# Soybean productivity, stability, and adaptability through mixed model methodology

**Jeniffer Santana Pinto Coelho Evangelista**[1] **Rodrigo Silva Alves**[1] **Marco Antônio Peixoto**[1]
**Marcos Deon Vilela de Resende**[2] **Paulo Eduardo Teodoro**[3] **Felipe Lopes da Silva**[1]
**Leonardo Lopes Bhering**[1*]

[1]Universidade Federal de Viçosa (UFV), Viçosa, MG, Brasil. E-mail: leonardo.bhering@ufv.br. *Corresponding author.
[2]Embrapa Café, Universidade Federal de Viçosa (UFV), Viçosa, MG, Brasil.
[3]Universidade Federal de Mato Grosso do Sul (UFMS), Chapadão do Sul, MS, Brasil.

**ABSTRACT**: *The genotype × environment (G×E) interaction plays an essential role in phenotypic expression and can lead to difficulties in genetic selection. Thus, the present study aimed to estimate genetic parameters and to compare different selection strategies in the context of mixed models for soybean breeding. For this, data referring to the evaluation of 30 genotypes in 10 environments, regarding the grain yield trait, were used. The variance components were estimated through restricted maximum likelihood (REML) and genotypic values were predicted through best linear unbiased prediction (BLUP). Significant effects of genotypes and G×E interaction were detected by the likelihood ratio test (LRT). Low genotypic correlation was obtained across environments, indicating complex G×E interaction. The selective accuracy was very high, indicating high reliability. Our results showed that the most productive soybean genotypes have high adaptability and stability.*
**Key words**: *REML/BLUP, plant breeding, genotype × environment interaction, genetic selection, Glycine max.*

## Produtividade, estabilidade e adaptabilidade da soja por meio de metodologia de modelo misto

**RESUMO**: *A interação genótipo × ambiente (G × E) desempenha um papel essencial na expressão fenotípica e pode provocar dificuldades na seleção genética. Assim, o presente estudo teve como objetivo estimar parâmetros genéticos e comparar diferentes estratégias de seleção no contexto de modelos mistos para melhoramento da soja. Para isso, foram utilizados dados referentes à avaliação de 30 genótipos em dez ambientes, referentes à característica produtividade de grãos. Os componentes de variância foram estimados pela* máxima verossimilhança restrita (REML) *e os valores genotípicos foram preditos* pela melhor previsão imparcial linear (BLUP). *Efeitos significativos dos genótipos e interação G × E foram detectados pelo teste da razão de verossimilhança (LRT). Correlação genotípica baixa foi obtida entre os ambientes indicando interação G × E do tipo complexa. A acurácia seletiva foi muito alta, indicando alta confiabilidade. Os resultados mostraram que os genótipos de soja mais produtivos apresentam alta adaptabilidade e estabilidade.*
**Palavras-chave**: *REML/BLUP, Glycine max, interação genótipo × ambiente, melhoramento de plantas, seleção genética.*

## INTRODUCTION

Soybean [*Glycine max* (L.) Merrill] is the fourth most widely grown crop in the world and it is a source of oil, protein, and raw material for biodiesel production (SILVA et al., 2017). Currently, Brazil consolidates as the mainly world soybean producer and exporter (USDA, 2020), even though other countries, such as EUA, China, and India, are representative in soybean production. Actually, the soybean is spread cultivated from low to high latitudes (LIU et al., 2017). In this scenario, the genotype × environment (G×E) interaction plays an essential role in phenotypic expression and can lead to difficulties in genetic selection.

The G×E interaction is characterized by the differential response of genotypes to a change in environment (ALVES et al., 2020). The environment influences on the phenotypic expression, in order to provide the G×E interaction, is one of the biggest challenges that breeders deal with in plant breeding. As a result, a specific environment can affect the genetic the selection; however, in practice, the genetic material generally is recommended to several environments.

