AGRONOMY (AGRONOMIA)

Revista Brasileira de Ciências Agrárias
ISSN (on line) 1981-0997
v.18, n.3, e3103, 2023
Recife, PE, UFRPE. www.agraria.pro.br
DOI: 10.5039/agraria.v18i3a3103 - Protocol 3103
Subm. 23/01/2023 • App. 28/08/2023 • Pub. 17/10/2023

Grain yield adjusting efficiency in common bean genotypes

Paulo Henrique Cerutti[1*], Luan Tiago dos Santos Carbonari[1], Rita Carolina de Melo[1], Altamir Frederico Guidolin[1], Jefferson Luís Meirelles Coimbra[1]

[1] Universidade do Estado de Santa Catarina, Lages, SC, Brasil. E-mail: paulohcerutti@gmail.com; carbonari.luan@gmail.com; rita_carol_mel@hotmail.com; altamirguidolin@gmail.com; coimbrajefferson@gmail.com

ABSTRACT: The objective of this study was to verify the efficiency of using the plant stand covariate to adjust the average grain yield of common bean genotypes. Sixteen years of experiments were considered. In all years, the genetic treatments were randomized in field in a completely randomized block design. The main variable grain yield and the covariate plant stand per observation unit were evaluated. The information was submitted to analysis of variance and covariance. In approximately half of the trials (43.75%), the mean square of treatments was significant in both analyses, indicating the small improvement of the model when plant stand was included as a covariate. This information is confirmed by estimate of covariate efficiency, since in just two years (2017 and 2015) the adjustments were effective (268 and 203%, respectively). In addition, an association was observed between the average grain yield and adjustment efficiency of -0.60. Thus, the covariate was useful in years when the genotypes showed low productive performance, possibly caused by adverse environmental conditions. These conditions are responsible for plants heterogeneity number in the observation units.

Key words: confounding; covariance analysis; experimental techniques; *Phaseolus vulgaris* L.; VCU trials; yield potential

Eficiência no ajuste do rendimento de grãos em genótipos de feijão

RESUMO: O objetivo deste estudo foi verificar a eficiência da utilização da covariável estande de plantas no ajuste do rendimento médio de grãos em genótipos de feijão. Foram considerados 16 anos de experimentos. Em todos os anos, os tratamentos genéticos foram aleatorizados a campo sob delineamento blocos completos casualizados. Foi avaliada a variável principal rendimento de grãos e a covariável estande de plantas por unidade de observação. As informações foram submetidas a análise de variância e covariância. Aproximadamente na metade dos ensaios (43,75%), o quadrado médio de tratamentos foi significativo em ambas as análises, indicando a pequena melhoria do modelo ao se incluir o estande de plantas como covariável. Esta informação é comprovada com a estimativa da eficiência relativa da covariável, já que em apenas dois anos (2017 e 2015) os ajustes foram efetivos (268 e 203%, respectivamente). Além disso foi observado uma associação entre rendimento médio de grãos e a eficiência de ajuste de -0,60. Desse modo, a covariável foi útil em anos que os genótipos apresentaram baixo desempenho produtivo, possivelmente provocado por condições de ambiente adversas. Estas condições são responsáveis pela heterogeneidade do número de plantas nas unidades de observação.

Palavras-chave: análise de covariância; confundimento; técnicas experimentais; *Phaseolus vulgaris* L.; ensaios de VCU, potencial produtivo

* Paulo Henrique Cerutti - E-mail: paulohcerutti@gmail.com (Corresponding author)
Associate Editor: Alberto Cargnelutti Filho

Introduction

Common bean are an economically and socially important crop worldwide. The main reason for this is the amount of protein, vitamins, and minerals (iron and zinc) contained in its grains (Blair et al., 2013). This boosted cultivation in different regions around the world. Based on this relevance, the development of common bean genotypes with improved characteristics for grain yield (kg ha^{-1}) has always been one of the main objectives of the genetic improvement of this crop (Katuuramu et al., 2020).

Once a genotype has been developed by breeding programs, its agronomic performance must be proven before it can be made available to farmers. In Brazil, this proof is obtained by including genotypes with registration potential in the Value for Cultivation and Use trials (VCU). However, during the execution of these comparative trials, the grain yield of the genetic treatments can be affected by variations in the number of plants available in each observation unit used to compose this agriculturally important trait. The number of plants available becomes an important management factor, affecting crop growth and development by modifying the light environment intercepted by the canopy and competition for water and nutrients (Ricaurte et al., 2016).

In addition, in many situations, using the principle of local control alone, the breeder may not be successful in circumventing characteristics from the environment that act on the response variable (Morgan & Rubin, 2012). This situation is common in VCU trials with bean crops, which have a high number of treatments per block, ranging from 20 to 28. Among the ways to get around this problem, you can increase the number of observations of each treatment in each block (observational error), use a homogeneous experimental area, appropriate trial management or use statistical control, applying analysis of covariance, known as ANCOVA (Gomez & Gomez, 1985).

ANCOVA examines variances and covariances between response variables so that treatment effects are adjusted for. This is due to its characteristic of reducing experimental error, which reduces the variations attributed to chance within blocks and increases them between blocks (Yang & Juskiw, 2011). The analysis of covariance is based on the measurement of two or more response variables in which no experimental control has been applied. In this way, these variables become independent or covariates (Steel & Torrie, 1960). By using covariates, the means of each treatment are adjusted without the influence of variations arising from the covariate. In addition to the average adjustment in the value of each treatment, the addition of covariates improves the experiments coefficient of determination (R^2), so that most of the variation in the response variable is explained by the statistical model (Zhao et al., 2020). This improves the conclusions to be drawn when exploring treatment effects (Piana et al., 2007; Leppink, 2018).

