

Mega-environment analysis of maize breeding data from Brazil

Francielly de Cássia Pereira^{1*}, Magno Antonio Patto Ramalho¹, Marcio Fernando Ribeiro de Resende Junior², Renzo Garcia Von Pinho³

¹Universidade Federal de Lavras – Depto. de Biologia, Av. Central s/n – C.P. 3037 – 37200-900 – Lavras, MG – Brasil.

²University of Florida – Horticultural Sciences Dept., 2550 Hull Road – 32611 – Gainesville, FL – USA.

³Universidade Federal de Lavras – Depto. de Agricultura, Av. Central s/n – 37200-900 – Lavras, MG – Brasil.

*Corresponding author <franciellyfcp@gmail.com>

Edited by: Alencar Xavier

Received October 01, 2020

Accepted December 22, 2020

ABSTRACT: The development and recommendation of single cross maize hybrids (SH) to be used in extensive land areas (mega-environments), and in different crop seasons requires many experiments under numerous environmental conditions. The question we asked is if the data from these multi-environment experiments are sufficient to identify the best hybrid combinations. The aim of this study was to critically analyze the phenotype data of experiments of yield, established by a large seed producing company, under a high level of imbalance. Data from evaluation of 2770 SH were used from experiments conducted over four years, involving the first and second crop seasons, in 50 locations of different years and regions of Brazil. Different types of analysis were carried out and genetic and non-genetic components were estimated, with emphasis on the different interactions of the SH with the environments. Results showed that the coincidence of common hybrids in these experiments is normally small. The estimates of the correlations between of the hybrids coinciding in the environments two by two is of low magnitude. The hybrid × crop season interaction was always expressive; however, the interactions of hybrids and other environmental variables were also important. Under these conditions, alternatives were discussed for making with the information obtained from the experiments, can be more efficient on the process to obtain new hybrids by companies.

Keywords: genotype × environment interaction, unbalance data, hybrid recommendation process, variance components, plant breeding

Introduction

Two maize crop seasons are common in Brazil per year. The first crop occurs from Sept to Dec, while the second season is from Jan to Apr. The environmental conditions between these two crop seasons are quite distinct in relation to temperature and rain distribution. In addition, farmers' use of technology in maize growing is quite diversified. This makes selection of hybrids for recommendation under these different conditions a much greater challenge than that in temperate regions, for example.

In order for a breeder to be successful in the identification of hybrids adapted to the mega-environment of maize growing, the hybrids obtained annually must be broadly evaluated. Clearly, these evaluations will only be successful if the experiments are conducted in the greatest number of environments possible. Experience in respect shows that secure recommendation was only possible through middle-term results coming from hundreds of replications (Troyer, 1996; Gaffney et al., 2015). However, companies obtain thousands of hybrids annually, which makes testing in multiple replications difficult. In this way, the same hybrid will rarely be evaluated in all the environments, resulting in a excessive imbalance data and consequently hindering the decision making at the time of recommendation.

In many situations, these experiments are used to evaluate the possibility of employing genomic selection in prediction of potentially superior hybrid combinations. In this sense, the more accurate the model is, the greater the association of future performance of

the hybrid will be through the response of the genotyped line. Previous experiences show that the effect of the hybrid × environment interaction greatly complicates the prediction process. Because of this interaction, the responses of the hybrids do not coincide in the diverse environments evaluated. An alternative is to include this effect in the predictive models to obtain more accurate information (Lado et al., 2016; Ferrão et al., 2018; Dias et al., 2018a; Montesinos-López et al., 2019; Krause et al., 2020). The question is whether the hybrid × environment interaction component obtained from highly unbalanced experiments can contribute to the predictive models.

Thus, the purpose of the present study was to analyze the phenotypic data from yield experiments of different crop years and seasons, estimate genetic and phenotypic parameters under a high level of imbalance and comment the impact of these conditions on breeder decisions to selection maize hybrids.

Materials and Methods

Genetic material, experimental design, and environments

Grain yield data ($t\ ha^{-1}$), kindly provided by a Brazilian company of hybrid maize cultivars, were used in this study. These data were obtained over four years with two crop seasons per year including numerous locations in the central and southern regions of Brazil (Figure 1). During this period, 2770 SH of maize were evaluated. These hybrids originated from crosses of 447 lines coming from different tropical, subtropical,

and temperate regions around the world. Due to the breeding program in question being a line introgression program, the number of SH common to the two crop seasons, as well as the number of experiments and of treatments evaluated per experiment in each location, was quite variable (Table 1 and Figure 2).

Randomized block (RBD) and incomplete block (IBD) experimental designs were used for evaluation of the hybrids, with two or three replications. The plots consisted of four 5 m rows with a 0.7 m between-row spacing. Different experiments were set up within each crop season in the same location. The experiments within each location were connected through check varieties in common since the hybrids evaluated in each experiment were not necessarily the same. Additional information regarding the number of hybrids, the locations, experiments, replications, and experimental design adopted in each crop season is provided in Table

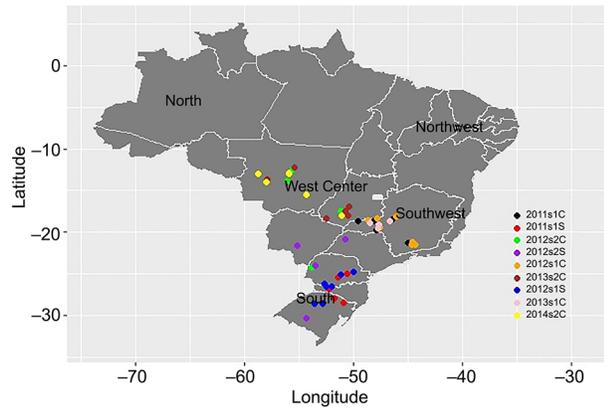


Figure 1 – Map of Brazil, showing the locations where experiments for evaluation of single cross hybrids of maize were conducted in the central and southern regions. Points in black, red, orange, blue, and pink correspond to experiments set up in the first crop season; points in green, purple, brown and yellow correspond to experiments set up in the second crop season. Crop seasons are labeled by their year of sowing (2011 – 2014), followed by their crop season (first crop, s1, or second crop, s2) and region (West Center, C, or South, S).

1. Each crop season was identified by an abbreviation that corresponds to the year of sowing (2011, 2012, 2013, or 2014), followed by the crop season (first crop, s1, or second crop, s2) and region (West Center, C, or South, S) (Figure 1 and Table 1).