The mixed models is the main method for evaluating G×E interaction (LI et al., 2017)the

environment, and the differential sensitivity of certain genotypes to different environments, also known as genotype by environment (G × E. This method has features to model the heterogeneity of genetic variances and correlations between environments, as well as to model spatial trends in individual trials (BERNARDO, 2020)summarized the alternatives available at that time and noted that all of these approaches could be classified as multiplicative models. Recently, mixed model approaches have become popular for the analysis of series of variety trials. There are numerous reasons for their use, including the ease with which incomplete data (not all varieties in all trials. The ability of genotypes to respond differently over a wide range of environmental conditions may be an essential factor in a breeding program. In this context, the study of stability and adaptability becomes relevant. A genotype is considered stable when it presents slight variations in the general deviation when evaluated under various environmental conditions, and is considered adapted when it has responsiveness to environmental improvement (RESENDE, 2004).

In the context of mixed models, simultaneous selection for productivity, stability and adaptability can be performed by harmonic mean of relative performance of genetic values method, HMRPGV (RESENDE, 2004). Currently, the HMRPGV method has been applied in several crops for breeding purpose (GONÇALVES et al., 2014; ALVES et al., 2018genetic selection is carried out based on several traits, which can be genetically correlated. In this case, selection bias may occur if these traits are analyzed individually. Thus, the present work aimed to evaluate the applicability and efficiency of multiple-trait best linear unbiased prediction (BLUP; SOUZA et al., 2020). This procedure allows the simultaneous selection via the three parameters mentioned and presents the following advantages: (i) it considers the genotypic effects as random and therefore provides genotypic and non-phenotypic adaptability and stability; (ii) it allows analyzes with unbalance data; (iii) it enables to use a non-orthogonal designs; (iv) it allows to deal with variance heterogeneity; (v) it allows to consider correlated errors within environments; (vi) it provides already discounted (penalized) genotypic values of the instability; (vii) it does not depend on the estimation and interpretation of other parameters; (viii) it eliminates the noise of the G×E interaction as it considers the heritability of these effects; (ix) it generates results on the magnitude or scale of the evaluated trait itself; and, (x) it allows to compute the genetic gain with the selection by the three attributes simultaneously (RESENDE, 2007).

Obtaining reliable estimates of genetic parameters based on their evaluation of phenotypic adaptability and stability are paramount in conducting soybean breeding programs. These estimates assist the selection process and serve as a theoretical reference in cultivars recommendation. Thus, the present study aimed to estimate genetic parameters and to compare different selection strategies in the context of mixed models for soybean breeding.

## MATERIALS AND METHODS

Thirty improved lines soybean, belonging to the 6-7 relative maturity group, were evaluated at 10 municipalities, allocated in the soybean macro-region 2 (micro-regions 201, 202 and 204), in the 2013/2014 harvest. The experiments were conducted in a randomized block design with three replications. Each plot consisted of four lines of 5 m, spaced 0.5 m between lines and between plots. At maturity, the two central lines were harvested, totalizing an area of 5 m$^2$. The trait evaluated was the grain yield (kg ha$^{-1}$), correcting to 13% of moisture. Seed control and pest control management followed the technical recommendations for soybean cultivation at each site.

The variance components were estimated through restricted maximum likelihood, REML (PATTERSON & THOMPSON, 1971) and genotypic values were predicted through best linear unbiased prediction, BLUP (HENDERSON, 1975). The statistical model associated with the evaluation of genotypes in randomized complete block design in several sites, with one observation per plot, was given by the following equation:

$$y = Xf + Zg + Ti + e \text{ ,}$$

where $y$ is the vector of phenotypic data, $f$ is the vector of replication-environment combinations (assumed to be fixed), which contemplates the effects of environment and replication within environment, added to the overall mean, $g$ is the vector of the genetic effects (assumed to be random) ($g \sim N(0, \hat{\sigma}_g^2)$, where $\hat{\sigma}_g^2$ is the genotypic variance), $i$ is the vector of G×E interaction effects (random) ( $i \sim N(0, \hat{\sigma}_i^2$ ), where $\hat{\sigma}_i^2$ is the G×E interaction variance), and $e$ is the vector of residuals (random) ($e \sim N(0, R)$, where $R$ represents a matrix of residual variances). Capital letters X, Z, and T represent the incidence matrices for $f$, $g$, and $i$, respectively.