Despite the potential employability of ANCOVA in agricultural experiments, questions still arise regarding the efficiency of using one or more covariates in the statistical model. For example, under what experimental conditions is ANCOVA really useful for adjusting treatment averages in a given experiment? What is the effectiveness and what are the gains of including a covariate in the experimental statistical model in line competition trials? Therefore, the objective of this study was to verify the efficiency of using ANCOVA, considering the assumptions of the statistical model in comparative grain yield trials in common bean cultivation, with adjustment of treatment averages by plant stand per observation unit.

Materials and Methods

In order to verify the efficiency of using ANCOVA, Value for Cultivation and Use trials (VCU) were conducted over 16 years in partnership with the Empresa de Pesquisa Agropecuária e Extensão Rural de Santa Catarina (EPAGRI). All these tests were carried out in the field in the municipality of Lages, SC, Brazil, on the premises of the Universidade do Estado de Santa Catarina (UDESC), at the Centro de Ciências Agroveterinárias (CAV), whose coordinates are 27º 47' and 50º 18', at 950 m above sea level. The area where the experiment was carried out has soil of the type Cambissolo Húmico Alumínico Léptico, with a clayey texture. According to the Climate Atlas of the Southern Region of Brazil, the average air temperature is 15.7 °C and rainfall is approximately 1,500 mm per year. Over the 16 years the VCUs have been running, between 20 and 28 common bean genotypes have been grown, comprising promising lines for grain yield and commercial cultivars (control). The variation in the number of genotypes in each year of the trial is due to the availability of each breeding program that provides lines for these trials.

In all the experiments, the genotypes were randomized in the field in a complete block design with four replications. Each experimental unit consisted of four 4 m rows, 0.5 m apart. The sowing density was 15 seeds per linear meter. The observation unit was formed by the two central rows of each experimental unit, in order to remove the effects of borders. Throughout the trials, management practices related to fertilization, weed control and insect pests were carried out in accordance with technical recommendations for growing beans (Fancelli & Dourado Neto, 2007; CQFS-RS/SC, 2016). Two response variables were evaluated, one primary and one secondary. The main response variable, grain yield (kg ha^{-1}), was obtained by mechanically threshing the plants harvested in the observation unit, followed by correcting the moisture content of the grains to 13% using the forced ventilation oven method at 60 °C and then measuring their grain mass. The secondary response variable, plant stand (ETD), was measured by the total number of plants harvested in each observation unit and used to estimate grain yield.

A descriptive analysis of the variables was carried out in all the trials for the genotype experimental factor. The assumptions regarding the mathematical model tested were: i) additivity of the effects, using Tukey non-additivity test; ii) normality of the residuals, using the Shapiro-Wilk test; and, iii) homogeneity of variances, using Levene test. The ANCOVA application considered a statistical model without and with the plant stand covariate. The model without the inclusion of the covariate represents the univariate analysis of variance for the complete block design, where $Y_{ij} = \mu + block_i + genotype_j + \varepsilon$. The model with the covariate is a combination of analysis of variance and regression analysis, described as $Y_{ij} = \mu + block_i + genotype_j + b(x_{ij} - x) + \varepsilon_{ij}$. Where, Y represents the phenotypic value of a given response variable measured in a unit of observation; μ the overall average of the test; $block_i$ the i-th level of the block factor; $genotype_j$ the j-th level of the genotype factor; b the regression coefficient; x is the covariate; and, x is the overall average of the covariate. Univariate analysis of variance (ANOVA) and analysis of covariance (ANCOVA) were performed using the general linear procedure (proc glm). The relative efficiency (RE, %) of using the covariate was determined using Equation 1.

$$RE = \frac{\text{MS residue yield}}{\text{MS residue corrected} \times 1 + \left(\frac{\text{MS treat stand}}{\text{SQ stand residue}}\right)} \quad (1)$$

where: MS residue yield - mean square of the residual for the main response variable grain yield; MS residue corrected yield - mean square of the corrected residual, with the inclusion of the covariate plant stand in the model; MS treat stand - mean square of the treatment effect of the covariate plant stand; and; SQ stand residue - sum of squares of the covariates residual.

In order to check whether the relative efficiency of using a covariate in the model changes depending on whether the models assumptions are met, two data transformations of $\sqrt{x + 0.5}$ and Box & Cox (1964) $yt = y^{-\lambda} - 1/\lambda$ were carried out.

Pearson simple correlation estimates were obtained using the proc corr procedure, and significance was tested using the t-test. In all analyses, type I error was considered with a significance level of 0.05.

All statistical analysis procedures were carried out using SAS software in the academic version (SAS OnDemand for Academics).

Results and Discussion

Information regarding descriptive statistics for the main response variable grain yield and the covariate plant stand were presented in Table 1. For grain yield, the overall average grain yield of the genetic treatments was 1440 kg ha^{-1}. According to estimates by the Companhia Nacional de Abastecimento (Conab, 2022), for the month of August 2022, common bean productivity in the country was 1067 kg ha^{-1}. Over the years of evaluation, the average grain yield below the genetic potential of the vast majority of genotypes is possibly associated with environmental conditions, such as low rainfall or irregular and poorly distributed rainfall over the years. The bean crop needs 300 to 400 mm of accumulated rainfall over the course of its cycle, with full flowering being the most critical time for the crop (Sofi et al., 2018). The low production performance in the years 2015 to 2017 and in 2019 was related to low efficiency in weed control and irregular rainfall distribution, causing unevenness in the plant stand per observation unit and, consequently, a drop in grain yield.

