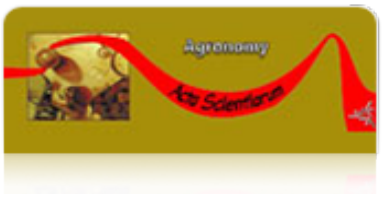

Acta Scientiarum. Agronomy
ISSN: 1679-9275
ISSN: 1807-8621
Editora da Universidade Estadual de Maringá - EDUEM

Yamamoto, Euriann Lopes Marques; Gonçalves, Manoel Carlos; Davide, Livia Maria Chamma; Rossoni, Diogo Francisco; Santos, Adriano dos

Spatial variability in evaluation experiments of corn genotypes in the state of Mato Grosso do Sul, Brazil

Acta Scientiarum. Agronomy, vol. 44, e55972, 2022
Editora da Universidade Estadual de Maringá - EDUEM

DOI: https://doi.org/10.4025/actasciagron.v44i1.55972

Acta Scientiarum

http://www.periodicos.uem.br/ojs/
ISSN on-line: 1807-8621
Doi: 10.4025/actasciagron.v44i1.55972

Spatial variability in evaluation experiments of corn genotypes in the state of Mato Grosso do Sul, Brazil

Euriann Lopes Marques Yamamoto[1*], Manoel Carlos Gonçalves[1], Livia Maria Chamma Davide[1], Diogo Francisco Rossoni[2] and Adriano dos Santos[3]

[1]Universidade Federal da Grande Dourados, Rua João Rosa Góes, 1761, 79825-070, Dourados, Mato Grosso do Sul, Brazil. [2]Universidade Estadual de Maringá, Maringá, Paraná, Brazil. [3]Empresa Brasileira de Pesquisa Agropecuária, Embrapa Agroenergia, Brasília, Distrito Federal, Brazil. *Author for correspondence. E-mail: euriann@outlook.com

ABSTRACT. Analysis of variance (ANOVA) is the most used procedure for comparing means between different groups. However, in some cases, disregarding the assumptions of ANOVA can lead to spatial dependence. In such cases, to ensure greater experimental precision, it is necessary to consider the study of spatial dependence. This study was carried out to compare the estimates of experimental precision of the traditional analysis of variance with those of the analysis of variance using an autoregressive (ANOVA-AR) model in corn experiments under different N conditions when evaluating grain yield. Data were obtained from 14 experiments using lattice designs conducted in 2012, 2014, and 2015 in the following counties in the Brazilian state of Mato Grosso do Sul: Caarapó, Dourados, Glória de Dourados, and Laguna Carapã. Of the 14 experiments, 7 were performed with N fertilization (ideal) and 7 experiments were performed under stressful conditions (zero or low). Both analyses were compared by considering estimates of reduction of the error mean square, coefficient of determination, F-value, and selective accuracy as well as the difference in the order of 25% of the genotypes of each experiment (from 13 to 56 genotypes, considering the size of the experiment). Differences in the error mean square and genotype mean square were slightly more evident in 1, 2, 3, 4, 5, 6, and 11 experiments but the use of ANOVA-AR did not promote major changes. The analysis of variance with an autoregressive model provided parameter values of experimental precision similar to those expressed by traditional analysis of variance. There was no difference in terms of correlated errors in experiments under different N conditions.

Keywords: *Zea mays* L.; nitrogen; autoregressive models; analysis of variance.

Received on September 24, 2020.
Accepted on April 17, 2021.

Introduction

The assessment of genotypes in corn experiments in environments with different nitrogen (N) levels has constantly been discussed, because this fertilizer is essential for plants. Without N plants cannot make DNA, RNA, proteins, enzymes, and many other components (Galembeck, Galembeck, & Santos, 2020). Nitrogen deficiency can reduce plant production, especially in maize crop (Morris et al., 2018).

However, the high commercial cost of N fertilizers, excessive application, nitrate leaching losses and contamination of groundwater has led researchers to apply strategies to minimize environmental impacts and reduce the production cost of the crop (Su, Ahmad, Ahmad, & Han 2020). The inoculation of maize with diazotrophic bacteria (Alves, Sobral, & Reis, 2020) has resulted in the development of genotypes with better performance under low N conditions (Ertiro et al., 2020).

Experiments that deposit N in the soil should be studied because this action strategy can interfere with obtaining better experimental precision. In genetic breeding programs, the development and recommendation of productive genotypes for stressful environments are some of the main objectives. To achieve this, it is necessary that the plan of experiments and procedures of statistical analyses used, such as the use of analysis of variance (ANOVA) and the comparison of means, should be appropriate.

Around 1925, Fisher proposed ANOVA, which is the decomposition of the total variance observed in the experiment into known sources of variation (Wahid, Latiff, & Ahmad, 2017). To ensure that the results of this analysis are considered precise, the variability of the error is reduced to the maximum extent possible to ensure safety in the results.

To correctly use ANOVA, the following assumptions must be met: additivity of the effects in the mathematical model, homogeneity of error variances, normality and independence of residuals. In case of violation of any of these assumptions, appropriate alternative statistical procedures must be used, namely: data transformation, nonparametric statistics or procedures that consider the known distribution of the response variable. Whenever possible, the independence of errors must be established by randomization or local control. However, this independence can be violated due to the existence of correlation between neighboring plots, thereby characterizing a situation of spatial autocorrelation or spatial dependence (Rossoni & Lima, 2019).

The presence of adjacent plots under spatial autocorrelation may influence the accurate selection and genetic gains of genotypes, thereby promoting the success or failure of a genetic breeding program (Bernadeli et al., 2021). Therefore, spatial statistical tools should be used to select genotypes for their real performance as verified by Duarte and Vencovsky (2005) and Bernadeli et al. (2021) in soybean genotypes, and by Silva et al. (2016) in their assessment of the efficiency of spatial methods in evaluating the yield of common bean families.

Statistical techniques based on spatial modeling are useful in experiments where the spatial dependence between errors is detected because the efficiency of treatment contrast estimators does not exclusively depend on the variation of the residual but on the positioning of experimental plots throughout geographic coordinates. Rossoni and Lima (2020) proposed ANOVA with the spatial correlation component ρ and found that the spatial factor provided higher experimental precision to the simulated data set.

