this combination of the \(i\) and \(j\) levels of the two factors; \(\mu\) is the vector of constants of the multivariate linear model given by \(\mu = [\mu_1, \mu_2, \ldots, \mu_p]\), \(\alpha_i\) is the vector of effects of the \(i\)th hybrid given by \(\alpha_i = [\alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{ip}]\), \(\beta_j\) is the vector of effects of the \(j\)th location given by \(\beta_j = [\beta_{j1}, \beta_{j2}, \ldots, \beta_{jp}]\), \(\delta_{ij}\) is the vector of effects of the interaction between the \(i\)th hybrid and the \(j\)th location given by \(\delta_{ij} = [\delta_{i1}, \delta_{i2}, \ldots, \delta_{ip}]\), and \(\varepsilon_{ijk}\) is the vector of effects of the nonobservable experimental error corresponding to observation \(Y_{ijk}\). The manova package of the software R v 3.0 [17] was used for MANOVA.

The genotype-ideotype distance index (GIDI) for selection was obtained using the data at the plot level of the seven traits assessed in the 25 hybrids. Thus, for each plot under assessment, an index was obtained based on the generalized Mahalanobis distance, using the matrix model according to the following expression:

\[
D_{G}^{2} = \text{diag} \left( IS^{-1} I' \right),
\]

where the term \( IS^{-1} I' \) is written in the following manner:

\[
\begin{bmatrix}
(Y_{1111} - Y_{11}) & \cdots & (Y_{nl11} - Y_{nl}) \\
\vdots & \ddots & \vdots \\
(Y_{1her} - Y_{11}) & \cdots & (Y_{nherr} - Y_{nt})
\end{bmatrix}
\]
in which $D_{Ge}^2$ is the matrix of the distances of $h$ ($h = 1, 2, \ldots, 25$) hybrids in relation to the ideotype $\ell$; $Y_{ij\ell}$ is the observation of trait $i$ ($i = 1, 2, \ldots, n$) where $n = 7$, of hybrid $j$ ($j = 1, 2, \ldots, h$), at location $k$ ($k = 1, 2, \ldots, e$) where $e = 6$, in replication $l$ ($l = 1, 2$); $V_{1}, V_{2}, \ldots, V_{n}$ are the variances of the residues (QMer of the multivariate analysis of variance) of the $n$ traits assessed and $\text{COV}_{12}, \text{COV}_{22}, \ldots, \text{COV}_{(n-1)n}$ are the covariances among the residues of these traits.

After obtaining the GIDI for each plot, joint analysis of variance was performed to verify the presence of the $G \times E$ interaction for the GIDI and for grain yield. Once the presence of the $G \times E$ interaction ($F$ test significant) was observed, analysis of adaptability and stability was performed for the two variables, which allowed the adaptation and stability of each hybrid under testing to be measured, for the GIDI and for yield. This evaluation was made using the GGE biplot method. As the interpretation of the GIDI is the opposite of yield; that is, the lower the value of the index, the better the performance, the value of each plot was subtracted from 2000. This value was determined using the same principle used in the determination of ideotype of yield, that is, using the next thousand value above the greatest value of the index considering all the plots of all the trails.

Graph interpretation was thereby able to be performed in the same manner as yield. GGE biplot analysis was performed by means of the package GGBiplotGUI of the software R v 3.0 [17].

Analysis by means of the GGE biplot method was performed as presented by de Oliveira et al. [18], considering the simplified model for two principal components:

$$Y_{ij} - \mu_j = \lambda_1 Y_{1i} + \alpha_1 j + \lambda_2 Y_{2i} + \alpha_2 j + \rho_j + \epsilon_{ij},$$

in which $Y_{ij}$ is the mean value of hybrid $i$ in location $j$, for the GIDI or yield; $\mu_j$ is the mean value of the location $j$; $\lambda_1 Y_{1i}$ is the first principal component (PCA 1) of the effect of genotypes ($G$) + interaction ($G \times A$); $\lambda_2 Y_{2i}$ is the second principal component (PCA 2) of the effect of genotypes ($G$) + interaction ($G \times A$); $\lambda_1$ and $\lambda_2$ are the eigenvalues associated with the PCA 1 and with the PCA 2; $Y_{1i}$ and $Y_{2i}$ are the scores of the PCA 1 and of the PCA 2, respectively, for genotypes; $\alpha_1 j$ and $\alpha_2 j$ are the scores of the PCA 1 and of the PCA 2, respectively, for environments; $\rho_j$ is the residual of the genotype $x$ environment interaction, also known as “noise”, corresponding to the principal components not retained in the model; $\epsilon_{ij}$ is the residual of the model with normal distribution, mean value zero, and variance $\sigma^2/r$ (where $\sigma^2$ is the variance of the error between plots for each environment and $r$ is the number of replications).