Table 1. Summary of the descriptive analysis of mean (x) and standard deviation (sd), considering the main response variable grain yield (YIELD, kg ha^{-1}) and the covariate plant stand (ETD), for the 16 years of Value for Cultivation and Use trials (VCU) in common bean cultivation.

Year	Grain yield (kg ha^{-1})		Plant stand	
	$\bar{x}$	sd	$\bar{x}$	sd
2007	2375	493	70.95	9.40
2008	2603	626	66.66	7.45
2009	2207	546	58.63	10.92
2010	1667	558	46.98	12.23
2011	1130	532	46.39	16.69
2012	2622	709	61.64	11.14
2013	2214	745	44.90	9.39
2014	1717	518	51.87	10.10
2015	220	233	25.78	11.96
2016	314	250	37.50	10.00
2017	172	151	40.14	22.91
2018	1286	512	32.20	12.15
2019	495	355	26.48	19.53
2020	1044	363	44.96	11.36
2021	1517	594	73.28	11.94
2022	1460	740	51.09	14.56
General	1440	989	48.71	19.94

The range of descriptive statistics related to the standard deviation for the grain yield variable was 151 to 745 kg ha^{-1}. The standard deviation indicates the variation in grain yield of the genotypes around the general average of the trials. The magnitude and significance of deviations serve as an estimate of the predictability of genotypes in the face of environmental conditions. In plant breeding, the aim is to develop a genotype with an average grain yield that is higher than that of the control (highest average) and that has as small a deviation as possible and is not significant. This information serves to predict behavior in the face of environmental changes, which are recurrent in VCUs (Torga et al., 2016). Genotypes with predictable performance are agronomically advantageous, as they can maintain grain yields even in less than ideal soil and climate conditions (Nascimento & Souza, 2022).

The averages for the covariate plant stand per observation unit ranged from 26 to 73. This indicates that although the researcher establishes the same number of seeds per linear meter in all experimental units, factors such as the growth habit of the genotypes, soil fertility, water availability, and the incidence of pests and diseases that affect the bean crop, interfere with the number of plants available for harvest (Soratto et al., 2017). With regard to growth habit, the VCU trials include type I (determinate), type II (indeterminate shrub growth), and type III (indeterminate prostrate growth). For these types of habits, there are recommendations for densities ranging from 10 to 15 seeds per linear meter (Barbosa & Gonzaga, 2012). Considering the estimated standard deviation, the range of plants per observation unit was 7.45 to 22.9. As there are variations in plant density, interspecific competition effects are accentuated, resulting in an imbalance of water and nutritional resources by the plants within the plots, promoting faulty grain yield estimates (Tokatlidis, 2017).

Based on the possible influence of the number of plants harvested per observation unit on the main grain yield variable, two analyses were carried out for each year of the VCU trial. One represents analysis of variance (ANOVA) and the other ANCOVA (analysis of covariation). In addition to the factors controlled in the experiment, in this case blocks and bean genotypes (lines and cultivars), ANCOVA considers a secondary response variable which becomes a covariate or independent variable. In general terms, this analysis makes it possible to remove one degree of freedom from the residual and incorporate this degree of freedom into the sources of variation controlled in the experiment. By reducing one degree of freedom of the residual, the precision of the comparison between treatments tends to increase, making the researchers inference more reliable. This allows treatment averages to be adjusted according to the number of plants harvested. This is because ANCOVA is a combination of analysis of variance and regression analysis (Yang & Juskiw, 2011).

As a requirement for carrying out the analysis of variance, the assumptions regarding the statistical model were represented (Table 2). It is important to note that if the researcher wishes to explore the effects of treatments in all experiments, these assumptions must be respected individually. In 31.2% of the VCU years, the assumptions of additivity of effects, normal distribution of residuals and homogeneity of variances were not met. Failure to meet the assumption related to the normality of the residuals was also observed in comparative performance trials with the bean crop (Storck et al., 2011). If the variances between treatments are homogeneous, the linear relationship between them and the covariate is positively influenced, resulting in a reduction in the residual estimate by ANCOVA. The non-additivity of the statistical model may be related to the erroneous execution of the test for this purpose or the presence of discrepant values obtained from the observation units. The number of plants harvested in each observation unit directly affects the estimates of Tukey non-additivity test. In soybean cultivation experiments, a tendency to overestimate the F-statistic values was observed in situations where the samples were small (less than five plants harvested per observation unit) (Souza et al., 2023). In the VCUs trials with the bean crop,

Table 2. Assumptions of the statistical model regarding the terms of the equation verified for the 16 years of Value for Cultivation and Use trials (VCUs) for the common bean crop in Lages, SC, Brazil.

Year	Additivity		Normality		Homogeneity	
	F	Pr > F	W	Pr < W	F	Pr > F
2007	2.77	0.100	0.97	0.037	1.46	0.116
2008	0.78	0.379	0.98	0.093	1.52	0.072
2009	0.01	0.915	0.98	0.100	1.20	0.253
2010	1.68	0.197	0.98	0.230	0.88	0.624
2011	1.71	0.193	0.98	0.205	1.61	0.053
2012	4.30	0.040	0.97	0.168	1.33	0.162
2013	0.12	0.725	0.97	0.075	2.20	0.011
2014	0.25	0.621	0.98	0.410	1.91	0.030
2015	3.41	0.066	0.87	0.001	2.05	0.004
2016	4.27	0.040	0.98	0.052	2.79	0.001
2017	23.21	0.001	0.91	0.001	1.25	0.210
2018	0.24	0.623	0.98	0.773	2.20	0.011
2019	73.4	0.001	0.87	0.001	1.52	0.082
2020	0.13	0.721	0.97	0.356	0.97	0.542
2021	4.88	0.028	0.98	0.097	0.83	0.711
2022	0.62	0.431	0.97	0.021	0.92	0.575

these variations were observed in the number of plants collected in the observation units, as shown in Table 2.