ANOVA-AR was described by Long (1996) and its basic premise involves the transformation of autocorrelated observations into uncorrelated observations, i.e., after detecting the spatial correlation in the variable of interest, ANOVA-AR removes this correlation and makes the data independent in relation to space, thereby facilitating appropriate statistical inferences. The ANOVA that considers this spatial dependence can be performed using autoregressive models. When location information is provided in these models, it enables estimating whether there is spatial dependence between plots.

The present study aimed to compare the estimates of high experimental precision (error mean square, F-value, coefficient of determination, and precision) of traditional ANOVA (independent errors) with those of ANOVA using an autoregressive model (ANOVA-AR) (correlated errors) in the analysis of 14 corn experiments under different levels of N conditions. It also aimed to verify whether N accumulation is associated with correlated errors between experiments under N fertilization conditions (ideal) and experiments under stressful conditions (low) by evaluating grain yield.

Material and methods

The data studied were obtained from experiments conducted during the second harvest period in 2012, 2014, and 2015 in four different counties in the state of Mato Grosso do Sul, Brazil: Caarapó, Dourados, Glória de Dourados, and Laguna Carapã. Of the 14 experiments conducted under contrasting N conditions, 7 were under N fertilization conditions (ideal) while the remaining 7 were under stressful conditions (low). In 2012, 4 experiments were conducted in a 12 × 12 simple lattice design in the cities of Caarapó and Dourados, where each city had one environment under N fertilization conditions (ideal) and one environment under stressful conditions (low). In 2014, 6 experiments were conducted in a 7 × 7 simple lattice design in the counties of Dourados, Glória de Dourados, and Laguna Carapã, which also had one ideal environment and one low N environment. Finally, in 2015, four experiments were conducted in a 15 × 15 triple lattice design in the cities of Caarapó and Dourados, also with one ideal environment and one low N environment in each county (Table 1).

The plot size was the same for all experiments with a line of 5 m and a space variation according to the experiment (Table 1), aiming for an ideal population of 55.000 plants per hectare. All experiments used urea as source of N. The amount of fertilizer applied in each experiment was based on the study's objective for each experiment (Table 2).

The decision to apply N or not depended on the researcher when implementing the experiment. Fritsche-Neto and Borém (2011) recommended minimum fertilization so that even the plant showing the stress condition, cannot overshadow the genetic variability in the genotypes.

The 49 genotypes used in Experiments 1 to 6 consisted of 42 progenies of half-siblings and 7 controls arranged in a 7 × 7 simple lattice design. In Experiments 7 to 10, 121 genotypes were used, including 110 top-cross hybrids, 5 base populations, and 6 controls, arranged in an 11 × 11 simple lattice design. In Experiments 11 to 14, 225 genotypes were evaluated, including 220 progenies of half-siblings and 5 controls arranged in a 15 × 15 lattice design with three replications. The variable used was grain yield in kg ha^{-1}, subjected to moisture correction of 13%.

Table 1. Information on the 14 corn experiments under N fertilization conditions (ideal) and stressful conditions (low).

Exp.	Lattice	Year	Local	Condition	NP	E	$\hat{r}_{\hat{g}g}$
1	7 × 7	2014	Ddos	Ideal	49	0.9	0.77
2	7 × 7	2014	Ddos	Low	49	0.9	0.83
3	7 × 7	2014	GDdos	Ideal	49	0.9	0.31
4	7 × 7	2014	GDdos	Low	49	0.9	0.43
5	7 × 7	2014	LCrpã	Ideal	49	0.9	0.73
6	7 × 7	2014	LCrpã	Low	49	0.9	0.30
7	12 × 12	2012	Crpó	Ideal	144	1.0	0.78
8	12 × 12	2012	Crpó	Low	144	1.0	0.76
9	12 × 12	2012	Ddos	Ideal	144	0.9	0.70
10	12 × 12	2012	Ddos	Low	144	0.9	0.60
11	15 × 15	2015	Crpó	Ideal	225	0.9	0.65
12	15 × 15	2015	Crpó	Low	225	0.9	0.72
13	15 × 15	2015	Ddos	Ideal	225	0.9	0.71
14	15 × 15	2015	Ddos	Low	225	0.9	0.71

NP: number of progenies. E: spacing. $\hat{r}_{\hat{g}g}$: Accuracy. Ddos: Dourados. GDdos: Glória de Dourados. LCrpã: Laguna Carapã.

Table 2. Information about basal or topdressing fertilizer and amount of urea fertilizer in the 14 maize experiments.

	Fertilizer		Amount of Urea Fertilizer	
Experiment	Basal	Topdressing	Low	Ideal
1 to 6	30 kg ha^{-1}	90 kg ha^{-1}	30 kg ha^{-1}	120 kg ha^{-1}
7 to 10	20 kg ha^{-1}	100 kg ha^{-1}	20 kg ha^{-1}	120 kg ha^{-1}
11 to 14	0 kg ha^{-1}	120 kg ha^{-1}	0 kg ha^{-1}	120 kg ha^{-1}

To work with spatial autocorrelation, the "spatial autoregressive" (SAR) model was used. This model was proposed by Long (1996) and its main objective is to transform autocorrelated observation in uncorrelated observations.

First, we defined the proximity pattern of the neighborhood region. For this work, we adopted the proximity pattern of the first order, that is, only adjacent plots to the reference plot were considered. The basic requirements for obtaining the adopted proximity pattern considered the radius in which the highest correlation was obtained [ρ(h)] and the lowest Akaike Information Criterion (AIC) value. After the definition of these two points, the proximity pattern was adopted in relation to the radius (Scolforo et al., 2016). More information about proximity pattern can be found at Gumpetz et al. (1997).

The autoregressive model SAR was described by Griffith (1988): $Y = \rho WY + X\beta + \varepsilon$, where *Y*: *nx1* vector of observed values; *p:* spatial autoregressive parameter; *W*: *nx1* matrix with neighborhood spatial weight assignments; *X*: *nxp* matrix of incidence of fixed effects; β: *px1* parameter vector; ε: *nx1* vector of errors assigned to each observation.

The matrix W was obtained by multiplying the matrices D and C. The matrix C is binary and has *n* x *n* dimensions. This describes the adjacent neighborhood of distance between the experimental plots, and its size varied according to the radius adopted for each existing experiment. The matrix D is a diagonal matrix with the element $1/k_i$, where k_i is the sum of values of the line *i* of the matrix C.

