BRAGANTIA

Boletim Técnico da Divisão de Experimentação e Pesquisas

INSTITUTO AGRONÔMICO

Vol. 12 Campinas, Julho-Setembro de 1952 N.os 7-9

ANÁLISE DE UMA EXPERIÊNCIA DE ROTAÇÃO [1]

FRANK YATES

Sc. D. — F.R.S., Chefe do Departamento de Estatística da Estação Experimental de Rothamsted — Inglaterra

1 - INTRODUÇÃO

Quando um experimento de rotação de culturas inclui rotações de duração diversa, de forma que não estejam representadas tôdas as fases de cada uma dessas rotações, as comparações entre as produções médias de uma cultura individual poderão estar afetadas pela variação anual da produção dessa cultura. Como exemplo, consideraremos um experimento planejado para a comparação do comportamento do milho e do algodão, plantados contìnuamente, com o de uma rotação bienal de milho com algodão. Dentre os canteiros incluídos em cada um dos blocos, um é cultivado contìnuamente com milho, outro com algodão, e um terceiro é ocupado pelas culturas de milho e algodão em anos alternados. Esquemàticamente, teremos :

Ano		*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Sequência de culturas	1	A	A	A	A	A	A	A	A	A	A
	2	M	M	M	M	M	M	M	M	M	M
	3	M	A	M	A	M	A	M	A	M	A

onde : A = algodão ; M = milho.

Caso a vantagem da rotação venha a ser avaliada pela comparação das produções médias do algodão e do milho nas seqüências 3 com 1 e 3 com 2, respectivamente, essas comparações serão afetadas pela variação anual na produção da cultura considerada. Assim, se os anos ímpares 1, 3, etc., fôssem, em média, anos de má produção tanto para o milho como para o algodão, pareceria que o comportamento do milho, na rotação, se teria mostrado pior relativamente ao milho plantado continuadamente do que realmente foi. Por outro lado, o resultado obtido para o algodão pareceria melhor do que aquêle que realmente deveria ser. No caso em que fôssem os anos ímpares desfavoráveis ao algodão e, os pares, ao milho, o resultado da rotação seria aparentemente melhor do que o que realmente foi, em relação às culturas contínuas.

[1] Êste trabalho foi executado durante uma visita realizada a Campinas, patrocinada pela Fundação Rockefeller e Fundo de Pesquisas do Instituto Agronômico. O autor agradece ao Eng.º Agr.º C. G. Fraga Jr. e Prof. W. L. Stevens, que verificaram os detalhes e acompanharam a impressão do mesmo. A versão portuguêsa foi feita por C. G. Fraga Jr.

Essa dificuldade, em uma experiência dêsse tipo, pode ser contornada de maneira bastante simples, comparando-se a produção média do milho na seqüência 3 com a produção média dos anos ímpares na seqüência 2 e, de maneira semelhante, usando, para o algodão, os anos pares da sequência 1. Por êste processo, conseguiremos extrair tôda a informação contida no experimento, nos casos em que nosso interêsse consista em comparar canteiros contendo uma mesma cultura, em um mesmo ano. Convém, no entanto, notar que o experimento seria muito melhor, caso tivesse sido incluída a outra fase da seqüência 3, isto é, a seqüência AMAM..., pois isso tornaria possível fazer, todos os anos, comparações com as duas séries de culturas continuadas. Não seria de esperar que essa modificação resultasse sòmente em uma redução do êrro experimental, pois ela também permitiria comparar o comportamento das duas culturas, todos os anos, de maneira que as variações nos diferentes anos (devidas às condições climáticas durante o ano), que favoreceriam a rotação relativamente às culturas permanentes, seriam eliminadas com maior rapidez.

Realmente, a regra de que tôdas as fases de cada uma das rotações devem ser incluídas no experimento constitui, atualmente, cousa consagrada no delineamento dos experimentos de rotação. O delineamento dêsses experimentos foi discutido por Stevens (1).

Neste trabalho, descrevemos a análise de um experimento no qual não foram representadas tôdas as fases das rotações, e onde não é aplicável o processo simples para contornar a dificuldade apresentada no parágrafo anterior. Essa discussão não deve ser considerada como uma justificativa para o uso, em futuros experimentos, de um delineamento dêste tipo, o qual aumenta excessivamente os cálculos necessários à análise dos resultados. Além disso, é ainda de maior importância considerar que um bom delineamento melhora muito a eficiência do experimento, a exatidão, a clareza e a compreensão das conclusões.

Devido ao longo tempo necessário à execução de um experimento de rotação, importante se torna extrair dos experimentos, em andamento ou já concluídos, tôdas as informações possíveis, mesmo nos casos em que seu delineamento pudesse ter sido muito melhor.

2 - DESCRIÇÃO DO EXPERIMENTO

O experimento considerado neste trabalho foi planejado com o objetivo de investigar o efeito de diversas rotações de culturas na conservação da fertilidade [1]. Foi instalado na Estação Experimental Central (fazenda Santa Elisa), em Campinas, durante os anos agrícolas de 1936/7 a 1943/4. Foram incluídas dez rotações diferentes, abrangendo as culturas de algodão, milho, feijão e leguminosa (esta última cultivada para produção de massa), assim como culturas continuadas de milho e algodão. Essas rotações acham-se discriminadas no quadro 1. No ensaio foram incluídas quatro repetições de uma única fase de cada uma das rotações. A distribuição dos canteiros nas repetições foi sistemática ; mas, na análise que se segue, foi considerada como sendo em blocos casualizados.