Reducing the number of plants in the crop row has a proportional influence on grain production per plant, resulting in more vigorous plants with greater branching and, consequently, more grains per plant (Soratto et al., 2017). Suitable ways of correcting the model's lack of additivity include transformations of √x, √x + 1, log x, and log x + 1. It should be noted that the type of transformation to be adopted depends on the nature of the response variable being measured. In addition to the transformation technique, the use of more appropriate designs compared to the randomized block design, the use of variable characteristics as covariates or even as supplementary experimental factors (type of growth habit, high or low seed vigor), can circumvent the problem of additivity, as long as this is based on experimental planning (Silva, 2020).

To check that ANCOVA could be carried out, the three main assumptions needed to validate this analysis were represented (Table 3). These three assumptions relate to: i) existence of a linear relationship between the main response variable (grain yield) and the covariate (plant stand); ii) regression coefficient (β) the same for all treatments; and, iii) independent effect of the covariate on the treatments.

Table 3. Assumptions made when carrying out the analysis of covariance (ANCOVA) for the 16 years of Value for Cultivation and Use trials (VCUs) in the common bean crop in Lages, SC, Brazil.

ANCOVA assumptions	Attended	Not attended
Linear relationship between the main variable and the covariate	14	2
Regression coefficient (β) equal for treatments	2	14
Independent effect of the covariate on the treatments	12	4

According to Table 3, the assumption of a linear relationship between the response variables was met in 87.5% of the years. This indicates that the number of plants harvested has a significant and non-zero relationship with grain yield over the 14 years that the VCU trials have been conducted. With each increase in the number of plants, grain yield also changed. This scenario was completely opposite to that seen for the assumption associated with the regression coefficient, since in 14 years of experiments, this assumption was met in only two (12.5%). The vast majority of the experiments (12 years) met the assumption of independence between the covariate and the treatments, indicating that the covariate was not affected by the genetic treatments considered. A covariate can first be used to adjust the means of the main response variable for differences in the values of the independent variable, but this adjustment is only reliable if the regression coefficient is common to all treatments (Steel & Torrie, 1960). ANCOVA provides a test comparing the sums of squares of the errors resulting from two models simultaneously (one model from analysis of variance and the other from regression analysis). The experiments carried out in 2017 and 2010 were the only ones that met the three assumptions of the analysis of covariance. In the other years, one or other assumption was not met, possibly generating implications or low effectiveness of using the plant stand covariate to adjust treatment averages.

The ANOVA and ANCOVA for the years of VCU trials are shown in Table 4. The analyses show that in seven years (43.75%) both the ANOVA and ANCOVA showed a significant effect for the genotype factor applied in each trial. This fact indicates that the analyzes converged to the same result. In addition, it may indicate that if the growing conditions during the bean cycle are ideal, ANCOVA has little effect on the correction of treatment averages, considering the plant stand covariate, compared to ANOVA.

In a recent experiment using ANCOVA to evaluate the use of soil resources by plants of the species *Polygonum criopolitanum* and *Carex thunbergii*, the researchers observed that a significant and apparent reduction in the experimental error estimate was only possible by including two covariates in the experiment mathematical model. Both species are multiplied vegetatively (rhizomes). In this way, the authors considered the size and mass of the rhizomes as covariates when setting up the experiment, significantly improving the accuracy of comparisons between treatments (Huangfu et al., 2022). In another experiment to evaluate the effect of soil biota on the growth and development of native plants in Minnesota, USA, the authors considered plant height as a covariate. This is because plants with larger initial statures tend to benefit compared to smaller plants, influencing the treatments applied (Stein & Mangan, 2020).

The two examples of the use of analysis of covariance highlight the importance of defining a covariate before the experiment is carried out. This definition can occur due to prior knowledge of the response variables that influence the expression of other variables through historical research data, as was the case with the plant height, size, and rhizome weight variables in the experiments mentioned above. Therefore, the analysis of covariance is improved when it is possible to quantitatively measure one or more response variables, prior to the allocation of treatments, and thus acts on sources of variation not controlled in the experimental design (Gomez & Gomez, 1985; Santos et al., 2014).

By applying ANCOVA, adjustments were observed in the averages of the genotypes for grain yield in all the years of evaluation. This was represented by Pearson simple linear correlation analysis (r) between treatment averages before

Table 4. Summary of the analysis of variance (ANOVA) and covariance (ANCOVA), considering the main response variable grain yield (kg ha^{-1}) and the covariate plant stand (ETD), for the 16 years of Value for Cultivation and Use trials (VCUs) in the common bean crop in Lages, SC, Brazil.