Statistical analyses

For better characterization of the dataset, the overall mean per crop season and the variation among mean values of the SH in the different experiments and, subsequently, in the locations were estimated. Considering only the data from the 2011s1C crop season for the purpose of making inferences regarding what occurs among locations in the same crop season, genetic

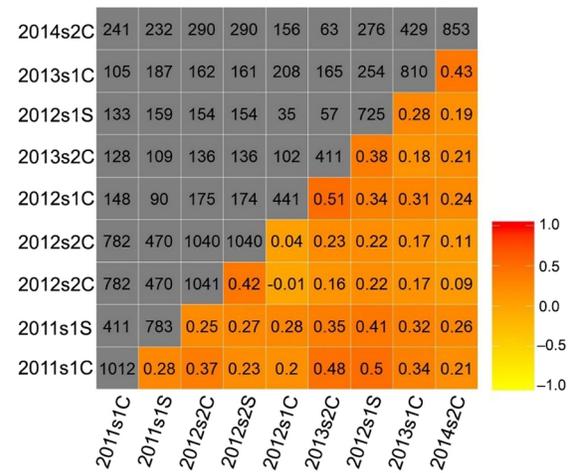


Figure 2 – The diagonal of the heat plot corresponds to the number of hybrids evaluated in each crop season. Upper diagonal heat plot indicates the number of genotypes in common that were evaluated in the pairs of environments while lower diagonal indicates the estimates of the correlation between the mean values of the SH in common in the crop seasons two by two. Crop seasons are labeled by their year of sowing (2011 – 2014), followed by their crop season (first crop, s1, or second crop, s2) and region (West Center, C, or South, S).

Table 1 – Description of the crop season, year, abbreviation, experimental design, number of single cross hybrids (SH), locations, experiments (EXP) and replications (REP) in each crop season.

Crop season	Year	Abbreviation	Design	SH	Location	EXP	REP
West Center 1 st crop	2011/2012	2011s1C	RBD*	1012	5	28	3
South 1 st crop	2011/2012	2011s1S	RBD	783	5	22	3
West Center 2 nd crop	2012	2012s2C	RBD	1041	6	24	3
South 2 nd crop	2012	2012s2S	RBD	1040	4	22	3
West Center 1 st crop	2012/2013	2012s1C	IBD	441	7	13	3
South 1 st crop	2012/2013	2012s1S	IBD	725	7	19	2
West Center 2 nd crop	2013	2013s2C	IBD	411	6	11	3
West Center 1 st crop	2013/2014	2013s1C	IBD	810	4	20	2
West Center 2 nd crop	2014	2014s2C	IBD	853	6	21	2

*Randomized block (RBD) and incomplete block (IBD) experimental designs.

variance (σ_g^2) among the hybrids evaluated, variance of the error (σ_e^2), and heritability (h^2) were estimated in each location using the following statistical model: σ^2

$$y = X\tau + Z_g u_g + Z_{ge} u_{ge} + Z_b u_b + e \quad (1)$$

where y is the vector of phenotypic observations; τ is the vector of fixed effect of the experiment; u_g is the vector of random genotypic effects of hybrids, with $u_g \sim N(0, \sigma_g^2 I_g)$; u_{ge} is the vector of random effects of the hybrid by experiment interaction, with $u_{ge} \sim N(0, \sigma_{ge}^2 I_{ge})$; u_b is the vector of random effect of replication within experiments, with $u_b \sim N(0, \sigma_b^2 I_b)$; e is the vector of random errors, with $e \sim N(0, \sigma_e^2 I_n)$; X , Z_g , Z_{ge} , and Z_b are the incidence matrices associated with the vectors τ , u_g , u_{ge} , and u_b ; σ_g^2 , σ_{ge}^2 , σ_b^2 , and σ_e^2 are the variance components associated with the vectors u_g , u_{ge} , u_b , and e ; and I , I_g , I_{ge} , I_b , and I_n are the identity matrices associated with the vectors u_g , u_{ge} , u_b , and e . In each crop season, the grain yield data were analyzed through a mixed models approach considering the model according to the experimental design adopted:

$$y = X\tau + Z_g u_g + Z_{gl} u_{gl} + Z_b u_b + e \quad (2)$$

where y is the vector of phenotypic observations; τ is the vector of fixed effects (experimental RBD: location and experiment within location; experimental IB: location, experiment within location, and replication within experiment and location); u_g is the vector of random genotypic effects of hybrids, with $u_g \sim N(0, \sigma_g^2 I_g)$; u_{gl} is the vector of random effects of the hybrid by location interaction, with $u_{gl} \sim N(0, \sigma_{gl}^2 I_{gl})$; u_b is the vector of random effects (experimental RBD: block within experiment and location; experimental IB: block within replication, experiment, and location, with $u_b \sim N(0, \sigma_b^2 I_b)$); e is the vector of random errors, with $e \sim N(0, R)$; X , Z_g , Z_{gl} , and Z_b are the incidence matrices associated with the vectors τ , u_g , u_{gl} , and u_b ; σ_g^2 , σ_{gl}^2 , σ_b^2 , and σ_e^2 are the variance components associated with the vectors u_g , u_{gl} , u_b , and e ; and I_g , I_{gl} , I_b , and I_n are the identity matrices associated with the vectors u_g , u_{gl} , u_b , and e . The residual (co)variance matrix, with the aim of modeling the effect of location within each crop season, adopted a diagonal block variance structure, using the identity matrix ($R = \bigoplus_{l=1}^l \sigma_e^2 I_{n_l}$).

Previously, alternative methodologies of unstructured variance covariance matrix were tested to try to model the genetic correlation between all environment pairs. These matrices allow a better understanding of the genetic structure and evaluate the stability of genotypes in mega-environments. To this end, the genetic and residual effects were considered as $u_g \sim MVN(0, \sigma_g^2 I \otimes \Sigma_1)$ and $e \sim MVN(0, \sigma_e^2 I \otimes R_1)$. Σ_1 is the VCOV matrix for the additive genetic effects in the l environments and R_1 represents the VCOV matrix for the residual effects in the l environments. In this case, the main environment effects were implicitly modeled and

an unstructured form for the genetic Σ_1 and residual R_1 VCOV matrix was assumed. Because of the large number of sites evaluated in each season (> 5) the convergence of these unstructured matrices was difficult. In this way, the diagonal block variance structure was adopted, as described above, to model the genetic and residual effects in this study.