[1] O presente ensaio foi iniciado pela antiga Secção de Experimentação do Serviço de Algodão e prosseguido pela atual Secção de Algodão do Instituto Agronômico de Campinas.

QUADRO (*Table*) 1.—Rotações (*Rotations*)

Tratamentos (*treatments*)	Anos Agrícolas (*crop years*)							
	1936/7	37/38	38/39	39/40	40/41	41/42	42/43	43/44
1	A	A	A	A	A	A	A	A
2	M	M	M	M	M	M	M	M
3	M	A	M	A	M	A	M	A
4	L	A	L	A	L	A	L	A
5	M	L	M	L	M	L	M	L
6	L	A	M	L	A	M	L	A
7	L	M	A	L	M	A	L	M
8	I	A	I	A	I	A	I	A
9	A	F_1	A	F_1	A	F_1	A	F_1
10	F_1	M	F_1	M	F_1	M	F_1	M
11	F_2	A	F_2	A	F_2	A	F_2	A
12	M	F_2	M	F_2	M	F_2	M	F_2

A — Algodão (*cotton*). M — Milho (*maize*). L — Leguminosa: Feijão de porco (*legume*). I — Milho e Mucuna (*mixture of maize and mucuna*). F_1 — Feijão das águas (*beans sown in rainy season*). F_2 — Feijão das águas e da sêca (*beans sown in rainy and dry seasons*).

As produções anuais de algodão e milho (totais de 4 canteiros) constam nos quadros 2 e 3. Pela observação dêsses quadros é possível concluir que as médias, tanto do algodão como do milho, para o total de anos, nas diversas rotações, estão afetadas por variações de produção devidas a diferenças entre anos. Assim, 1939/40 foi um ano mau, tanto para o algodão como para o milho e, como conseqüência, ficaram reduzidas as produções médias de algodão nas rotações 1, 3, 4, 8 e 11 e as de milho nas rotações 2 e 10, em relação às rotações restantes.

QUADRO 2.—Produções de algodão (*yields of cotton*). Totais de quatro canteiros, em kg (*totals of four plots, kg*). Área do canteiro (*area of one plot*) $=160$ m²

Rotação (*rotation*)	Anos (*years*)								Média (*mean*) 37/8-43/4	Média ajustada (*adjusted mean*)
	1936/7	37/38	38/39	39/40	40/41	41/42	42/43	43/44		
1 A (Contínuo)	54	92	74	24	75	67	58	45	62.1	60.9
3 MA	–	92	–	38	–	62	–	60	63.0	66.7
4 LA	–	96	–	48	–	63	–	67	68.5	72.2
6 MLA	–	88	–	–	83	–	–	69	80.0	72.2
7 LMA	–	–	75	–	–	60	–	–	67.5	62.6
8 IA	–	96	–	55	–	78	–	74	75.8	79.5
9 AF_1	41	–	76	–	72	–	72	–	76.7	65.6
11 F_2A	–	71	–	34	–	67	–	49	55.2	59.0
		±7.6	±5.5	±3.8	±4.7	±4.5	±3.8	±5.2		

Os erros-padrão incluídos nos quadros 2 e 3 foram calculados separadamente a partir das produções de cada ano. Fornecem, por conseguinte, estimativas válidas do êrro (salvo o que se refere à falta de casualização) para a comparação das rotações em um único ano. Como uma mesma rotação ocupou os mesmos canteiros durante tôda a experiência, os erros que afetam uma rotação em diferentes anos estarão possìvelmente correlacionados, e, por conseguinte, as diferenças entre um mesmo par de rotações,

QUADRO 3.—Produções de milho (*yields of maize*). Totais de quatro canteiros, em kg (*totals of four plots, kg*). Área do canteiro (*area of one plot*) = 160 m²

Rotação (*rotation*)	Anos (*years*)								Média (*mean*) 37/38-43/4	Média ajustada (*adjusted mean*)
	1936/7	37/38	38/39	39/40	40/41	41/42	42/43	43/44		
2 M (Contínuo)	232	182	185	116	145	123	134	110	142	145
3 AM	225	—	187	—	192	—	161	—	180	179
5 LM	221	—	199	—	219	—	184	—	201	199
6 LAM	—	—	207	—	—	215	—	—	211	212
7 ALM	—	220	—	—	215	—	—	272	236	227
8 IA	130	—	160	—	171	—	186	—	172	171
10 F_1M	—	190	—	126	—	170	—	175	165	172
12 F_2M	238	—	168	—	182	—	153	—	168	166
		±10.9	±8.4	±10.5	±9.6	±7.9	±8.0	±16.4		

em anos diferentes, estão até certo ponto afetadas pelos mesmos erros e não podem ser consideradas como variaveis independentes.

A estimação dos diversos componentes do êrro será discutida adiante. Por enquanto, será suficiente notar que ao menos algumas das diferenças de produções, encontradas nos quadros 2 e 3, representam efeitos reais. A diferença na produção de algodão na rotação milho-algodão e na de algodão plantado continuadamente, no ano agrícola de 1943/4, por exemplo, é de 15 kg, o que é 2,1 vêzes o êrro-padrão ($5{,}1 \sqrt{2} = 7{,}2$). De maneira semelhante, a diferença das produções correspondentes de milho, em 1942/3, é de 27 ± 11 kg. É, por conseguinte, visível a vantagem da rotação milho-algodão sôbre as culturas continuadas, nos últimos anos do experimento.

