FRACTIONAL FACTORIAL DESIGNS 1/32 (4^5) IN BLOCKS OF SIXTEEN UNITS

Armando Conagin[1], Luís Alberto Ambrósio[2]

[1]Scientific Researcher VI, retired, Instituto Agronômico - APTA, Campinas, SP, Brasil. CP, 28 – CEP 13001-970.
[2]Scientific Researcher V, Centro de Bovinos de Leite, Instituto de Zootecnia - APTA, Nova Odessa, SP, Brasil. CP. 60, CEP 13460-000, ambrosio@iz.sp.gov.br.

ABSTRACT

The full factorial and fractional factorial designs are widely used in the several areas of experimental sciences. The use of three or more levels in the factorial design is common in the agronomic research, especially in fertilizer studies. This paper presents a set of 16 types of fractional factorial 1/32(4^5), with 32 treatments, in blocks of sixteen units, allowing the balance of each level of each factor with the levels of the other factors inside each block. For the complete quadratic model, covariances are present; if not include the linear interactions between factors the design is orthogonal. This design may be very useful in soils with low fertility, for example the Savanna (Cerrado) soils, in the initial basic research for nutritional requirements of different crops as maize, peanut, rice, soybeans, sugarcane, fruitculture, forestry, and pasture. Therefore it allows the simultaneous studies of N, P, K, Ca/Mg and micronutrients and enables economic analysis of the responses of the fertilizers, lime and micronutrients to guarantee an adequate productivity. Other agronomic informations as spacing, populations, fractionality of nitrogen, etc may be associated with the nutritional requirements in the research. Instead of 1024 different treatments for the complete factorial it needs only 32 treatments to give the basic information required.

Key words: statistical method, design of experiment, fractional factorial design, surface response.

DELINEAMENTO FATORIAL FRACIONADO 1/32 (4^5) EM BLOCOS DE 16 UNIDADES

RESUMO

Os delineamentos com esquemas fatoriais completos ou fracionados são amplamente usados nas várias áreas das ciências experimentais. O uso de três ou mais níveis nos fatores é comum na pesquisa agronômica, especialmente em ensaios de adubação. Este trabalho apresenta 16 tipos de fatoriais fracionados 1/32(4^5) com 32 tratamentos, em 2 blocos de 16 unidades, permitindo o balanceamento de cada nível de cada fator com os níveis dos demais fatores, dentro de cada bloco. No modelo quadrático completo aparecem covariâncias. O delineamento fica ortogonal se não forem incluídas as interações lineares entre fatores. Este delineamento pode ser muito útil na fase inicial de pesquisas sobre nutrição de plantas em solos de baixa fertilidade, por exemplo nos solos do Cerrado, para as culturas de milho, amendoim, arroz, soja, cana-de-açúcar, fruticultura, silvicultura e pastagens. Isto porque permite o estudo simultâneo de N, P, K, Ca/Mg e micronutrientes e possibilita a análise econômica das respostas dos fertilizantes, calcário e micronutrientes. Outros fatores agronômicos como espaçamento, densidade de plantas, etc podem ser associados aos fatores nutricionais nos ensaios. O uso do fatorial 4x4x4x4x4 completo

requereria 1024 tratamentos diferentes o que é inviável. Já o delineamento proposto 1/32 (4^5) requer 32 tratamentos em 2 blocos, o que é bem mais viável.

Palavras-chave: métodos estatísticos, delineamento experimental, experimento fatorial fracionado, superfície de resposta.

INTRODUCTION

Fisher (1947) and Yates (1937) proposed the use of factorial and fractional factorial types of design and presented a great number of different types. Cochran and Cox (1957) present in theirs book a great collection of factorial and fractional factorial designs.

Initially the fractional designs were extensively used in agronomic field research and mainly in fertilizer research as 2^2, 2^3, 2^4, and 2^5 designs with replications when necessary, to guarantee enough degrees of freedom for the error. In chemical and chemical industry where the treatments change the industrial process in use, experiments of the types 2^3 or 2^4 used only one replication and then the test of the linear factors using the interactions pooled as the residual error were used by Davies (1954); Bennett & Franklin (1954) presented many examples of these types for use in the chemical industry research.

