

ANALYSING $(1/125)5^5$ FRACTIONAL FACTORIAL DESIGNS AND SOME CONSIDERATIONS

Armando CONAGIN¹

Décio BARBIN²

Clarice Garcia Borges DEMÉTRIO²

Rafael de Andrade MORAL²

- **ABSTRACT:** *In this paper we discuss some aspects of $(1/125)5^5$ fractional factorial designs and show how to analyze a simulated data set. We present the analysis of variance, the estimated coefficients of the regression model, the determination of maximum response and the economical analysis. We also discuss how to allocate the 25 treatments of $(1/5)5^3$ fractional factorials in completely randomized, incomplete block or incomplete latin square designs and the 25 treatments of $(1/25)5^4$ fractional factorials in incomplete block design. We present the R program to perform the analysis in the Appendix.*
- **KEYWORDS:** *Factorial experiments; high-order interactions; confounding; simulation.*

1 Introduction

In the past, it was common to study the influence of a particular factor on a phenomenon, using several dosages of the factor, while keeping the remaining factors that were supposed to have influence on the phenomenon at the same level. Fisher (1935) introduced the concept of factorial experiments for which the treatments are all combinations of the levels of all factors and, generally, the number of treatments is equal to the product of the numbers of levels of the factors in the experiment. Yates (1937) presented a comprehensive review of the simpler factorial designs and a description of the appropriate methods of analysis.

Initially, factorials included two or three factors with only two dosages each. When the numbers of the factors under investigation were four or more, the fractional factorial designs were proposed to reduce the number of treatments. Considering a full 2^5 factorial, for example, 32 treatments are generated. To produce a $(1/2)2^5$ fractional

¹ Instituto Agronômico de Campinas – IAC, CEP: 13012-970, Campinas, SP, Brazil.

² Universidade de São Paulo – USP, Escola Superior de Agricultura “Luiz de Queiroz” - ESALQ, Departamento de Ciências Exatas, CEP: 13418-900, Piracicaba, SP, Brazil. E-mail: decio.barbin@usp.br; clarice.demetrio@usp.br; rafael.moral@usp.br

factorial with 16 treatments the highest interaction ABCDE was supposed null, for which the main effects and all the interactions but ABCDE are estimated, perfectly.

Later, the factorial designs were extended to more complicated ones with increasing number of treatments, such as 3^3 , 4^3 , $2 \times 3 \times 4$, 5^2 , etc., aiming to better estimate the shape of the response surface, by including linear and quadratic effects and interactions between two, three factors, etc. If the factors have three or more levels the fractional factorial designs can be used to fit response surface models (Box and Draper, 1987; Khuri and Cornell, 1996; Myers, Montgomery, Anderson-Cook, 2009). The shape of the response surface is very important for fertilizer experiments to determine the maximum response and to obtain the economical analysis to maximize profit.

In general, there is a great deal of redundancy in factorial experiments, in that high-order interactions are likely to be negligible and some variables may not affect the response at all (Box, Hunter and Hunter 2005, Brien, 2010a, b, Montgomery, 2012). Solutions to overcome these problems are the use of incomplete block designs and fractional factorial designs. Conagin and Jorge (1977, 1982a) proposed a $(1/5)^3$ fractional factorial to be used in fertilizer experiments with one application in Caetano et al (2013). Andrade and Noletto (1986) and Conagin and Jorge (1982b) presented $(1/2)^3$ and $(1/4)^4$ fractional factorials to be used in experiments to study the fertility of “cerrado” soils. Primavesi et al (2004) used the $(1/2)^3$ fractional factorial in one experiment with oats. Conagin et al. (2014) discussed some aspects of $5^{k-(k-2)}$ fractional factorial designs, where k is the number of factors, with only 25 treatments involving two to six quantitative factors, with the purpose of using them on experiments on poor soil areas like those of “cerrado” or “savanna”.