Year	ANOVA			ANCOVA			r
	MST	MSR	F	MST	MSR	F	
2007	227150	230663	0.98	252958	222730	1.14	0.98
2008	314596	355418	0.89	307812	349973	0.88	0.99
2009	693980*	212728	3.26	443103*	154182	2.87	0.99
2010	641270*	209108	3.07	376360*	160736	2.34	0.98
2011	518713*	201664	2.57	219359*	148010	1.48	0.99
2012	1098748*	346606	3.17	944601*	335381	2.82	0.99
2013	496236	575018	0.86	216931	327138	0.66	0.97
2014	635376*	125408	5.07	610786*	109038	5.60	0.99
2015	130285*	44924	2.90	57928*	21931	2.64	0.95
2016	139573*	45612	3.06	103717*	42224	2.46	0.99
2017	21451	13625	1.57	4837	5018	0.96	0.93
2018	310221	247772	1.25	281627	240339	1.17	0.99
2019	80621	81911	0.98	38910	50302	0.77	0.97
2020	161831	114262	1.42	86312	83601	1.03	0.96
2021	248671	228362	1.09	227119	174757	1.30	0.97
2022	690674	492618	1.40	503548	452528	1.11	0.99

* Significant at the 0.05 level by the F test; MST - mean square of treatment; MSR - mean square of the residual; r - Pearson simple linear correlation coefficient between the treatment averages without the covariate versus the averages adjusted for the covariate.

and after adjustment by stand, as shown in Table 4. The correlation estimates ranged from 0.93 to 0.99 and were all significant at the 0.05 level using the t-test, indicating that there is a strong association between the observed and adjusted averages. This situation highlights the fact that when a covariate is used in the statistical model, the treatment averages are adjusted, but there were no drastic changes in the ranking of the genotypes according to their average performance. In this type of analysis, it is common to only adjust the averages of the genotypes, without drastic and significant changes to the ranking, since the change is made based on the average of the covariate. The use of covariates, such as soil tension and texture, temperature, radiation incidence and rainfall were effective in adjusting and ranking wheat genotypes, in genotype × environment interaction studies with this crop (Rincent et al., 2019) .

In addition to estimating the corrected value of the mean square of treatment and residue using ANCOVA, the relative efficiency (RE, %) of applying a covariate to the experiment model was estimated using the equation presented in the material and methods section, also for each year of VCU (Figure 1). The equation expresses the relationship between the mean square of the residual of the main response variable (in this case grain yield), and the mean square of the residual after adjustment for the covariate, plus the treatment effect.

In general terms, the efficiency of ANCOVA is achieved by comparing the variance of the treatment effect with and without adjusting for the covariate (Yang & Juskiw, 2011). The average relative efficiency over the 16 years was 137.4%. The extreme values were observed in 2007 and 2017 (100.2 and 268.7%, respectively). This observation of extreme values was obtained by ranking the relative efficiency calculated for each year in descending order. In addition to 2017, the experiment carried out in 2015 showed a relative efficiency value of 203%. In experiments with maize, relative efficiency

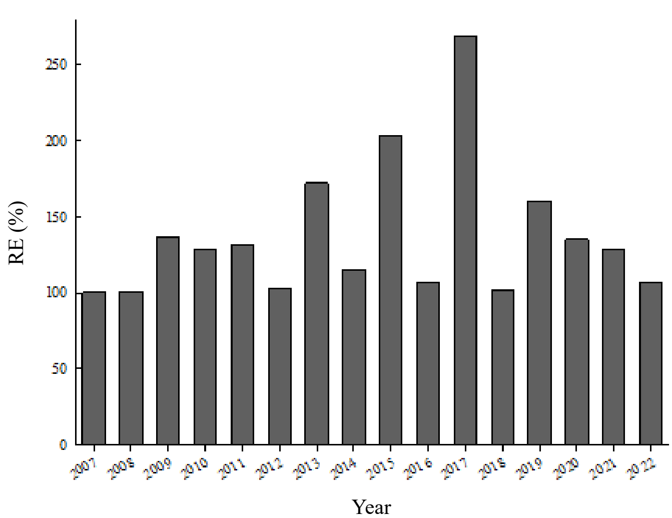

Figure 1. Relative efficiency (RE, %) of the use of the covariate plant stand in Value for Cultivation and Use trials (VCUs), considering the main response variable grain yield (kg ha^{-1}), for the common bean crop between 2007 and 2022.

was detected with ANCOVA of 555%. In this case, the authors concluded that the analysis of covariance in a randomized complete block design allowed for greater precision in comparing treatments compared to experiments with up to 23 repetitions (Yang & Juskiw, 2011).

The effectiveness of the use of ANCOVA considering the covariate plant stand per observation unit proved to be of little relevance, with only two years being effective (2017 and 2015), whose effectiveness exceeded 100%. The higher relative efficiency in these years is possibly associated with the low average grain yield and high variation in plant stand, previously represented by descriptive statistics. Based on this, considering the assumptions of the analysis of variance, data transformations were applied to the years of experiments that did not meet the assumptions of additivity of the effects, normality of the residues, and homogeneity of variance of the treatments. To this end, the transformations of $\sqrt{x} + 0.5$ and Box & Cox (1964) $yt = y^{\lambda} - 1/\lambda$ were considered. These transformations are commonly cited in the literature among a wide family of transformations. They are relatively useful tools for improving data normality and equalizing variances, making it possible to apply parametric and non-parametric tests to compare treatment effects (Osborne, 2010). The transformations and the fulfillment of the assumptions for two years of VCU cultivation (2012 and 2016) were represented.

The experiment carried out in 2012 showed a lack of additivity in the model, and in 2016 there were flaws in both the additivity and normality of the residuals. The lambda values obtained were 0.80 (2012) and 0.35 (2016). As shown in Table 5, after transforming the data using both transformations, we found that additivity was met in 2012 (p = 0.089; 0.614), and additivity and normality of the residuals in 2016 (p = 0.340; 0.595; 0.611; 0.499), respectively for

Table 5. Representation of the ANOVA and ANCOVA assumptions after the transformation of the grain yield response variable.

