ARTICLE

Optimal plot size for the evaluation of phenotypic traits in soybean seedlings

Tiago de Avila Silva[1], Eder Matsuo[1*], Ithalo Coelho de Sousa[2], Silvana da Costa Ferreira[1], Paulo Roberto Cecon[3], Luiz Felipe Queiroz Noronha[1]

Crop Breeding and Applied Biotechnology
25(4): e52822545, 2025
Brazilian Society of Plant Breeding.
Printed in Brazil
http://dx.doi.org/10.1590/1984-70332025v25n4a50

Abstract: *Phenotypic traits with discriminatory capacity have been studied in recent years. However, scientific studies often employ experimental plots of varying sizes. Therefore, the objective of this study was to estimate the optimal plot size for the evaluation of phenotypic traits in soybean seedlings. Two blank trials were conducted, each consisting of 120 basic experimental units. Each basic experimental unit consisted of a single plant grown in a pot, which was evaluated during the juvenile stage. Coefficients of variation and optimal plot sizes were estimated using a segmented linear model with plateau response, a segmented quadratic model with plateau response, and the modified maximum curvature method. Model fit quality was assessed based on the adjusted coefficient of determination, Akaike information criterion (AIC), and Bayesian information criterion (BIC). Plot size recommendations were obtained, but there are variations in the optimal plot size for the different traits evaluated in two soybean cultivars.*

Keywords: Glycine max*, uniformity trial, basic experimental units, experimental precision*

INTRODUCTION

Soybean (*Glycine max* L. Merr.) was cultivated on 47.4 million hectares during the 2024/2025 growing season, resulting in an average grain yield of 3,527 kg ha^{-1} and a total production of 167.3 million tons of grain (CONAB 2025). According to Oda et al. (2015), part of this remarkable success is attributed to genetic improvement programs conducted by various research institutions and Brazilian universities. In the breeding process, the genetic material undergoes several stages, and after selecting the line to be released as a cultivar, it must be registered and protected with the Ministry of Agriculture, Livestock and Food Supply (MAPA) (Matsuo et al. 2021). To be granted protection, a candidate cultivar must meet a requirement; it must be distinct, uniform, and stable (DUS) (Viana 2013, MAPA 2017). A distinct cultivar is one that is clearly distinguishable from any other whose existence is officially recognized at the time of the protection request (Ferraz and Campos et al. 2009).

In the field of agricultural experimentation, there is often interest in determining the size, shape, and number of replications of experimental plots. This is because these factors represent key issues that must be addressed to reduce experimental error arising from plot heterogeneity, thereby maximizing the information obtained from an experiment (Steel and Torrie 1960).

***Corresponding author:**
E-mail: edermatsuo@ufv.br

Received: 14 August 2025
Accepted: 08 September 2025
Published: 29 October 2025

[1] Universidade Federal de Viçosa, Rodovia BR 230, km 7, s/n, Campus Rio Paranaíba, 38810-000, Rio Paranaíba, MG, Brazil
[2] Universidade Federal de Rondônia, Rua Rio Amazonas, 351, Jardim dos Migrantes, Campus Ji-Paraná, 76900-726, Ji-Paraná, RO, Brazil
[3] Universidade Federal de Viçosa, Avenida P. H. Rolfs, s/n, Campus Universitário, 36570-900, Viçosa, MG, Brazil

Storck and Uitdewilligen (1980) reported that one of the basic problems for researchers is determining the size, shape, and number of replicates of experimental plots, aiming to increase experimental efficiency by reducing experimental error. From a statistical point of view, the best plot size will be the one that provides the least accidental variation, if it does not affect the precision of the experiment (Ferreira 2000). Plot size and shape should be determined for each crop and each location where climatic and soil conditions differ from those previously determined (Oliveira and Estefanel 1995).

Experimental error, which consists of the variation between plots that received the same treatment, must be minimized so that smaller differences between treatment means are identified as significant, that is, they are not attributed to chance variations (Pimentel-Gomes 2009, Storck et al. 2016). This minimization is due to the fact that the coefficient of variation reduces gradually, and in a non-linear manner, with the increase in plot size (Cargnelutti Filho et al. 2023a, Cargnelutti Filho et al. 2023b).

Most methods used to estimate the optimal size of experimental plots are based on blank trials, also known as uniformity trials, in which the entire experimental area is planted with a single cultivar using identical cultivation practices and without the application of treatments (Storck et al. 2000). Moreover, by using an appropriate plot size and genetic material with lower variability, it is possible to minimize experimental error and, consequently, increase the precision of inferences (Cargnelutti Filho et al. 2011). Furthermore, the use of six plants per plot in the assessment of anthracnose severity would allow a 33% reduction in the cost of experiments of this nature and would make it possible to assess a greater number of genotypes (Mencalha et al. 2024).

Phenotypic traits have been emphasized in identifying differences between cultivars (Nogueira et al. 2008, Alves et al. 2019, Hanyu et al. 2020). However, in experimental design, especially concerning the size of the experimental unit, different plot sizes have been employed, meaning that the basic unit has been composed of varying numbers of plants. No studies in the literature reported plot size for traits measured in the juvenile stage of soybean. This is important for directing future work. Thus, the objective was to estimate the optimal plot size for different soybean seedling traits.