In this paper, we review some aspects of $(1/125)^5$ fractional factorial designs and show how to analyze a simulated data set following this design. We present the analysis of variance, the estimated coefficients of the regression model, the determination of maximum response and the economical analysis. We also discuss how to allocate the 25 treatments of $(1/5)^3$ fractional factorials in completely randomized, incomplete block or latin square designs and the 25 treatments of $(1/25)^4$ fractional factorials in completely randomized and incomplete block designs. We present the R program to perform the analysis in the Appendix.

2 Methodology

The method suggested by Cochran and Cox (1957), Davies (1950) and John (1971) produces $(1/125)^5$ fractional factorial designs by superimposing three of the four orthogonal 5×5 latin squares with the addition of two adequate columns of treatments (the 5×5 combinations of levels of factors A and B). These four orthogonal 5×5 latin square types (I, II, III and IV) are presented in Fisher and Yates (1938). The four basic types of the fractional factorial $(1/125)^5$, corresponding to the fractions AB , AB^2 , AB^3 and AB^4 of the interaction $A \times B$, having as central point (3,3,3,3,3), are presented in Table 2 of Conagin et al. (2014), following Wu and Yamada (2000), and reproduced in Appendix A of this paper (Table 1A). It is not difficult to show that it is an orthogonal array.

To analyze the data from a fractional factorial design we consider the classical linear model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e},$$

where \mathbf{y} is the $n \times 1$ vector of observations, \mathbf{X} is the $n \times p$ design matrix, $\boldsymbol{\beta}$ is the $p \times 1$ parameter vector and \mathbf{e} is $n \times 1$ the error vector, normally distributed with mean $\mathbf{0}$ and variance-covariance matrix $\sigma^2 \mathbf{I}$, $\mathbf{e} \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$. The vector of estimated parameters is given by

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

with $E(\hat{\boldsymbol{\beta}}) = \mathbf{0}$ and $\text{Var}(\hat{\boldsymbol{\beta}}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$. The analysis of variance can be done in the usual way (Steel and Torrie, 1981, Draper and Smith, 1966, Montgomery, 2012) using statistical packages such as R (R Core Team, 2015) and SAS (SAS Institute, Cary NC, 2015), for example. After estimating the vector of parameters, $\boldsymbol{\beta}$, we get the predicted values $\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}}$.

For a quantitative factor in a five-level design, for example, the quadratic regression model without interaction for $(1/125)5^5$ fractional factorials, using orthogonal polynomials can be expressed as

$$E(Y_i) = \beta_0 + \beta_{11}P_{11i} + \dots + \beta_{15}P_{15i} + \beta_{21}P_{21i} + \dots + \beta_{25}P_{25i} \quad (i = 1, \dots, 25),$$

where β_0 is the intercept, β_{1k} and β_{2k} , $k = 1, \dots, 5$, are, respectively, the linear and quadratic parameters for the k^{th} factor, P_{1ki} and P_{2ki} are, respectively, the values of the

linear and quadratic polynomials, $P_1 = x$ and $P_2 = x^2 - \frac{n^2 - 1}{12}$, with $x = \frac{X - \bar{X}}{q}$, for

coded values $X = 1, 2, 3, 4, 5$, $q = 1$, the distance between adjacent coded levels, \bar{X} , the mean of the values of X and n , the number of levels of X . Then the values assumed by P_1 and P_2 are $(-2, -1, 0, 1, 2)$ and $(2, -1, -2, -1, 2)$, respectively, for each $X (= 1, 2, 3, 4, 5)$. The estimated values of the polynomial regression are given by

$$\hat{\beta}_0 = \bar{Y} \quad \text{and} \quad \hat{\beta}_{jk} = \frac{\sum_{i=1}^{25} P_{ji} y_i}{\sum_{i=1}^{25} P_{ji}^2}, \quad i = 1, \dots, 25, j = 1, 2, k = 1, \dots, 5,$$

with $\text{Var}(\hat{\beta}_0) = \frac{\sigma^2}{25}$, $\text{Var}(\hat{\beta}_{jk}) = \frac{\sigma^2}{\sum_{i=1}^{25} P_{ji}^2}$ and $\text{Cov}(\hat{\beta}_{jk}, \hat{\beta}_{j'k'}) = 0$.