Soon the statisticians and agronomist researchers started using factorials as 3x3, 3x3x3, 3x3x2, 3x3x4 etc in order to test the linear, quadratic and linear interactions effects.

These designs made possible the economic analysis of the experiments in order to obtain the best solutions. In 1951 Box and Wilson proposed a family of designs in the paper "On the experimental attainment of optimum conditions"; in 1952 Box published the paper "Multifactor design of first order" in which the model was a linear model with various factors. Later on Box and Hunter and others (1957) extended these type of central composite designs (CCD) producing CCD´s orthogonal, orthogonal divided in blocks, rotatable, orthogonal rotatable, etc in order to obtain an analysis of a second order design for various factors from 2 to 8 factors and more. These designs had been use in cases in which external factors as temperature, pressure etc, are better controlled in the laboratories, in the industry, etc. Myers (1971) presented in his book "Response Surface Methodology" a great number of different designs factorials, fractional factorials and different types of Central Composit Design, of first and second order types. Due to the ecological variation (climate, soil, rainfall, etc) the factorials and fractional factorials continued been used in agronomic experiments and in fertilizer research. In the psychology area Winer, Brown and Michels (1991) presented good examples in this field of research.

Haaland (1981) presented examples of use of factorials and fractional factorials with and without axial axes in the area of immunology and biotechnology research.

Aiming economic analysis of the fertilizer experiments the economist had pointed the necessity of use of more than tree levels of each factor to better describe the response surface and allow the determination of extreme points and the optimum economic response.

In Brazil, Andrade & Noleto (1986) presented a $1/2(4^3)$ group of design in 2 blocks of 16 treatments. Conagin et al. (1997) presented a design of $1/2(4^3)$ in 4 blocks of eight treatments. Conagin & Jorge (1982) presented types of $1/5(5^3)$. Recently Conagin & Ambrósio (2003 and 2006) proposed a group of $1/8(4^4)$ factorial design in two blocks. This paper presents 16 types of $1/32(4^5)$ fractional factorial design in 2 blocks of 16 different treatments.

MATERIAL AND METHODS

From a special group of 1/8(4^4), presented by Conagin & Ambrosio (2006) was derived 16 types of 1/32(4^5) in two blocks of sixteen units. They are presented in Table 1.

Yet, by permutation of the four levels of the fifth factor and permutation of columns is possible to increase greatly the number of these design; nevertheless, we believe that this sixteen design presented in Table 1 permit a good fit in the analyses of multiple regression, the economic analyses in fertilizer research. Due to the small numbers of treatments the interactions are correlated with the linear terms. How interactions in fertilizer experiments are less important than the linear and quadratic terms we prefer to recommend reduced model without interactions.

The parametric statistical model to 4^5 factorial design in randomized blocks without interaction is:

$$Y = \beta_0X_0 + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4 + \beta_5X_5 + \beta_{11}X_{11} + \beta_{22}X_{22} + \beta_{33}X_{33} + \beta_{44}X_{44} + \beta_{55}X_{55} + \beta_bX_b + \varepsilon$$

The levels adopted for fertilizer studies had been 0, 1, 2, and 3, but in other circumstances in which all levels may be present the level may be changed to 1, 2, 3, and 4; for the linear coefficients the levels used were: -1.5, -0.5, +0.5, and +1.5 with $\lambda_1 = 1$; for the quadratic coefficients the levels were: +1, -1, -1, +1, in which $\lambda_2 = 1$; for the two blocks the coefficients were -0.5 and +0.5.

RESULTS AND DISCUSSION

Several experiments was simulated using the design (a) in which we introduced coefficients of variation of 5.0%, 8.0%, 11.0% and 14.0%; we analyzed the experiments by proc REG and proc RSREG of SAS (1990).

The analysis of variance of the experiment (a) with the model without interaction with coefficient of variation (CV) of 5%, is presented below, as an example.

The stationary point obtained was a maximum.