As comparações dêste tipo podem ser feitas com qualquer par de rotações que contenha, ao menos em alguns anos, uma mesma cultura. Quando isso se verifica, é em geral possível obter alguma informação adicional, mesmo nos anos em que as culturas nessas rotações são diferentes. O mesmo não se verifica com rotações que não possuam cultura comum em um mesmo ano ; estas não podem ser comparadas pelo processo indicado.

Um conjunto de estimativas, adequado a tôdas as rotações, pode ser obtido com o ajustamento de constantes pela utilização do método de quadrados mínimos. O processo adequado será indicado no próximo capítulo. Para efetuar comparações, incluímos as médias ajustadas resultantes da análise, as quais constam na última coluna dos quadros 2 e 3. Os ajustamentos resultaram em modificações substanciais das produções médias, particularmente no caso do algodão.

3 - ANÁLISE PELO PROCESSO DE QUADRADOS MÍNIMOS

Vamos supor que as produções das diversas rotações, nos diferentes anos, incluídas nos quadros 2 e 3, são formadas por um par de constantes aditivas, representando uma delas o efeito da rotação (considerado o mesmo todos os anos), ao passo que a outra representa o efeito do ano (considerado o mesmo para tôdas as rotações) mais um resíduo, o qual representa a diferença entre a produção obtida e a soma dessas constantes. A análise por quadrados mínimos conduz à obtenção de valores dessas constantes, os quais tornam mínima a soma dos quadrados dos resíduos, fornecendo, assim, estimativas das diferenças entre rotações. Estas estimativas serão mais precisas, quando os erros das produções das diversas rotações, nos diversos anos, forem distribuídos normal e independentemente.

Outra maneira de abordar êste assunto consiste em considerar a análise pelos quadrados mínimos resultando em ajustamentos das produções médias, de forma que forneçam uma compensação pelo fato de essas médias serem obtidas a partir de resultados encontrados em anos diferentes (quadros 2 e 3). Esta maneira de considerar torna fácil ver como devem ser obtidos os valores das constantes pelo processo das aproximações sucessivas (2). Êste processo é mais simples que o da obtenção das equações matemáticas, e ainda usualmente requer, com dados do tipo considerado, cálculos menos trabalhosos.

No caso em aprêço, uma inspeção dos resultados indica que as produções correspondentes ao milho e algodão, provenientes de plantações continuadas, parecem apresentar uma deterioração relativamente às rotações. Por essa razão, incluímos nas duas análises (para algodão e para milho) a constante *b*, a fim de representar a deterioração anual, de forma que seja tomada em consideração e estimada a magnitude dessa deterioração. Para o algodão, as constantes correspondentes às diversas rotações são r_1, r_2, etc., e as que representam os anos são s_1, s_2, etc., de forma que o esquema das constantes para o algodão é o representado no quadro 4 (indicamos ùnicamente a representação para as duas primeiras linhas, pois é óbvia a maneira de estendê-la às demais, o que é desnecessário para os cálculos).

QUADRO 4.—Esquema das constantes para o algodão (*Scheme of constants for cotton*)

Rotação (*Rotation*)	Anos (*Years*)						
	1	2	3	4	5	6	7
1	r_1+s_1+3b	r_1+s_2+2b	r_1+s_3+b	r_1+s_4	r_1+s_5-b	r_1+s_6-2b	r_1+s_7-3b
3	r_3+s_1	----------	r_3+s_3	----------	r_3+s_5	----------	r_3+s_7

Os cálculos estão indicados no quadro 5. Uma única casa decimal parece ser suficiente para os ajustamentos. A inspeção dos resultados permite verificar que as maiores diferenças são as devidas a anos e, por essa razão, iniciamos o cálculo por essas constantes, desprezando as outras. As constantes a considerar são simplesmente as médias anuais, as quais se acham, **com uma única decimal,** na linha 1, numerada no quadro 5. Seus valores devem ser cuidadosamente verificados. Calcula-se, em seguida, a coluna 2, adicionando-se as médias da linha 1, correspondentes às células ocupadas em cada uma das fileiras, e subtrai-se o total dessa linha.

Assim :

$$+37{,}6 = \text{soma da totalidade da linha } 1 - 435$$
$$+\ 3{,}9 = 89{,}2 + 39{,}8 + 66{,}2 + 60{,}7 - 252,$$

QUADRO 5.—Aproximações sucessivas na solução por quadrados mínimos, para algodão (*Sucessive approximations in the least square solution for cotton*)