The sums of squares are obtained by

$$SSq = \hat{\boldsymbol{\beta}}' (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y} = \mathbf{y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

and for linear and quadratic regressions, using orthogonal polynomials, are calculated by

$$LSSq = \frac{\left(\sum_{i=1}^{25} P_{1i} y_i\right)^2}{\sum_{i=1}^{25} P_{1i}^2} \quad \text{and} \quad QSSq = \frac{\left(\sum_{i=1}^{25} P_{2i} y_i\right)^2}{\sum_{i=1}^{25} P_{2i}^2}.$$

To illustrate how to perform the analysis, we simulated a data set for corn yield using as treatments a type-III $(1/125)5^5$ fractional factorial (see Table 1A, in the Appendix A reproduced from Conagin et al, 2014), in a completely randomized design, with a quadratic linear predictor with the features, based on practical results of previous experiments, presented in Table 1 and Table 2. It was assumed that, without any fertilizer, the yield was around 3000kg/ha, for a population of 50000 plants per hectare. So, for example, to generate the value for treatment 6, that is, 24514, we got

$$\mu_6 = 3000 + 450 + 600 + 600 + 210 + 450 + 300 + 270 + 0 + 90 + 160 = 6130,$$

and $y_6 = \mu_6 + e_6$, where e_6 was simulated from a $N(0, \sigma^2)$ and σ^2 was obtained by $\sigma^2 = 0.05 \frac{\sum_{i=1}^{25} \mu_i}{25}$.

Table 1 - Dosage (D) levels of the factors and their effects on the yield for data simulation

Factors	Dosages					
Nitrogen (N)	30	45	6	75	9	kg/ha
Phosphorus (P_2O_5)	30	45	6	75	9	kg/ha
Potassium (K_2O)	30	40	5	60	7	kg/ha
Limestone (Ca)	1.	1.	2	2.	3	t/ha
Population of plants (Pop)	50	55	6	65	7	thousand plants/ha

Table 2 - Effects of the dosages on the yield

Factors	Linear					Quadratic				
	1	2	3	4	5	1	2	3	4	5
N	300	450	600	750	900	60	300	600	300	0
P_2O_5	240	360	480	600	720	60	270	540	270	0
K_2O	120	240	360	480	600	50	180	380	160	0
Ca	210	330	450	570	690	90	210	390	210	0
Pop	0	150	300	450	600	0	160	350	160	0

The corresponding coded dosages X were obtained by using $(D - 15)/15$ for N and P_2O_5 , $(D - 20)/10$ for K_2O , $(D - 0.5)/0.5$ for limestone and $(D - 45)/5$ for population

of plants. The corn yield can be seen in the R (R Core Team, 2015) program presented in the Appendix.

Regarding the convenience of including or not the two-factor interactions in the model, it is possible to demonstrate that the effect of the linear x linear interaction of every two factors is not orthogonal to the other effects. This becomes clear by looking at the calculation of the contrast $(N4 - N2)(P4 - P2) + (N5 - N1)(P5 - P1)$ or $[44443 + 22223 - 42452 - 24514] + [55553 + 11113 - 51421 - 15245]$ in which the last three numbers represent the levels of K, Ca and Population. This contrast, besides estimating the interaction $N_L P_L$, include differences in dosages of potassium, limestone and population that are confounded with the true values of the interactions. Additionally, it uses 10 degrees of freedom, leaving only four degrees of freedom for the error. This would contribute for decreasing the efficiency of the tests for the regression coefficients. So the interaction was not included in the model, what is a reasonable assumption for poor soils (Conagin et al., 2014).