Models with Identity variance (IDV) and diagonal (DIAG) residual error variance structures (RESENDE et al. 2014), as presented below,

$$IDV = \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix} \sigma_e^2 = \begin{bmatrix} \sigma_e^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_e^2 \end{bmatrix}, \text{ and}$$

$$DIAG = \begin{bmatrix} \sigma_{e1}^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_{e10}^2 \end{bmatrix}$$

were tested by the Bayesian information criterion, BIC (SCHWARZ, 1978), given by:

$$BIC = -2LogL + pLog[n - r(x)],$$

where $LogL$ is the logarithm of the restricted maximum likelihood function, $p$ is the number of estimated parameters, $n$ is the number of observations, and $r(x)$ is the rank of the incidence matrix of fixed effects. The significance of the random effects of the statistical model was tested by the LRT (RAO, 1973) using chi-square statistics with one degree of freedom and a probability level of 1%.

Phenotypic variance ( $\hat{\sigma}_{phen}^2$ ), heritability of the total genetic effects ($h_g^2$), coefficient of determination of G×E interaction effect ( $c_i^2$ ), the genotypic correlation across environments ( $r_{gl}$ ), and selective accuracy ($r_{\hat{g}g_i}$) , were obtained by the following expressions (RESENDE et al., 2014):

$$\hat{\sigma}_{phen}^2 = \hat{\sigma}_g^2 + \hat{\sigma}_i^2 + \hat{\sigma}_{res}^2,$$

$$h_g^2 = \frac{\hat{\sigma}_g^2}{\hat{\sigma}_{phen}^2},$$

$$c_i^2 = \frac{\hat{\sigma}_i^2}{\hat{\sigma}_{phen}^2},$$

$$r_{gl} = \frac{\hat{\sigma}_g^2}{\hat{\sigma}_g^2 + \hat{\sigma}_i^2}, \text{ and}$$

$$r_{\hat{g}g_i} = \sqrt{1 - \frac{PEV_i}{\sigma_g^2}};$$

where $\hat{\sigma}_{res}^2$ is the residual variance and is the prediction error variance.

For the adaptability and stability analyses the following statistics were used: harmonic mean of the genetic values (HMGV), associated with the concept of stability; relative performance of the genetic values (RPGV), associated with the concept of adaptability; and HMRPGV, associated with the concepts of adaptability, stability, and productivity, simultaneously.

The predicted genetic value ($GV_{ij}$) which is the genetic value of genotype $i$ at environment $j$, HMGV, RPGV, and HMRPGV, of each genotype, were obtained by the following expressions:

$$GV_{ij} = m_j + g_i,$$

$$HMGV_i = \frac{n}{\sum_{j=1}^{n} \frac{1}{GV_i}},$$

$$RPGV_i = \frac{1}{n} \frac{\sum_{j=1}^{n} GV_i}{m_j}, \text{ and}$$

$$HMRPGV_i = \frac{n}{\sum_{j=1}^{n} \frac{1}{RPGV_i}};$$

where $g_i$ is the genetic effect of the genotype $i$, $n$ is the number of environments ($n = 10$), and $m_j$ is the mean productivity in environment $j$.

The direct selection were obtained by: $m_j + g_i + ge_{ij}$, where $m_j$ is the mean productivity in environment $j$, is the effect of genotype $i$, and $ge_{ij}$ is the effect of genotype $i$ × environment $j$ interaction). Further, the indirect selection for all environments were measured by: (i) $\mu + g_i$, (where: $\mu$ is the general productivity mean and $g_i$ is the effect of genotype $i$); (ii) considering stability and productivity (HMGV), (iii) adaptability and productivity (RPGV); and, (iv) stability, adaptability and productivity simultaneously (HMRPGV). The selection gain (SG) and SG in percentage, from direct and indirect selection, were obtained by the following expressions:(RESENDE, 2015):