The variance components associated with the random effects were obtained using the residual maximum likelihood method (REML) (Patterson and Thompson, 1971) and their significance levels were verified by the likelihood ratio test. To make inferences regarding the occurrence of interaction, the correlation (r_{qs}) among the mean values of the SH coinciding in the q and s crop seasons was estimated. The estimator used was similar to that presented by Steel et al. (1997):

$$r_{qs} = \frac{\text{cov}(\overline{HS}_{iq} \times \overline{HS}_{is})}{\sqrt{\sigma_{HS_{iq}}^2 \times \sigma_{HS_{is}}^2}} \quad (3)$$

where \overline{HS}_{iq} is the mean of the single cross hybrid i in crop season q ; \overline{HS}_{is} is the mean of the single cross hybrid i in crop season s ; $\sigma_{HS_{iq}}^2$ is the variance of the single cross hybrid i in crop season q ; and $\sigma_{HS_{is}}^2$ is the variance of the single cross hybrid i in crop season s .

In addition, the effect of the hybrid \times crop season interaction was also verified through the coincidence of the genotypes selected based on the mean of two environments (considering different combinations of year, region, and sowing time) in relation to selection based on the mean of each environment individually. For that purpose, the maize yield data from each location within the crop seasons were fitted regarding the effect of blocking and replication, according to the design adopted in each situation, to obtain the EBLUE (Empirical Best Linear Unbiased Estimation) estimates. These estimates were used to carry out the individual analyses of each crop season and also the combined analyses of the environments two by two. Using the EBLUP (Empirical Best Linear Unbiased Prediction) predictions, the ten best hybrids in the mean of the environments and also in each one of the environments were selected.

Furthermore, in each crop season, the correlation between the mean value and the EBLUP of the SH was estimated, and the estimates of heritability were obtained using the following estimators:

Standard method using the expression presented by Falconer and Mackay (1996) that presupposes balanced data and independent genetic effects:

$$h^2 = \frac{\sigma_g^2}{\sigma_g^2 + \frac{\sigma_{g \times l}^2}{l} + \frac{\sigma_e^2}{l \times r}} \quad (4)$$

Holland et al. (2003) cited by Piepho and Möhring (2007) recommended for cases of imbalance and fixed effects of genotype:

$$h_H^2 = \frac{\sigma_g^2}{\sigma_g^2 + \bar{v}_{EBLUE} / 2} \quad (5)$$

where \bar{v}_{EBLUE} means the mean variance of the difference of the fitted mean values of two treatments (EBLUE).

Cullis et al. (2006), also estimated in cases of imbalance and random effect of genotype:

$$h_C^2 = 1 - \frac{\bar{v}_{EBLUP}}{2\sigma_g^2} \quad (6)$$

where \bar{v}_{EBLUP} means the mean variance of the difference of two EBLUPs.

Finally, combined analysis was carried out involving the single cross hybrids common to two crop seasons using different alternatives. The crop seasons chosen were 2011s1C/2012s2C and 2013s1C/2014s2C. These crop seasons were evaluated in different years and also differed in experimental accuracy and in the number of hybrids that coincided among them. In all the alternatives adopted, selection was made of the 20 best SH common to the two crop seasons chosen (2011s1C/2012s2C and 2013s1C/2014s2C).

Combined analysis based on the EBLUE estimates that were obtained for the hybrids in each environment was performed a) involving only the SH common to the two environments and common residual variance; b) considering all the SH present in the two environments, i.e., also those that were eliminated in the year, crop season, or location and common residual variance; c) involving only the SH common to the two environments and considering a residual variance different for each environment using the diagonal matrix, and d) considering all the SH present in the two environments and residual variance different for each environment.

Results

The dataset evaluated is typical of breeding programs with the objective of line introgression to obtain new SH. As many lines do not adapt well to Brazilian climate conditions, an imbalance was observed in the number of SH, locations, experiments, and replications evaluated over the crop seasons (Table 1).

The number of SH that were repeated among the crop seasons varied widely. Comparing the 2011s1S crop season and the 2012s1S crop season, of the 783 SH evaluated in the first year, only 159 proceeded, i.e., high selection intensity was applied and only 20 % of the SH evaluated in 2011s1S were allocated to the experiments in the following crop season. An even more complex scenario was observed between the 2012s2C and 2013s2C crop seasons, where only 13 % of the hybrids evaluated in 2012s2C advanced to the 2013s2C crop season. The same observation is valid for other years when comparing crop seasons and/or regions, and it becomes clear that there is great difficulty in evaluating data from different crop seasons in a combined manner (Figure 2).

The overall mean of the crop seasons ranged from 3.9 to 9.7 (t ha⁻¹), and the first crop seasons (8.3 t ha⁻¹) were 1.56 times higher yielding than the second crop seasons (5.3 t ha⁻¹), regardless of the region evaluated. The variation among the crop seasons was high; for example, for the West Center region, the yield in the first crop season in 2011/2012 was 58 % greater than in the second crop season. However, in 2012/2013, this superiority was much lower, only 18 %, but returned to a higher level in 2013/2014 at 51 %, once more showing the discrepancy among the crop seasons evaluated (Table 2).

Within each crop season, a greater variation in mean yield of the SH tested was observed among the experiments than among the locations. The amplitude of variation of the experiments in relation to the mean was up to 110 %, as is the case of the 2011s1S crop season ($[(11.2 - 2.7) / 7.6] = 1.10$). The variation of the lower and upper limits among the locations in relation to the overall mean was under 55 %, except in the 2011s1S and 2014s2C crop seasons. It is important to highlight that the greater variation among experiments is a result of the effect of locations and also of the different SH evaluated among the experiments (Table 2).

The grain yield data were fitted through the mixed models/REML approach. The estimates of correlations between the mean of the SH and their EBLUPs was greater than the 0.9 involving the crop seasons of more recent years. It follows that under these conditions, an

Table 2 – Overall mean, lower limit (LL), upper limit (UL) of mean grain yield of single cross maize hybrids in the experiments and environments of each crop season in t ha⁻¹ and correlation between the mean value and the EBLUP of each hybrid in the nine crop seasons evaluated (r).

Crop Season	LL and UL of the mean yield of the experiments	LL and UL of the mean yield of the locations	Overall mean of the crop seasons	r
2011s1C	6.0 – 9.9	6.5 – 9.3	8.5	0.80
2011s1S	2.7 – 11.1	4.8 – 9.8	7.6	0.77
2012s2C	3.4 – 8.2	4.2 – 7.6	6.0	0.69
2012s2S	2.7 – 5.6	3.3 – 5.2	3.9	0.62
2012s1C	4.8 – 7.6	5.7 – 7.6	7.2	0.94
2012s1S	6.3 – 10.3	7.9 – 11.4	9.7	0.90
2013s2C	3.6 – 5.9	4.6 – 7.2	5.6	0.96
2013s1C	5.9 – 9.2	6.6 – 10.1	8.5	0.99
2014s2C	1.9 – 7.3	2.9 – 7.9	5.7	0.97

optimal association between the mean values and the EBLUPs of the SH was obtained (Table 2).