To evaluate the performance of the design we did the analysis using the adapted basic parameter and added the random part utilizing values with CV of 5.0%, 8.0%, 11.0% and 14.0%. The values obtained showed growing discrepancies between the calculated and the estimated values from the regression when coefficient of variation increased; the results are presented in Table 6. Below CV of 8.0% all experiments showed satisfactory approximation between calculated and estimated values and produced points of maximum. For CV of 11.0% and 14.0% the discrepancy increased and appeared saddle points.

This alert us to use satisfactorily this designs for laboratory, greenhouse, screening fertilizer experiments and in all cases in experimental sciences when the expected CV normally is low and bellow CV of 10%. For higher CV's saddle points appear more frequently.

To calculate the regression of the experiments we may use SAS, Minitab, and other statistical software. In brazilian literature we may use SANEST and Pimentel-Gomes & Garcia (2002), for example.

Besides the possible utilization in nutritional research in plants and animals, in fertilizer experiments for studies in poor soils (Cerrado, Savanna), it may be used in other areas, in metallurgical research, in studies viewing to get special types of steel, in plastic research aiming new products, in cement research and others areas in which several factors with four levels are desirable, and in which only β_i's and β_{ii}'s effects are of major interest.

Table 1. Types of fractional factorial design $1/32(4^5)$.

a	b	c	d	e	f	g	h
			Block I				
00030	30000	03000	11100	00300	01110	10110	11010
12310	11230	31120	23020	23110	22300	02230	30220
23120	22310	12230	30230	31220	33020	23300	02330
31200	03120	20310	02310	12030	10230	31020	23100
01111	10111	11011	12221	11101	21221	22121	22211
13231	31321	23131	20301	32311	02031	30201	03021
22001	02201	00221	33111	20021	13311	11331	31131
30321	23031	32301	01031	03231	30101	03011	10301
02222	20222	22022	13332	22202	31332	33132	33312
10102	01012	10102	21212	01012	12122	21212	12122
21332	32132	33212	32002	13322	03202	00322	20032
33012	13302	01332	00122	30132	20012	12002	01202
03303	00333	30033	10013	33003	11003	01103	00113
11023	21103	02113	22133	10213	32213	13223	21323
20213	12023	21203	31323	02123	23133	32313	13233
32133	33213	13323	03203	21333	00323	20033	32003
			Block II				
22212	12222	21222	33322	22122	23332	32332	33232
30132	33012	13302	01202	01332	00122	20012	12002
01302	00132	30012	12012	13002	11202	01122	20112
13022	21302	02132	20132	30212	32012	13202	01322
23333	32333	33233	30003	33323	03003	00303	00033
31013	13103	01313	02123	10133	20213	12023	21203
00223	20023	22003	11333	02203	31133	33113	13313
12103	01213	10123	23213	21013	12323	21233	32123
20000	02000	00200	31110	00020	13110	11310	11130
32320	23230	32320	03030	23230	30300	03030	30300
03110	10310	11030	10220	31100	21020	22100	02210
11230	31120	23110	22300	12311	02230	30220	23020
21121	22111	12211	32231	11221	33221	23321	22331
33201	03321	20331	00311	32031	10031	31001	03101
02031	30201	03021	13101	20301	01311	10131	31011
10311	11031	31101	21021	03111	22101	02211	10221

Table 1. Types of fractional factorial design 1/32(4^5).