The spatial parameter ρ varies between −1 and +1. Positive values indicate positive spatial autocorrelation, wherein high (low) values tend to group close to high (low) values, indicating an effect of contagion or overflow. Negative values indicate negative spatial autocorrelation, wherein high (low) values tend to be located at low (high) values, presenting as a situation of dissimilarity between the variable value and plot location. Therefore, the higher the value of the ρ parameter in the module, the greater the autocorrelation, which can either be positive or negative (Almeida, 2012).

The maximum likelihood (ML) method was used to estimate the ρ parameter of the SAR model. A solution for ML estimation of spatial autocorrelation models was initially proposed by Ord (1975) cited by Rossoni and Lima (2019). Therefore, after estimating the ρ parameter, adjustments to the observed data based on the following equation were necessary:

$$Y_{adj} = Y - (\hat{\rho}WY - \hat{\rho}\beta_0),$$

where: *Y*:*nx1* vector of observed values; ρ^: estimation of the spatial autoregressive parameter; W: *nxn* matrix with neighborhood spatial weight assignments; β_0: mean of observed values; and Y_{adj}: *nx1* vector of adjusted values.

The adjusted value of response variable (Y_{adj}) remove spatial variability by considering observations in the same proximity. When $p\hat{}$ is equal to zero, no spatial structure is detected and ANOVA-AR results are identical to those of traditional ANOVA. Therefore, Y_{adj} is the grain yield adjusted for spatial correlation, allowing appropriate statistical inferences.

After obtaining Y_{adj}, ANOVA-AR was generated (Table 3). The corrected total sum of squares (SQTcorrected) was obtained by the difference between the total sum of squares (SQT) of ANOVA of the unadjusted data and the total sum of squares of the adjusted data (SQTadj), as described in the equation: $SQT_{corrected} = SQT - SQT_{adj}$.

Table 3. Analysis of variance with an autoregressive model (ANOVA-AR).

FV	GL	SQ	QM	F
Block	n - 1			
Parameter	k - 1	SQP_{adj}	QMP_{adj}	QMP_{adj}/QME_{adj}
Error	$(n-1)(k-1)$	SQE_{adj}	QME_{adj}	
Total corrected	nk - 1	$SQT_{corrected}$		

To compare the selective efficiency between the traditional ANOVA and the proposed ANOVA-AR, we evaluated the following statistics provided by the two analysis: the coefficient of determination (R^2), the accuracy value ($\hat{r}_{\hat{g}g}$), the F value, and the value of the error mean square, which by its reduction promotes greater precision by reducing the total variability of the experiment. Selective accuracy, in genetic evaluation, correlates the real genotype value with the predicted value from the experiment. It is a correlation, so its range from 0 to 1 and accuracy values should be more appropriate when values are closer to unity or 100%. According to Resende and Duarte (2007), it is estimated by the expression: $\hat{r}_{\hat{g}g} = \left(1 - {}^{1}/_{F}\right)^{0,5}$.

Further, Tukey's comparison test of means was applied using traditional ANOVA and ANOVA-AR for assessing the significance at 5% level of probability, to determine the genotypes that presented the highest yields under a selection intensity of 25%, considering both analyses. The R software was applied in the analyses using geoR packages (Ribeiro & Diggle, 2001) and spdep (Bivand et al., 2018) (R Core Team Development, 2020).

Results and discussion

After obtaining the matrix W, the value of the autocorrelation coefficient ρ for all experiments was calculated. The ρ parameter estimates and radius used for each experiment are shown in Table 4.

Table 4. Estimates of the ρ parameter, likelihood ratio test (LRT), and the adopted radius for each of the 14 corn experiments under nitrogen fertilization conditions (ideal) and stressful conditions (low).

Local	Condition	Exp.	ρ	p-value	Radius	AIC[(1)]
Ddos	Ideal	1	−0.50	0.15	7	1616.95
Ddos	Low	2	−1.25	0.01	7	1545.12
GDdos	Ideal	3	−0.78	0.08	9	1655.00
GDdos	Low	4	−0.96	0.02	7	1638.70
LCrpã	Ideal	5	−0.49	0.16	7	1557.95
LCrpã	Low	6	−0.45	0.06	5	1602.18
Crpó	Ideal	7	−0.21	0.54	15	4047.83
Crpó	Low	8	−0.07	0.77	11	3992.87
Ddos	Ideal	9	−0.25	0.28	9	4117.27
Ddos	Low	10	−0.21	0.62	19	4073.99
Crpó	Ideal	11	0.37	0.00	3	11567.93
Crpó	Low	12	−0.04	0.85	53	11434.84
Ddos	Ideal	13	−0.23	0.54	37	11371.16
Ddos	Low	14	−0.11	0.84	159	11430.78

[(1)]AIC: Akaike Information Criterion. Exp.: experiment. Ddos: Dourados. GDdos: Glória de Dourados. LCrpã: Laguna Carapã.

Experiments 2, 3, 4, and 6 presented significant ρ parameter values ($p \leqslant 0.1$), indicating spatial autocorrelation. When the ρ parameter tends to unity, the research should be interpreted with care. It is necessary to interpret this value together with other estimates like the effect of block and error mean square (QME) because high values of ρ parameters can create the impression that the correlation is bigger when QME and block effect values are really small (Piepho, Mohring, Pflugfelder, Hermann, & Williams, 2015). Experiment 11 indicated a significant ρ parameter value ($p \leqslant 0.1$) and a value greater than zero (positive spatial autocorrelation), which suggests that plots with higher values for yield tended to cluster together. The farther from zero the value of the ρ parameter was, the greater was the spatial variability detected by the proximity pattern (Rossoni & Lima, 2019).

The observation of autocorrelated or dependent plots showed that the experimental error was highly similar when the plots were closer to each other (Andrade et al., 2020). Considering that the error between the plots is dependent and not independent, as per classical statistics (Duarte & Vencovsky, 2005), the breeder can use spatial statistical analysis to accurately select the genotypes.