$$SG = \frac{\sum_{i=1}^{n} GV}{p},$$

$$SG (\%) = \frac{M_s - M_0}{M_0} * 100,$$

where is the genotypic value, $p$ is the number of selected genotypes, $M_s$ is the productivity mean of selected genotypes, and $M_0$ is the mean productivity of the evaluated population. The SG was calculated considering a selection intensity of 20% (six genotypes). The concordances between the selection strategies were calculated using the Kappa coefficient (K) (COHEN, 1960), given by:

$$K = \frac{A - C}{D - C},$$

where: $A$ is the number of matching selected genotypes, $C$ is the number of selected genotypes due to chance ($C = bD$, where b is the selection intensity) and $D$ is the number of selected genotypes (six).

The variance components, genetic and non-genetic parameters, and genotypic values obtained using the Selegen-REML/BLUP software (RESENDE, 2016). The model accounting for different residual variance structures were tested using ASReml software (GILMOUR et al., 2015).

## RESULTS AND DISCUSSION

Traditional analyses of multi-environment trials, such as additive main effects and multiplicative interaction (AMMI) (GAUCH, 1992) and genotype + G×E biplot (GGE biplot) (YAN et al., 2000) environment (E, are based on the analysis of variance

(ANOVA), thus assuming fixed genotype effect and IDV residual variance (homoscedasticity) (LI et al., 2017). In the context of mixed model methodology (REML/BLUP), besides the effect of genotypes being advantageously treated as random, the needed of homoscedasticity can be circumvent, that is, REML/BLUP deal with DIAG residues (SMITH et al., 2005). This methodology allows; therefore, to model the G×E interaction, using for this purpose, different (co) variances structures. Thus, it is important to consider the structure of residual variance (for example, identity variance and diagonal) to select the model that provides the best fit (SMITH et al., 2005).

In this sense, several criteria for model selection were proposed, especially BIC (SCHWARZ, 1978). With this criterion, is implicit that there is a model that describes the relationship between the variables involved and the criterion tries to maximize the probability of choosing the true model. Thus, the BIC indicates the selection of parsimonious models, that is, the model that involves the fewest possible parameters to be estimated and that explains the behavior of the response variable well (RESENDE et al., 2014)the Akaike (AIC. In this study, the BIC values for grain yield trait were 11606.40 and 11620.32 for models with IDV and DIAG residual variance, respectively. Therefore, the best model was the one with IDV residual variance (lowest BIC). Thus, this model was adopted to estimate variance components and to genotypic values. Simply put, this result means that the residuals having the same scatter.

The variance components have a great importance in plant breeding, since the population and breeding strategy to be used depend on information that can be obtained from these components (RESENDE, 2015). According to the LRT, the random effects of the model (genetic and G×E interaction effects) were significant (p<0.01) for the grain yield trait (Table 1). The significance of the estimation of genetic and G×E

interaction effects via LRT demonstrates the existence of genetic variability and presence of G×E interaction. Similar results are reported for the grain yield trait in soybean crop (SHAW et al., 2016, WHALEY & ESKANDARI, 2019).

The determination coefficients indicate the amount of each effects where explained in relation to phenotypic variance. For instance, genotypic and G×E interaction effects explained, respectively, 18% (heritability) and 22% of phenotypic variance (Table 1). According to the classification presented by RESENDE (2015), the heritability of individual plots in the broad sense, that is, the total genetic effect (0.18) was of moderate magnitude. The weak genotypic correlation across environments (0.45) indicates the presence of complex type G×E interaction, which is problematic for the breeder. Thus, efficient selection methods are needed to capitalize on the favorable effects or circumvent the adverse effects of the G×E interaction (RESENDE et al., 2014). The higher the accuracy in evaluating a genotype, the greater the reliability of its predicted genotypic value. According to RESENDE & DUARTE (2007), the accuracy obtained in this study (0.90) is classified as very high, therefore, this result indicated a scenario very favorable to selection and recommendation.