		37/8	38/9	39/40	40/1	41/2	42/3	43/4		(2)	(3)	(6)	(7)	(10)	(11)	Total +2,3 b
Rot.	31	6	3	5	3	6	2	6	130	—59,9	+1,8	—12,0	+0,1	—2,9	+0,1	
1	7	92	74	24	75	67	58	45	435	+37,6	—5,4	+3,4	—0,5	+1,9	—0,3	—6,2 r_1
3	4	92		38		62		60	252	+3,9	—1,0	—1,8	+0,4	—0,9	+0,2	—0,4 r_3
4	4	96		48		63		67	274	—18,1	+4,5	—1,8	+0,4	—0,9	+0,2	+5,1 r_4
6	3	88			83			69	240	—13,4	+4,5	—1,9	+0,6	—0,1	+0,0	+5,1 r_6
7	2		75			60			135	+6,2	—3,1	+1,9	—1,0	+0,8	—0,4	—4,5 r_7
8	4	96		55		78		74	303	—47,1	+11,8	—1,8	+0,4	—0,9	+0,2	+12,4 r_8
9	3		76		72		72		220	—3,3	+1,1	+5,2	—1,7	+2,8	—0,9	—1,5 r_9
11	4	71		34		67		49	221	+34,9	—8,7	—1,8	+0,4	—0,9	+0,2	—8,1 r_{11}
		535	225	199	230	397	130	364	2080	+0,7		+1,4		+1,8		
	(1)	89,2	75	39,8	76,7	66,2	65	60,7								
	(4)	+11,1	—3,8	+3,0	+0,2	—3,7	—7,9	+0,3	—0,8							
	(5)	—1,8	+1,3	—0,6	—0,1	+0,6	+4,0	—0,0								
	(8)	+2,9	—2,4	+1,5	—1,6	—0,3	—3,0	+0,5	—2,4							
	(9)	—0,5	+0,8	—0,3	+0,5	+0,0	+1,5	—0,1								
Total		86,9 s_1	77,1 s_2	38,9 s_3	77,1 s_4	66,8 s_5	70,5 s_6	60,6 s_7								

etc. Note-se que as produções correspondentes às rotações 3, 4, 8 e 11, ocorrem nos mesmos anos e, por conseguinte, têm o mesmo total de médias. O total da coluna 2 (quadro 5), deveria ser zero, sendo a discrepância encontrada devida aos erros de aproximação no cálculo das médias. Seu total é de

$6(89,2) + 3(75) + \ldots - 2080 = +0,7$, o que fornece uma prova exata das operações.

As médias da coluna 3 foram obtidas a partir da coluna 2, dividindo-se pelo número de células em cada linha e **invertendo os sinais.** Também êstes resultados da coluna 3 deverão ser verificados cuidadosamente. As médias, nessa coluna, representam as primeiras aproximações aos valores das diferenças entre rotações, tomando em conta as diferenças entre anos. A deterioração de b, na rotação 1, é calculada usando-se a fórmula da regressão linear

$$b = -\frac{Sy(x - \bar{x})}{S\,(x - \bar{x})^2}$$

e tomando-se a diferença da regressão calculada a partir da linha 1 (+) e a das produções da rotação 1 (—). A soma de produtos assim obtida é de $+3(89,2) + 2(75) + 1(39,8) - 1(66,2) - 2(65) - 3(60,7) - 3(92) - 2(74) - 1(24) + 1(67) + 2(58) + 3(45) = -50,9$.

Êste valor é indicado na coluna 2. O divisor é $3^2 + 2^2 + 1^2 + 1^2 + 2^2 + 3^2 = 28$. Dividindo e trocando sinais, obtém-se $+1,8$, valor êste anotado na coluna 3.

Calculam-se agora os ajustamentos para as médias da linha 1, a partir dos valores da coluna 3. O múltiplo apropriado de **b** corresponde a cada um dos anos foi incluído. Assim, o total para o primeiro ano é de

$(+3)(+1,8) - 5,4 - 1,0 + 4,5 + 4,5 + 11,8 - 8,7 = +11,1$,

enquanto que, para o correspondente ao segundo ano, obtemos

$(+2)\,(+1,8) - 5,4 - 3,1 + 1,1 = -3,8$.

Escrevemos êsses totais na linha 4. Sua soma é igual a

$$7(-5,4) + 4(-1,0) + \ldots\ldots = -0,8.$$

Os ajustamentos da linha 5 são obtidos dividindo, como anteriormente, os resultados da linha (4) pelo número de células ocupadas e trocando os sinais. Estas são correções das médias anuais, e se tornaram necessárias devido à inclusão da regressão e às diferenças entre rotações. Assim, uma segunda aproximação à constante s_1, do primeiro ano, é dada por $89,2 - 1,8 = 87,4$.

O processo é agora repetido, fazendo-se uso dos ajustamentos da linha 5, a fim de obter os totais e médias das colunas 6 e 7. Os valores da coluna 7 são os ajustamentos, correspondentes aos valores da coluna 3, das diferenças entre rotações. Outra repetição completa dêste ciclo permite obter as linhas 8 e 9 e as colunas 10 e 11.

Os valores da coluna 11 são suficientemente pequenos para indicar a desnecessidade de maior aproximação. Os ajustamentos das colunas 3, 7 e 11, são agora somados, fornecendo os totais da última coluna, os quais constituem as estimativas finais das produções nas diferentes rotações.

Para a apresentação dos resultados, poderia ser conveniente adicionar a essas estimativas a produção média do experimento, 67,1.

De maneira semelhante, os totais das linhas 1, 5 e 9 fornecem as estimativas das produções médias anuais, livres do efeito das rotações.

