v. 4, n. 2, 2024

Arts, Linguistics, Literature and Language
Research Journal

EVALUATION OF THE PAPADAKIS TECHNIQUE APPLIED TO METHODS FOR ESTIMATING THE SIZE OF EXPERIMENTAL PLOTS IN SOYBEAN CROPS

Joaquim Francisco Mazunga
Universidade Federal de Lavras, (UFLA), Lavras-MG
https://lattes.cnpq.br/2657250678869799

Augusto Ramalho de Morais
Universidade Federal de Lavras, (UFLA), Lavras-MG
https://lattes.cnpq.br/8214476768978726

Renato Ribeiro de Lima
Universidade Federal de Lavras, (UFLA), Lavras-MG
https://lattes.cnpq.br/7486689110603841

Abstract: The experimental plot size must be adequate to reduce experimental error and increase precision, hence there is a need to have more efficient experiments. And, to determine the sizes of experimental plots, many methods are found in the literature that present adequate results, but some methods have presented unsatisfactory results, with excessively small or large sizes. To overcome this problem, an alternative is to use the Papadakis technique. Therefore, the objective of this work is to evaluate the use of the Papadakis method for estimating the size of experimental plots for application in soybean cultivation. To estimate the size of plots, the modified maximum curvature method, the segmented linear model method with plateau and the maximum curvature coefficient of variation method were used considering the original data and the data adjusted by the Papadakis method. It was found that the use of the Papadakis method provides a good fit in estimating plot sizes of experimental plots using estimates from the modified maximum curvature methods and the maximum curvature coefficient of variation method.
Keywords: Experimental planning, experimental precision, moving average, maximum curvature, data fit.

INTRODUCTION

Experimental plot is the unit in which the treatment is randomly applied in order to provide experimental data that must reflect its effect. It is also considered the smallest portion of the experimental material on which treatments are evaluated (STORCK; BISOGNIN; OLIVEIRA, 2006). In experimental planning, one of the most important aspects to consider is the definition of the experimental unit or plot. The choice is made to minimize experimental error so that the plot is as uniform as possible, so that it reflects the effect of the applied treatments and has good precision over the applied area.

To avoid errors due to heterogeneity, the variability of the experimental units in the area in question must be known, using results from previous research. Steel, Torres and Dickey (1997) presented three ways to control experimental error, thus avoiding erroneous conclusions about the effect of treatments. The first is through the experimental design, which consists of planning the test aiming to control natural variation. The second way is based on the use of concomitant observations, with which covariance analysis is performed in trials with fixed-effect treatments in which the dependent variable is affected by one or more independent variables. The third is related to the size and shape of the plots, as, in general, the smallest plot size compatible with the treatments and the greatest number of replications possible must be used, in restricted areas (CARGNELUTTI-FILHO et al., 2015).

When determining plot size, it is common to use a uniformity test. The uniformity test consists of that experiment in which there is only a single genetic material throughout the experimental area, subjecting the entire area to identical cultivation practices, without the use of treatments. To determine the optimal plot size, determination methods are used that seek to estimate the most convenient sizes to reduce the experimental size.

As an alternative for reducing experimental error, the Papadakis method (1937) significantly reduces the error due to an adjustment of spatial variability, in which the covariate for use in the covariance analysis is obtained from the experimental errors themselves. The methodology consists of a covariance analysis in which an environmental index is used as a covariate, correcting the production of each plot (response variable) by the average effect of neighboring plots, with the environmental index being obtained by averaging the residuals. of neighboring plots.

Therefore, the objective of this work is to evaluate the use of the Papadakis method for estimating the size of experimental plots for use in soybean cultivation. Specifically, the objectives include: i) estimation of plot size using weight data per plant without and with transformation using the Papadakis technique; ii) compare the modified maximum curvature methods, the segmented linear method with plateau response and the maximum curvature coefficient of variation model applied to data without and with application of the Papadakis technique.

METHODOLOGY

MATERIAL AND METHODS

The data used in this work come from an experiment carried out on the experimental farm of ``Universidade Federal de Lavras` (UFLA), located in the municipality of Lavras, MG, at the geographic coordinates of 20°14′ south latitude, 45°00′ west longitude and average latitude of 918 meters. According to Koppen's international climate classification, the region's climate is type Cwa, tropical, temperate rainy, with dry winter, rainy summer and temperature of the hottest month greater than 22°C (DANTAS; CARVALHO; FERREIRA, 2007).