MATERIAL AND METHODS

Two uniformity trials were set up in a greenhouse and conducted at the Federal University of Viçosa, Rio Paranaíba *Campus* (lat 19° 11' 39'' S, long 46° 14' 37'' W, and alt 1133 m asl). Seeds with a diameter of 6.0 mm were manually planted at a depth of 2 cm in a substrate (2/3 soil: 1/3 organic matter) placed in 3 dm^3 capacity pots. In trial A, seeds from the cultivar MG/BR 46 (Conquista) were used, with planting carried out in October and evaluations completed in November 2023. In trial B, seeds from the cultivar BRS752S were used, with planting conducted in November and evaluations completed in December 2023. Throughout the installation, conduction, and evaluation period of the Trial, daily temperatures were measured using a digital thermometer, recording average minimum and maximum temperatures of 18.6 ºC and 45.4 ºC (Trial A), and 18.2 ºC and 47.2 ºC (Trial B). The plants were managed according to Sediyama (2009).

Each plant grown in a pot was considered a basic unit (bu), so there were 120 basic units for each trial. Thus, from the blank trial map (120 pots distributed in 10 columns and 12 rows) the plants duly identified in the rows and columns, numbered and the basic units could be grouped in 15 different ways, as described in Table 1.

The response variables were epicotyl length (EL), measured on the main stem as the distance between the cotyledonary node and the node of the unifoliate leaves; internode length (IL), measured on the main stem as the distance between the node of the unifoliate leaves and the node of the first trifoliate leaf; petiole length of the first trifoliate leaf (PL); and seedling height (SH). Measurements were taken using a millimeter ruler, with values expressed in centimeters, at the V2, V3, and V4 developmental stages, according to Fehr and Caviness (1977) (herein referred to as V2, V3, and V4, respectively).

The coefficients of variation, $CV_{(x)}$ in %, were estimated in the blank trials based on the different plot sizes and calculated as follows:

$$CV_{(x)\%} = \left(\frac{s_x}{\overline{x}}\right) \times 100$$

Where: $CV_{(x)\%}$ = Coefficient of variation in %, $\overline{x}$ = Mean of the measured characteristic in plots composed of *n* basic units, and s_x = Sample standard deviation measured in plots composed of *n* basic units.

Table 1. Grouping of 120 basic units

Number of plants (basic units)	**Grouping shape**
120	1 rectangular plot of 10 columns and 12 rows
60	2 rectangular plots of 10 columns and 6 rows
40	3 rectangular plots of 10 columns and 4 rows
30	4 rectangular plots of 10 columns and 3 rows
24	5 rectangular plots of 2 columns and 12 rows
20	6 rectangular plots of 10 columns and 2 rows
15	8 rectangular plots with 5 columns and 3 rows
12	10 rectangular plots with 2 columns and 6 rows
10	12 rectangular plots with 10 columns and 1 row
8	15 rectangular plots with 2 columns and 4 rows
6	20 rectangular plots of 2 columns and 3 rows
5	24 rectangular plots of 5 columns and 1 row
4	30 square plots of 2 columns and 2 rows
3	40 rectangular plots of 1 column and 3 rows
2	60 rectangular plots of 1 column and 2 rows

The optimal plot sizes were estimated using three models: the Segmented Linear Model with Plateau Response (SLMPR), the Segmented Quadratic Model with Plateau Response (SQMPR), and the Modified Maximum Curvature Method (MMCM), as described below:

Segmented Linear Model with Plateau Response (SLMPR)

This model consists of two segments: the first represents an increasing or decreasing straight line (depending on the value of β_1 in the equation presented below) up to a certain constant *P*, which is the value of Y at the point corresponding to the plateau. The second segment refers to the plateau, where Y assumes a constant value from that point onward (Ferreira 2006). In the present study, the theory of linear plateau models was applied in the context of determining the dimensionality of optimal plot sizes. For this purpose, the following model was adopted:

$$CV_{(x)} = \begin{cases} \beta_0 + \beta_1 X_i + \varepsilon_i & if\ X \le X_0 \\ CVP\% + \varepsilon_i & if\ X > X_0 \end{cases}$$

Where: $CV_{(x)}$ is the coefficient of variation among the totals of plots of size X, X_1 is the size of the grouped basic experimental units (bu), X_0 is the optimal plot size at which the linear model transitions to a plateau, in relation to the x-axis, $CVP_{\%}$ is the coefficient of variation, in %, at the point corresponding to the plateau (at the junction of the segments), β_0 and β_1 is the coefficient of variation, in %, at the point corresponding to the plateau (at the junction of the segments), and ε_1 is the random error associated with the corresponding $CV_{(x)}$, assumed to be independent and normally distributed with zero mean and constant variance (σ^2_ε).

Segmented Quadratic Model with Plateau Response (SQMPR)

The quadratic regression model with plateau response was adapted from Silva et al. (2012) and is defined as follows:

$$CV_{(x)} = \begin{cases} \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \varepsilon_i & if\ X \le X_0 \\ CVP\% + \varepsilon_i & if\ X > X_0 \end{cases}$$

Where $CV_{(x),\%}$ is the coefficient of variation among the totals of plots of size X, in %, X_1 is the size of the grouped basic experimental units (bu), X_0 is the optimal plot size at which the quadratic model transitions to a plateau, in relation to the x-axis, $CVP_{\%}$ is the coefficient of variation, in %, at the point corresponding to the plateau (at the junction of the segments), β_0 , β_1 and β_2 are parameters to be estimated for the first segment (quadratic), and ε_1 is the random error associated with the corresponding $CV_{(x)}$, assumed to be independent and normally distributed with zero mean and constant variance (σ^2_ε).