In all the crop seasons, hybrids were observed with discrepant performance in relation to the set evaluated. The 2012s1S crop season exhibited the widest amplitude of variation, which was associated with the highest mean yield value (9.7 t ha⁻¹). In that crop season, the EBLUPs of the hybrids ranged from -4.0 to 3.1. In the 2012s2C crop season, lower amplitude of variation and a mean value of 6.0 t ha⁻¹ was found, with EBLUPs ranging from -0.7 to 1.0 (Table 2 and Figure 3A).

To more easily make inferences regarding what happens among the locations within the same crop season, each location in the 2011s1C crop season was analyzed in detail. In that year, a total of 28 different experiments were set up; however, the number of experiments evaluated in each location was different. The mean yield variation of the experiments in each location was at most 44 % (6.1 to 8.8 t ha⁻¹), and the overall mean of the five locations evaluated was from 6.5 to 9.2 t ha⁻¹. Variation was observed in the magnitude of the estimates of genetic variance and of standard heritability among the five locations. The variance of the hybrid × experiment interaction was greater than the genetic variance in all the locations evaluated in this season, except for one of locations where both variances were similar.

The estimates of the genetic variance components (σ_g^2) involving all the experiments and locations obtained in each crop season were significant by the likelihood ratio test, indicating the presence of genetic variability and the possibility of selection among the hybrids. The magnitude of the variance of the hybrid by environment interaction in relation to genetic variation was expressive, reflecting the hybrid performance that did not coincide across the environments. The $\sigma_{g \times l}^2 / \sigma_e^2$ ratio ranged from 0.12 (2013s1C) to 4.47 (2012s2C) among the crop seasons, accentuating what was commented.

The expressive existence of the hybrid × environment interaction within each crop season was also found through the estimates of the correlations (r_{qs}) of the mean performance of the SH coinciding in the crop seasons two by two. The estimates of r_{qs} across the combinations of crop seasons ranged from -0.01 to 0.51. In the pairs of environments 2012s2C/2012s1C, 2012s2S/2012s1C, 2012s2C/2014s2C, and 2012s2S/2014s2C, in which the lowest estimates of r were observed, high intensity of selection applied was also found in the SH evaluated from one crop season to another (Figure 2).

The presence of the interaction was also highlighted by the coincidence among the ten best SH selected in the mean of the EBLUPs of the two crop seasons in relation to their relative performance in each crop season. When the same region and year of evaluation were considered, i.e., the response of the first and second crop season, the coincidence varied between the sowing times. As expected, the greatest coincidence in most cases was in the first crop season. However, even under these conditions, in the 2011/2012 and 2013/2014 crop years, the coincidence was less than 50 %. These results, once more showing, the difficulty of moving toward recommendation of new SH involving different regions, sowing times, and crop years.

As the experiments were unbalanced, the heritability (h^2) in the mean of the hybrids within each crop season was estimated considering three strategies: i) standard method according to Falconer and Mackay (1996), ii) method according to Cullis et al. (2006), and iii) method according to Holland et al. (2003), cited by Piepho and Möhring (2007). Strategy i, prescribing the use of balanced data, ranged from 0.46 to 0.69 among the crop seasons, and as expected, was always superior to the estimates of h^2 obtained by the other strategies, except in the 2013s1C crop season (Figure 3B).

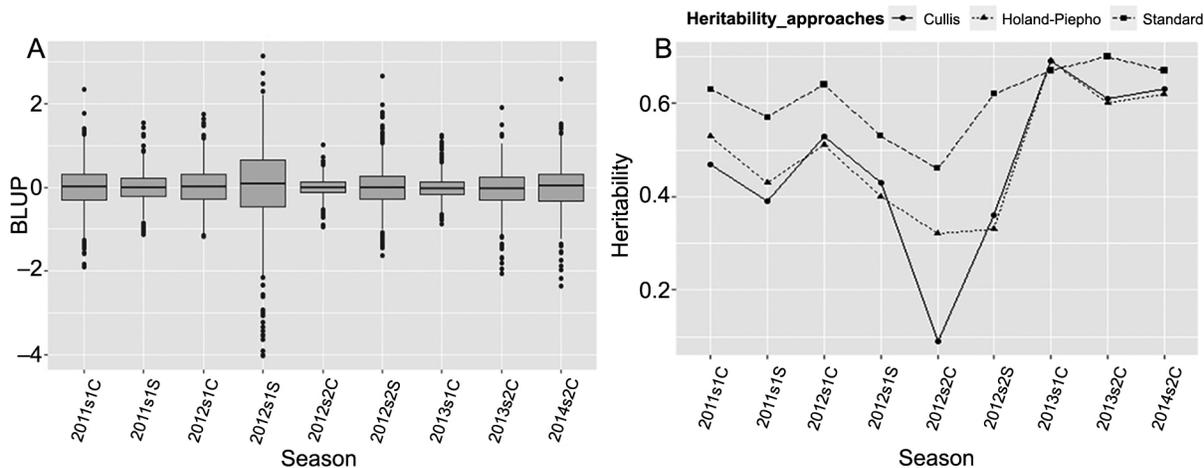


Figure 3 – A) Boxplot representing the variation between the upper and lower EBLUP limit of the predictions of the single cross hybrids evaluated in each crop season and possible outliers. B) Heritability estimates for grain yield over nine crop seasons using three different methods (standard, Cullis, and Holland-Piepho).

The estimates of heritabilities proposed by Cullis et al. (2006) and by Holland et al. (2003); Piepho and Möhring (2007) were always of similar magnitude, except in the 2012s2C crop season, and ranged from 0.09 to 0.69 among the crop seasons evaluated. However, the coincidence in the estimates of h^2 of the three strategies was greater in the experiments conducted in recent years (Figure 3A).

In carrying out combined analysis involving the SH common to two or more environments, there are some alternatives. One is involving only the SH in common and the other would be considering all the SH, i.e., also those that were eliminated in some environments. It is also possible to carry out the analyses considering residual variance in common or heterogeneous residual variance. In this study, combined analysis was performed considering the combinations of the 2011s1C/2012s2C and 2013s1C/2014s2C crop seasons. These pairs were chosen as they consisted of data from the first and second crop seasons of different years that also differed in experimental accuracy and in the number of coinciding hybrids.