The experiment was conducted in a randomized block design (DBC), with three replications, with treatments consisting of 10 soybean cultivars, each plot was formed by four rows of five meters in length, spaced 0.5 meters apart. In each plot, a 4x48 network of basic units (BU) was created, corresponding to the four rows of the plot and the 48 plants within each row, forming a total of 192 UB. Each UB was formed by a plant and each UB was evaluated for grain production, in grams (g) per plant.

The 192 UB arranged in four rows and 48 columns were used, different sample sizes were simulated, which are formed by x_1 basic units of width (rows, $x_1 = 1, 2$ *e* 4) and for x_2 basic units of length ($x_2 = 1, 2, 4, 6, 8, 12, 24$). Sample sizes were simulated by grouping adjacent or contiguous UBs, so that $x_1 * x_2 = x$ corresponding to the sample size, with x UB.

Based on the grouped values, the following quantities were estimated: n number of samples with x UB $\left(n = \frac{192}{x}\right)$; m_x average of samples with x UB V_x variance between samples with x UB $CV_x = \frac{V_x}{X^2}$ variance by UB or reduced variance; CV_x coefficient of variation among samples of x UB $\left(CV_x = \frac{\sqrt{VU_x}}{m_x}.100\right)$. Table 1 illustrates the BU grouping structure for the formation of different plot sizes.

Structure	Format (X1*X2)	Portion size (UEB)	Total number (UEB)
1	1X1	1	192
2	1X2;2X1	2	96
3	1X3	3	64
4	1X4;4X1;2X2;	4	48
5	1X6;2X3	6	32
6	1X8;2X4;4X2	8	24
7	1X12;2X6;4X3	12	16
8	1X16;2X12;4X4	16	12
9	1X24;2X12;4X6	24	8
10	2X16;4X8	32	6
11	4X12	48	4

Table 1: Grouping structure of basic structures (UB), including formats, plot size in UB and total number of UBs, with UBs consisting of plants for each soybean cultivar.

GOODNESS-OF-FIT MEASURES AND EVALUATION CRITERIA

To verify the efficiency of data transformation by the Papadakis method using analysis of variance and to compare the different methods, the following goodness-of-fit measures were used: (i) standard error of estimates, the smaller the standard error indicates the better precision; (ii) residual standard error, which is obtained by the

Arts, Linguistics, Literature and Language Research Journal ISSN 2764-1929 DOI https://doi.org/10.22533/at.ed.929422419032

square root of the mean square of the error and the smaller the residual standard error is an indicator of improved quality of fit; (iii) adjusted coefficient of variation (R^2_a) was calculated using the expression (RENCHER; SCHAALJE, 2008). $R_a^2 = 1 - \left(\frac{n-1}{n-p}\right)\frac{SQR}{SQT}$ on what; n is the number of observations; p is the number of model parameters; SQR is the sum of squared residuals; SQT is the total sum of squares given to the mean. The higher your estimate, the better the model. (iv) the Akaike information criterion (AIC), is a characteristic used to compare the quality of model fit using the maximum likelihood function.

$AIC = -2logL(\hat{\Theta}) + 2p$, on what $L(\hat{\Theta})$ is the value of the maximum likelihood function of the model in $(\hat{\Theta})$ and p is the number of model parameters.