As regras acima seguidas, em relação aos sinais, podem ser resumidas da forma seguinte :

Excetuando-se as médias iniciais, que constituem as primeiras aproximações ao primeiro conjunto de constantes, os totais a partir dos quais foram calculadas as primeiras aproximações e os ajustamentos subseqüentes tiveram, por conveniência, o sinal contrário ao dessas aproximações e ajustamentos. Assim, temos êsses resultados **mais a** respectiva função das primeiras aproximações ou dos ajustamentos subseqüentes para outros conjuntos de constantes, **menos** (quando necessário) a função correspondente das produções atuais. Por conseguinte, os **sinais são invertidos**, quando, para obter as aproximações e ajustamentos, calculamos as médias.

Convém notar que os cálculos poderiam ter sido um pouco abreviados pela fusão das linhas correspondentes às rotações 3, 4, 8 e 11, cujas produções ocorrem nos mesmos anos, calculando-se uma constante para a média dessas rotações. A tabela de operações, indicando os múltiplos das aproximações e ajustamentos a serem considerados (os quais poderiam substituir as produções no quadro 5), apresentaria o algarismo 4 nessa linha e 1 em todos os outros lugares correspondentes a produções. A soma de totais seria escrita na coluna correspondente aos totais e a constante obtida seria a média das constantes correspondentes a essas quatro rotações. Os desvios destas constantes, em relação à sua média, seriam os mesmos que os das rotações originais, relativos à média destas. Caso não houvesse necessidade de incluir a regressão, poderíamos reunir, de forma semelhante, os anos agrícolas de 1937/8 e 1943/4. No caso de tabelas extensas, possuindo um certo grau de simetria interna, as combinações de resultados do tipo acima considerado podem reduzir muito o trabalho a executar.

4 - CÁLCULO DO ÊRRO RESIDUAL

Em uma análise pelo processo de quadrados mínimos, a soma dos quadrados dos resíduos, após o ajustamento das constantes, fornece uma estimativa dos erros que afetam as observações. No próximo capítulo, discutiremos o significado dessa estimativa, pois, por enquanto, estamos ùnicamente interessados no processo de seu cálculo.

A soma de quadrados do resíduo, na análise por quadrados mínimos, é obtida a partir da fórmula geral

$$S(y^2) - b_1 S(x_1y) - b_2 S(x_2y) - \ldots\ldots \qquad (A),$$

na qual y representa os valores originais, b_1, b_2, ... as constantes ajustadas e x_1, x_2, ... os coeficientes dessas constantes nas expressões individuais correspondentes a cada um dos y. Assim, no quadro 4, o coeficiente de r_1 é $+1$ para todos os y da fileira 1 e zero para os outros lugares. Por conseguinte, $S(x_1y)$ correspondente a r_1 é igual ao total da fileira 1, isto é, 435. Também para s_1, $S(x_1y) = 535$, etc. De maneira semelhante, os coeficientes correspondentes à constante de deterioração b, são +3, +2, +1, 0, —1, —2, —3 para a primeira linha e zero em tôdas as outras, de forma que $S(x_by)$ é igual a +130. Por conseguinte, a redução na soma de quadrados, a qual inclui a correção usual na análise da variância para a média, é dada por ... —6,2(435) — 0,4(252) + 5,1(274) + ... + 86,9 + 535 + 77,1(225) + ... + 2,3(+130) = 148208,3.

A fórmula (A), a partir da qual essa quantidade foi calculada, sòmente é correta quando os valores de b são soluções de quadrados mínimos. Caso a solução tenha sido obtida por aproximações sucessivas, existem pequenos erros residuais, os quais não afetam substancialmente as estimativas das constantes, mas introduzem um êrro considerável na soma de quadrados, calculada a partir de (A). Desde que êsses erros são constantes para todos os s ou todos os r, é possível fazer um ajustamento. Caso $\bar{\delta}s$ represente a diferença da média ponderada dos valores s(+) para a média das produções originais (—), como os pesos são iguais ao número de observações em cada coluna, a quantidade a ser adicionada à redução, na soma de quadrados acima considerada, é $-\bar{\delta}$ sT, onde T é o total de tôdas as produções.

Neste caso,

$\bar{\delta}s = \left[6(86{,}9) + 3(77{,}1) + \ldots\right] / 31 - 67{,}09677 = +0{,}12581$ e, por conseguinte, $-0{,}12581\ (2080) = -261{,}7$ deve ser adicionado ao total acima obtido. Para os valores de **r**, a média ponderada é —0,18064 e, por conseguinte, a correção é +375,7.

A correção usual para a média é 139561,3 e a redução na soma de quadrados, devida ao ajustamento de outras constantes além da média, é 148208,3 — 261,7 + 375,7 — 139561,3 = 8761,0. Uma parte dessa soma de quadrados é devida à variação nas médias anuais, e a soma de quadrados a ela correspondente é calculada da maneira usual. (Caso sejam utilizadas as médias anuais, é necessário usar maior número de casas decimais do que as que constam do quadro 5.)

A análise da variância completa está representada no quadro 6. Por conveniência, os quadrados médios relativos ao milho também foram incluídos nesse quadro. As somas de quadrados foram divididas por 4, a fim de corresponderem a um canteiro como base. Os graus de liberdade para o conjunto de constantes ajustadas, no caso do algodão, são 14, isto é, 7 correspondentes ao número de anos menos um, o qual corresponde à média geral, mais 8 para as rotações menos 1, que representa um contraste redundante (existem sòmente 7 contrastes independentes), mais um para a deterioração.