	Assumptions	$\sqrt{x} + 0.5$	$y^{\lambda} - 1/\lambda$
	2012		
ANOVA	Additivity	0.089	0.614
	Normality	0.116	0.001
	Homogeneity	0.001	0.437
ANCOVA	Linearity[1]	0.001	0.049
	β common[2]	0.083	0.380
	Independence[3]	0.366	0.882
	RE (%)[4]	102	104
	2016		
ANOVA	Additivity	0.340	0.595
	Normality	0.611	0.499
	Homogeneity	0.010	0.020
ANCOVA	Linearity	0.001	0.001
	β common	0.780	0.079
	Independence	0.926	0.918
	RE (%)	111	113

[1] Linear relationship between the main variable and the covariate; [2] Linear regression coefficient equal for all treatments; [3] Independent effect of the covariate on the treatments; [4] Relative efficiency of use of the covariate plant stand.

√x + 0.5 and Box & Cox (1964). The interpretation of the probabilities shows that the transformations were useful for the assumptions of additivity and normality. However, when looking at the relative efficiency values for the use of the covariate, there were no marked improvements. This is because in 2012 and 2016, considering the original data, relative efficiencies of 102 and 106% were obtained, respectively. According to Table 5, after the transformations the efficiencies were 102 and 104% for 2012 and 111 and 113% for 2016. On this basis, it was possible to state that even after transforming the data, the plant stand covariate did not provide a considerable gain to the statistical model.

One of the reasons that may be associated with the effectiveness of ANCOVA is the model itself for adjusting treatment averages

$$\hat{y} = \overline{y} \times \beta\left(x - \overline{x}\right).$$

The regression coefficient is multiplicative of the differences between the variances of the covariate. Thus, the correct application and maximum efficiency of the covariate occurs when this coefficient stipulated for testing the linear relationship between a dependent variable and an independent variable is the same for all treatments (Yang & Juskiw, 2011). In the two years in which the covariate was most effective, the regression coefficient was the same for all treatments, which justifies its efficiency.

Another reason linked to effectiveness may be related to the denominator part of the expression that calculates relative efficiency, which is influenced by the reduction in the mean square and sum of squares of the residual after correction with the covariate and by the mean square of the treatment of this covariate used in the model. In 2007, the variation in the number of plants per observation unit was 53.1%, while in 2017 it was 95.0%, with this year variation being higher in comparison. Based on this information, ANCOVA was effective when the variation in the covariate was high, thus making it necessary to correct the averages of the treatments according to the number of plants available to make up the average grain yield per observation unit. In addition, there was a significant association between the efficiency of use of the covariate in relation to grain yield and the number of plants harvested. Significant simple correlations were found between the covariates use efficiency and grain yield (-0.60) and the number of plants harvested per observation unit (-0.47). This information is consistent, since the years 2017 and 2015, which showed the greatest deviation in grain yield from the general average (-1268 and -1220), were the years in which the covariate was effective in adjusting treatment averages.

The low average yield shown by the crop in some years may be related to the rainfall during the period. From 2007 to 2022, 466 mm of average rainfall was observed, with extremes between 271 and 670 (information obtained from NASA POWER Global Meteorology, Surface Solar Energy and Climatology Data Client for R - Sparks, 2018). However, it is known that unfortunately this rainfall is not uniform over the course of the crop cycle. Another obstacle observed during the cultivation cycle is infestation by weeds, including Brachiaria plantaginea, *Cynodon dactylon*, and *Eleusine indica*, which are frequently observed in VCU trials in Southern Brazil. Bean crops are negatively influenced by weeds, especially grasses, due to their pressure on the emergence phase of the soil seed bank, rapid initial development, and accelerated biomass production, compared to legumes (D'Amico-Damião et al., 2020; Parker et al., 2020).

As shown, the low frequency of effective use of a covariate in VCU trials may be due to factors related to: the number of repetitions and treatments per repetition, the presence of uncontrolled environmental characteristics, and genotype × environment interaction. In VCU trials with the bean crop, many treatments are conducted per block (20 - 28), due to the rules drawn up for the execution of these experiments. In this sense, it is common for there to be variations attributed to uncontrolled sources within the replications, altering the productive position of the genotypes due to the inflated estimate of the experimental error (Cargnelutti Filho et al., 2009).

Blocking becomes effective when there is a predictable pattern of variability in the area where the experiment is set up. In this situation, the shape of the plots and the orientation of the blocks stand out. What is often observed is the incorrect allocation of blocks in relation to the variation gradient, and experimental units with an inappropriate size and format for each cultivation condition where the VCUs are implemented. To overcome this problem, more robust experimental designs can be used, such as incomplete blocks. This type of design is common in the field of genetic improvement, as it allows the use of a large number of treatments, compared to the complete block design. Furthermore, dividing the experimental area into smaller blocks increases homogeneity within each block (Gomez & Gomez, 1985).

The use of more appropriate experimental designs aims to reduce the estimate of experimental error. The experimental error of a response variable is estimated as a function of the variation in the measured values of this variable that are attributed to uncontrolled characteristics, and can come from different sources of variation, among them: variation attributed to the data collection and recording process itself, uneven reproduction of the experimental conditions imposed by the researcher in each experimental unit, possible interaction between environmental characteristics and experimental conditions and other sources of variation not controlled by the researcher (Silva, 2020). In addition, there is another source of experimental error, arising from the confounding of treatment effects (factors) with unknown effects (uncontrolled).