The graph accuracy of the identification methods of megaenvironments and winning genotypes was tested by the cross validation procedure proposed by Gabriel [19]. For that purpose, the PRESSm and PRESScorr statistics were used to measure the discrepancy between the observed and predicted values and the predictive correlation [20]. This cross validation analysis was performed by means of PROC IML of the statistical package of the software SAS v 9.0 [21].

The GGE biplot graph was constructed as of the data from breakdown of the empirical mean values, graphically presenting the hybrid with the best development, as detailed by de Oliveira et al. [18]. The environments were grouped by means of the graphic approach called “Which Won Where” by Yan et al. [11]. Thus, the environments that share the same winning genotype remain in the same group.

The five hybrids with greatest adaptability and stability for the GIDI and yield were identified and selected by means of their distance in relation to the “ideal genotype” which was plotted in the center of a “target” in the biplot. That way, the hybrids plotted nearer the center of this “target” are those that have the best combination between adaptability and stability as presented by Balestre et al. [22].
Table 2: Estimates obtained for PRESS_m (sum of squared prediction error) and PRECORR_m (predictive accuracy) in cross validation of the GGE1 and GGE2 models (biplot with 1 and 2 principal components, resp.) for the GIDI and for yield (ton ha$^{-1}$).

| Model | GIDI PRESS_m | GIDI PRECORR_m | Yield PRESS_m | Yield PRECORR_m |
|-------|--------------|----------------|---------------|-----------------|
| GGE1  | 36324.678    | 0.8086993      | 1.0547498     | 0.8656868       |
| GGE2  | 21597.326    | 0.8913184      | 1.0179156     | 0.8708956       |

3. Results and Discussion

The G × E interaction was significant ($P < 0.01$), showing that the performance of the hybrids was not consistent in the different assessment environments in regard to the GIDI (data not shown). The results of analysis of variance for grain yield were similar to those obtained in the GIDI analysis. The genotype and environment sources of variation and their interaction were significant ($P < 0.01$). The coefficient of variation was very low (3.3%), which shows the high experimental precision of the trials.

The mean value of the hybrids for grain yield (kg ha$^{-1}$) and for the GIDI (already subtracted from 2000, to make its interpretation similar to that of yield), involving the six locations, are shown in Table 1.

As may be observed, the highest yielding hybrids are not necessarily those that have the greatest values of the GIDI. This is due to the fact that the highest yielding hybrids are not always the best in the other traits assessed and considered in the GIDI.

The results of cross validation of the GGE biplot model for the GIDI and for grain yield are shown in Table 2.

In this study, only the results of cross validation for the first two principal components were presented since the reduced model was chosen due to the difficulty of evaluating the biplot in more than two dimensions.

It may be seen that for the GGE2 model, the predictive accuracy (PRECORR_m) was high for the two variables analyzed. In addition, the values were close when the two variables are compared. de Oliveira et al. [18] and Balestre et al. [22], evaluating the stability and adaptability of grain yield in maize and rice, obtained a PRECORR of 0.8617 and 0.90, respectively. These values are very near those obtained in this study. The GGE2 model was therefore accurate for the two variables.

Subsequently, the biplot was obtained and the “Which Won Where” approach was carried out to identify the environments or megaenvironments, evaluating the GIDI (Figure 1). In this figure, the hybrids were identified from G1 to G25.

The first result to be highlighted is the high explanatory power of the sum of squares of $G + G × E$ presented by the two principal components of the biplot (PCA 1 and PCA 2). The two components added together accounted for 92.14%, showing that the analysis was very efficient. In regard to grouping, two groups were obtained, identified on the graph...
Figure 3: GGE biplot graph, evaluating grain yield (t ha$^{-1}$), with the grouping of the six locations (Vacaria, RS; Abelardo Luz, SC; Candoi, PR; Canoinhas, SC; Castro, PR; and Ponta Grossa, PR), identified from E1 to E6, respectively, the gray line represents the polygon formed by genotypes evaluated.

The five hybrids nearest to the “ideal genotype” were hybrids 21, 23, 24, 22, and 25, in that order. Comparing this selection with that obtained by the analysis with the GIDI (hybrids 12, 21, 19, 8, and 22), it may be seen that only hybrids 21 and 22 were common to the two selections. There was also a change in the position of hybrids 21 and 22 comparing the two selections. To diagnose the reason for this difference, the measurements of each hybrid were obtained for each one of the seven variables assessed considering the six locations.