Modified Maximum Curvature Method (MMCM)

The plot size was estimated using the Modified Maximum Curvature Method, proposed by Lessman and Atkins (1963). According to this method, the relationship between the coefficient of variation in % ($CV_{(x),\%}$) and the plot size with X basic units is described by the following model:

$$CV_{(x),\%} = \beta_0 X^{-\beta 1}$$

Where: β_0 and β_1 are the parameters to be estimated.

Based on the curvature function defined by this model, the value of the abscissa at the point of maximum curvature, X_0, was determined using the equation below. This value corresponds to the estimated optimal experimental plot size (Meier and Lessman 1971).

$$X_0 = \left[\frac{\beta_0^2 \beta_1^2 (2\beta_1 + 1)}{\beta_1 + 2}\right]^{\left(\frac{1}{2\beta_1+2}\right)}$$

Where X_0 is the abscissa value corresponding to the point of maximum curvature, and β_0 and β_1 are the estimates of the parameters β_0 and β_1 of $CV_{(x),\%}$, obtained through nonlinear regression methods, as described by Bates and Watts (1988).

The evaluation of the model fit quality was performed based on the adjusted coefficient of determination (Silva 2023), the Akaike Information Criterion (AIC) (Akaike 1974), and the Bayesian Information Criterion (BIC) (Schwarz 1978), as described below:

Adjusted Coefficient of Determination (R^2_{Adj})

The adjusted coefficient of determination can be used for model comparison (Silva 2003). The author presents the following equation:

$$R^2_{Adj} = \frac{R^2(n-1) - p + 1}{n - p}$$

Where: R^2_{Adj} is the adjusted coefficient of determination, n is the number of observations, and p is the number of model parameters.

Akaike Information Criterion (AIC)

The AIC criterion allows the application of the principle of parsimony in model selection. In other words, according to this criterion, the most parameterized model is not necessarily the best. Lower AIC values indicate a better fit (Akaike 1974). Akaike (1974) defined his information criterion as:

$$AIC = -2logL(\hat{\theta}) + 2(k)$$

Where: $logL(\hat{\theta})$ is the maximized log-likelihood function and k is the number of parameters.

Bayesian Information Criterion (BIC)

Just like the AIC, the BIC criterion takes into account the degree of parameterization of the model. Similarly, the lower the BIC value, the better the fit of the model (Schwarz 1978):

$$BIC = k\ log(n) - 2logL(\hat{\theta})$$

Where: k is the number of parameters to be estimated, n is the number of observations in the sample, and $L(\hat{\theta})$ is the likelihood of the fitted model.

Subsequently, the confidence interval for the mean was estimated at a 95% confidence level, considering the optimal plot size values for all variables and models, in 3 distinct conditions: optimal plot size data for trial A, for trial B, and all optimal plot size values. The data were analyzed using the R software (R Core Team 2023), in R, through the packages "easynls" (Arnhold 2017), "easyreg" (Arnhold 2018), "biotools" (Silva et al. 2017, Silva 2021) and "soilphysics " (Silva and Lima 2015, 2017, Lima et al. 2016, Lima et al. 2021).

RESULTS AND DISCUSSION

In Experiment A, which used the cultivar MG/BR 46 (Conquista), the evaluation of epicotyl length at the V2 stage showed that the segmented linear model with plateau response (SLMPR) and the Modified Maximum Curvature Method

(MMCM) had significant parameters according to the t-test ($p < 0.05$). At the V3 stage, the methods with significant parameters were SLMPR and MMCM, and at the V4 stage, the methods with significant parameters were SLMPR, the Segmented Quadratic Model with Plateau Response (SQMPR), and MMCM (Table 2).

In the analysis of the models, it was found that all models, in both experiments, had adjusted coefficient of determination values greater than 0.75. Furthermore, when analyzing AIC and BIC within each trait and within each experiment, it was observed that the estimated values for MMCM and SQMPR were lower than those for SLMPR. For the cultivar MG/BR 46 (Conquista), the optimal experimental plot sizes at the V2 stage were 12.9 and 3.1 plants, respectively for SLMPR and MMCM. At the V3 stage, the optimal plot sizes were 13.0 and 3.4 plants, respectively for SLMPR and MMCM. Finally, at the V4 stage, the optimal experimental plot sizes were 10.8, 15.2, and 6.8 plants, respectively for SLMPR, SQMPR, and MMCM.

For the evaluations of epicotyl length conducted with the cultivar BRS752S, used in Experiment B, the methods for analyzing the coefficients of variation at the V2 and V4 stages that showed significant parameters according to the t-test ($p < 0.05$) were SLMPR, SQMPR, and MMCM. In contrast, for the evaluation conducted at the V3 stage, only the SLMPR and MMCM methods showed significant parameters (Table 2). At the V2 stage, the optimal experimental plot sizes were 8.5, 10.5, and 6.2 plants, respectively for SLMPR, SQMPR, and MMCM. For the evaluations conducted at the V3 stage, the optimal plot sizes were 8.2 and 6.1 plants, respectively for SLMPR and MMCM. The results for the V4 stage indicated that the optimal experimental plot sizes were 10.2, 12.5, and 4.8 plants, respectively for SLMPR, SQMPR, and MMCM.