Comparing the accuracy values obtained by the experiments (Table 1) with the spatial autocorrelation coefficient (Table 4), we observed that Experiment 11 presented moderate accuracy ($\hat{r}_{\hat{g}g}$=0.65), and spatial analysis had greater efficiency when errors showed spatial dependence. Campos et al. (2016) assessed the efficiency of spatial analysis using geostatistics to classify common bean families and concluded that in experiments with moderate experimental precision, spatial analysis presents higher efficiency in the classification of common bean families.

Other authors in literature have used spatial statistical analysis involving first-order autoregressive models which are separable in two dimensions (Resende & Sturion, 2003; Maia, Siqueira, Carvalho, Peternelli, & Latado, 2013), geostatistical models (Campos et al., 2016; Silva et al., 2016), Papadakis methods, and moving averages (Candido, Perecin, Landell, & Pavan, 2009). However, ANOVA-AR is yet to be used. This statistical procedure is an informative and easy-to-use tool, and studies on its use in agricultural research with large crops are scarce.

Local control and randomization are sometimes inadequate experimental procedures for avoiding spatial dependence. Accordingly, spatial statistical analysis, when associated with ANOVA, can promote increased efficiency of these factors as well as higher experimental accuracy (Andrade et al., 2020).

When analyzing the coefficient of the spatial autoregressive parameter (ρ) individually, we observed that most experiments did not have spatial dependence (Table 4). However, the conclusion that there may or may not be an autocorrelation between the plots should be made based on an aggregate of other statistical information. Therefore, we chose to analyze the error mean square (QME) to verify whether the adoption of ANOVA-AR was sufficient at promoting lower mean squared error than the traditional ANOVA.

According to Steel and Torrie (1980), the precision of an experiment is closely related to the amplitude of the experimental error. Therefore, regardless of the error in an experimental unit being small, it will be reflected in the value of the error mean square (QME) of the ANOVA of the experiment.

On comparing traditional ANOVA with ANOVA-AR in Experiment 1, a 6.64% reduction in experimental error was observed, which is composed of uncontrolled variations in the experiment and expressed by reduction of the error mean square (QME) (Table 5). The mean square of the replication showed a 38.02% improvement in ANOVA-AR. The decrease in variability may have been caused by 11.75% inflation of the mean square of blocks within repetitions. This situation should be considered because the effect that was majorly attributed to replications was in fact attributed to the effect of the blocks within replications, thereby minimizing the effect of the error in smaller portions of blocks.

The non-exclusion of the variation derived from the blocks was consistent with the proposal by Resende and Sturion (2003), who used spatial analysis with an autoregressive structure in yerba mate and found that the autoregressive model did not remove the variation between the blocks. This result showed that the design used was essential to control the variation of the block effect.

ANOVA-AR is expected to be more efficient at detecting differences between treatment means, providing an F-value higher than that generated by traditional ANOVA. Therefore, by analyzing Experiment 1, it was verified that the F-value of ANOVA-AR was slightly higher than that of traditional ANOVA.

According to Banzatto and Kronka (2013), the F-test of the ANOVA was only valid when the errors were considered homogeneous, independent, and normally distributed. Assuming the presence of correlated errors, the traditional ANOVA would violate the assumption of independence between errors and present an incorrect F-value.

Table 5. Traditional analysis of variance and analysis of variance using autoregressive models (ANOVA-AR), with the parameters of coefficient of determination (R^2) and accuracy ($\hat{r}_{\hat{g}g}$) for Experiments 1 to 6 under nitrogen fertilization conditions (ideal) and stressful conditions (low).

ANOVA (Exp. 1)					ANOVA-AR (Exp. 1)			
FV	QM	F	R^2	$\hat{r}_{\hat{g}g}$	QM	F	R^2	$\hat{r}_{\hat{g}g}$
Rep	216712	0.31	0.78	0.77	134324	0.21	0.80	0.79
Gen	1718319	2.49**			1733064	2.69**		
Rep*	517849	0.75			586806	0.91		
Erro	691086				645212			
ANOVA (Exp. 2)					ANOVA-AR (Exp. 2)			
FV	QM	F	R^2	$\hat{r}_{\hat{g}g}$	QM	F	R^2	$\hat{r}_{\hat{g}g}$
Rep	1050229	2.83	0.83	0.83	5230776	16.87*	0.85	0.84
Gen	1194629	3.22**			1078434	3.48**		
Rep*	444319	1.20			386945	1.25		
Erro	371240				310020			
ANOVA (Exp. 3)					ANOVA-AR (Exp. 3)			
FV	QM	F	R^2	$\hat{r}_{\hat{g}g}$	QM	F	R^2	$\hat{r}_{\hat{g}g}$
Rep	2599640	2.52	0.64	0.31	6368422	6.69*	0.66	0.35
Gen	1142018	1.11			1087636	1.14		
Rep*	656347	0.64			685274	0.72		
Erro	1031464				951370			
ANOVA (Exp. 4)					ANOVA-AR (Exp. 4)			
FV	QM	F	R^2	$\hat{r}_{\hat{g}g}$	QM	F	R^2	$\hat{r}_{\hat{g}g}$
Rep	1591374	1.73	0.66	0.43	3870803	4.80*	0.68	0.46
Gen	1125771	1.23			1021103	1.27		
Rep*	663004	0.72			648658	0.81		
Erro	918791				805594			
ANOVA (Exp. 5)					ANOVA-AR (Exp. 5)			
FV	QM	F	R^2	$\hat{r}_{\hat{g}g}$	QM	F	R^2	$\hat{r}_{\hat{g}g}$
Rep	344245	0.91	0.76	0.73	441667	1.25	0.77	0.75
Gen	804229	2.11*			798551	2.26**		
Rep*	361079	0.95			386900	1.09		
Erro	379983				353401			
ANOVA (Exp. 6)					ANOVA-AR (Exp. 6)			
FV	QM	F	R^2	$\hat{r}_{\hat{g}g}$	QM	F	R^2	$\hat{r}_{\hat{g}g}$
Rep	1864695	2.96	0.65	0	4078257	7.35*	0.68	0
Gen	579563	0.92			546758	0.99		
Rep*	1014879	1.61			1054670	1.90		
Erro	630696				554927			

FV: source of variation. QM: mean square. Rep: replication. Gen: genotype. Rep*: blocks within replications.

Regarding the coefficient of determination (R^2), which indicates the ability of the linear model to adjust the data, the value in ANOVA-AR was slightly higher (2.5%) compared to the traditional ANOVA, and there was also a slight improvement in the accuracy value.