PET PROCEDURES

The Papadakis method: the procedure was carried out in each of the UBs that are arranged in four rows (lines) of 48 columns with $i = 1,2,3,4$ lines and $j = 1,2,...,48$ columns, so that Y12 corresponds to the observed value of the grain weight of the plant located in line 1 and column 2. Initially, the residues of each UB were estimated by the expression: $R_{ij} = Y_{ij} - \hat{Y}_{i+}$ on what R_{ij} is the value of the residual estimated in row i and column j; Y_{ij} is the value observed in row i and column j; $\hat{Y}_{i+}$ is the average of the UB in the i-th crop line. Waste values R_{ij} They are positioned according to the initial arrangement of the values of each UB. Then, the average residuals are estimated, which are the values of the covariates or environmental index, in each experimental unit, using the expression: $C_{ij} = \frac{[R_{ij}+R_{(i-1,j)}+R_{(i+1,j)}]}{3}$ in which, C_{ij} is the value of the covariate associated with the portion of row i and column j. The adjustment of the original value of each installment using covariance analysis is done by the expression: $Y^*_{(ij)}=Y_{ij}-\beta(C_{ij} - \overline{C}_j)$ on what $Y^*_{(ij)}$ is the corrected variable in row i and column j; Y_{ij} is the original variable in row i and column j; $\overline{C}_j$ is the average of the covariate in the crop line. With the values of the original observations (Y_{ij}) and transformed values $Y^*_{(ij)}$ the necessary procedures were carried out to estimate the plot size using the following three estimation methods.

MODIFIED MAXIMUM CURVATURE METHOD

To explain the relationship between the coefficients of variation and plot size, a function developed by Lessman and Atkins (1963) was used, using the expression: $CV_{(x)} = \frac{A}{X^B} + \varepsilon_x$ on what; $CV_{(x)}$ is the value of the coefficient of variation observed between the plots; X is the number of grouped UBs. A is the coefficient of variation of a UB in a plot; B is the soil heterogeneity index; ε_x is the associated error in the $CV_{(x)}$ considered independent and normally distributed with zero mean and constant variance.

The optimal plot size was determined using the estimator: $\hat{X}_0 = \left[\frac{\hat{A}^2\hat{B}^2(2B+1)}{\hat{B}+2}\right]^{\frac{1}{(2+2B)}}$ on what; $\hat{X}_0$ is the optimal plot size in number of UB; $\hat{A}$ is the estimate of parameter A; $\hat{B}$ is the estimate of parameter B.

SEGMENTED LINEAR MODEL METHOD WITH PLATEAU RESPONSE

The determination of the optimal size by this method is carried out first by determining the coefficient of variation given by the expression: $CV_{(X)} = \begin{cases}\beta_0 + \beta_1 X + \varepsilon_x; & se\ X \le X_0 \\ CVP + \varepsilon_x; & se\ X > X_0\end{cases}$ on what; $CV_{(X)}$ is the observed value of the coefficient of variation or other measure of variability between sample totals of size x; X is the number of grouped UBs; X_0 is the sample size, in UB; CVP is the coefficient of variation of the point corresponding to the plateau; β_0 is the intercept of the linear segment; β_1 is the angular coefficient of the linear segment; ε_x is the error associated with $CV_{(X)}$ considered as normally distributed, independent with zero mean and constant variance. The optimal sample size was determined by the expression: $\hat{X}_0 = \frac{\widehat{CVP}}{\hat{\beta}_1}$ on what; $\hat{X}_0$ is the sample size, in UB; $\widehat{CVP}$ is the estimate of the CVP parameter; $\hat{\beta}_0$ is the parameter estimate β_0; and $\hat{\beta}_1$ is the parameter estimate β_1.

MAXIMUM CURVATURE COEFFICIENT OF VARIATION METHOD

With adjusted and unadjusted data, the optimal plot size was estimated using the maximum curvature coefficient of variation method, obtained by the expression: $\hat{X}_0 = \frac{10\sqrt[3]{2(1-\hat{\rho}^2)S^2\bar{Y}}}{\bar{Y}}$ on what; $\hat{X}_0$ is the optimal plot size; S^2 is the variance of the crop line; $\overline{Y}$ is the average of the plants in the crop line; $\hat{p}_2$ is the first-order spatial autocorrelation coefficient estimated by the equation: $\hat{\rho}^2 = \frac{\sum_2^n (R_{ij})(R_{(i-1,j)})}{\sum_1^n (R_{ij})^2}$ on what; R_{ij} is the residue of parcel i in line j. The coefficient of variation in the optimal plot size is calculated by the expression: $CV_{X_0} = \frac{100\sqrt{2(1-\hat{\rho}^2)S^2/\bar{Y}}}{\sqrt{X_0}}$ on what; CV_{X0} is the coefficient of variation of the optimal plot size.