For internode length (Table 3), in the cultivar MG/BR 46 (Conquista), all three models used (SLMPR, SQMPR, and MMCM) showed significant parameters according to the t-test ($p < 0.05$) across all three evaluation stages. The optimal plot sizes, considering the best-fitting model, were 8.6 (by MMCM), 7.4 (by SQMPR), and 7.6 (by SQMPR), respectively for the evaluations at V2, V3, and V4 in the experiment with cultivar MG/BR 46 (Conquista). In Trial B, with the cultivar BRS752S, the model parameters were significant, except for SQMPR at the V2 stage. The optimal plot sizes were 7.5 (by MMCM), 9.3 (by SQMPR), and 8.9 (by SQMPR), respectively for the evaluations at V2, V3, and V4.

Table 2. Estimates of the parameters (β_0, β_1 and β_2), adjusted coefficient of determination (R^2_{Adj}), Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), optimal plot size (X_0) and coefficient of variation in % at the plateau point ($CVP_{\%}$) through three estimation models (SLMPR, SQMPR and MMCM) for epicotyl length (EL) evaluated at three stages (V2, V3 and V4) and in two distinct trials (A and B)[1]

Trial A (MG/BR 46 (Conquista))									
	Models[2]	β_0	β_1	β_2	R^2_{Adj}	**AIC**	**BIC**	X_0	$CVP_{\%}$
EL_V2	SLMPR	14.8500**	-0.7208**		0.70	63.04	65.60	12.9	14.8
	SQMPR	26.9089**	-1.6353*	0.0565	0.75	60.48	63.04	14.5	15.1
	MMCM	26.8018**	-0.1827**		0.76	59.11	61.03	3.1	21.7
EL_V3	SLMPR	14.4167**	-0.7746**		0.74	62.55	65.11	13.0	14.4
	SQMPR	29.6923**	-2.5703*	0.1125	0.81	58.68	61.23	11.4	15.0
	MMCM	27.5634**	-0.2003**		0.80	58.35	60.27	3.4	21.5
EL_V4	SLMPR	17.3000**	-2.0946**		0.88	70.39	72.94	10.8	17.3
	SQMPR	43.1384**	-3.4380**	0.1133*	0.90	67.65	70.21	15.2	17.0
	MMCM	46.5862**	-0.3174**		0.91	65.33	67.25	6.8	25.3
Trial B (BRS752S)									
EL_V2	SLMPR	16.5250**	-2.6186**		0.86	70.37	72.93	8.5	16.5
	SQMPR	45.2101**	-5.4652**	0.2612*	0.89	65.96	68.52	10.5	16.6
	MMCM	41.7635**	-0.3114**		0.88	67.37	69.28	6.2	23.6
EL_V3	SLMPR	16.3625**	-2.6786**		0.75	78.58	81.14	8.2	16.3
	SQMPR	48.6769**	-7.2333*	0.4089	0.81	74.81	77.36	8.8	16.7
	MMCM	41.0752**	-0.3129**		0.79	75.33	77.25	6.1	23.3
EL_V4	SLMPR	15.7143**	-1.5269**		0.85	63.96	66.52	10.2	15.7
	SQMPR	35.3079**	-3.1194**	0.1247*	0.88	60.56	63.12	12.5	15.8
	MMCM	34.4888**	-0.2549**		0.89	58.28	60.20	4.8	23.0

[1] ** and * significant at 1% and 5% probability, respectively, by the t-test; [2] Models: SLMPR = Segmented Linear Model with Plateau Response, SQMPR = Segmented Quadratic Model with Plateau Response and MMCM = Modified Maximum Curvature Method.

Table 3. Estimates of the parameters (β_0, β_1 and β_2), adjusted coefficient of determination (R^2_{Adj}), Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), optimal plot size (X_0) and coefficient of variation in % at the plateau point ($CVP_{\%}$) through three estimation models (SLMPR, SQMPR and MMCM) for internode length (IL) evaluated at three stages (V2, V3 and V4) and in two distinct trials (A and B)[1]

Trial A (MG/BR 46 (Conquista))									
	Models[2]	β_0	β_1	β_2	R^2_{Adj}	**AIC**	**BIC**	X_0	$CVP_{\%}$
	SLMPR	26.7125**	-4.0086**		0.85	83.58	86.13	8.6	26.7
IL_V2	SQMPR	68.8271**	-7.4961**	0.3333*	0.87	80.80	83.36	11.2	26.6
	MMCM	65.6820**	-0.3010**		0.88	79.71	81.62	8.6	34.3
	SLMPR	20.5100**	-5.3000**		0.79	78.25	80.81	5.8	20.5
IL_V3	SQMPR	61.4084**	-11.0809**	0.7484*	0.84	74.75	77.31	7.4	20.4
	MMCM	45.6352**	-0.2858**		0.79	77.57	79.49	6.4	26.8
	SLMPR	21.0667**	-4.6000**		0.85	72.58	75.14	6.3	21.0
IL_V4	SQMPR	60.4275**	-10.3789**	0.6858*	0.88	69.23	71.79	7.6	21.1
	MMCM	44.6580**	-0.2611**		0.76	78.08	79.99	6.0	27.9
Trial B (BRS752S)									
	SLMPR	20.4600**	-1.4514**		0.76	84.82	87.38	15.5	20.4
IL_V2	SQMPR	52.1030**	-4.2205**	0.1470	0.83	79.95	82.50	14.4	21.8
	MMCM	55.5665**	-0.3022**		0.90	72.12	74.04	7.6	30.1
	SLMPR	15.9111**	-2.2900**		0.86	59.84	62.40	7.4	15.9
IL_V3	SQMPR	37.5610**	-4.6342**	0.2485*	0.89	56.96	59.52	9.3	15.9
	MMCM	31.4482**	-0.2289**		0.77	65.70	67.61	4.2	22.6
	SLMPR	16.1222**	-2.8800**		0.87	62.77	65.33	7.0	16.1
IL_V4	SQMPR	42.2493**	-5.8675**	0.3302*	0.90	59.26	61.81	8.9	16.1
	MMCM	34.2629**	-0.2551**		0.75	71.14	73.06	4.8	22.9