In Experiment 2 (Table 5), reduction of the QME in ANOVA-AR was by 16.49%, favoring the reduction of experimental error. When using ANOVA-AR, the variability of the mean square of blocks within replications was reduced (12.91%), and the mean square of the replication showed high inflation of its value (20.08%). Further, there was decreased variability of the genotype mean square (QMG) when using ANOVA-AR (9.73%). This fact should be highlighted because the genetic variability of progenies was slightly reduced when considering ANOVA-AR. This situation should be considered in plant breeding programs because the genetic variability of progenies was not as high as that indicated by the QMG of the traditional ANOVA, although the F-test was unaffected.

This information is important for both germplasm banks and breeders. In a genetic breeding program, it is crucial to understand the genetic variability of corn genotypes. If the level of genetic variability (QMG) is inflated (by a factor other than the genetic factor) when conducting experiments, several genotypes may be mistakenly selected.

In Experiments 3, 4, and 6 (Table 5), it was observed that although ANOVA-AR did not favor the presence of significance among the genotypes of those experiments, the QME in ANOVA-AR was slightly lower for Experiment 3 (7.76%), Experiment 4 (12.32%), and Experiment 6 (12.01%). The QMG in ANOVA-AR was also reduced for Experiments 3, 4, and 6 by 4.76, 9.29, and 5.66%, respectively.

Scolforo et al. (2016) evaluated the effect of the spatial autoregressive approach in the perennial candeia tree species (*Eremanthus erythropappus*). Using the traditional ANOVA, no significant differences were detected between the different fertilization treatments. However, when considering the ANOVA-AR approach, the treatments were significant. The F-value for treatments increased from 1.84 traditional ANOVA to 2.07 in ANOVA-AR, providing a higher accuracy of F-test estimates.

In Experiment 5 (Table 5), the QME in ANOVA-AR showed a 6.99% reduction in experimental error. The coefficient of determination in ANOVA-AR was similar (0.77) compared to the traditional ANOVA (0.76); regarding the QMG, the significance for genotypes decreased from 5 to 1%. In this case, the increased power of significance of the test (from 5 to 1%) may increase the probability of a Type II error (which would be to accept that the genotypes are similar, when in fact they are different) (Mcintosh, 2015). This situation occurred only in Experiment 5 and should be evaluated with caution because it may lead to incorrect inferences.

The finding of spatial autocorrelation can negatively influence the comparison of genotypes (Duarte, 2005). According to Es and Es (1993) and Legendre et al. (2002), in experiments with autocorrelation, statistical tests related to contrasts between treatments whose plots were separated by small distances have a higher probability of a Type II error (accept that genotypes are similar, when in fact they are different); whereas plots that were separated by large distances have a higher probability of a Type I error (accept that genotypes are different when they are similar). Therefore, the presence of spatial autocorrelation should be considered because it directly influences the selection of promising genotypes.

In Experiments 7, 8, 9, and 10 (Table 6), the QME reduction was lower (0.95, 0.53, 2.61, and 0.39%, respectively). When using ANOVA-AR, the F-value for genotypes was slightly higher than that observed using traditional ANOVA. The F-value was strictly related to selective accuracy, which is a statistical parameter used in cultivar evaluation experiments. The higher the F-value for genotypes, the greater the selective accuracy of the experiment (Resende & Duarte, 2007).

Table 6. Traditional analysis of variance and analysis of variance using autoregressive models (ANOVA-AR) with the parameters of coefficient of determination (R^2) and accuracy ($\hat{r}_{\hat{g}g}$) for Experiments 7–10 under nitrogen fertilization conditions (ideal) and stressful conditions (low).

ANOVA (Exp. 7)					ANOVA-AR (Exp. 7)			
FV	QM	F	R^2	$\hat{r}_{\hat{g}g}$	QM	F	R^2	$\hat{r}_{\hat{g}g}$
Rep	45538594	56.48**	0.82	0.78	64428616	80.68**	0.82	0.78
Gen	2087760	2.59**			2073818	2.60**		
Rep*	3010922	3.73**			3062640	3.83**		
Erro	806295				798621			
ANOVA (Exp. 8)					ANOVA-AR (Exp. 8)			
FV	QM	F	R^2	$\hat{r}_{\hat{g}g}$	QM	F	R^2	$\hat{r}_{\hat{g}g}$
Rep	8021795	12.54**	0.79	0.76	9005439	14.15**	0.79	0.76
Gen	1505982	2.35**			1508295	2.37**		
Rep*	2718174	4.25**			2738969	4.30**		
Erro	639703				636376			
ANOVA (Exp. 9)					ANOVA-AR (Exp. 9)			
FV	QM	F	R^2	$\hat{r}_{\hat{g}g}$	QM	F	R^2	$\hat{r}_{\hat{g}g}$
Rep	8168374	7.48**	0.76	0.70	12648744	11.89**	0.76	0.70
Gen	2138145	1.96**			2115299	1.98**		
Rep*	3718101	3.40**			3888137	3.65**		
Erro	1092528				1064023			
ANOVA (Exp. 10)					ANOVA-AR (Exp. 10)			
FV	QM	F	R^2	$\hat{r}_{\hat{g}g}$	QM	F	R^2	$\hat{r}_{\hat{g}g}$
Rep	12535574	14.03**	0.72	0.60	18691927	21.00**	0.73	0.6
Gen	1394921	1.56*			1386270	1.56*		
Rep*	2603535	2.91**			2685574	3.02**		
Erro	893314				889819			

FV: source of variation. QM: mean square. Rep: replication. Gen: genotype. Rep*: blocks within replications.

Regarding the coefficient of determination and accuracy values for Experiments 7, 8, 9, and 10 there was no variation observed between analysis.

The QME in ANOVA-AR for Experiment 11 (Table 7) reduced the experimental error by 11.99%. In other experiments, the reduction was small or non-existent: 0.05% for Experiment 12; 0.38% for Experiment 13, and 0% for Experiment 14. Regarding the QMG, the level of genetic variability reduced by 9.02% for Experiment 11; 0.13% for Experiment 12, and 0.23% for Experiment 14. For the coefficient of determination,

there was no difference between the analyses for Experiments 11, 13, and 14 (Table 7). Only Experiment 12 showed a slight improvement of the coefficient of determination in ANOVA-AR (R^2 = 0.63) compared with that in traditional ANOVA (R^2 = 0.62), with no differences in the accuracy values between both analyses.