For this type of experiment in general, the experimenter is interested in determining the dosages that give the maximum production, that is, to find the levels X_1, X_2, X_3, X_4, X_5 that optimize the predicted response. This point (called stationary point), if it exists, will be the set of X_1, X_2, X_3, X_4, X_5 for which the partial derivatives $\partial \hat{y} / \partial X_1 = \partial \hat{y} / \partial X_2 = \partial \hat{y} / \partial X_3 = \partial \hat{y} / \partial X_4 = \partial \hat{y} / \partial X_5 = 0$ and represent a point of maximum response if the matrix of second partial derivatives is negative definite (Montgomery, 2012). For the discussed example,

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_{11}P_{11i} + \dots + \hat{\beta}_{15}P_{15i} + \hat{\beta}_{21}P_{21i} + \dots + \hat{\beta}_{25}P_{25i} \quad (i = 1, \dots, 25)$$

with $P_1 = X - 3$ and $P_2 = X^2 - 6X + 7$, where $X (= 1, 2, 3, 4, 5)$ is the coded dosage level. Then, the stationary point X_1, X_2, X_3, X_4, X_5 will be obtained by

$$\frac{\partial \hat{y}}{\partial X_k} = \hat{\beta}_{1k} + \hat{\beta}_{2k}(2X_k - 6) \Rightarrow X_k = \frac{6\hat{\beta}_{2k} - \hat{\beta}_{1k}}{2\hat{\beta}_{2k}}, \quad (k = 1, \dots, 5).$$

Additionally, $\frac{\partial^2 \hat{y}}{\partial X_k^2} = 2\hat{\beta}_{2k}$ and $\frac{\partial^2 \hat{y}}{\partial X_k \partial X_{k'}} = 0$ and as the matrix of second partial derivatives is a diagonal matrix in this case, it will be negative definite if all the elements of the diagonal are negative, that is if all $\hat{\beta}_{2k}$ are negative.

Besides finding the maximum response, the experimenter could be interested in finding the point X_1, X_2, X_3, X_4, X_5 that gives the maximum profit for the farmer, that is, optimize the profit function. The profit function is defined as the difference between the revenue function and the cost function. Assuming that the unitary price of the product is p_y and that the unitary cost of the dosage D_k of the factor is p_{D_k} , the profit function (E), without considering fixed costs, is

$$E = yp_y - \sum_{k=1}^5 D_k p_{D_k}$$

Then,

$$\frac{\partial E}{\partial X_k} = [\hat{\beta}_{1k} + \hat{\beta}_{2k}(2X_k - 6)]p_y - \frac{dD_k}{dX_k} p_{D_k} \Rightarrow X_k = \frac{\frac{dD_k}{dX_k} p_{D_k} + (6\hat{\beta}_{2k} - \hat{\beta}_{1k})p_y}{2\hat{\beta}_{2k}p_y}, (k = 1, \dots, 5)$$

and again the matrix of second partial derivatives is a diagonal matrix that will be negative definite if all the elements of the diagonal are negative, that is if all $\hat{\beta}_{2k}$ are negative.

3 Results and discussion

The results for the analysis of variance for the simulated data, using orthogonal polynomials, are presented in Table 3 and show, at 10% level, evidence of significance for all linear and quadratic effects.

Table 3 - Analysis of variance for the generated data using the values from Table 1

Source	DF	S.S.	M.S.	F	Pr > F
N (Linear)	1	827927	827927	11.01	0.0051
P (Linear)	1	909361	909361	12.09	0.0037
K (Linear)	1	851512	851512	11.32	0.0046
Ca (Linear)	1	1177192	1177192	15.65	0.0014
Pop (Linear)	1	901019	901019	11.98	0.0038
N (Quadratic)	1	749651	749651	9.97	0.0070
P (Quadratic)	1	1471460	1471460	19.57	0.0006
K (Quadratic)	1	235596	235596	3.13	0.0985
Ca (Quadratic)	1	255130	255130	3.39	0.0868
Pop (Quadratic)	1	420593	420593	5.59	0.0330
Error	14	1052771	75198		