[1] ** and * significant at 1% and 5% probability, respectively, by the t-test; [2] Models: SLMPR = Segmented Linear Model with Plateau Response, SQMPR = Segmented Quadratic Model with Plateau Response and MMCM = Modified Maximum Curvature Method.

Table 4. Estimates of the parameters (β_0, β_1 and β_2) adjusted coefficient of determination R^2_{Adj}), Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), optimal plot size (X_0) and coefficient of variation in % at the plateau point ($CVP_{\%}$) through three estimation models (SLMPR, SQMPR and MMCM) for petiole length (PL) evaluated at three stages (V2, V3 and V4) and in two distinct trials (A and B)[1]

Trial A (MG/BR 46 (Conquista))									
	Models[2]	β_0	β_1	β_2	R^2_{Adj}	**AIC**	**BIC**	X_0	$CVP_{\%}$
	SLMPR	24.0625**	-3.3800**		0.91	72.54	75.10	9.0	24.0
PL_V2	SQMPR	61.1125**	-6.2380**	0.2624**	0.94	68.21	70.77	11.9	24.0
	MMCM	59.0849**	-0.2961**		0.89	74.98	76.90	7.9	32.0
	SLMPR	13.6111**	-3.5600**		0.88	63.07	65.62	6.5	13.6
PL_V3	SQMPR	43.5028**	-7.1984**	0.4337*	0.91	58.82	61.37	8.3	13.6
	MMCM	35.4151**	-0.3357**		0.88	63.18	65.10	5.6	19.8
	SLMPR	14.5400**	-5.3600**		0.83	69.86	72.42	5.2	14.5
PL_V4	SQMPR	60.0377**	-15.7387**	1.3665*	0.88	64.40	66.95	5.8	14.7
	MMCM	35.0956**	-0.3208**		0.74	74.78	76.70	5.5	20.3
Trial B (BRS752S)									
	SLMPR	33.6444**	-9.4800**		0.87	91.47	94.02	6.3	33.6
PL_V2	SQMPR	112.6488**	-19.9292**	1.2585*	0.89	88.91	91.47	7.9	33.7
	MMCM	88.1406**	-0.3394**		0.81	95.39	97.31	11.2	38.8
	SLMPR	22.0400**	-7.3100**		0.88	76.61	79.17	5.6	22.0
PL_V3	SQMPR	74.3119**	-14.0762**	0.9432**	0.91	71.74	74.30	7.5	21.8
	MMCM	55.1907**	-0.3304**		0.81	82.22	84.14	7.8	27.9
	SLMPR	20.4778**	-3.7900**		0.77	76.48	79.03	6.6	20.4
PL_V4	SQMPR	57.0482**	-9.8041*	0.661	0.82	72.58	75.13	7.4	20.7
	MMCM	42.3368**	-0.2518**		0.78	74.95	76.87	5.6	27.3

[1] ** and * significant at 1% and 5% probability, respectively, by the t-test; [2] Models: SLMPR = Segmented Linear Model with Plateau Response, SQMPR = Segmented Quadratic Model with Plateau Response and MMCM = Modified Maximum Curvature Method.

For the petiole length of the first trifoliate leaf (Table 4) in Trial A, which used the cultivar MG/BR 46 (Conquista), all three models (SLMPR, SQMPR, and MMCM) showed significant parameters according to the t-test ($p < 0.05$) across the three evaluation stages.

Regarding the optimal plot size, the best-fitting model was SQMPR, which indicated 11.9, 8.3, and 5.8 as the optimal plot sizes for evaluations conducted at V2, V3, and V4, respectively. In Trial B (with cultivar BRS752S), the model parameters were significant, except for SQMPR at the V4 stage. The optimal plot sizes were 7.9 (by SQMPR), 7.5 (by SQMPR), and 5.6 (by MMCM), respectively for the evaluations at V2, V3, and V4.

For plant height, all three models used (SLMPR, SQMPR, and MMCM) showed significant parameters according to the t-test ($p < 0.05$) for the evaluations conducted at stages V2, V3, and V4 (Table 5). It was found that SQMPR was the best-fitting model across both Trials and all three developmental stages. The optimal plot sizes were 6.8 (V2), 12.8 (V3), and 6.3 (V4) for the cultivar MG/BR 46 (Conquista), and 7.9 (V2), 6.4 (V3), and 8.2 (V4) for the cultivar BRS752S.

The 95% confidence intervals, i.e., $CI(\mu)_{95\%}$ for μ of the optimal plot size were $6.94 \leq \mu \leq 9.07$ for Trial A, $6.55 \leq \mu \leq 8.54$ for Trial B, and $7.02 \leq \mu \leq 8.45$ for Trials A and B combined. The 95% confidence intervals indicate that there is a 95% level of confidence that this numerical range contains μ, meaning that 95% of such intervals constructed in the same way would include μ within their bounds (Cecon et al. 2012).