A soma residual de quadrados, calculada por êsse processo, é de 169,8, valor êste não muito exato devido aos erros acima mencionados. Um valor exato pode ser obtido calculando-se os resíduos e fazendo-se a soma de seus quadrados e, apesar de não ser êste cálculo realmente necessário para a estimativa do êrro, fornece êle um bom contrôle da análise como um todo. Um outro exame dos resultados é, às vêzes, aconselhável, a fim de se verificar se não existem outros efeitos que deixaram de ser considerados na primeira análise.

QUADRO 6.—Análise da variância pelo método de quadrados mínimos (*Least squares analysis of variance*)

Fonte de Variação (*source of variation*)	Algodão (*cotton*)			Milho (*maize*)		
	Graus de liberdade (*degrees of freedom*)	Soma de quadrados (*sum of squares*)	Quadrado médio (*mean squares*)	Graus de liberdade (*degrees of freedom*)	Soma de quadrados (*sum of squares*)	Quadrado médio (*mean squares*)
Constantes (*constants*)	14	2190,2		14	7914,6	
Anos (*years*)	6	1843,0	307,2	6	2443,1	407,2
Rotações e deterioração (*rotations and deterioration*)	8	347,2	43,4	8	5471,5	683,9
Resto (*remainder*)	16	169,5	10,6	13	969,4	74,6
Total	30	2359,7		27	8884,0	
Resto (dos resíduos) (*remainder from residuals*)	16	161,7	10,11	13	954,5	73,42

Constam no quadro 7 os resíduos encontrados. Aquêle correspondente à rotação 1 (1937/8), por exemplo, é dado por

$$92 - 86{,}9 + 6{,}2 - 3(2{,}3) = +\ 4{,}4.$$

QUADRO 7.—Resíduos (*residuals*)

Tratamentos (*treatments*)	Anos (*years*)							Total
	37/8	38/9	39/40	40/1	41/2	42/3	43/4	
ALGODÃO (*cotton*)								
1	+4,4	—1,5	—11,0	+4,1	+8,7	—1,7	—2,5	+0,5
3	+5,5		—0,5		—4,4		—0,2	+0,4
4	+4,0		+4,0		—8,9		+1,3	+0,4
6	—4,0			+0,8			+3,3	+0,1
7		+2,4			—2,3			+0,1
8	—3,3		+3,7		—1,2		+1,0	+0,2
9		+0,4		—3,6		+3,0		—0,2
11	—7,8		+3,2		+8,3		—3,5	+0,2
Total	—1,2	+1,3	—0,6	+1,3	+0,2	+1,3	—0,6	+1,7
MILHO (*maize*)								
A — 2	+7,6	—11,0	—1,0	+6,5	+5,8	—19,4	+11,6	+0,1
B — 3		—7,1		—7,2		+14,3		0,0
C — 5		+1,6		—13,5		+12,0		+0,1
D — 6		+6,1			—6,0			+0,1
E — 7	+10,0			+18,0			—27,8	+0,2
F — 8		+12,2		+6,1		—18,4		—0,1
G — 10	—15,1		+1,7		—1,1		+14,1	—0,4
H — 12		—0,4		—9,5		+10,0		+0,1
Total	+2,5	+1,4	+0,7	+0,4	—1,3	—1,5	—2,1	+0,1

Pode-se verificar que as somas das fileiras e colunas como é de esperar são sempre aproximadamente nulas. O quociente da soma de quadrados, quando dividida por 4, é 161,7, o qual verifica satisfatòriamente o resultado obtido pelo processo usual.

A inspeção dos resíduos não revela a existência de efeitos adicionais, não incluídos na primeira análise.

5 - COMPONENTES DO ÊRRO

Os erros a que as produções dos canteiros estão sujeitas, podem subdividir-se em três componentes :

a) Um componente independente de um ano para outro e de um canteiro para outro.

b) Um componente que é constante em relação a anos, mas é independente de um canteiro para outro.

c) Uma variação real de um ano para outro, das diferenças entre rotações, a qual é atribuída ao efeito das diferenças climáticas nos diferentes anos, sôbre os efeitos das diversas rotações.

Esta subdivisão dos erros equivale a considerar os erros, nas produções dos canteiros, como sendo formados de três partes aditivas,

$$e_1 + e_2 + e_3,$$

as quais se distribuem independentemente, com variâncias v_1, v_2, v_3. O valor de e_1 pode variar para cada canteiro e para cada ano, ao passo que e_2, correspondente a um determinado canteiro, terá o mesmo valor todos os anos e e_3 terá um mesmo valor em todos os canteiros de uma dada rotação, em um determinado ano.

Os componentes a e b podem, teòricamente, ser suplementados por outros componentes, os quais representariam diferenças de um canteiro para outro, em relação à tendência linear ou aos componentes de grau mais elevado da regressão em anos. Os resultados práticos indicam que após a remoção do componente constante a, o resto da variação entre canteiros pode ser considerado virtualmente independente de um ano para outro, salvo, talvez, em experiências de duração muito longa.