The multiple regression expressed, as a function of polynomial coefficients, is

$$\hat{y}_i = 6064.24 + 128.68P_{11i} + 134.86P_{12i} + 130.50P_{13i} + 153.44P_{14i} + 134.24P_{15i} - 103.49P_{21i} - 144.99P_{22i} - 58.01P_{23i} - 60.37P_{24i} - 77.51P_{25i} \quad (i = 1, \dots, 25), \quad (1)$$

which for $P_1 = X - 3$ and $P_2 = X^2 - 6X + 7$, where $X (= 1, 2, 3, 4, 5)$ is the coded dosage level, becomes

$$\hat{y}_i = 908.48 + 749.59X_{1i} + 1004.77X_{2i} + 478.59X_{3i} + 515.67X_{4i} + 599.33X_{5i} - 103.49X_{1i}^2 - 144.99X_{2i}^2 - 58.01X_{3i}^2 - 60.37X_{4i}^2 - 77.51X_{5i}^2 \quad (i = 1, \dots, 25), \quad (2)$$

and, for example, the predicted values that include the center of the design are given by $\hat{y}_{11113} = 4391$, $\hat{y}_{22223} = 6039$, $\hat{y}_{33333} = 6953$, $\hat{y}_{44443} = 7134$ and $\hat{y}_{55553} = 6581$ kg/ha.

Then, the stationary point for equation (1) is $X_{1M} = 3.62$, $X_{2M} = 3.46$, $X_{3M} = 4.12$, $X_{4M} = 4.27$ and $X_{5M} = 3.86$ with a predicted maximum response of $\hat{y}_M = 7,253.35$ kg/ha. The corresponding natural dosage values are $D_{1M} = 69.3$ kg of N/ha, $D_{2M} = 67.0$ kg of P_2O_5 /ha, $D_{3M} = 61.2$ kg of K_2O /ha, $D_{4M} = 2.6$ ton of limestone/ha and $D_{5M} = 64.3$ thousand plants/ha.

Now, assuming that the prices are: $p_y = \$0.55$ /kg of corn, $p_{D_1} = \$4.90$ /kg N (ammonia sulphate, 20%), for is $p_{D_2} = \$4.80$ /kg P_2O_5 (single superphosphate, 18%), $p_{D_3} = \$2.70$ /kg K_2O (potassium chloride, 55%), $p_{D_4} = \$0.20$ /kg calcium (dolomite limestone) and $p_{D_5} = \$7.00$ /thousand seeds, the stationary point for the profit equation is $X_{1E} = 2.98$, $X_{2E} = 3.01$, $X_{3E} = 3.70$, $X_{4E} = 4.27$ and $X_{5E} = 3.46$ with a predicted economical response of $\hat{y}_E = 7,157.21$ kg/ha. The corresponding natural dosage values are $D_{1E} = 59.6$ kg of N/ha, $D_{2E} = 60.2$ kg of P_2O_5 /ha, $D_{3E} = 57.0$ kg of K_2O /ha, $D_{4E} = 2.6$ ton of limestone/ha and $D_{5E} = 62.3$ thousand plants/ha. The profit, without considering the fixed costs, is then calculated by

$$E = 7,157.21(0.55) - 59.6(4.90) - 60.2(4.80) - 57.0(2.70) - 2.6(0.20) - 62.3(7.00),$$

which gives a net gain of \$2,764.83.

Additional Considerations

In this paper, it was assumed that the 25 treatments of the $(1/125)5^5$ fractional factorial design were allocated following a completely randomized design. We now discuss how to allocate the treatments according to randomized incomplete block or incomplete latin square designs.

We begin by considering the $(1/5)5^3$ fractional factorial design. The resulting 25 treatments can be obtained by superimposing three orthogonal 5×5 latin squares or by using a selected design (I, II, III or IV) of Table 2 of Conagin et al. (2014) by elimination of two digits. These treatments can be allocated in a completely randomized design. Instead, if four orthogonal 5×5 latin squares are superimposed or Table 2 of Conagin et al. (2014) is used, the first three digits can be used to form the 25 treatments and the fourth digit can be used as the number of the incomplete block with five treatments each. If the first three digits are used for composing the treatment, the fourth digit as row number and the fifth as column number, the 25 treatments can be allocated in an incomplete latin square design.

In a similar way, the 25 treatments for a $(1/25)5^4$ fractional factorial design can be obtained from Table 2 of Conagin et al. (2014), by using four digits of the numbers, for a

completely randomized design or additionally by using the fifth digit the 25 treatments can be allocated in an incomplete block design with five treatments each.