For the combination of 2011s1C/2012s2C that involves the first and second crop seasons in the 2011/2012 crop year, the coincidence of the 20 best SH, which is what most interests breeders, changed significantly. Of the 20 best SH ranked in the analysis involving all the hybrids common to the two crop seasons, only five remained when the analysis was performed considering only the SH common to the two crop seasons. This information is valid for both cases, when the homogeneous or heterogeneous residual variance is considered.

Different results were observed for the combination of the first and second crop seasons in the 2013/2014 crop year (2013s1C/2014s2C combination) in relation to previous crop seasons. Coincidence in identification of the 20 best SH, involving all the SH or only the SH common to two crop seasons, was total upon using the same residual variance. However, when the residual variance used was heterogeneous, though the coincidence was high, it was not total (15 SH in 20 SH).

Discussion

The challenge common to all companies is evaluating a large number of SH annually for the purpose of recommending those that have the best performance for farmers. The information coming from these evaluations is often unbalanced in relation to the number of SH, of replications, of experiments, and of locations, which may compromise the choice of the best hybrid. In addition, the breeder needs to deal with the hybrid \times environment interaction in seeking greater reliability in future recommendations since this interaction is a complicating factor in the performance of the SH evaluated.

The results obtained in this analysis showed considerable substitution of hybrids among the different environmental conditions (Figure 2). This low coincidence among the hybrids evaluated is explainable because if a determined SH evaluated did not have good results under certain conditions, why reevaluate it under other conditions? Other plausible explanations would be the difficulty of continuing to evaluate a SH that has some agronomic trait other than grain yield that would make its future recommendation unviable, as well as the lack of adaptation of the lines to tropical conditions, which impedes the production of hybrid seeds in large quantity for evaluation of the SH in different environments. Thus, it is expected above all between crop seasons that the low coincidence among the SH evaluated is a reality that likely will not change.

The low coincidence among the hybrids evaluated in the different environments makes it difficult to estimate genetic and phenotypic parameters, especially genetic variance and, above all, the SH \times environment interaction. This difficulty has been reported in the literature by diverse authors in recent years (Smith et al., 2001; Möhring and Piepho, 2009; Smith et al., 2015; Nuvunga et al., 2015; Silva et al., 2019).

To deal with unbalanced data, some proposals have been implemented more recently for analysis of experiments with plants using, for example, analysis in two steps, in which weighting is considered in the second step in accordance with the number of replications, with the experimental design, and with residual variance (Smith et al., 2001; Möhring and Piepho, 2009; Welham et al., 2010; Piepho et al., 2012). Other alternatives are multiplicative models (Smith et al., 2015; Nuvunga et al., 2015), sequential analysis, which considers all the hybrids evaluated in the previous generations (Piepho and Möhring, 2006) and models that consider the use of heterogeneous residual variance (Edwards and Jannink, 2006; So and Edwards, 2011; Orellana et al., 2014; Hu et al., 2014; Andrade et al., 2015; Silva et al., 2019).

Due to the wide variation in analytical possibilities for unbalanced experiments in multi-environments, in this study, individual analyses were initially performed in each location within each crop season. Due to the great volume of information, we chose to present only the results in reference to the 2011s1C crop season. In this analysis, the importance of the SH \times experiment interaction was clear, even in a single location. This was possible because some SH were present in more experiments.

After that, the yield data from each crop season were fitted through mixed models regarding the block and location effect, seeking to obtain the best estimates of the genetic value of each hybrid. In conditions as observed in this study, wherein a huge number of hybrids were evaluated in many environments across the years, the use of approaches like unstructured VCOV matrices and factorial analytic models have been adopted, once these structures allow to model different

genetic variance to each site and different covariances between pairs of environments evaluated (Smith et al., 2002; Burgueño et al., 2012; Krause et al., 2020; Oliveira et al., 2020).

In this study, unstructured VCOV structures were tested to better understand the correlation between environments by including the genotype by environment interaction in the model. However, models that include unstructured VCOV matrices shows computational difficulties to converge due a huge number of parameters to be estimated under high unbalance level. The factor analytic structure is an alternative approach to deal with these limitations (Smith et al., 2002; Kelly et al., 2007; Dias et al., 2018b). Due to the difficulty of convergence of models and computational limitations, in this study we adopted VCOV simpler structures to model the effect of location within each crop season. Despite that, always is possible, it is important to assume models with more complex structures, as previously mentioned.

In situations in which the dataset exhibits considerable imbalance, a way of checking the fit of the model is to correlate the mean values and the EBLUPs of the SH. In general, the estimates of correlation in most of the crop seasons evaluated were high, especially in more recent years, showing that in many situations, when the experiments are well conducted, the mean can be considered a good indication of the performance of the SH, even under unbalanced conditions (Figure 2).

A significant effect of the SH × environment (locations and experiments within locations) interaction was found in all the crop seasons evaluated. In these cases, the SH × environment interaction component was greater than the genetic variance component in most of the crop seasons. This is very frequent in most of the situations in which various hybrids are evaluated in the same crop season (Tonk et al., 2011; Nzuve et al., 2013; Ndhlela et al., 2014; Mengesha et al., 2019). In the conditions evaluated, the significant effect of the interaction is expected due to the expressive environmental variation of numerous factors, such as climate, soil fertility, and management practices that occurs in the different locations in which maize experiments are conducted (Noia Junior et al., 2019; Embrapa, 2020).

It should be emphasized that environmental variations under tropical and subtropical conditions are more expressive than those normally observed under temperate conditions. This environmental variation is even more challenging since it is largely unpredictable (Eeuwijk et al., 2016). Given this situation, the great challenge of breeders is identifying hybrids that are more adapted and stable under these growing conditions. For that purpose, numerous methods have been proposed in the literature over the past fifty years (Eberhart and Russel, 1966; Wricke and Weber, 1986; Gauch and Zobel, 1988; Piepho, 1997; Yan et al., 2000; Smith et al., 2015; Nuvunga et al., 2015); most recently, the use of mixed models has been proposed above all, according to a survey performed by Eeuwijk et al. (2016).

In addition, in Brazil, variation in environmental factors in the second crop season is much more expressive than in the first, especially due to drought stress or heat stress (Andrea et al., 2019; Andrea et al., 2018), and so a difference in mean yield between the crop seasons is expected. In spite of that, this yield difference has diminished through the choice of more adapted hybrids and the use of greater technology in crop fields. The great challenge for seed production companies currently has been identifying hybrids adapted to both growing conditions. The results obtained using this dataset show that finding a hybrid with wide adaptation to different climatic regions is a challenging factor for breeders because of the enormous contribution of the SH × crop season interaction (Figure 2).