As observed and expected, direct selection $(u_j + g_i + ge_{ij})$ (Table 2) has always led to greater gains with selection. However, when selecting via $u_j + g_i + ge_{ij}$ (where G×E interaction is captured) recommendations are restricted to each environment (GONÇALVES et al., 2014). The selection via HMGV, RPGV, and HMRPGV led to equal gains among them and similar those obtained via $u_. + g_i.$ (Table 2). In this way, the results showed that most productive soybean genotypes have high adaptability and stability. The HMGV classifies genotypes by genotypic values and stability, the lower the standard deviation of genotypic performance between environments, the higher the

Table 1 - Deviance, likelihood ratio test (LRT), variance components, and determination coefficients for the grain yield trait evaluated in 30 soybean genotypes in ten environments.

| Effect | Deviance | LRT | Variance component | Determination coefficient |
|---|---|---|---|---|
| Genotype | 11641.30 | 55.22[**] | 43889.85±14240.27 | 0.18±0.05 |
| G×E interaction | 11638.30 | 52.23[**] | 53205.46±9405.88 | 0.22±0.04 |
| Residue | - | - | 147229.36±8645.61 | - |
| Complete model | 11586.10 | - | 244324.67 | - |

[**]Significant at 1% probability by chi-square test. The null hypothesis was that the complete and reduced models do not differ from each other.

Table 2 - Selection gains (SG): direct for each environment (E); and indirect via $u. + g_{i.}$, harmonic mean of genetic values (HMGV), relative performance of the genetic values (RPGV), and harmonic mean of the relative performance of the genetic values (HMRPGV).

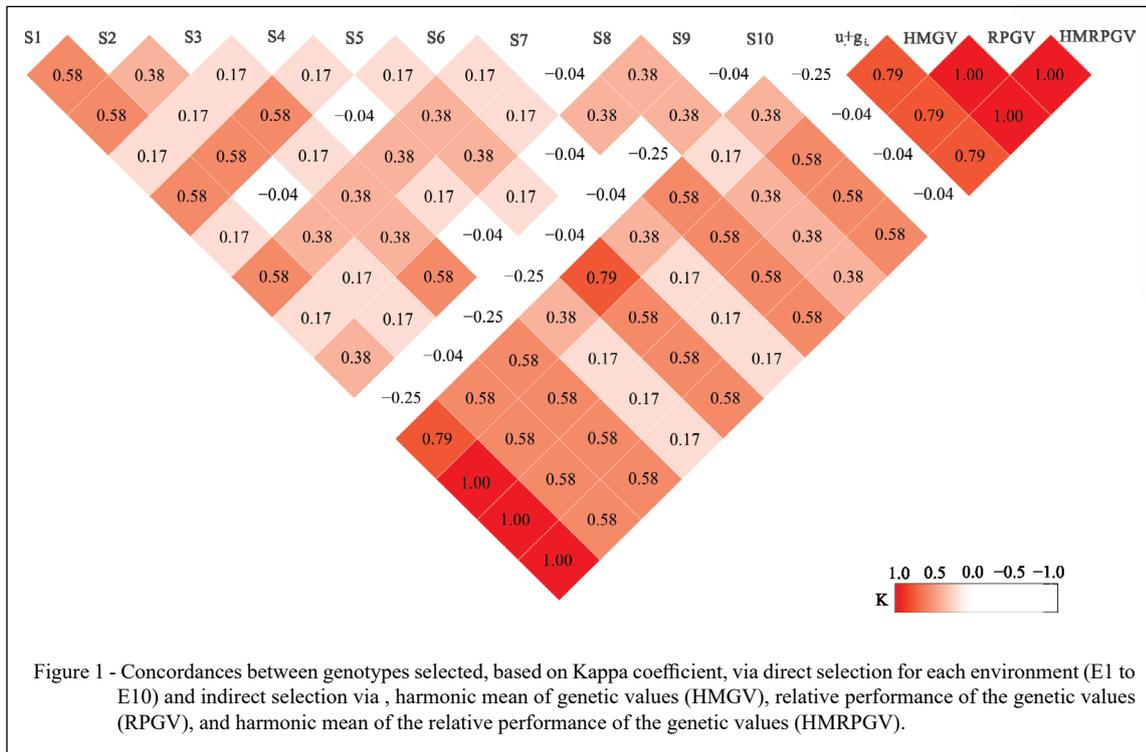| Environment | SG (kg ha$^{-1}$) | SG (%) |
| --- | --- | --- |
| E1 | 331.10 | 15.72 |
| E2 | 327.37 | 6.83 |
| E3 | 282.23 | 15.67 |
| E4 | 274.95 | 4.74 |
| E5 | 301.03 | 10.70 |
| E6 | 418.88 | 6.87 |
| E7 | 286.87 | 8.37 |
| E8 | 229.82 | 8.24 |
| E9 | 361.94 | 20.41 |
| E10 | 307.31 | 5.36 |
| $u. + g_{i.}$ | 197.65 | 5.32 |
| HMGV | 190.90 | 5.14 |
| RPGV | 190.90 | 5.14 |
| HMRPGV | 190.90 | 5.14 |