QUADRO 8.—Análise da variância parcial (*Partial analysis of variance*)

Fonte de variação (*variation source*)	Algodão (*cotton*) Graus de liberdade (*degrees of freedom*)	Algodão (*cotton*) Quadrado médio (*mean square*)	Algodão (*cotton*) Valor esperado (*expectation*)	Milho (*maize*) Graus de liberdade (*degrees of freedom*)	Milho (*maize*) Quadrado médio (*mean square*)	Milho (*maize*) Valor esperado (*expectation*)
Blocos (*blocks*)	3	28,23	----------	3	123,14	----------
Rotação (*rotation*)	4	72,50	----------	4	218,64	----------
Rotações × blocos (*rotations × blocks*)	12	10,76	$v_1 + 4v_2$	12	31,09	$v_1 + 3v_2$
Anos (*years*)	3	528,06	----------	2	124,52	----------
Anos × blocos (*years × blocks*)	9	19,46	----------	6	69,92	----------
Anos × rotações (*years × rotations*)	----------	----------	----------	----------	----------	----------
Deterioração (*deterioration*)	1	17,62	----------	1	193,16	----------
Resíduo (*remainder*)	11	12,24	$v_1 + 4v_3$	7	58,14	$v_1 + 4v_3$
Ano × rot. × blocos (*years × rot. × bl.*)	36	6,15	v_1	24	10,79	v_1
Total	79	----------	----------	59	----------	----------

Quando, para uma cultura, existem produções de todos os canteiros, todos os anos, os diversos componentes podem ser isolados em uma análise da variância. Isto não é possível neste caso, mas uma análise parcial da variância pode ser feita para a cultura de algodão nas rotações 1, 3, 4, 8 e 11, as quais estão representadas nos quatro anos agrícolas de 1937/8, 1939/40, 1941/2 e 1943/4. Análise semelhante pode ser feita para o milho, abrangendo as rotações 2, 3, 5, 8 e 12, nos três anos 1938/9, 1940/1 e 1942/3. Os resultados dessas análises acham-se incluídos no quadro 8.

A fim de se tomar em consideração a deterioração devida à cultura contínua em relação às rotações, o grau de liberdade adequado foi isolado do componente anos x rotações. O quadrado médio é obtido a partir da tabela de "anos x rotações", usando-se os multiplicadores do quadro 9, a fim de calcular o valor Q.

QUADRO 9.—Multiplicadores para o cálculo da deterioração relativa (*Multipliers for calculating the relative deterioration*)

Algodão (*cotton*)					Milho (*maize*)			
Rotação (*rotation*)	37/38	39/40	41/42	43/44	Rotação (*rotation*)	38/39	40/41	42/43
1	+12	+4	—4	—12	2	+4	0	—4
3, etc.	—3	—1	+1	+3	3, etc.	—1	0	+1
Total	0	0	0	0	Total	0	0	0

A contribuição do ano 1937/8 para o algodão é equivalente a mais 12 vêzes a produção da rotação 1, menos 3 vêzes as produções das outras quatro rotações. Cada valor na tabela é o total da produção de quatro canteiros e, assim, para o algodão, a soma dos quadrados dos coeficientes dos canteiros individuais, incluídos em Q, é de

$$4(2)\left[12^2 + 4(3^2) + 4^2 + 4(1^2)\right] = 1600.$$

Encontrou-se, para Q, o valor 167,9, e a soma de quadrados é, por conseguinte, $167{,}9^2/1600 = 17{,}62$. Para o milho, o divisor é

$$4(2)\left[4^2 + 4(1^2)\right] = 160.$$

O quadrado médio no caso do algodão, relativo a "rotações x blocos x anos" é uma estimativa de v_1; o quadrado médio de "rotações x blocos", estima $v_1 + 4v_2$ (pois há quatro anos incluídos na análise), e, o quadrado médio de "rotações x anos" (excluindo-se o componente de deterioração) é uma estimativa de $v_1 + 4v_3$ (pois existem quatro canteiros para cada rotação). O mesmo se dá com o milho, salvo quanto ao quadrado médio correspondente a "rotações x blocos", o qual é uma estimativa de $v_1 + 3v_2$.

Torna-se, assim, clara a relação entre a análise da variância parcial, incluída no quadro 8, e a análise da variância por quadrados mínimos do quadro 6. Os 11 e 7 graus de liberdade, correspondentes ao resíduo dos componentes de "anos x rotações", são uma parte, respectivamente, dos 16 e 13 graus de liberdade dos componentes residuais encontrados no quadro 6.

Será melhor usar os valores do quadro 6 (dos resíduos) para estimar o componente 3 da variância. As estimativas dos três componentes da variância são, por conseguinte :

ALGODÃO	MILHO
$v_1 = 6{,}15$	$v_1 = 10{,}79$
$v_2 = 1{,}15$	$v_2 = 6{,}77$
$v_3 = 0{,}99$	$v_3 = 15{,}66$

Em relação ao componente 1, os componentes 2 e 3 são substancialmente maiores para o milho do que para o algodão.

6 - ERROS DAS COMPARAÇÕES ENTRE ROTAÇÕES

Caso uma dada rotação contenha uma determinada cultura durante *p* anos e o número de repetições, cada ano, seja de *q*, sendo usados os mesmos canteiros todos os anos, a variância do êrro, aplicável à média sôbre todos os *p* anos, e apropriada para comparações com outras rotações contendo a mesma cultura que a considerada na primeira rotação, todos os *p* anos, é de

$$\frac{1}{pq} v_1 + \frac{1}{q} v_2 + \frac{1}{p} v_3 = \frac{1}{pq} (v_1 + q v_3) + \frac{1}{q} v_2.$$

O quadrado médio do resíduo na análise por quadrados mínimos tem uma esperança igual a $v_1 + qv_3$. A variância da média de cada rotação deve, por conseguinte, ser aumentada de $v_2/4$, independentemente do número de anos que entrem no cálculo da média.