i	j	k	l	m	N	o	p
Block I							
00130	30010	13000	01300	11200	01120	20110	12010
12210	11220	21120	22110	23320	22330	32230	33220
23320	22330	32230	33220	30030	33000	03300	00330
31000	03100	00310	10030	02110	10210	11020	21100
01311	10131	31011	13101	12021	21201	02121	20211
13031	31301	03131	30311	20101	02011	10201	01021
22101	02211	10221	21021	33211	13321	21331	32131
30221	23021	22301	02231	01331	30131	33011	13301
02022	20202	02022	20202	13132	31312	13132	31312
10302	01032	30102	03012	21012	12102	01212	10122
21232	32122	23212	12322	32302	03232	30322	23032
33112	13312	11332	31132	00222	20022	22002	02202
03203	00323	20033	32003	10313	11033	31103	03113
11123	21113	12113	11213	22233	32223	23223	22323
20013	12003	01203	00123	31123	23113	12313	11233
32333	33233	33323	23333	03003	00303	00033	30003
Block II							
22312	12232	31222	23122	33022	23302	02332	30232
30032	33002	03302	00332	01102	00112	10012	11002
01102	00112	10012	11002	12212	11222	21122	22112
13222	21322	22132	32212	20332	32032	33202	03322
23133	32313	13233	31323	30203	03023	20303	02033
31213	13123	21313	12133	02323	20233	32023	23203
00323	20033	32003	03203	11033	31103	03113	10313
12003	01203	00123	20013	23113	12313	11233	31123
20200	02020	20200	02020	31310	13130	31310	13130
32120	23210	12320	21230	03230	30320	23030	32300
03010	10300	01030	30100	10120	21010	12100	01210
11330	31130	33110	13310	22000	02200	00220	20020
21021	22101	02211	10221	32131	33211	13321	21331
33301	03331	30331	33031	00011	10001	01001	00101
02231	30221	23021	22301	13301	01331	30131	33011
10111	11011	11101	01111	21221	22121	22211	12221

Table 2. Analysis of the Regression Model without interactions.

Source of variation	D.F.	S. Square	M. Square	F test	Prob.
Model	10	5712847	571285	11.24	0.000
Error	21	1067140	50816		
Corrected Total	31	6779987			

Mean = 4241.4; Root MSE = 225.4; R square = 0.843; Adj. R Sq. = 0.768

The test of parameter estimates are presented in Table 3.

Table 3. Test of Parameters estimates for example with CV equal to 5% with the simulated values presented in Table 6.

Variable	DF	Parameter Estimate	Error est.	t for H_0: $\beta=0$	Prob. > | t |
Intercept	1	4241.38	38.85	106.43	0.000
X_1	1	195.00	35.64	5.49	0.000
X_2	1	184.60	35.64	5.18	0.000
X_3	1	69.90	35.64	1.96	0.063
X_4	1	121.85	35.64	3.42	0.002
X_5	1	131.33	35.64	3.68	0.001
X_{11}	1	-110.62	39.85	-2.78	0.011
X_{22}	1	-99.94	39.85	-2.51	0.020
X_{33}	1	-39.25	39.85	-0.99	0.335
X_{44}	1	-98.50	39.85	-2.47	0.022
X_{55}	1	-91.19	39.85	-2.29	0.032

In this simulated experiments we ignored the blocks effect.

The matrix **X´X** inverse had the coefficients (C):$C_{0,0} = 0.03125$, $C_{1,1} = C_{2,2} = C_{4,4} = C_{5,5} = 0.025$, $C_{11,11} = C_{22,22} = C_{33,33} = C_{44,44} = C_{55,55} = 0.03125$. All $Cij = Cji = 0$.

The matrix of covariances (Cov) of estimates was: Cov_{00}=1,508.01; $Cov_{1,1} = Cov_{2,2} = Cov_{3,3} = Cov_{4,4} = Cov_{5,5}$ = 1,270.41; $Cov_{11,11} = Cov_{22,22} = Cov_{33,33} = Cov_{44,44} = Cov_{55,55}$ = 1,588.01.

Using RSREG from SAS (1990) that uses the complete model, the results are presented in Table 4.

Table 4. Analysis of Regression of the complete model.

Source of variation	D.F.	S. Square	R. Square	F test	Prob.
Linear	5	5826447	0.781	317.70	0.000
Quadratic	5	1541667	0.2067	84.06	0.000
Cross Product	10	50794	0.0068	1.385	0.300
Regress	20	7418908	0.9946	101.1	0.000

Doing the canonical analysis we obtained the eigenvalues and eigenvectors presented in Table 5.

Table 5. Canonical Analysis of Response Surface (based on coded data).