In VCU experiments, the unknown characteristics related to: i) seed of the genotypes (purity, health, germination, and vigor); ii) environment (soil, climate, incidence of insects, diseases, and invasive plants); iii) cultivation techniques (fertilization and disease control); and, iv) of the process of collecting and measuring information. The control of experimental techniques can aid when considering, for example, the use of healthy seeds, fertilization, pest control and the collection and measurement of information in an appropriate and efficient manner, in accordance with the recommendations for the crop. Statistical control, on the other hand, can act to define one or more covariates before the experiment is carried out (Silva, 2020). This definition has not yet been used in VCU trials in bean cultivation. Thus, measuring the covariates germination percentage and seed vigor of each genotype, seed mass, as well as the history of weed incidence, soil fertility levels per block and soil compaction in the experimental area can increase the efficiency of ANCOVA in VCU trials with the bean crop, as long as they are considered in the design stage of the experiments.

Conclusion

The covariate plant stand per observation unit was only effective in adjusting treatment averages in two VCU years (2015 and 2017). The trials with average grain yields lower than the general average showed greater use of the covariate. The relative utilization efficiency of the covariate showed an association of -0.60 with the average grain yield.

Acknowledgments

We would like to thank the Fundação de Amparo à Pesquisa e Inovação do Estado de Santa Catarina (FAPESC) for granting scholarships and financial support for research projects.

Compliance with Ethical Standards

Author contributions: Conceptualization: PHC; LTSC; JLMC; AFC; RCM; Data curation: PHC; JLMC; Formal analysis: PHC; LTSC; JLMC; Funding acquisition: JLMC; AFG; Investigation: PHC; LTSC; RCM; Methodology: PHC; RCM; Project administration: JLMC; AFG; Resources: PHC; JLMC; AFG; Software: PHC; JLMC; Supervision: JLMC; AFG; Validation: PHC; RCM; JLMC; Visualization: PHC; LTSC; JLMC; AFG; RCM; Writing – original draft: PHC LTSC; Melo, R. C.; Writing – review & editing: PHC; JLMC; AFG

Conflict of interest: The authors declare that they have no conflict of interest.

Funding source: The Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and the Fundação de Âmparo à Pesquisa e Inovação de Santa Catarina (FAPESC).

Literature Cited

Barbosa, F. R.; Gonzaga, A. C. O. Informações técnicas para o cultivo do feijoeiro-comum na Região Central-Brasileira: 2012-2014. Santo Antônio de Goiás: Embrapa Arroz e Feijão, 2012. 247p. (Embrapa Arroz e Feijão. Documentos, 272). https://ainfo.cnptia.embrapa.br/digital/bitstream/item/61388/1/seriedocumentos-272.pdf. 19 Jan. 2023.

Blair, M. W.; Izquierdo, P.; Astudillo, C.; Grusak, M. A. A. legume biofortification quandary: Variability and genetic control of seed coat micronutrient accumulation in common beans. Frontiers in Plant Science, v. 4, n. 7, p.1–14, 2013. https://doi.org/10.3389/fpls.2013.00275.

Box, G. E. P.; Cox, D. R. An analysis of transformations. Journal of the Royal Statistical Society Series B: Statistical Methodology, v. 26, n. 2, p.211-243, 1964. https://doi.org/10.1111/j.2517-6161.1964.tb00553.x.

Cargnelutti Filho, A.; Ribeiro, N.D.; Storck, Lo. Número de repetições para a comparação de cultivares de feijão. Ciência Rural, v.39, n.9, p. 2419–2424, 2009. https://doi.org/10.1590/S0103-84782009000900006.

Companhia Nacional de Abastecimento - Conab. Acompanhamento da safra brasileira de grãos. Brasília: Conab, 2022. 81p. (v.10, n.3, Safra 2022/23, terceiro levantamento). https://www.conab.gov.br/info-agro/safras/graos/boletim-da-safra-de-graos/item/download/45427_510a1459a11d50bd53a250780fea86a4. 20 Dec. 2022.

CQFS-RS/SC. Manual de adubação e de calagem para os estados do Rio Grande do Sul e Santa Catarina. Porto Alegre: Universidade Federal do Rio Grande do Sul; Núcleo Regional Sul, 2016. 376p.

D'Amico-Damião, V.; Barroso, A. A. M.; Alves, P. L. C. C.; Lemos, P. L. Intercropping maize and succession crops alters the weed community in common bean under no-tillage. Pesquisa Agropecuaria Tropical, v. 50, e65244, 2020. https://doi.org/10.1590/1983-40632020V5065244.

Fancelli, A. L.; Dourado Neto, D. Produção de feijão. Piracicaba: Os Autores, 2007. 386p.

Gomez, K. A., Gomez, A. A. Statistical procedures for agricultural research. New York: John Wiley & Sons, 1985. 704p.

Huangfu, C.; Zhang, L.; Hui, D. Science of the Total Environment Density-dependent plant-soil feedbacks of two plant species affected by plant competition. Science of the Total Environment, v. 807, part. 2, e150908, 2022. https://doi.org/10.1016/j.scitotenv.2021.150908.

Katuuramu, D. N.; Luyima,G.B.; Nkalubo, S.T.; Wiesinger, J. A.; Kelly, J. D.; Cichy, K. A. On-farm multi-location evaluation of genotype by environment interactions for seed yield and cooking time in common bean. Scientific Reports, v. 10, n. 1, e3628, 2020. https://doi.org/10.1038/s41598-020-60087-2.

Leppink, J. Analysis of Covariance (ANCOVA) vs. Moderated Regression (MODREG): Why the Interaction Matters. Health Professions Education, v. 4, n. 3, p. 225–232, 2018. https://doi.org/10.1016/j.hpe.2018.04.001.