HMGV value. Thus, selection based on this statistic implies simultaneous selection for productivity and stability (RESENDE & DUARTE 2007). Therefore, indirect selection via HMGV is the most suitable strategy for known unfavorable environments, since for this type of environment genotypes with high stability are desirable.

The RPGV classifies genotypes by genetic values expressed as a ratio of the overall mean of each environment (RESENDE, 2007). Thus, selection based on this RPGV implies simultaneous selection for productivity and adaptability (RESENDE, 2007). RPGV selection is the most appropriate strategy for known favorable environments because the selected genotypes have greater responsiveness with the improvement of the environment. Finally, HMRPGV combines HMGV and RPGV statistics and implies simultaneous selection for productivity, stability and adaptability; and therefore encompasses favorable and unfavorable environments. According to the Kappa coefficient, the concordances between the selection strategies ranged from -0.04 to 1.00 (Figure 1). The HMGV, RPGV, and HMRPGV methods showed perfect agreement with each other. Moderate to near perfect agreement in ranking genotype was found among the environments 1, 2, 3, 5, 7, and 9 with the HMGV, RPGV, and HMRPGV methods. For environment 10, insignificant concordances, except with the environment 8, were observed.

From the six genotypes selected via HMGV, RPGV, and HMRPGV five were among the six genotypes selected via $u. + g_{i.}$, so these methods showed high agreement. This result demonstrated the efficiency of the HMGV, RPGV, HMRPGV, and $u. + g_{i.}$ methods in the selection of soybean genotypes. GONÇALVES et al. (2014), evaluating sugarcane clones, found that four of the five clones selected via $u. + g_{i.}$ were the same selected by the HMGV, RPGV, and HMRPGV methods and, by the three methods, the same clones were selected. COLOMBARI-FILHO et al. (2013), evaluating elite rice genotypes, reported 84% of coincidence in the selection by the $u. + g_{i.}$ and HMRPGV methods, and verified that the elite genotypes selected via HMRPGV were the same ones selected via RPGV. These authors concluded that the HMRPGV method is an appropriate tool for the selection of stable, adapted and high yield potential rice genotypes.

COLOMBARI-FILHO et al. (2013), evaluating genetic progress, showed a significant increase in stability and adaptability of rice genotypes developed during the period from 1996 to 2010. According to those authors, this fact is due to the introduction of new genotypes in the germplasm bank of the Embrapa breeding program, provided by the USA and France, in 1996. This allowed new crosses; and consequently, with the selection over the years it provided more adapted and stable genotypes. The company GDM seeds, holder of the soybean genotypes under evaluation in this study, originated in 1982, in Argentina. Today, the company also manages soybean breeding programs in several countries in South America, US and Canada. With its growth,

Figure 1 - Concordances between genotypes selected, based on Kappa coefficient, via direct selection for each environment (E1 to E10) and indirect selection via , harmonic mean of genetic values (HMGV), relative performance of the genetic values (RPGV), and harmonic mean of the relative performance of the genetic values (HMRPGV).

there were introductions of new genotypes to the germplasm bank and high selection intensities, making it possible to obtain genotypes with high adaptability, stability and productivity, simultaneously.