In a study with four soybean cultivars, Matsuo et al. (2012) reported that the optimal sample size for epicotyl length, evaluated at the V3 stage, ranged from 58 to 73 for the mean, and from 29 to 70 for the variance. According to the authors, the optimal sample size for the mean showed little variation among the four genotypes studied, meaning that the sample sizes for epicotyl length differed by only 8 plants among the genotypes. However, for the optimal minimum number required to represent the variance, variations of up to 35 plants were observed among genotypes for epicotyl length. Accordingly, the optimal minimum sample sizes of plants to be measured in order to represent the mean

Table 5. Estimates of the parameters (β_0, β_1 and β_2), adjusted coefficient of determination (R^2_{Adj}), Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), optimal plot size (X_0) and coefficient of variation in % at the plateau point ($CVP_{\%}$) through three estimation models (SLMPR, SQMPR and MMCM) for seedling height (SH) evaluated at three stages (V2, V3 and V4) and in two distinct trials (A and B)[1]

Trial A (MG/BR 46 (Conquista))									
	Models[2]	β_0	β_1	β_2	R^2_{Adj}	**AIC**	**BIC**	X_0	$CVP_{\%}$
SH_V2	SLMPR	18.1300**	-8.0100**		0.92	67.45	70.00	5.2	18.1
	SQMPR	70.8502**	-15.4706**	1.1299**	0.94	63.94	66.50	6.8	17.9
	MMCM	48.6294**	-0.3613**		0.77	82.53	84.45	7.3	23.6
SH_V3	SLMPR	15.3286**	-1.2880**		0.86	57.85	60.41	10.4	15.3
	SQMPR	32.0209**	-2.5923**	0.1010*	0.90	53.31	55.87	12.8	15.3
	MMCM	31.0731**	-0.2274**		0.87	56.29	58.21	4.1	22.4
SH_V4	SLMPR	15.5900**	-6.0200**		0.87	66.86	69.42	5.1	15.5
	SQMPR	58.4805**	-13.5351**	1.0674*	0.89	64.49	67.05	6.3	15.5
	MMCM	37.1653**	-0.3154**		0.72	77.21	79.13	5.7	21.4
Trial B (BRS752S)									
SH_V2	SLMPR	14.1111**	-5.3900**		0.91	70.24	72.80	6.5	14.1
	SQMPR	60.6636**	-11.7116**	0.7385**	0.94	63.57	66.13	7.9	14.2
	MMCM	50.0245**	-0.4537**		0.87	74.23	76.14	7.9	19.6
SH_V3	SLMPR	12.5182**	-3.0500**		0.83	48.90	51.45	4.8	12.5
	SQMPR	30.9885**	-5.7860**	0.4508*	0.85	46.96	49.52	6.4	12.4
	MMCM	20.5786**	-0.1771**		0.67	56.94	58.86	2.5	17.5
SH_V4	SLMPR	15.1778**	-3.0200**		0.86	61.34	63.89	6.5	15.1
	SQMPR	40.7504**	-6.1906**	0.3752*	0.89	58.60	61.16	8.2	15.2
	MMCM	31.5546**	-0.2504**		0.73	69.79	71.71	4.4	21.7

[1] ** and * significant at 1% and 5% probability, respectively, by the t-test; [2] Models: SLMPR = Segmented Linear Model with Plateau Response, SQMPR = Segmented Quadratic Model with Plateau Response and MMCM = Modified Maximum Curvature Method.

epicotyl length were 49, 43, 51, and 49 for the cultivars BRS Valiosa RR, Água-Marinha RR, UFVS 2010, and NK 7059 RR, respectively. For representing the variance of epicotyl length, the optimal minimum sample sizes were 70, 29, 64, and 70 for BRS Valiosa RR, Água-Marinha RR, UFVS 2010, and NK 7059 RR, respectively (Matsuo et al. 2012).

By fixing the evaluation stage of the traits, it was observed that the optimal plot size varied according to the values obtained in the two Trials (A and B). This may have occurred because different cultivars were used in each Trial, i.e., in Trial A, seeds from MG/BR 46 (Conquista) were used, while in Trial B, seeds from the cultivar BRS752S were used. According to Cargnelutti Filho et al. (2009), there is a variability in the sample size estimate among genotypes in relation to the number of nodes per plant and among experiments regarding plant height at maturation, number of nodes per plant, number of branches per plant, and number of pods per plant. Specifically, regarding phenotypic traits evaluated early in seedling development, Camargos et al. (2019) reported that soybean cultivars exhibit distinguishability for hypocotyl and epicotyl length when evaluated at stages V2 and V3. These results are like those found by Matsuo et al. (2012), who reported the existence of genetic variability among genotypes, which enables successful selection of promising materials. Additionally, the magnitude of genetic parameter estimates characterized the influence of genetic components on the phenotypic expression of hypocotyl and epicotyl length in different soybean genotypes (Matsuo et al. 2012). Furthermore, soybean genotypes differ in hypocotyl length, epicotyl length, and plant height, evaluated at stages V2 and V3, and the magnitude of the estimated genetic parameters (genotypic determination coefficient and CVg/CVe ratio) demonstrated the genetic influence on the phenotypic expression of hypocotyl length, epicotyl length, and plant height at stages V2 and V3 (Hanyu et al. 2020).