	Eigenvectors				
Eigenvalues	X_1	X_2	X_3	X_4	X_5
-164.752966	0.053407	-0.036744	-0.325956	-0.003877	0.943152
-171.827929	-0.110884	0.394312	0.850510	-0.097564	0.315178
-207.266014	0.71330	-0.576269	0.315922	-0.239188	0.045357
-250.530500	0.681370	0.696193	-0.195412	0.081715	-0.078659
-331.806366	0.108396	-0.162476	0.179982	0.962589	0.053692

Table 6. Results of experiments that were calculated from the model with CV of 5.0%, 8.0%, 11.0% and 14.0%; the estimated values were obtained from the regression and which the model utilized are without interactions.

Treatments	Parameters Values	CV = 5%		CV = 8%		CV = 11%		CV = 14%	
		Calculated	Estimated	Calculated	Estimated	Calculated	Estimated	Calculated	Estimated
Block I									
00030	3100	3146	3112	3507	3388	3289	3150	3387	2917
12310	4350	4183	4261	4498	4170	5171	4565	4855	4823
23120	4550	4400	4502	4352	4599	4457	5014	5227	5096
31200	3900	4089	3937	4194	3955	2900	3291	3722	3842
01111	3875	3781	3912	3508	3900	3987	3983	4147	3504
13231	4750	4745	4615	4697	4674	5061	5082	4339	4682
22001	3900	4332	4242	4355	3986	3521	3436	4142	3715
30321	4375	4461	4298	4707	4712	4658	4446	3706	4143
02222	4500	4532	4420	4433	4407	5245	4907	4052	4179
10102	3550	3774	3757	3780	3590	3649	3109	3097	3194
21332	4675	4671	4764	4637	4739	4235	4597	5460	4831
33012	4475	5121	4652	4501	4804	4762	4748	4703	4049
03303	3900	4236	3904	3975	3736	4483	4262	3477	3392
11023	4225	4390	4383	3869	4150	4587	4219	3846	4547
20213	4390	4196	4291	4073	4213	4654	4185	5102	4556
32133	4925	4596	4811	4531	4818	4413	4587	4837	4811
Block II									
22212	4925	4940	4911	5035	4738	4705	4863	5614	4972
30132	4350	4224	4293	4925	4604	3949	3963	4064	3868
01302	3700	3514	3786	3510	3625	3158	3534	2275	3004
13022	4425	4225	4603	4853	4573	4496	5000	3737	4408
23333	5075	4892	4883	4964	4850	5573	5324	5239	5218
31013	4275	4260	4431	4451	4381	4302	3967	4041	4189
00223	3925	3876	3800	4240	3882	3540	4229	3436	3763
12103	4125	4393	4275	3531	3804	3125	3733	4469	4137
20000	3075	3337	3360	3229	3327	2248	2553	2829	3194
32320	4700	4523	4554	4876	4792	4549	4866	5001	4981
03110	3825	3641	3768	3955	3877	4673	4505	3949	3787
11230	4300	4245	4132	3810	4117	4582	4096	4553	4717
21121	4600	4989	4647	5067	4622	4603	4492	3866	4814
33201	4350	4069	4420	4180	4513	4765	4277	3256	3807
02031	3925	4023	3995	3845	4047	4076	4033	3623	3438
10311	4025	3920	4006	3491	4090	3736	4145	4512	3985

CONCLUSIONS

The 16 proposed designs permit the study of 5 factors in 4 levels.

The level of each factor is present in 8 treatments and is balanced to the levels of the other factors.

In the reduced model (without interactions) each coefficient of the model is tested with 20 degrees of freedom by the t test with satisfactory precision; in case of the complete model (with interaction) we have 10 degrees of freedom for the residual, and some covariances may be of high values.

The designs present a better spatial distribution of the 32 points that permit a better determination of the shape of the response surface.

The complete factorial requires 1024 different treatments that turn out the complete design, impracticable; the 32 treatments with better spatial spread of the treatments than other possible combinations are easily performed.

If the interactions are absent, the matrix **X´X,** the inverse and the covariances are diagonal. With the complete model exist covariances among the factors.

BIBLIOGRAPHIC REFERENCES

ANDRADE, D. F.; NOLETO, A. 1986. Exemplos de fatoriais fracionados (1/2) 4^3 para ajuste de modelos polinomiais quadráticos. **Pesquisa Agropecuária Brasileira**, 21(6): 677-680.