## CONCLUSION

The genetic parameters are easily estimated by mixed model methodology, being the very high accuracy values an indicative of model reliability in the genotypes ranking. The most prominent strategy for selection accounting for soybean productivity, stability, and adaptability is the HMRPGV.

## ACKNOWLEDGEMENTS

## DECLARATION OF CONFLICT OF INTERESTS

The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

## AUTHORS' CONTRIBUTIONS

All authors contributed equally for the conception and writing of the manuscript. All authors critically revised the manuscript and approved of the final version.

## REFERENCES

ALVES, R. S., et al. Multiple-trait BLUP: a suitable strategy for genetic selection of Eucalyptus. **Tree Genetics and Genomes**. v.14, n.7. 2018. Available from: <https://link.springer.com/article/10.1007/s11295-018-1292-7>. Accessed: Mar. 01, 2019.

BERNARDO, R. Reinventing quantitative genetics for plant breeding: something old, something new, something borrowed, something BLUE. **Heredity**. 2020. Available from: <https://www.nature.com/articles/s41437-020-0312-1>. Accessed: May, 02, 2020. doi: 10.1038/s41437-020-0312-1.

COHEN, J., 1960. A coefficient of agreement for nominal scales, Educational and psychological measurement. **Sage Publications**. v.10, n.1. 1960. Available from: <https://journals.sagepub.com/doi/abs/10.1177/001316446002000104>. Accessed: Mar. 05, 2019. doi: 10.1177/001316446002000104.

COLOMBARI-FILHO, J. M., et al. Upland rice breeding in Brazil: a simultaneous genotypic evaluation of stability, adaptability and grain yield. **Euphytica**. v.192, p.117–129. 2013. Available from:

<https://link.springer.com/article/10.1007/s10681-013-0922-2>. Accessed: Apr. 06, 2020.

GAUCH, H. G. **Statistical analysis of regional yield trials**: AMMI analysis of factorial designs. Elsevier Science Publishers, 1992, 278p.

GILMOUR, A. R., et al. **ASReml user guide release 4.1 structural specification**. Hemel hempstead VSN Int. 2015, 432p.

GONÇALVES, G. M. Breeding new sugarcane clones by mixed models under genotype by environmental interaction. **Scientia Agricola**. v.71, p.66–71. 2014 Available from: <https://www.scielo. br/scielo.php?pid=S0103-90162014000100009&script=sci_arttext>. Accessed: Dec. 21, 2019.

HENDERSON, C. R. Best linear unbiased estimation and prediction under a selection model. **Biometrics**. v.31, n.2, p.423–447. 1975. Available from: <https://www.jstor.org/stable/2529430>. Accessed: May, 01, 2020.

LI, Y., et al. Genotype by environment interactions in forest tree breeding: review of methodology and perspectives on research and application. **Tree Genetics and Genomes**. v.13. n.60. 2017. Available from: <https://link.springer.com/article/10.1007%2Fs11295-017-1144-x>. Accessed: Oct. 12, 2019. doi: 10.1007/s11295-017-1144-x.

LIU, X., et al. Genetic variation of world soybean maturity date and geographic distribution of maturity groups. **Breeding Science**. v.67, n.3, p.221-232. 2017. Available from: <https://www.jstage. jst.go.jp/article/jsbbs/67/3/67_16167/_article/-char/en>. Accessed: Jan. 01, 2019. doi: 10.1270/jsbbs.16167.

PATTERSON, H. D. & THOMPSON, R. Recovery of inter-block information when block sizes are unequal. **Biometrika**. v.58, p.545–554. 1971. Available from: <https://www.jstor. org/stable/2334389?seq=1>. Accessed: Mar. 04, 2020. doi: 10.2307/2334389.