In the present study, the effect of the interaction on SH performance can be observed through estimation of the correlation between the mean of the hybrids in common across the crop seasons (Figure 2). Another option for the study of the interaction with greater importance for breeders was the coincidence of the hybrids selected considering two or more environments. The low magnitudes of the estimates of correlation and the low coincidences observed show that in most of the cases evaluated, the response to the interaction was of a complex nature and in some cases it was probably not even linear, making identification of the best hybrid difficult. Results similar to these are discussed by Eeuwijk et al. (2016) through graph illustrations involving the yield of the genotype and environmental quality.

The heritability (h^2) estimate is a key parameter in plant breeding because it is associated with predictive measurement of success in selection. It has been estimated by the ratio between the part of genetic variance exploited by the genotypes evaluated and the phenotypic variance of the selection unit applied (Falconer and Mackay, 1996; Bernado, 2010). However, with the increased use of mixed models to attenuate the effects of unbalanced data, new options of h^2 estimates have been proposed (Cullis et al., 2006; Piepho and Mohring, 2007; Schmidt et al., 2019).

In this study, h^2 was estimated by three procedures. Especially in the first crop seasons, the estimates did not greatly coincide. However, in more recent crop seasons, there was greater coincidence. It should be emphasized that, as was expected in all cases, the absolute value of h^2 in the standard method was superior to the other two (Cullis et al., 2006; Schmidt et al., 2019) (Figure 3A). This occurs because in the standard method, phenotypic variance is estimated considering that there is no variation in the number of replications and locations, for example. This discrepancy in the estimates of h^2 has also frequently been observed in other conditions (Piepho and Mohring, 2007; Schmidt et al., 2019).

Regardless of the method used, the h^2 estimates, in most cases, were considered of medium magnitude, which is a favorable condition for selection of SH, based

on the overall mean of each crop season. It should be emphasized that when selection is made among SH, all the genetic variance is used, i.e., additive, dominant, and epistatic variance (Hallauer and Miranda Filho, 1998; Souza Junior, 2007).

The proposal of using all the hybrids in analyses and not only those in common in both crop seasons, as proposed by Piepho and Möhring (2006), did not prove to be effective in relation to the use of only the SH in common. The fact of considering homogeneous residual variance or not leads to a difference above all when the h^2 are of lower magnitude and, therefore, it is difficult to decide in the latter case from the results obtained in this study that the use of heterogeneous residual variance is more appropriate, since there is no way to prove which ranking is more trustworthy. In the literature, however, there are numerous reports that the use of heterogeneous variance is more advisable (Edwards and Jannink, 2006; So and Edwards, 2011; Orellana et al., 2014; Hu et al., 2014; Andrade et al., 2015; Silva et al., 2019).

From the above, it is clear that the possibility of selecting general hybrids for different growing seasons is very difficult. The possibilities of identifying hybrids that stand out under both conditions can be increased when using experiments with a smaller number of hybrids, with check varieties that are common to the experiments, with more replications and evaluations in the greatest number of locations possible, as has frequently been reported in the literature (Troyer, 1996; Cooper et al., 2014; Gaffney et al., 2015).

In the current period of "plant breeding 4.0", the need for evaluating hybrids considering various replications is not disregarded. In addition, the proposal considers the use of other information, such as climate, soil, geographic coordinates, phenological data, molecular markers, and the possibilities that exist in current analytical terms to identify the SH with best performance (Wallace et al., 2016; Ersoz et al., 2019; Ramstein et al., 2019). Obtaining accurate experiments is especially important in the molecular marker validation phase. Without accurate experiments, it is impossible to find trustworthy associations between the phenotype and the molecular marker.

More recently, the use of genomic selection models, including the effect of the genotype \times environment interaction, have frequently been reported as a tool to accelerate the selection process and improve the accuracy between the predicted value and observed value in breeding programs (Cuervas et al., 2016; Lado et al., 2016; Ferrão et al., 2018; Dias et al., 2018a; Montesinos-López et al., 2019; Monteverde et al., 2019; Ames and Bernardo, 2020). Nevertheless, it is clear that the interaction information will only effectively contribute to improve the predictive capability of the models if the interaction component used includes not only the genetic variation but also the future possibilities of environmental variation.

The analysis presented here, were carried out to better understand what happens with this data set. In addition, provide subsidies for the genomic prediction study, within will be carried out in a subsequent step, based on the genotyping of the parental lines of the hybrids evaluated in this study. Studies published recently in the literature, involving the prediction of hybrids under different environmental conditions, suggest that the inclusion of the component genotype \times environment interaction in genomic prediction models, may improve hybrids predictions if the environmental component is reliable (Krause et al., 2020; Oliveira et al., 2020). The question remains, given that the experiments are very unbalanced, if the component of the interaction to be used in the model will be able to improve its predictive capability, since, as found in this study, the component of the hybrid \times environment interaction is very expressive.

Therefore, it should be highlighted that in breeding programs of any species, the most important step is the final evaluation of the lines/hybrids. Recommendation of a cultivar with low accuracy of evaluation is a huge risk, not only economically, but also for the image of the company. The risk will only be reduced if, as already emphasized, the experiments are not only conducted in various environments, but are also as accurate as possible.

Acknowledgments

This work was financed in part by the Coordination for the Improvement of Higher Level Personnel (CAPES) – Financial Code 001, and by the Brazilian National Council for Scientific and Technological Development – CNPq / MCTIC, through grant of productivity to the authors. The authors would like to thank the professor Marcio Balestre (*in memoriam*), who contributed brilliantly to the development of this work.

Authors' Contributions

Conceptualization: Pereira, F.C.; Ramalho M.A.P.; Von Pinho, R.G. **Data acquisition:** Von Pinho, R.G. **Data analysis:** Pereira, F.C.; Resende, M.F.R. **Design of methodology:** Pereira, F.C.; Resende, M.F.R.; Ramalho, M.A.P.; Von Pinho, R.G. **Software development:** Pereira, F.C.; Resende, M.F.R. **Writing and editing:** Pereira, F.C.; Ramalho, M.A.P.; Von Pinho, R.G.