Table 7. Traditional analysis of variance and analysis of variance using autoregressive models (ANOVA-AR) with the parameters of coefficient of determination (R^2) and accuracy ($\hat{r}_{\hat{g}g}$) for Experiments 11-14 under nitrogen fertilization conditions (ideal) and stressful conditions (low).

ANOVA (Exp. 11)					ANOVA-AR (Exp. 11)			
FV	QM	F	R^2	$\hat{r}_{\hat{g}g}$	QM	F	R^2	$\hat{r}_{\hat{g}g}$
Rep	9584311	6.95**	0.52	0.65	3019663	2.49	0.52	0.66
Gen	2359352	1.71**			2146574	1.77**		
Rep*	1538158	1.12			942869	0.78		
Erro	1377644				1212450			
ANOVA (Exp. 12)					ANOVA-AR (Exp. 12)			
FV	QM	F	R^2	$\hat{r}_{\hat{g}g}$	QM	F	R^2	$\hat{r}_{\hat{g}g}$
Rep	63210039	63.47**	0.62	0.72	67154948	67.46**	0.63	0.72
Gen	2067028	2.07**			2064431	2.07**		
Rep*	1842967	1.85**			1830482	1.84**		
Erro	995959				995479			
ANOVA (Exp. 13)					ANOVA-AR (Exp. 13)			
FV	QM	F	R^2	$\hat{r}_{\hat{g}g}$	QM	F	R^2	$\hat{r}_{\hat{g}g}$
Rep	753043	0.83	0.56	0.71	1350675	1.49	0.56	0.71
Gen	1826001	2.00**			1826554	2.02**		
Rep*	1253417	1.38			1275792	1.41		
Erro	909343				905870			
ANOVA (Exp. 14)					ANOVA-AR (Exp. 14)			
FV	QM	F	R^2	$\hat{r}_{\hat{g}g}$	QM	F	R^2	$\hat{r}_{\hat{g}g}$
Rep	43543287	44.00**	0.60	0.70	47208293	47.71**	0.60	0.7
Gen	1966316	1.98**			1961632	1.98**		
Rep*	1613745	1.63**			1623675	1.64**		
Erro	989601				989510			

FV: source of variation. QM: mean square. Rep: replication. Gen: genotype. Rep*: blocks within replications.

Typically, a reduction in the variability of the factors of interest was not always observed (mean square of replication, mean square of blocks within replications, and genotype mean square). In the 14 experiments analyzed, there was a reduction in the experimental error via the QME. However, in 7 experiments, the difference in the error mean square and in the genotype mean square was slightly more evident (Experiments 1, 2, 3, 4, 5, 6, and 11) and in the other experiments, the use of ANOVA-AR against traditional ANOVA did not promote major changes (Experiments 7, 8, 9, 10, 12, 13, and 14).

Rossoni and Lima (2019) used ANOVA-AR in simulated experiments with spatial dependence and verified a reduction of the variability for the QME, the mean square of blocks (QMB), and mean square of treatment (QMT), concluding that spatial statistical analysis decreased the overall variability of the experiments.

In all experiments with low N conditions (2, 4, 6, 8, 10, 12, and 14), there was an increase in the mean square of the replication in ANOVA-AR compared with that in traditional ANOVA. ANOVA-AR allowed the detection of an increase in the internal spatial variability of the replication. This increase can be explained because traditional ANOVA cannot detect variability in the mean square of the replication when N was not applied to the soil (or if the amount applied was considerably smaller than required). This was more evident in ANOVA-AR because the value of the mean square of the replication increased.

The significance of the mean square of the replication from the traditional ANOVA to ANOVA-AR was verified in Experiments 2, 3, 4, and 6. These experiments were arranged in a 7 × 7 simple lattice design. Therefore, the significance for this factor proved the need to adopt a lattice design for detecting the spatial variability of replications.

In most experiments analyzed, ANOVA-AR demonstrated a slight improvement in the F-value and was unable to confer higher experimental quality for genotype selection (Resende & Duarte, 2007). Scolforo et al. (2016) found that ANOVA-AR was efficient at identifying significant differences between genotypes of the candeia tree species compared with the traditional ANOVA, which did not identify significant differences. However, this situation was not observed in the current study.

To evaluate the effect of using autoregressive models in ANOVA, we analyzed its effect on the order selection of 25% of the genotypes of each experiment.

For Experiments 1, 2, 3, 4, 5, and 6 with a selection intensity of 25% (13 out of 49 genotypes) of the most productive genotypes, approximately 31, 46, 46, 69, 54, and 54% of genotypes, respectively, were affected by the difference in the order selection on comparing traditional ANOVA with ANOVA-AR (data not shown). Experiment 6 is noteworthy because genotype 26 would be excluded from selection if just the traditional ANOVA was considered. Using the spatial approach, Duarte and Vencovsky (2005) analyzed soybean genotypes and found that the spatial effect favored the selection of genotypes that would not be selected by the traditional method.

It is worth noting that solely a difference in the order selection between the two analyses is insufficient to ensure an advantage of one analysis over the other. However, in Experiment 6, in addition to the different order selection, there was the exclusion of a genotype in the selection made by traditional ANOVA. ANOVA-AR may favor the selection of genotypes that could be excluded in traditional ANOVA.

For experiments 7, 8, 9, and 10 with a selection intensity of 25% (36 out of 144 genotypes) of the most productive genotypes, approximately 39, 47, 69, and 39% of genotypes, respectively, had their order selection changed when ANOVA-AR was considered (data not shown). In these experiments, genotype 38 was eliminated from the selection of the best genotypes in Experiment 8 and genotype 116 in Experiment 9. Heinz, Mota, Gonçalves, Viegas Neto, and Carlesso (2012) analyzed 144 partially endogamic lines to obtain N efficient hybrids, and genotype 38 was selected among the 3 best lines with potential to be used in genetic breeding programs.

For experiments 11, 12, 13, and 14 with a selection intensity of 25% (56 out of 225 genotypes) of the most productive genotypes, approximately 96, 29, 57, and 25% of genotypes, respectively, had their order selection changed between the analyses (data not shown). In Experiment 11, genotypes 22, 172, and 173 were discarded by traditional ANOVA.