RAO, C. R., **Linear statistical inference and its applications**. Wiley New York. 1973.

RESENDE, M. D. V. Software Selegen-REML/BLUP: A useful tool for plant breeding. **Crop Breeding and Applied Biotechnology**. v.16, p.330–339. 2016. Available from: <https://www.scielo.br/scielo.php?script=sci_arttext&pid=S1984-70332016000400330>. Accessed: Mar. 05, 2020. doi: 10.1590/1984-70332016v16n4a49.

RESENDE, M. D. V. **Genética quantitativa e de populações**. Suprema, Visconde do Rio Branco 463. 2015.

RESENDE, M. D. V. **Métodos estatísticos ótimos na análise de experimentos de campo**, Embrapa Florestas. Embrapa Florestas, Colombo, PR, 2004.

RESENDE, M. D. V. & DUARTE, J. B. Precisão e controle de qualidade em experimentos de avaliação de cultivares. **Pesquisa Agropecuária Tropical**. v.37, n.3, p.182–194. 2007. Available from: <https://www.revistas.ufg.br/pat/article/view/1867>. Accessed: Mar. 06, 2020. doi: 10.5216/pat.v37i3.1867.

RESENDE, M. D. V., et al. **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. 2014.

SCHWARZ, G. Estimating the dimension of a model. **The Annals of Statistics**. v.6, n.2, p.461–464. 1978. Available from: <https://www.jstor.org/stable/2958889?seq=1>. Accessed: Mar. 21, 2020.

SHAW, E. J., et al. Effect of genotype, environment, and genotype × environment interaction on tocopherol accumulation in soybean seed. **Crop and Science**. v.56, p.40–50. 2016. Available from: <https://acsess.onlinelibrary.wiley.com/doi/full/10.2135/cropsci2015.02.0069>. Accessed: Mar. 22, 2020. doi: 10.2135/cropsci2015.02.0069.

SILVA, F. L., et al. 2017. **Soybean Breeding**. Springer. 211p.

SMITH, A. B., et al. The analysis of crop cultivar breeding and evaluation trials: An overview of current mixed model approaches. **The Journal of Agricultural Science**. v.143, p.449–462. 2005. Available from: <https://www.cambridge.org/core/journals/journal-of-agricultural-science/article/analysis-of-crop-cultivar-breeding-and-evaluation-trials-an-overview-of-current-mixed-model-approaches/149BDADEE1D04E52616E74A1191A8F69>. . Accessed: Mar. 27, 2020. doi: 10.1017/S0021859605005587.

SOUZA, B. M., et al. Genotype-by-environment interaction in *Corymbia citriodora* (Hook.) K.D. Hill, &amp; L.A.S. Johnson progeny test in Luiz Antonio, Brazil. **Forestry Ecology and Management**. v.460. n.15, 117855. 2020. Available from: <https://www.sciencedirect.com/science/article/abs/pii/S0378112719324363>. Accessed: Mar. 26, 2020. doi: 10.1016/j.foreco.2019.117855.

US Departament of Agriculture, 2020. **Oilseeds**: World markets and trade. USDA. Available from: <https://fdc.nal.usda.gov/index.html>. Accessed: Mar. 23, 2020.

WHALEY, R., & ESKANDARI, M. Genotypic main effect and genotype-by-environment interaction effect on seed protein concentration and yield in food-grade soybeans (*Glycine max* (L.) Merrill). **Euphytica**. v.215, n.33. 2019. Available from: <https://link.springer.com/article/10.1007/s10681-019-2344-2>. Accessed: Mar. 25, 2020. doi: 10.1007/s10681-019-2344-2.

YAN, W., et al. Cultivar evaluation and mega-environment investigation based on the GGE biplot. **Crop and Science**. v.40, p.597–605. 2000. Available from: <https://acsess.onlinelibrary.wiley.com/doi/abs/10.2135/cropsci2000.403597x>. Accessed: Mar. 26, 2020. doi: 10.2135/cropsci2000.403597x.