Garcia and Júnior [13] also used the GIDI estimated by means of the Euclidean distance for selection of maize hybrids and, to this, they incorporated the parameters of adaptability and stability, $b$ and $R^2$, respectively. This procedure led to the selection of stable and adapted hybrids only for year yield (kg ha$^{-1}$). Euclidean distance may only be applied to traits which are mutually independent, which was not the case of the traits used by the authors, with a negative effect on the selection obtained by them. For that reason, these authors recommended that the correlation between the variables assessed must be considered in the selection index. This correction was performed in this study since the GIDI was estimated by means of the Mahalanobis distance, which is weighted by the matrix of variances and covariances. In addition, the evaluation of adaptability and stability was performed with the GIDI estimated at the plot level, allowing the selection of adapted and stable hybrids for all the variables analyzed.

4. Conclusion

The evaluation of adaptability and stability of the GIDI led to the selection of hybrids that combine adaptability and
stability for most of the traits assessed. Use of it is more practical than analyzing each trait separately.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] H. F. Smith, “A discriminant function for plant selection,” Annals Eugenics, vol. 7, no. 3, pp. 240–250, 1936.
[2] C. A. Brim, H. W. Johnson, and C. C. Cockerham, “Multiple selection criteria in soybeans,” Agronomy Journal, vol. 51, no. 1, pp. 42–46, 1959.
[3] O. Kempthorne and A. W. Nordskog, “Restricted selection indices,” Biometrics, vol. 15, no. 1, pp. 10–19, 1959.
[4] J. Pešek and R. J. Baker, “Desired improvement in relation to selection indices,” Canadian Journal of Plant Science, vol. 49, no. 6, pp. 803–804, 1969.
[5] G. C. C. Tai, “Index selection with desired gains,” Crop Science, vol. 17, no. 1, pp. 182–183, 1977.
[6] O. S. Smith, A. R. Hallauer, and W. A. Russell, “Use of index selection in recurrent selection programs in maize,” Euphytica, vol. 30, no. 3, pp. 611–618, 1981.
[7] R. C. Elston, “A weight free index for the purpose of ranking of selection with respect to several traits at a time,” Biometrics, vol. 19, no. 1, pp. 85–97, 1963.
[8] G. Wricke and W. E. Weber, Quantitative Genetics and Selection in Plant Breeding, Walter de Gruyter, New York, NY, USA, 1986.
[9] N. N. Mulamba and J. J. Mock, “Improvement of yield potential of the Eto Blanco maize (Zea mays L.) population by breeding for plant traits,” Egyptian Journal of Genetic and Cytology, vol. 7, no. 1, pp. 40–51, 1978.
[10] J. Crossa, “Statistical analyses of multilocation trials,” Advances in Agronomy, vol. 44, no. 1, pp. 55–85, 1990.
[11] W. Yan, L. A. Hunt, Q. Sheng, and Z. Szlavnicz, “ Cultivar evaluation and mega-environment investigation based on the GGE biplot,” Crop Science, vol. 40, no. 3, pp. 597–605, 2000.
[12] W. Yan, M. S. Kang, B. Ma, S. Woods, and P. L. Cornelius, “GGE biplot vs. AMMI analysis of genotype-by-environment data,” Crop Science, vol. 47, no. 2, pp. 643–655, 2007.
[13] A. A. F. Garcia and C. L. S. Júnior, “Comparação de índices de seleção não paramétricos para a seleção de cultivares,” Bragantia, vol. 58, no. 2, pp. 253–267, 1999.
[14] E. Farshadfar, “Incorporation of AMMI stability value and grain yield in a single non-parametric index (GSI) in bread wheat,” Pakistan Journal of Biological Sciences, vol. 11, no. 14, pp. 1791–1796, 2008.
[15] Agroceres, Guia Agroceres de Sanidade, Sementes Agroceres, São Paulo, Brazil, 1996.
[16] D. F. Ferreira, “ Análise de variância multivariada,” in Estatística Multivariada, pp. 329–331, UFLA, Lavras, Brazil, 2nd edition, 2011.
[17] R Development Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2011.
[18] R. L. de Oliveira, R. G. von Pinho, M. Balestre, and D. V. Ferreira, “Evaluation of maize hybrids and environmental stratification by the methods AMMI and GGE biplot,” Crop Breeding and Applied Biotechnology, vol. 10, no. 3, pp. 247–253, 2010.
[19] K. R. Gabriel, “Le biplot-outil d’exploration de données multidimensionnelles,” Journal de la Société Française de Statistique, vol. 143, no. 3–4, pp. 5–55, 2002.
[20] C. T. S. Dias and W. J. Krzanowski, “Model selection and cross validation in additive main effect and multiplicative interaction models,” Crop Science, vol. 43, no. 3, pp. 865–873, 2003.
[21] SAS Institute, SAS/STAT User’s Guide, SAS Institute, Cary, NC, USA, 2000.
[22] M. Balestre, V. B. dos Santos, A. A. Soares, and M. S. Reis, “Stability and adaptability of upland rice genotypes,” Crop Breeding and Applied Biotechnology, vol. 10, no. 4, pp. 357–363, 2010.