References

- Ames, N.C.; Bernardo, R. 2020. Genomewide predictions as a substitute for a portion of phenotyping in maize. *Crop Science* 60: 181-189. <https://doi.org/10.1002/csc2.20082>
- Andrade, V.; Gonçalves, F.M.A.; Nunes, J.A.R.; Botelho, C.E. 2015. Statistical modeling implications for coffee progenies selection. *Euphytica* 207: 177-189. <https://doi.org/10.1007/s10681-015-1561-6>

- Andrea, M.C.S.; Dallacort, R.; Barbieri, J.D.; Tieppo, R.C. 2019. Impacts of Future Climate Predictions on Second Season Maize in an Agrosystem on a Biome Transition Region in Mato Grosso State. *Revista Brasileira de Meteorologia* 34: 335-347.
- Andrea, M.C.S.; Boote, K.J.; Sentelhas, P.C.; Romanelli, T.L. 2018. Variability and limitations of maize production in Brazil: potential yield, water-limited yield and yield gaps. *Agricultural Systems* 165: 264-273. <https://doi.org/10.1016/j.agsy.2018.07.004>
- Burgueño, J.G.; Campos, G.; Weigel, K.; Crossa, J. 2012. Genomic prediction of breeding values when modeling genotype \times environment interaction using pedigree and dense molecular markers. *Crop Science* 52: 707-719. <https://doi.org/10.2135/cropsci2011.06.0299>
- Cooper, M.; Messina, C.D.; Podlich, D.; Totir, L.R.; Baumgarten, A.; Hausmann, N.J.; Graham, G. 2014. Predicting the future of plant breeding: complementing empirical evaluation with genetic prediction. *Crop and Pasture Science* 65: 311. <https://doi.org/10.1071/cp14007>
- Cullis, B.R.; Smith, A.B.; Coombes, N.E. 2006. On the design of early generation variety trials with correlated data. *Journal of Agricultural, Biological, and Environmental Statistics* 11: 381-393. <https://doi.org/10.1198/108571106x154443>
- Dias, K.O.D.G.; Gezan, S.A.; Guimarães, C.T. 2018a. Improving accuracies of genomic predictions for drought tolerance in maize by joint modeling of additive and dominance effects in multi-environment trials. *Heredity* 121: 24-37. <https://doi.org/10.1038/s41437-018-0053-6>
- Dias, K.O.D.G.; Gezan, S.A.; Guimarães, C.T.; Parentoni, S.N.; Guimarães, P.E.O.; Carneiro, N.P. 2018b. Estimating genotype \times environment interaction for and genetic correlations among drought tolerance traits in maize via factor analytic multiplicative mixed models. *Crop Science* 58: 72-83. <https://doi.org/10.2135/cropsci2016.07.0566>
- Eberhart, S.A.; Russell, W.A. 1966. Stability parameters for comparing varieties. *Crop Science* 6: 36-40.
- Euwijk, F.; Bustos-Korts, D.; Malosetti, M. 2016. What should students in plant breeding know about the statistical aspects of genotype \times environment interactions? *Crop Science* 56: 2119-2140. <https://doi.org/10.2135/cropsci2015.06.0375>
- Edwards, J.W.; Jannink, J.L. 2006. Bayesian modeling of heterogeneous error and genotype \times environment interaction variances. *Crop Science* 46: 820-833. <https://doi.org/10.2135/cropsci2005.0164>
- Empresa Brasileira de Pesquisa Agropecuária [Embrapa]. 2020. Embrapa production systems = Sistemas de produção Embrapa. Available at: <https://www.spo.cnptia.embrapa.br/> [Accessed July 6, 2020] (in Portuguese).
- Ersöz, E.S.; Martin, N.F.; Stapleton, A.E. 2019. On to the next chapter for crop breeding: convergence with data science. *Crop Science* 60: 639-655.
- Falconer, D.S.; Mackay, T.F.C. 1996. *Introduction to Quantitative Genetics*. 4ed. Longman, London, UK.
- Ferrão, L.F.V.; Marinho, C.D.; Patricio, R.; Munoz, P.R.; Resende Jr, M.F.R. 2018. Integration of dominance and marker \times environment interactions into maize genomic prediction models. *Crop Science* 59: 1-12. <https://doi.org/10.1101/362608>
- Gaffney, J.; Schussler, J.; Löffler, C.; Cai, W.; Paszkiewicz, S.; Messina, C.; Cooper, M. 2015. Industry-scale evaluation of maize hybrids selected for increased yield in drought-stress conditions of the US belt. *Crop Science* 55: 1608-1618. <https://doi.org/10.2135/cropsci2014.09.0654>
- Gauch, H.G.; Zobel, R.W. 1988. Predictive and postdictive success of statistical analyses of yield trials. *Theoretical and Applied Genetics* 76: 1-10. <https://doi.org/10.1007/BF00288824>
- Hallauer, A.R.; Miranda Filho, J.B. 1998. *Quantitative Genetics in Maize Breeding*. 2ed. Iowa State University Press, Ames, IA, USA.
- Hu, X.; Yan, S.; Li, S. 2014. The influence of error variance variation on analysis of genotype stability in multi-environment trials. *Field Crops Research* 156: 84-90.
- Kelly, A.M.; Smith, A.B.; Eccleston, J.A.; Cullis, B.R. 2007. The accuracy of varietal selection using factor analytic models for multi-environment plant breeding trials. *Crop Science* 47: 1063-1070. <https://doi.org/10.2135/cropsci2006.08.0540>
- Krause, M.D.; Olímpio Dias, K.G.; Santos, J.P.R.; Oliveira, A.A.; Guimarães, L.J.M.; Pastina, M.M.; Margarido, G.R.A.; Garcia, A.A.F. 2020. Boosting predictive ability of tropical maize hybrids via genotype by environment interaction under multivariate GBLUP models. *Crop Science* 60: 3049-3065. <https://doi.org/10.1002/csc2.20253>
- Lado, B.; Barrios, P.G.; Quincke, M.; Silva, P.; Gutiérrez, L. 2016. Modeling genotype \times environment interaction for genomic selection with unbalanced data from a wheat breeding program. *Crop Science* 56: 2165-2179. <https://doi.org/10.2135/cropsci2015.04.0207>
- Mengesha, W.; Menkir, A.; Meseka, S. 2019. Factor analysis to investigate genotype and genotype \times environment interaction effects on pro-vitamin A content and yield in maize synthetics. *Euphytica* 215: 180. <https://doi.org/10.1007/s10681-019-2505-3>
- Möhring, J.; Piepho, H.P. 2009. Comparison of weighting in two-stage analysis of plant breeding trials. *Crop Science* 49: 1977-1988.
- Monteverde, E.; Gutierrez, L.; Blanco, P.; Vida, F.P.; Rosas, J.E.; Bonnacerrere, V. 2019. Integrating molecular markers and environmental covariates to interpret genotype by environment interaction in rice (*Oryza sativa* L.) grown in subtropical areas. G3 9: 1519-1531. <https://doi.org/10.1534/g3.119.400064>
- Montesinos-López, O.A.; Montesinos-López, A.; Tuberosa, R.; Maccaferri, M.; Sciara, G.; Ammar, K.; Crossa, J. 2019. Multi-trait, multi-environment genomic prediction of durum wheat with genomic best linear unbiased predictor and deep learning methods. *Frontiers in Plant Science* 10: 1311. <https://doi.org/10.3389/fpls.2019.01311>
- Ndhlela, T.; Herselman, L.; Magorokosho, C.; Setimela, P.; Mutimaamba, C.; Labuschagne, M. 2014. Genotype \times environment interaction of maize grain yield using AMMI biplots. *Crop Science* 54: 1992-1999. <https://doi.org/10.2135/cropsci2013.07.0448>
- Nuvunga, J.; Oliveira, L.; Pamplona, A.; Silva, C.; Lima, R.R. 2015. Factor analysis using mixed models of multi-environment trials with different levels of unbalancing. *Genetics and Molecular Research* 14: 14262-14278. <http://dx.doi.org/10.4238/2015.November.13.10>