The use of the segmented linear model with plateau response and the modified maximum curvature method allows us to conclude that it is possible to recommend the optimal experimental plot size for arabica coffee, based on the tested data and for the edaphoclimatic conditions and varieties used (Brioschi Junior et al. 2020). The optimal experimental plot size is specific to each situation, crop, and cultivar and should always be estimated when the conditions differ from those already evaluated (González et al. 2022). Moreover, the comparison of results should be analyzed with caution due to the different methods used to determine the plot size, environmental differences, the distinct management of uniformity trials, and the different cultivars (Cargnelutti Filho et al. 2023c).

CONCLUSIONS

For the cultivar MG/BR 46 (Conquista), in a greenhouse, the optimal plot size for epicotyl length was 3.1, 3.4, and 6.8 plants; for internode length, it was 8.6, 7.4, and 7.6 plants; for petiole length, it was 11.9, 8.3, and 5.8 plants; and for plant height, it was 6.8, 12.8, and 6.3 plants, respectively, for the measurements taken at the development stages V2, V3, and V4, considering the best-fitting model used.

For the cultivar BRS752S, in a greenhouse, the optimal plot size for epicotyl length was 10.5, 8.8, and 4.8 plants; for internode length, it was 7.6, 9.3, and 8.9 plants; for petiole length, it was 7.9, 7.5, and 5.6 plants; and for plant height, it was 7.9, 6.4, and 8.2 plants, respectively, for the measurements taken at the development stages V2, V3, and V4, considering the best-fitting model used.

ACKNOWLEDGEMENTS

This work was carried out with the support of CNPq, which granted a scientific initiation scholarship to the author L.F.Q.N. (PIBIC UFV / CNPq 2024 2025) and with the support of Fapemig, which granted a scientific initiation scholarship to the author T.Á.S. (PIBIC UFV / FAPEMIG 2023-2024). In addition, the authors thank Prof. Sebastião Martins Filho for his contributions.

DATA AVAILABILITY

The datasets generated and/or analyzed during the current research are available from the corresponding author upon reasonable request.

REFERENCES

Akaike H (1974) A new look at the statistical model identification. **IEEE Transactions on Automatic Control 19**: 716-723.

Alves GF, Nogueira JPG, Machado-Junior R, Ferreira SC, Nascimento M and Matsuo E (2019) Stability of the hypocotyl length of soybean cultivars using neural networks and traditional methods. **Ciência Rural 49**: e20180300.

Arnhold E (2017) **easynls**: Easy nonlinear model. R package version 5.0. Available at <https://CRAN.R-project.org/package=easynls>.

Arnhold E (2018) R-environment package for regression analysis. **Pesquisa Agropecuária Brasileira 53**: 870-873.

Bates DM and Watts D (1988) **Nonlinear regression analysis and its applications**. John Wiley, New York, 365p.

Brioschi Junior D, Guarçoni RC, Alixandre FT, Pereira LL, Sousa DG, Marcate JPP, Favarato LF, Sousa LHBP, Fornazier MJ and Filete CA (2020) Optimal size of experimental plot to evaluate physico-chemical characteristics of arabic coffee. **Revista Ifes Ciência 6**: 3-11.

Camargos T, Campos N, Alves GF, Ferreira SC and Matsuo E (2019) The effect of soil volume, plant density and sowing depth on soybean seedlings characters. **Agronomy Science and Biotechnology 5**: 47-58.

Cargnelutti Filho A, Bubans VE, Somavilla FM, Costa SL, Dumke GE and Osmari LF (2023c) Optimal plot size in wheat with comparison of three methods. **Revista Ciencia Agronômica 54**: e20218267.

Cargnelutti Filho A, Evangelista DHR, Gonçalves ECP and Storck L (2009) Sample size of the characters in soybeans genotypes. **Ciência Rural 39**: 983-991.

Cargnelutti Filho A, Loro MV, Ortiz VM, Andretta JA, Reis MB, Schuller BR, Denardin Junior JC and Konrad M (2023a) Optimal plot size in chickpea comparing three estimation methods. **Journal of Environmental Analysis and Progress 8**: 307-318.

Cargnelutti Filho A, Silveira DL, Bubans VE, Loro MV, Somavilla FM and Ortiz VM (2023b) Optimal plot size in white oat with comparison of three methods. **Semina-Ciencias Agrarias 44**: 171-184.

Cargnelutti Filho A, Toebe M, Burin C, Casarotto G and Lúcio ADC (2011) Methods for estimating the optimum experimental plot size of corn hybrids single, triple and double. **Ciência Rural 41**: 1509-1516.

Cecon PR, Silva AR, Nascimento M and Ferreira A (2012) **Métodos estatísticos**. Editora UFV, Viçosa, 229p.

CONAB - Companhia Nacional de Abastecimento (2025) **Monitoring the Brazilian grain harvest. 2024/25 harvest – sixth survey**. Available at <https://www.conab.gov.br>. Accessed on March 25, 2025.

cSchwarz GE (1978) Estimating the dimension of a model. **Annals of Statistics 6**: 461-464.

Fehr WR and Caviness CE (1977) **Stages of soybean development**. Iowa State University of Science and Technology, Ames, 12p. (Information Bulletin, 80).