CONAGIN, A.; BARBIN, D.; DEMÉTRIO, C. G. B.; MORAL, R. A. Analisando fatoriais fracionados $(1/125)5^5$ e algumas considerações. *Rev. Bras. Biom.*, São Paulo, v.33, n.4, p.471-483, 2015.

- RESUMO: Neste artigo, discutem-se alguns aspectos do fatorial fracionado $(1/125)5^5$ e mostra-se como analisar um conjunto de dados simulados. Apresentam-se a análise de variância, os coeficientes estimados do modelo de regressão, a determinação da dose de máxima resposta e a dose econômica. Discute-se, também, como alocar os 25 tratamentos de fatoriais fracionados $(1/5)5^3$, usando-se os delineamentos completamente casualizado, casualizado em blocos incompletos e quadrado latino incompleto e os 25 tratamentos de fatoriais fracionados $(1/25)5^4$, usando-se os delineamentos completamente casualizado e casualizado em blocos incompletos. Apresenta-se no Apêndice o programa R para a análise.
- PALAVRAS-CHAVES: Experimentos fatoriais; interações de alta ordem; confundimento; simulação.

References

ANDRADE, D. F.; NOLETO, A. Exemplo de fatoriais fracionados $1/2(4^3)$. Ajuste de modelos polinomiais quadráticos. *Pesquisa Agropecuária Brasileira*, Brasília, v.21, n.6, p.677-682, 1986.

BOX, G. E. P.; DRAPER, N. R. *Empirical Model-Building and Response Surfaces*. John Wiley & Sons, New York, 1987.

BOX, G. E. P.; HUNTER, J. S.; HUNTER, W. G. *Statistics for Experimenters: Design, Innovation, and Discovery*. Wiley, 2005.

BRIEN, C. J. *Factorial experiments*. Disponível em: <<http://chris.brien.name/ee2>>. Accessed in Dec. 2015.

BRIEN, C.J. (2010b) *Factorial designs at two levels*. Disponível em: <<http://chris.brien.name/ee2>>. Accessed in: Dec. 2015.

CAETANO, L. C. S.; VENTURA, J. A.; COSTA, A. F. S.; GUARÇONI, R. C. Efeito da adubação com nitrogênio, fósforo e potássio no desenvolvimento, na produção e na qualidade de frutos do abacaxi 'Vitória'. *Revista Brasileira de Fruticultura*, v.35, n.3, p.1-5, 2013.

COCHRAN, W. G.; COX, G. M. *Experimental design*. 2.ed., John Wiley & Sons, Inc., 1957.

CONAGIN, A.; JORGE, J. P. N. Delineamentos $1/5(5^3)$. *Bragantia*, v.36, n.1, p.1-4, 1977.

CONAGIN, A.; JORGE, J. P. N. Delineamentos $(1/5)(5 \times 5 \times 5)$ em blocos. *Bragantia*, v.41, n.1, p.1-3, 1977.