BENNETT, C.A.; FRANKLIN, N.L. 1954. **Statistical Analysis in Chemistry and the Chemical. Industry**. New York: John Wiley & Sons, Inc. 724 p.

BOX, G. E. P. 1952. Multifactor design of first order. **Biometrika,** 39:49-57.

BOX G. E. P.; HUNTER J. S. 1957. Multifactor experimental designs for exploring response surfaces. **Annals of Mathematical Statistics**,28.

BOX, G.E.P.; WILSON, K. B. 1951. On the experimental attainment of optimum conditions. **Journal of the Royal Statistical Society**, B, Vol. 13:1-45.

COHRAN, W.G.; COX, G..M.1957. **Experimental designs**. 2ª Ed. New York, John Wiley, 611p.

CONAGIN, A.; AMBRÓSIO, L. A. 2003. Delineamentos $1/8(4^4)$ em Blocos de 16 unidades. Piracicaba, SP. **Revista de Agricultura**, v.78(2), 169-180.

CONAGIN, A.; AMBRÓSIO, L. A. 2006. Delineamentos $1/8(4^4)$ em blocos de 16 unidades. **Revista de Agricultura** (Piracicaba), v. 81, p. 187-201.

CONAGIN, A.; JORGE, J. P. N. 1982. Delineamentos 1/5(5x5x5) em blocos. **Bragantia**, 41:156-168.

CONAGIN, A.; NAGAI, V.; IGUE, T. 1997. Delineamento 1/2(4x4x4) em blocos de oito unidades. Instituto Agronômico, Campinas, São Paulo, Brasil. **Boletim Científico,** 36.

DAVIS, O.L. 1954. **Design and analysis of industrial experiments**. Oliver & Boyd and Hafner Publishing Company, New York, 636p.

FISHER, R.A. 1947. **The Design of Experiments**. Hafner Publ. Co. Inc., New York, 240p.

HAALAND, P. D. 1989. **Experimental Design in Biotechnology.** New York, Marcel Dekker, 259p.

MYERS, R.H. **Response Surface Methodology**, Allyn and Bacon, Boston, Mass, 246p.

MINITAB. 2000. **User's Guide 2: Data Analysis and Quality Tools. Release 13 for Windows**. Minitab Inc. USA.

PIMENTEL-GOMES, F.; CONAGIN, A. 1991. Experimentos de adubação. Planejamento e análise estatística. In OLIVEIRA, A. J. DE; GARRIDO, W.E.; ARAUJO, J.D. De; LOURENÇO, S. Coordenadores. **Métodos de Pesquisas em**

Fertilidade do Solo. Brasília, EMBRAPA-SEA, 392 p.

PIMENTEL-GOMES, F.; HENRIQUE GARCIA, C. 2002. **Estatística Aplicada a Experimentos Agronômicos e Florestais** (Com Uso de Programas SAS e SANEST). Biblioteca de Ciências Agrárias "Luiz de Queiroz", Vol. 11, FEALQ, 309p.

PIMENTEL-GOMES, F. 2000. **Curso de Estatística Experimental.** 14ª Edição, ESALQ-USP, Piracicaba, 477p.

SAS INSTITUTE. 1990. **SAS/STAT User's Guide: Statistics**. Release.6.04. Cary, N.C., USA. SAS Institute Inc, 1686p.

SAS INSTITUTE. 1999. **SAS/STAT User's Guide: Statistics**. Release 8. Cary, N.C., USA. SAS Institute, Inc.

WINER, B.J.; BROWN, D.R.; MICHELS, K.M. 1991. **Statistical Principles in Experimental Design.** McGraw-Hill Inc., 1057p.

YATES, F. 1937. **The Design and Analysis of Factorial Experiments**. Commonwealth Bureau of Soil Science, Hafner Publ. Co. England, 93 p. (Technical Communication, 35).

ZONTA, E. P.; MACHADO, A. A. 1987. **SANEST - Sistema de análise estatística para microcomputadores**. Pelotas: DMEC/IFM/UFPel, 138p.