COMPUTATIONAL RESOURCES

All statistical analyzes were carried out with routines developed in the statistical program R (R-CORE-TEAM, 2018), and Office Excel (WALKENBACH, 2010). All analyzes were carried out assuming a 5% probability of error.

RESULTS AND DISCUSSION

MODIFIED MAXIMUM CURVATURE METHOD (MMCM)

The estimated values of parameters (a) and (b) of the (MMCM) showed a small variation between the cultivars and the techniques used without and with the Papadakis method correction. The estimated values for (b) were higher when using the Papadakis correction, with an average value of 1.536 with Papadakis, and 1.497 without Papadakis, without the correction. The asymptotic standard errors of parameter estimates were slightly larger when using Papadakis.

The estimates of the optimal plot size had little variation in relation to those using the Papadakis correction technique, without the correction it was 7.14 UEB, and with the Papadakis correction the average was 7.01 UEB. Using the Papadakis method, a slight decrease in the optimal plot size was observed in relation to plot size values without using the method. Storck et al., (2008) using the Papadakis method with diverse environments and genetic resources, found that the method proved to be efficient in improving precision indicators. Cargnelutti-Filho, Storck and Lúccio (2003) used the environmental index estimated in five ways as a covariate and compared the conventional analysis with analysis using the Papadakis methodology using estimates of mean squares of errors, coefficient of variation, minimum significant difference from the test of Tukey, verified the Papadakis method improved experimental precision in relation to different forms of estimation.

SEGMENTED LINEAR MODEL WITH PLATEAU RESPONSE (MLRP) METHOD)

Using the (MLRP) method, no significant changes were found between parameter estimates considering data without and with adjustment of the Papadakis technique. Humada et al., (2018), studied plot size for sweet potato cultivation, using (MLRP) and found that the method obtains more appropriate plot sizes, because the behavior of the CV curve in relation to (X_0) tends to a trajectory, so that increasing the plot size produces a continuous gain in accuracy. In average terms, it was found that the Papadakis adjustment using the method (MLRP) does

Grow crops	Method	Parameters				X_0	R_a^2	AIC	DPR
		a	*ep*	b	*ep*				
C1	s/P	81,58	0,832	1,483	0,017	7,24	0,9975	64,230	1,281
	c/P	84,83	1,179	1,544	0,029	6,98	0,9954	80,860	1,205
C2	s/P	84,91	1,193	1,487	0,027	7,18	0,9940	74,510	1,209
	c/P	84,84	1,199	1,545	0,029	6,98	0,9952	81,570	1,225
C3	s/P	84,26	0,845	1,475	0,019	7,20	0,9973	61,840	0,868
	c/P	84,69	1,082	1,521	0,026	7,05	0,9961	76,490	1,107
C4	s/P	84,42	0,733	1,493	0,017	7,16	0,9981	57,520	0,752
	c/P	84,72	0,973	1,526	0,023	7,04	0,9968	71,490	0,996
C5	s/P	84,30	0,692	1,482	0,016	7,17	0,9983	54,640	0,711
	c/P	84,72	1,006	1,530	0,024	7,03	0,9966	72,940	1,029
C6	s/P	84,31	0,654	1,479	0,015	7,18	0,9985	51,590	0,671
	c/P	84,66	1,015	1,514	0,024	7,08	0,9961	73,580	1,040
C7	s/P	84,58	0,827	1,514	0,019	7,07	0,9977	62,150	0,836
	c/P	84,78	1,109	1,537	0,027	7,00	0,9956	75,380	1,134
C8	s/P	84,58	0,726	1,511	0,017	7,08	0,9981	56,930	0,744
	c/P	84,83	1,143	1,542	0,028	6,99	0,9957	79,511	1,169
C9	s/P	84,77	1,045	1,536	0,025	7,01	0,9964	75,198	1,069
	c/P	84,95	1,397	1,558	0,035	6,94	0,9936	89,009	1,427
C10	s/P	84,58	0,768	1,514	0,018	7,07	0,9980	59,902	0,786
	c/P	84,79	1,166	1,540	0,028	6,99	0,9955	80,440	1,192
average(s/P)		84,23	0,832	1,497	0,019	7,14	0,9974	61,851	0,893
average(c/P)		84,78	1,127	1,536	0,027	7,01	0,9957	78,127	1,152

Table 2: Estimates of parameters (a) and (b) with their respective asymptotic standard errors (ep), plot size (X_0), coefficient of determination (R_a^2), of the Akaike evaluation criterion (AIC) and the residual standard deviation (DPR) for data not adjusted by the Papadakis method (s/P) and adjusted (c/P) of weight per plant in ten soybean cultivars.

not significantly change the plot size results (X0) in relation to the results without the use of Papadakis with an average value estimated at 4.14 UEB in both techniques.