- CONAGIN, A.; JORGE, J. P. N. *Delineamentos $1/2(4^3)$ em blocos de 8 unidades*. Instituto Agronômico, Campinas, Boletim Científico n.36. 1982a.
- CONAGIN, A.; BARBIN, D.; ZOCCHI, S. S.; DEMÉTRIO, C. C. B. Fractional factorial designs for fertilizer experiments with 25 treatments in poor soils. *Revista Brasileira de Biometria*, v.32, n.2, p.180-189, 2014.
- DAVIES, O. L. *Design and analysis of industrial experiments*. Hafner Publishing Company, 1950.
- DRAPER, N. R.; SMITH, H. *Applied Regression Analysis*, New York. John Wiley & Sons, Inc., 1966.
- FISHER, R. A. *The Design of Experiments*. 2.ed. London, Oliver and Boyd, 1935.
- FISHER, R. A.; YATES, F. *Statistical Tables for Biological, Agricultural and Medical Research*, Copyright by Oliver and Boyd. London, 1938.
- JOHN, P. W. M. *Statistical Design and Analysis of Experiments*. Macmilan, 1971.
- KHURI, A. I.; CORNELL, J. A. *Response Surfaces*. 2ed. Marcel Dekker, New York, 1996.
- MONTGOMERY, D. C. *Design and Analysis of Experiments*. John Wiley, 2012.
- MYERS, R. H.; MONTGOMERY, D. C.; ANDERSON-COOK, C. M. *Response Surface Methodology. Process and Product Optimization Using Designed Experiment*. 3.ed. John Wiley & Sons, New Jersey, 2009.
- PRIMAVESI, A. C.; PRIMAVESI, O.; CANTARELLA, H.; GODOY, R. Resposta da aveia branca à adubação em latossolo vermelho-amarelo em dois sistemas de plantio. *Revista Brasileira de Zootecnia*. v.33, n.1, p.1-10, 2004.
- R Core Team 2015. R: *A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Disponível em: <<http://www.R-project.org/>>, 2015.
- SAS Institute Inc. *System of Microsoft Windows*, release 9, Cary, N.C. USA, 2015.
- STEEL, R.; TORRIE, J. *Principles and Procedures of Statistics*. Mc Grow Hill Int. Ed., 1981.
- YATES, F. *Design and Analysis of Factorial Experiments*. Tech. Comm. No. 35, Imperial Bureau of Soil Science, London, 1937.
- WU, C. F. J.; YAMADA, M. *Experiments. Planning Analysis and Parameter Design Optimization*. John Wiley, 2000.

Received in 08.09.2015

Approved after revised in: 24.11.2015

Appendix A

Table 1A. Four 5^{5-3} fractional factorial design obtained by the generated level sequences $(x_3, x_4, x_5, x_1, x_2)$, $(x_3, x_4, x_6, x_1, x_2)$, $(x_3, x_5, x_6, x_1, x_2)$ and $(x_4, x_5, x_6, x_1, x_2)$ from Table 1 of Conagin et al. (2014)

I	11113	22223	33333	44443	55553
	23414	34524	45134	51244	12354
	35215	41325	52435	13545	24155
	42511	53121	14231	25341	31451
	54312	15422	21532	32142	43252
II	11113	22223	33333	44443	55553
	23511	34121	45231	51341	12451
	35412	41522	52132	13242	24352
	42314	53424	14534	25144	31254
	54215	15325	21435	32545	43155
III	11113	22223	33333	44443	55553
	24514	35124	41234	52344	13454
	32415	43525	54135	15245	21355
	45311	51421	12531	23141	34251
	53212	14322	25432	31542	42152
IV	11113	22223	33333	44443	55553
	34515	45125	51235	12345	23455
	52414	13524	24134	35244	41354
	25312	31422	42532	53142	14252
	43211	54321	15431	21541	32151

Fonte: Conagin et al. (2014)