- Nzuve, F.; Githiri, S.; Mukunya, D.; Gethi, J. 2013. Analysis of genotype \times environment interaction for grain yield in maize hybrids. *Journal of Agricultural Science* 5: 2013. <https://doi.org/10.5539/jas.v5n11p75>
- Oliveira, A.A.; Resende Junior, M.F.R.; Ferrão, L.F.V.; Rampazo, R.R.; Guimarães, L.J.M.; Guimarães, C.T.; Pastina, M.M.; Margarido, G.R.A. 2020. Genomic prediction applied to multiple traits and environments in second season maize hybrids. *Heredity* 125: 60-72. <https://doi.org/10.1038/s41437-020-0321-0>
- Orellana, M.; Edwards, J.; Carriquiry, A. 2014. Heterogeneous Variances in Multi-Environment Yield Trials for Corn Hybrids. *Crop Science* 54: 1048-1056. <https://doi.org/10.2135/cropsci2013.09.0653>
- Patterson, H.D.; Thompson, R. 1971. Recovery of interblock information when block sizes are unequal. *Biometrika* 58: 545-554.
- Piepho, H.P. 1997. Analyzing genotype-environment data by mixed models with multiplicative terms. *Biometrics* 53: 761-766. <https://doi.org/10.2307/2533976>
- Piepho, H.; Mohring, J. 2006. Selection in cultivar – is it ignorable? *Crop Science* 46: 192-201.
- Piepho, H.P.; Möhring, J. 2007. Computing heritability and selection response from unbalanced plant breeding trials. *Genetics* 177: 1881-1888.
- Piepho H.P.; Mohring, J.; Schulz-Streeck, T.; Ogutu, J.O. 2012. A stagewise approach for the analysis of multi-environment trials. *Biometrical Journal* 54: 844-860.
- Ramstein, G.P.; Jensen, S.E.; Buckler, E.S. 2019. Breaking the curse of dimensionality to identify causal variants in Breeding 4. *Theoretical and Applied Genetics* 132: 559-567.
- Smith, A.B.; Ganesalingam, A.; Kuchel, H. 2015. Factor analytic mixed models for the provision of grower information from national crop variety testing programs. *Theoretical and Applied Genetics* 128: 55-72.
- Smith, A.; Cullis, B.R.; Thompson, R. 2002. Analyzing variety by environment data using multiplicative mixed models and adjustments for spatial field trend. *Biometrics* 57: 1138-1147. <https://doi.org/10.1111/j.0006-341X.2001.01138.x>
- Smith, A.B.; Cullis, B.R.; Gilmour, A. 2001. The analysis of crop variety evaluation data in Australia. *Australian and New Zealand Journal of Statistics* 43: 129-145.
- Schmidt, P.; Hartung, J.; Rath, J.; Piepho, H.P. 2019. Estimating broad-sense heritability with unbalanced data from agricultural cultivar trials. *Crop Science* 59: 525-536.
- Silva, C.P.; Oliveira, L.A.; Nuvunga, J.J.; Pamplona, A.K.A.; Balestre, M. 2019. Heterogeneity of variances in the bayesian AMMI. *Crop Science* 59: 2455-2472. <https://doi.org/10.2135/cropsci2018.10.0641>
- So, Y.-S.; Edwards, J. 2011. Predictive ability assessment of linear mixed models in multi-environment trials in corn. *Crop Science* 51: 542.
- Souza Junior, J.R.C.L. 2007. Improvement of allogamous species = Melhoria de espécies alógamas. p. 159-199. In: Nass, L.L., ed. *Plant genetic resources = Recursos genéticos vegetais*. Embrapa Recursos Genéticos e Biotecnologia, Brasília, DF, Brazil (in Portuguese).
- Steel, R.G.D.; Torrie, J.H.; Dickey, D.A. 1997. *Principles and Procedures of Statistics: A Biometrical Approach*. McGraw-Hill, New York: NY, USA.
- Tonk, F.A.; Ilker, E.; Tosun, M. 2011. Evaluation of genotype \times environment interactions in maize hybrids using GGE biplot analysis. *Crop Breeding Applied Biotechnology* 11: 1-9.
- Troyer, A.F. 1996. Breeding widely adapted, popular maize hybrids. *Euphytica* 92: 163-174. <https://doi.org/10.1007/bf00022842>
- Wallace, J.G.; Rodgers-Melnick, E.; Buckler, E.S. 2016. On the road to breeding 4.0: unraveling the good, the bad, and the boring of crop quantitative genomics. *Annual Review of Genetics* 52: 421-444. <https://doi.org/10.1146/annurev-genet-120116-024846>
- Welham, S.; Gogel, B.; Smith, A.; Thompson, R.; Cullis, B. 2010. A comparison of analysis methods for late-stage variety evaluation trials. *Australian and New Zealand Journal of Statistics* 52: 125-149.
- Wricke, G.; Weber, W.E. 1986. *Quantitative Genetics and Selection in Plant Breeding*. Walter de Gruyter, Berlin, Germany.
- Yan, W.; Hunt, L.; Sheng, Q.; Szlavnick, Z. 2000. Cultivar evaluation and mega-environment investigation based on the GGE biplot. *Crop Science* 40: 597-605.