Grow crops	Method	Parameters				CVP	X_0	R_a^2	AIC	DPR
		β_0	*ep*	β_1	*ep*					
C1	s/P	80,96	6,492	19,01	2,114	1,932	4,16	0,8712	161,08	6,292
	c/P	84,14	6,754	20,02	2,452	1,549	4,12	0,8710	163,09	6,546
C2	s/P	83,63	6,839	19,72	2,227	1,969	4,13	0,8659	163,69	6,629
	c/P	84,16	6,754	20,03	2,199	1,535	4,12	0,8711	163,08	6,545
C3	s/P	83,01	7,012	19,37	2,283	2,241	4,16	0,8581	164,87	6,797
	c/P	83,50	6,860	19,65	2,234	1,619	4,16	0,8666	164,04	6,649
C4	s/P	86,15	6,906	19,58	2,249	2,092	4,15	0,8628	164,15	6,788
	c/P	83,80	6,818	19,82	2,220	1,666	4,14	0,8681	163,54	6,608
C5	s/P	83,41	9,933	19,61	2,258	2,322	4,13	0,8609	164,34	6,720
	c/P	83,85	6,796	19,85	2,213	1,713	4,13	0,8678	163,57	6,588
C6	s/P	83,15	6,981	19,59	2,273	2,177	4,16	0,8597	164,67	6,766
	c/P	83,24	6,891	19,51	2,244	1,616	4,18	0,8654	164,04	6,679
C7	s/P	83,77	6,818	19,81	2,221	1,914	4,13	0,8669	163,54	6,608
	c/P	84,10	6,770	19,99	2,205	1,652	4,12	0,8689	163,34	6,562
C8	s/P	83,75	6,829	19,79	2,224	1,856	4,13	0,8674	163,51	6,619
	c/P	83,94	6,756	19,91	2,200	1,566	4,14	0,8702	163,21	6,548
C9	s/P	84,19	6,764	20,04	2,203	1,714	4,12	0,8714	163,03	6,556
	c/P	84,09	6,739	19,99	2,194	1,444	4,13	0,8714	163,11	6,531
C10	s/P	83,82	6,816	19,63	2,220	1,811	4,13	0,8681	163,42	6,607
	c/P	84,13	6,760	20,01	2,201	1,741	4,11	0,8695	163,24	6,552
average(s/P)		83,58	7,139	19,62	2,227	2,003	4,14	0,8652	163,63	6,638
average(c/P)		83,90	6,790	19,88	2,236	1,610	4,14	0,8690	163,43	6,581

Table 3: Parameter estimates (β_0) e (β_1) with their respective asymptotic standard errors (ep), coefficient of variation at the plateau point (CVP), plot size (X_0), adjusted coefficient of determination (R_a^2), of the Akakaike evaluation criterion (AIC) and the residual standard deviation (DPR) for data not adjusted by the Papadakis method (s/P) and adjusted (c/P) of weight per plant in ten soybean cultivars by the method (MLRP).

MAXIMUM CURVATURE COEFFICIENT OF VARIATION (MMCV) METHOD

Using the (MMCV) method, the values of plot size: 97 to 5.26 UEB. This shows that the model (MMCV) with data adjusted by the De Papadakis method shows a certain efficiency. This considering that the estimation of the smallest optimal parcel size is a more appropriate solution, in average terms, without the Papadakis correction, the optimal parcel size was adjusted to 4.76 UEB and, with Papadakis, it was 3.74 UEB. Lúcio et al., (2016), in evaluating the quality of experiments with lettuce cultivation, used the coefficient of variation method to obtain estimates of plot size, verifying that the use of covariance with the covariable estimated by the Papadakis method, improved the quality of experiments, making it possible to estimate smaller plot sizes.