From the 14 experiments, Experiment 11 presented a high and significant p autocorrelation coefficient value (p = 0.37 and p-value ⩽ 0.01; Table 4). In this experiment, the spatial autocorrelation was more clearly detected, with greater reduction of the QME and QMG using ANOVA-AR. Moreover, a greater number of genotypes were excluded in this experiment when compared to the classical approach. According to Candido et al. (2009), the adoption of a method that considers the spatial relationship between genotypes and yield should be previously evaluated because this strategy may lead to the success or failure of a genetic breeding program.

Using Spearman's correlation test (r), which correlated the results of order selection of genotypes by traditional ANOVA and ANOVA-AR, only Experiments 3, 4, and 11 showed low correlations (Table 8).

Table 8. Spearman's correlation between the order selection generated by traditional ANOVA and by ANOVA-AR for data collected from 14 corn experiments under nitrogen fertilization conditions (ideal) and stressful conditions (low).

Experiment	Correlation	Experiment	Correlation
1	0.34^{*}	8	0.57^{**}
2	0.45^{**}	9	0.20^{*}
3	0.23^{ns}	10	0.30^{**}
4	0.24^{ns}	11	0.12^{ns}
5	0.38^{**}	12	0.61^{**}
6	0.31^{*}	13	0.35^{**}
7	0.55^{**}	14	0.59^{**}

In these 3 experiments, there was no concordance in order selection between the two analyses with a selection intensity of 25% of genotypes. In the remaining experiments, a median correlation that indicated the presence of a certain change in the order was observed. Candido et al. (2009) used neighborhood analysis to evaluate sugar genotypes and found that using the Papadakis method and moving averages method 1, there was no change in order selection of genotypes. Therefore, the effect of spatial analysis in experiments could be inexpressible.

A correlation test between accuracy and correlation value found using traditional ANOVA and ANOVA-AR was also proposed, to confirm whether accuracy is indeed capable of generating information about the correct order selection of genotypes for selection purposes, as stated by Andrade et al. (2020). The value found (R = 0.63*) indicated that the higher the accuracy, the greater the similarity between the order selections of genotypes by the analyses studied.

ANOVA-AR was proposed for agricultural experiments with corn to promote the reduction of error variability. However, its efficiency cannot be generalized and its use does not replace the importance of the traditional ANOVA applied to agricultural crops. It may be recommended that ANOVA-AR should be used for verifying spatial dependence; if not, traditional statistical analysis should be used.

Conclusion

Analysis of variance using an autoregressive can favor the selection of genotypes that could be excluded in traditional analysis of variance. Analysis of variance using an autoregressive provided parameter values of experimental precision similar to those generated by the traditional analysis of variance. Analysis of variance using an autoregressive does not replace the use of traditional analysis of variance, but its use is recommended to verify the existence of spatial dependence. There was no difference in relation to correlated errors in experiments with and without N fertilization in evaluating grain yield.

Acknowledgements

This work was carried out with financial contributions granted by Coordination for the Improvement of Higher Education Personnel (CAPES). The authors declare that they have no conflicts of interests in the conduct and publishing of the manuscript.

References

Andrade, M. H. M. L., Fernandes Filho, C. C., Fernandes, M. O., Bastos, A. J. R., Guedes, M. L., Marçal, T. S., ... Zotarelli, L. (2020) Accounting for spatial trends to increase the selection efficiency in potato breeding. *Crop Science*, *60*(5), 2354-2372. DOI: https://doi.org/10.1002/csc2.20226

Almeida, E. S. (2012). *Econometria espacial aplicada*. Piracicaba, SP: Esalq-USP.

Alves, G. C, Sobral, L. F., & Reis, V. M. (2020). Grain yield of maize inoculated with diazotrophic bacteria with the applicationof nitrogen fertilizer. *Revista Caatinga*, *33*(3), 644-652. DOI: https://doi.org/10.1590/1983-21252020v33n307rc

Banzatto, D. A., & Kronka, S. N. (2013). *Experimentação agrícola* (4. ed). Jaboticabal, SP: Funep.

Bernadeli, A., Rocha, J. R. A. S. C., Borém, A., Lorenzoni, R., Aguiar, R., Silva, J. N. B., ... Costa, M. D. B. (2021) Modeling spatial trends and enhancing genetic selection: An approach to soybean seed composition breeding. *Crop Science*, *61*(2), 976-988. DOI: https://doi.org/10.1002/csc2.20364

Bivand, R., Altman, M., Anselin, L., Assunção, R., Berke, O., Bernat, A., ... Yu, D. (2018). *Package 'spdep'*. Retrieved on April 6, 2018 from https://cran.r-project.org/web/packages/spdep/spdep.pdf

Campos, J. F., Carneiro, A. P., Peternelli, L. A., Carneiro, J. E. S., Silva, M. J., & Cecon, P. R. (2016). Classificação de famílias do feijoeiro sob diferentes cenários de dependência espacial e precisão experimental. *Pesquisa Agropecuária Brasileira*, *51*(2), 105-111. DOI: https://doi.org/10.1590/S0100-204X2016000200002

Candido, L. S., Perecin, D., Landell, M. G., & Pavan, B. E. (2009). Análise de vizinhança na avaliação de genótipos de cana-de-açúcar. *Pesquisa Agropecuária Brasileira*, *44*(10), 1304-1311. DOI: https://doi.org/10.1590/S0100-204X2009001000014

Duarte, J. B. (2005). *Especificação do modelo de análise estatística e suas implicações na seleção de genótipos em plantas*. In 11st Simpósio de Estatística Aplicada à Experimentação Agronômica / 50th Reunião Anual da Região Brasileira da Sociedade Internacional de Biometria, Londrina. Anais, Universidade Estadual de Londrina. Retrieved on April 19, 2018 from https://www.agro.ufg.br/up/68/o/Seagro_2005.pdf

Duarte, J. B., & Vencovsky, R. (2005). Spatial statistical analysis and selection of genotypes in plant breeding. *Pesquisa Agropecuária Brasileira*, *40*(2), 107-114. DOI: https://doi.org/10.1590/S0100-204X2005000200002

Es, H. M. V., & Es, C. L. V. (1993). Spatial nature of randomization and its effect on the outcome of field experiments. *Agronomy Journal*, *85*(2), 420-428. DOI: https://doi.org/10.2134/agronj1993.00021962008500020046x

Ertiro, B. T., Labuschagne M, Olsen, M., Das, B., Prasanna, B. M., & Gowda, M. (2020). Genetic dissection of nitrogen use efficiency in tropical maize through genome-wide association and genomic prediction. *Frontiers in Plant Science*, *11*(474), 1-16. DOI: https://doi.org/10.3389/fpls.2020.00474

Fritsche-Neto, R., & Borém, A. (2011). *Melhoramento de plantas para as condições de estresses abióticos*. Viçosa, MG: UFV.