Morgan, K. L.; Rubin, D. B. Rerandomization to improve covariate balance in experiments. Annals of Statistics, v. 40, n. 2, p. 1263–1282, 2012. https://doi.org/10.1214/12-AOS1008.

Nascimento, M. R.; Souza, A. G. Adaptability and stability of black and purple bean genotypes. Pesquisa Agropecuária Tropical. v.52, e71631, 2022. https://.doi.org/10.1590/1983-40632022v5271631.

Osborne, J.W. Improving your data transformations: Applying the Box-Cox transformation. Practical Assessment, Research and Evaluation, v. 15, e12, 2010. https://doi.org/10.7275/qbpc-gk17.

Parker, T. A.; Palkovic, A.; Gepts, P. Determining the genetic control of common bean early-growth rate using unmanned aerial vehicles. Remote Sensing, v. 12, n. 11, p. 13–16, 2020. https://doi.org/10.3390/rs12111748.

Piana, F. C.B.; Silva, J.G.C.; Antunes, I.F. Ajuste do rendimento para a variação do estande em experimentos de melhoramento genético do feijão. Pesquisa Agropecuária Brasileira. v. 12, n. 1, p. 1687–1696, 2007. https://doi.org/10.1590/S0100-204X2007001200004.

Ricaurte, J; Michelangeli, J. A. C.; Sinclair, T.R.; Rao, I. M.; Beebe, S. E. Sowing density effect on common bean leaf area development. Crop Science, v. 56, n. 5, p. 2713–2721, 2016. https://doi.org/10.2135/cropsci2016.01.0056.

Rincent, R.; Malosetti, M.; Ababaei, B.; Touzy, G.; Mini, A.; Bogard, M.; Martre, P.; Le Gouis, J.; Van Eeuwijk, F. Using crop growth model stress covariates and AMMI decomposition to better predict genotype-by-environment interactions. Theoretical and Applied Genetics, v. 132, n. 12, p. 3399–3411, 2019. https://doi.org/10.1007/s00122-019-03432-y.

Santos, D.; Dal Col Lúcio, A.; Cargnelutti Filho, A.; Storck, L.; Lorentz, L. H.; Schabarum, D. E. Efeito de vizinhança e tamanho de parcela em experimentos com culturas olerícolas de múltiplas colheitas. Pesquisa Agropecuaria Brasileira, v. 49, n. 4, p. 257–264, 2014. https://doi.org/10.1590/S0100-204X2014000400003.

Silva, J. G. C. Experiment: conceptual basis. Journal of Experimental Agriculture International, v.42, n. 6, p. 7–22, 2020. https://doi.org/10.9734/jeai/2020/v42i630530.

Sofi, P. A.; Djanaguiraman, M.; Siddique, K. H. M.; Prasad, P. V. V. Reproductive fitness in common bean (*Phaseolus vulgaris* L.) under drought stress is associated with root length and volume. Indian Journal of Plant Physiology, v. 23, n. 4, p. 796–809, 2018. https://doi.org/10.1007/s40502-018-0429-x.

Soratto, R. P.; Catuchi, T. A.; De Souza, E. F. C.; Garcia, J. L. N. Plant density and nitrogen fertilization on common bean nutrition and yield. Revista Caatinga, v. 30, n. 3, p. 670–678, 2017. https://doi.org/10.1590/1983-21252017v30n315rc.

Souza, R. R.; Toebe, M.; Mello, A. C.; Bittencourt, K. C.; Toebe, I. C. D. Does soybean sample size impact Tukey's test for non-additivity? Ciência Rural, v. 53, n. 5, e20220181, 2023. https://doi.org/10.1590/0103-8478cr20220181.

Sparks, A. H. Nasapower: a NASA POWER global meteorology, surface solar energy and climatology data client for R. Journal of Open Source Software, v.3, n.30, p.1-3, 2018. https://doi.org/10.21105/joss.01035.

Steel, R.G.D,; Torrie, J.H. Principles and procedures of statistics. New York: Mcgraw-Hill, 1960. 481p.

Stein, C.; Mangan, S. A. Soil biota increase the likelihood for coexistence among competing plant species. Ecology, v. 101, n. 11, e03147, 2020. https://doi.org/10.1002/ecy.3147.

Storck, L.; Ribeiro, N.; Cargnelutti Filho, A. Precisão experimental de ensaios de feijão analisada pelo método de Papadakis Experimental precision in common bean trials using. Pesquisa Agropecuária Brasileira, v.46, n.8, p. 798–804, 2011. https://doi.org/10.1590/S0100-204X2011000800003.

Tokatlidis, I. S. Crop adaptation to density to optimise grain yield: breeding implications. Euphytica, v. 213, n. 4, e92, 2017. https://doi.org/10.1007/s10681-017-1874-8.

Torga, P. P.; Melo, P. G. S.; Pereira, H. S.; Faria, L. C.; Melo, L. C. Genetic potential of black bean genotypes with predictable behaviors in multienvironment trials. Genetics and Molecular Research, v 15, n. 4, e5049029, 2016. https://doi.org/10.4238/gmr15049029.

Yang, R. C.; Juskiw, P. Analysis of covariance in agronomy and crop research. Canadian Journal of Plant Science, v. 91, n. 4, p. 621–641, 2011. https://doi.org/10.4141/cjps2010-032.

Zhao, Y.; Potgieter, A. B.; Zhang, M.; Wu, B.; Hammer, G.L. Predicting wheat yield at the field scale by combining high-resolution Sentinel-2 satellite imagery and crop modelling. Remote Sensing, v. 12, n. 6, e1024, 2020. https://doi.org/10.3390/rs12061024.