Grow crops	Method	ρ	X_0	CV
C1	s/P	0,026	4,64	0,478
	c/P		3,78	0,526
C2	s/P	0,021	4,98	0,462
	c/P		3,67	0,533
C3	s/P	0,016	5,26	0,451
	c/P		3,90	0,518
C4	s/P	0,033	4,79	0,470
	c/P		3,93	0,516
C5	s/P	0,005	5,16	0,453
	c/P		3,71	0,530
C6	s/P	0,069	4,90	0,465
	c/P		3,83	0,522
C7	s/P	0,040	4,59	0,479
	c/P		4,11	0,505
C8	s/P	0,012	4,51	0,483
	c/P		3,67	0,533
C9	s/P	0,297	3,97	0,514
	c/P		2,98	0,588
C10	s/P	0,054	4,82	0,469
	c/P		3,86	0,521
average(s/P)			4,76	0,472
average(c/P)			3,74	0,529

Table 4: Estimates of the first-order spatial autocorrelation coefficient (ρ), plot size (X_0) using the Papadakis adjustment method (c/P) and without using it (s/P) and coefficient of variation (CV).

CONCLUSION

The application of the Papadakis method proved to be efficient in reducing the optimal experimental plot size for the MMCV and MMCM methods.

The MMCM method estimated larger optimal plot sizes in relation to the MLRP and MMCV methods.

THANKS

To the Coordination for the Improvement of Higher Education Persons (CAPES) for granting a scholarship.

To ``Universidade Federal de Lavras``, Department of Statistics and Agricultural Experimentation for accepting me to pursue a doctorate.

REFERENCES

AKAIKE, H. *A new look at the statistical model identification*. IEEE transaction on automatic control: Boston, v. 19, n. 6, p. 716-723, 1974.

CARGNELUTTI-FILHO, A. et al. *Tamanho de parcela e número de repetições em ervilha forrageira*. Ciência Rural, Santa Maria, v. 45, n. 7, p. 1174–1182, 2015.

CARGNELUTTIFILHO, A.; STORCK, L.; LÚCIO, A. D. *Ajustes de quadrado médio do erro em ensaios de competição de cultivares de milho pelo método de Papadakis*. Pesquisa Agropecuária Brasileira, Brasília, v. 38, n. 4, p. 467–473, 2003.

DANTAS, A. A. A.; CARVALHO, L. d.; FERREIRA, E. *Classificação e tendências climáticas em lavras, MG*. Ciência e Agrotecnologia, Lavras, v. 31, n. 6, p. 1862–1866, 2007.

HUMADA, G. G. et al. *Estimación del tamaño ótimo de parcela en experimentación con batata dulce*. Agrociência, Uruguai, v. 22, n. 2, p. 1–10, 2018.

LÚCIO, A. D. et al. *Método de papadakis e tamanho de parcela em experimentos com a cultura da alface*. Horticultura Brasileira, Brasília, v. 34, n. 1, p. 66–73, 2016.

PAPADAKIS, J. *Méthode statistique pour des experiences sur champ*. Thessalonike: Institut d'Amelionation des plantes, Salonique, 30p. (Bulletin 23)., 1937.

R-CORE-TEAM. *A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria:, 2018. URL https://www.R-project.org/.

STEEL, R. G.; TORRIE, J. H.; DICKEY, D. A. *Principles and procedures of statistics*. 3. ed. [S.l.]: Boston: WCB/ McGraw-Hill, 1997. 666 p.

STORCK, L.; BISOGNIN, D. A.; OLIVEIRA, S. J. R. *Dimensões dos ensaios e estimativas do tamanho ótimo de parcela em batata*. Pesquisa Agropecuária Brasileira, Brasília, v. 41, n. 6, p. 903–909, 2006.

STORCK, L. et al. *Utilização do método de Papadakis na melhoria da qualidade experimental de ensaios com soja*. Pesquisa Agropecuária Brasileira, v. 43, n. 5, p. 581–587, 2008.

WALKENBACH, J. *Excel 2010 power programming with* VBA. [S.l.]: New York: John Wiley & Sons, 2010. v. 6.