Ferraz e Campos SR, Machado VLS, Viana AAN and Azevedo ZMM (2009) Registro e proteção de cultivares. In Sediyama T (ed) **Tecnologias de produção e usos da soja**. Mecenas, Londrina, p. 235-245.

Ferreira DF (2006) **Uso de recursos computacionais**. UFLA, Lavras, 123p.

Ferreira PV (2000) **Estatística experimental aplicada à agronomia**. EDUFAL, Maceió, 419p.

González GGH, Mendoza CC and Humada BG (2022) Method of maximum curvature of the coefficient of variation in determining the optimal plot size in a field experiment with cassava. **Brazilian Journal of Animal and Environmental Research 5**: 3992-3998.

Hanyu J, Ferreira SC, Cecon PR and Matsuo E (2020) Genetic parameters estimate and characters analysis in phenotypic phase of soybean during two evaluation periods. **Agronomy Science and Biotechnology 6**: 1-12.

Lessman KJ and Atkins RE (1963) Optimum plot size and relative efficiency of lattice designs for grain sorghum yield tests. **Crop Science 3**: 477-481.

Lima RP, Silva AR and Silva A (2021) soilphysics: An R package for simulation of soil compaction induced by agricultural field traffic. **Soil and Tillage Research 206**: 104824.

Lima RP, Silva AR, Silva AP, Leao TP and Mosaddeghi MR (2016) soilphysics: an R package for calculating soil water availability to plants by different soil physical indices. **Computers and Eletronics in Agriculture 120**: 63-71.

MAPA - Ministério da Agricultura e Pecuária (2017) **Forms for plant variety protection**. Available at <https://www.gov.br/agricultura/pt-br/assuntos/insumos-agropecuarios/insumos-agricolas/protecao-de-cultivar/formularios-para-protecao-de-cultivares>. Accessed on August 14, 2025.

Matsuo E, Borém A and Sediyama T (2021) Desenvolvimento de cultivares. In Sediyama T, Matsuo E and Borém A (eds) **Melhoramento da soja no Brasil**. Mecenas, Londrina, p. 93-102.

Matsuo E, Sediyama T, Cruz CD, Oliveira RCT and Cadore LR (2012) Estimates of the genetic parameters, optimum sample size and conversion of quantitative data in multiple categories for soybean genotypes. **Acta Scientiarum Agronomy 34**: 265-273.

Meier VD and Lessman KJ (1971) Estimation of optimum field plot shape and size for testing yield in *Crambe abyssinica* Hochst. **Crop Science 11**: 648-650.

Mencalha J, Dias MA, Souza EA, Carneiro VQ, Pereira FAC, Carneiro PCS and Carneiro JES (2024) Optimization of the assessment of anthracnose severity in artificially inoculated common beans. **Crop Breeding and Applied Biotechnology 24**: e47552426.

Nogueira APO, Sediyama T, Cruz CD, Reis MS, Pereira DG and Jangarelli M (2008) New traits for differentiation of soybeans using discriminant analysis. **Ciência Rural 38**: 2427-2433.

Oda MC, Sediyama T, Matsuo E, Cruz, CD, Barros EG and Ferreira MFS (2015) Phenotypic and molecular traits diversity in soybean

launched in forty years of genetic breeding. **Agronomy Science and Biotechnology 1**: 1-9.

Oliveira PH and Estefanel V (1995) Tamanho e forma ótimos da parcela para avaliação de rendimento em experimentos com batata. **Ciência Rural 25**: 205-208.

Pimentel-Gomes F (2009) **Curso de estatística experimental**. FEALQ, Piracicaba, 451p.

R Core Team (2023) **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna. Available at <https://www.R-project.org/>.

Sediyama T (2009) **Tecnologias de produção e usos da soja**. Editora Mecenas, Londrina, 314p.

Silva AR (2021) **biotools**: Tools for biometry and applied statistics in agricultural science. R package version 4.2. Available at <https://cran.r-project.org/package=biotools>.

Silva AR (2023) **Estatística decodificada**. Blucher, São Paulo, 418p.

Silva AR and Lima RP (2015) soilphysics: an R package to determine soil preconsolidation pressure. **Computers and Geosciences 84**: 54-60.

Silva AR and Lima RP (2017) Determination of maximum curvature point with the R package soilphysics. **International Journal of Current Research 9**: 45241-45245.

Silva AR, Malafaia G and Menezes IPP (2017) Biotools: an R function to predict spatial gene diversity via an individual-based approach. **Genetics and Molecular Research 16**: 1-6.

Silva LFO, Campos KA, Morais AR, Cogo FD and Zambon CR (2012) Tamanho ótimo de parcela para experimentos com rabanetes. **Revista Ceres 59**: 624-629.

Steel RGD and Torrie JH (1960) **Principles and procedures of statistics**. McGraw-Hill, New York, 841p.

Storck L and Uitdewilligen WPM (1980) Estimativa para tamanho e forma de parcelas e número de repetições para experimentos com milho. **Agronomia Sulriograndense 16**: 269-282.

Storck L, Garcia DC, Lopes SJ and Estefanel V (2000) **Experimental vegetal**. UFSM, Santa Maria, 198p.

Storck L, Garcia DC, Lopes SJ and Estefanel V (2016) **Experimentação vegetal**. UFSM, Santa Maria, 200p.

Viana AAN (2013) Proteção de cultivares e comercialização de sementes. In Sediyama T (ed) **Tecnologia de produção de sementes de soja**. Mecenas, Londrina, p. 345-352.