Appendix B

R Program for the analysis of the simulated data

```
## 1. Input data
### 1.1. Response
Y <-
c(4298, 5992, 6791, 7252, 6688, 6419, 5729, 6603, 6300, 6441, 6186, 668
4,
6341, 5955, 5679, 5474, 5164, 5357, 5736, 6491, 5847, 6168, 5944, 6118,
5949)
### 1.2. Regressors
N <- c(1, 2, 3, 4, 5, 2, 3, 4, 5, 1, 3, 4, 5, 1, 2, 4, 5, 1, 2, 3, 5, 1, 2, 3, 4)
P <- c(1, 2, 3, 4, 5, 4, 5, 1, 2, 3, 2, 3, 4, 5, 1, 5, 1, 2, 3, 4, 3, 4, 5, 1, 2)
K <- c(1, 2, 3, 4, 5, 5, 1, 2, 3, 4, 4, 5, 1, 2, 3, 3, 4, 5, 1, 2, 2, 3, 4, 5, 1)
Ca <- rep(1:5, times=5)
Pop <- rep(c(3, 4, 5, 1, 2), each=5)
N2 <- N^2; P2 <- P^2; K2 <- K^2; Ca2 <- Ca^2; Pop2 <- Pop^2
## 2. Analysis
### 2.1. Computing orthogonal polynomials
linear <- function(x) x - 3
quadratic <- function(x) (x - 3)^2 - sum((x - 3)^2)/25
N11 <- linear(N)
P11 <- linear(P)
K11 <- linear(K)
Ca11 <- linear(Ca)
Pop11 <- linear(Pop)
N12 <- quadratic(N)
P12 <- quadratic(P)
K12 <- quadratic(K)
Ca12 <- quadratic(Ca)
Pop12 <- quadratic(Pop)

### 2.2. Total sum of squares
anova(lm(Y ~ 1))

### 2.3. Obtaining analysis of variance table (Table 2 of
the paper)
corn.fit1 <- lm(Y ~ N11 + P11 + K11 + Ca11 + Pop11 + N12 +
P12 + K12 + Ca12 + Pop12)
anova(corn.fit1)

### 2.4. Original dosage levels
Nd <- N*15 + 15; Pd <- P*15 + 15; Kd <- K*10 + 20
Cad <- Ca*0.5 + 0.5; Popd <- Pop*5 + 45
Nd2 <- Nd^2; Pd2 <- Pd^2; Kd2 <- Kd^2
```

```

Cad2 <- Cad^2; Popd2 <- Popd^2

### 2.5. Obtaining coefficients for Eq. (1)
(cf <- coef(corn.fit1))

### 2.6. Obtaining coefficients for Equation (2)
corn.fit2 <- lm(Y ~ N + P + K + Ca + Pop + N2 + P2 + K2 +
Ca2 + Pop2)
(cf2 <- coef(corn.fit2))

### 2.7. Predicted values
fitted(corn.fit2)

### 2.8. Stationary points for Eq. (1) and corresponding
dosage levels
maximum <- function(coefs) {
  c.lin <- coefs[1]
  c.quad <- coefs[2]
  Max <- (6*c.quad - c.lin)/(2*c.quad)
  return(Max)}
## coded dosages
(X1 <- maximum(cf[c(2,7)])); (X2 <- maximum(cf[c(3,8)]));
(X3 <- maximum(cf[c(4,9)]))
(X4 <- maximum(cf[c(5,10)])); (X5 <- maximum(cf[c(6,11)]))
## original scale dosages
(D1 <- X1*15 + 15); (D2 <- X2*15 + 15); (D3 <- X3*10 + 20)
(D4 <- X4*0.5 + 0.5); (D5 <- X5*5 + 45)

#### 2.8.1. Maximum response
X <- c(1, X1, X2, X3, X4, X5, X1^2, X2^2, X3^2, X4^2, X5^2)
(yM <- sum(cf2*X))

### 2.9. Profit (E) function
p <- c(4.9,4.8,2.7,0.2,7,0.55) # prices for N, P, K, Ca, Pop
and y

#### 2.9.1. Stationary points and dosage levels for profit
function
maximum.E <- function(price.D, coefs, q) {
  price.y <- p[6]
  c.lin <- coefs[1]
  c.quad <- coefs[2]
  Max <- (q*price.D + (6*c.quad -
c.lin)*price.y)/(2*c.quad*price.y)
  return(Max)}
## coded dosages
(X1E <- maximum.E(p[1], cf[c(2,7)], 15))

```

```

(X2E <- maximum.E(p[2], cf[c(3,8)], 15))
(X3E <- maximum.E(p[3], cf[c(4,9)], 10))
(X4E <- maximum.E(p[4], cf[c(5,10)], 0.5))
(X5E <- maximum.E(p[5], cf[c(6,11)], 5))
## original scale dosages
(D1E <- X1E*15 + 15); (D2E <- X2E*15 + 15); (D3E <- X3E*10 +
20)
(D4E <- X4E*0.5 + 0.5); (D5E <- X5E*5 + 45)

#### 2.9.2. Predicted economical response
XE <- c(1, X1E, X2E, X3E, X4E, X5E, X1E^2, X2E^2, X3E^2,
X4E^2, X5E^2)
(yE <- sum(cf2*XE))

#### 2.9.3. Expected profit
XP <- c(-D1E, -D2E, -D3E, -D4E, -D5E, yE)
(yP <- sum(X

```