Galembeck. F., Galembeck, A., & Santos, L. P. (2020). NPK: essentials for sustainability. *Química Nova, 42*(10), 1199-1207. DOI: https://doi.org/10.21577/0100-4042.20170441

Griffith, D. A. (1988). Estimating spatial autoregressive model parameters with commercial statistical package. *Geographical Analysis*, *20*(2), 176-186. DOI: https://doi.org/10.1111/j.1538-4632.1988.tb00174.x

Gumpertz, M. L., Graham, J. M., & Ristaino, J. B. (1997). Autologistic model of spatial pattern of phytophthora epidemic in bell pepper: effects of soil variables on disease presence. *Journal of Agricultural, Biological and Environmental Statistics*, *2*, 131-156. DOI: https://doi.org/10.4148/2475-7772.1350

Heinz, R., Mota, L. H. S., Gonçalves, M. C., Viegas Neto, A. L., & Carlesso, A. (2012). Seleção de progênies de meios-irmãos de milho para eficiência no uso de nitrogênio. *Revista Ciência Agronômica*, *43*(4), 731-739. DOI: https://doi.org/10.1590/S1806-66902012000400015

Legendre, P., Dale, M. R. T., Fortin, M. J., Gurevitch, J., Hohn, M., & Myers, D. (2002). The consequences of spatial structure for the design and analysis of ecological field surveys. *Ecography*, *25*(5), 601-615. DOI: https://doi.org/10.1034/j.1600-0587.2002.250508.x

Long, D. C. (1996). *Practical handbook of spatial statistics* (1st ed.). London, UK: CRC.

Maia, E., Siqueira, D. L., Carvalho, S. A., Peternelli, L. A., & Latado, R. R. (2013). Aplicação da análise espacial na avaliação de experimentos de seleção de clones de laranjeira Pêra. *Ciência Rural*, *43*(1), 8-14. DOI: https://doi.org/10.1590/S0103-84782012005000134

Mcintosh, M. S. (2015). Can analysis of variance be more significant? *Agronomy Journal*, *107*(2), 706-717. DOI: https://doi.org/10.2134/agronj14.0177

Morris, T. F., Murrell, T. S.; Beegle, D. B., Camberato, J. J., Ferguson, R. B., Grove, J., Ketterings, Q., Kyveryga, P. M., ... Yang, H. (2018). Strengths and limitations of nitrogen rate recommendations for corn and opportunities for improvement. *Agronomy Journal*, *110*(1), 1-37, 2018. DOI: https://doi.org/10.2134/agronj2017.02.0112

Ord, J. K. (1975). Estimation methods for models of spatial interaction. *Journal of the American Statistical Association*, *70*(349), 120-126. DOI: https://doi.org/10.2307/2285387

Piepho, H. P., Mohring, J., Pflugfelder, M., Hermann, W., & Williams, E. R. (2015). Problems in parameter estimation for power and AR(1) models of spatial correlation in designed field experiments. *Communications in Biometry and Crop Science*, *10*(1), 3-16.

R Development Core Team (2020). *R: A language and environment for statistical computing*. Viena, AT: R Foundation for Statistical Computing. Retrieved on Jan. 2, 2017 from https://www.R-project.org/

Resende, M. D. V., & Duarte, J. B. (2007). Precisão e controle de qualidade em experimentos de avaliação de cultivares. *Pesquisa Agropecuária Tropical*, *37*(3), 182-194.

Resende, M. D. V., & Sturion, A. (2003). Análise estatística espacial de experimentos via modelos mistos individuais com erros modelados por processos ARIMA em duas dimensões. *Revista de Matemática e Estatística*, *21*(1), 7-33.

Ribeiro, J., & Diggle, P. J. (2001). GeoR: A package for geostatistical analysis. *R-News*, *1-2*, 15-18.

Rossoni, D., & Lima, R. R. (2019). Autoregressive analysis of variance for experiments with spatial dependence between plots: A simulation study. *Brazilian Journal of Biometrics*, *37*(2), 244-257. DOI: https://doi.org/10.28951/rbb.v37i2.388

Scolforo, H. F., Scolforo, J. R. S., Mello, J. M., Ferraz Filho, A. C., Rossoni, D. F., Altoé, T. F., ... Lima, R. R. (2016). Autoregressive spatial analysis and individual tree modeling as strategies for the management of *Eremanthus erythropappus*. *Journal of Forest Research*, *27*(3), 595-603. DOI: https://doi.org/10.1007/s11676-015-0185-y

Silva, M. J., Carneiro, A. P. S., Feres, A. L. G., Carneiro, J. E. S., Santos, N. T., & Cecon, P. R. (2016). Spatial dependence in experiments of progeny selection for bean (*Phaseolus vulgaris* L.) yield. *Revista Ceres*, *63*(4), 477-485. DOI: https://doi.org/10.1590/0034-737X201663040007

Steel, R. G. D., & Torrie, J. H. (1980). *Principles and procedures of statistics* (2nd ed.). New York, NY: McGraw-Hill.

Su, W., Ahmad, S., Ahmad, I., & Han, Q. (2020). Nitrogen fertilization affects maize grain yield through regulating nitrogen uptake, radiation and water use efficiency, photosynthesis and root distribution. *PeerJ*, *8*, 1-21. DOI: https://doi.org/10.7717/peerj.10291

Wahid, Z., Latiff, A. I., & Ahmad K. (2017). Application of one-way ANOVA in completely randomized experiments. *Journal of Physics*, *949*(2017), 1-6. DOI: https://doi.org/10.1088/1742-6596/949/1/012017