Os erros-padrão calculados desta forma constam no quadro 10, o qual inclui as médias ajustadas para o algodão e milho, a partir da análise por quadrados mínimos, convertidas para 1000 kg por hectare.

QUADRO 10.—Produções médias ajustadas (*Adjusted mean yields*). 1000 kg por ha

Rotação (*rotation*)	Algodão (*cotton*)	Milho (*Maize*)
1 A contínuo	0,95 ± 0,05	
2 M contínuo	—	2,27 ± 0,13
3 MA	1,04 ± 0,06	2,80 ± 0,17
4 LA	1,13 ± 0,06	—
5 ML	—	3,11 ± 0,17
6 LAM	1,13 ± 0,06	3,31 ± 0,23
7 LMA	0,98 ± 0,08	3,55 ± 0,17
8 IA	1,24 ± 0,06	2,67 ± 0,17
9 AF_1	1,03 ± 0,06	
10 F_1M		2,69 ± 0,16
11 F_2A	0,92 ± 0,06	
12 MF_2		2,59 ± 0,17

Nota : Para as abreviações, consultar o quadro 1 (*For key see table 1*)

Quando se tratar de rotações em que a cultura a considerar não ocorre sempre no mesmo ano nas duas rotações, êsses erros-padrão fornecem apenas limites inferiores para os erros-padrão verdadeiros. No caso especial em que uma das rotações inclui uma cultura um número maior de anos que a outra, mas de maneira a apresenta-la sempre que a outra também apresente essa cultura, o êrro-padrão das comparações será $\sqrt{2}$ vêzes o êrro-padrão da segunda rotação. Em outros casos, não existe outro processo rápido para a obtenção do êrro-padrão exato. Os erros-padrão para tôdas as comparações são dados pela inversão da matriz dos coeficientes nas equações normais (3). No entanto, o trabalho a ser realizado seria proibitivo numa experiência como esta. Podemos, para certas comparações de interêsse especial, obter limites superiores pelo método a ser descrito no próximo parágrafo.

Suponhamos que, por exemplo, desejamos o êrro-padrão da diferença das produções de algodão nas rotações 6 e 8. Vamos admitir, por enquanto, que cada uma das produções do quadro 5 seja independente e tenha uma variância residual de σ^2.

O algodoeiro ocorre nas duas rotações em 1937/8 e em 1943/4 e a variância da diferença das médias nesses dois anos é $(1/2 + 1/2)\sigma^2 = \sigma^2$, de forma que o pêso a ser considerado é de $1/\sigma^2$. A produção da rotação 6 em 1940/1 pode ser comparada com a da rotação 1 nesse ano e as produções da rotação 1 podem ser comparadas com as da rotação 8 em 1939/40 e 1941/2. Nenhuma dessas comparações foi incluída na comparação anterior, de forma que as duas comparações consideradas são independentes. A variância da segunda comparação é $(1 + 1 + 1/2 + 1/2)\,\sigma^2 = 3\sigma^2$, de forma que o pêso é de $1/3\sigma^2$ e o pêso da média ponderada das duas comparações é $4/3\ \sigma^2$. A variância é, por conseguinte, $3/4\ \sigma^2$ ou $0{,}75\ \sigma^2$. No caso em aprêço, o componente independente da variação é $v_1 + 4v_3$ por canteiro, de forma que σ^2 deve ser substituído por $(v_1 + 4v_3)/4$. Também existe o componente adicional da variância 2, o qual aumenta a variância da diferença de $2v_2/4$. Um limite superior para a variância é dado, por conseguinte, por

$$0{,}1625\ (v_1 + 4v_3) + 0{,}5v_2 = 2{,}47.$$

Para as comparações, especialmente importantes, quando é necessário obter a melhor estimativa possível do desvio-padrão, esta pode ser encontrada por meio de uma solução por aproximação sucessiva, semelhante à utilizada para as estimativas. Substituem-se, primeiramente, os totais das duas rotações a serem comparadas por +1 e por —1 e, por zero, os totais de tôdas as outras rotações, os totais anuais e o da regressão. A diferença entre os valores finais das duas rotações consideradas, fornece uma quantidade que corresponde ao coeficiente de σ^2 obtido no parágrafo anterior. Para as rotações 6 e 8, os valores finais correspondentes às duas rotações são +0,29 e —0,38, dando uma variância devida aos componentes 1 e 3 de $0{,}69\ \sigma^2$, à qual deve ser comparado o limite superior $0{,}75\ \sigma^2$, já obtido e com o limite inferior de $(1/4 + 1/3)\ \sigma^2 = 0{,}58\ \sigma^2$. A variância adicional devida a v_2 deve ser introduzida em todos os casos, como foi acima explicado, e isto reduzirá as diferenças proporcionais entre os diversos valores.

No caso mais geral em que uma comparação do tipo $\mu_1 r_1 + \mu_2 r_2 + \mu_3 r_3 \ldots.$, com $\mu_1 + \mu_2 + \mu_3 + \ldots = 0$, precisa ser submetida a um teste, a seguinte extensão do método acima indicado pode ser considerada: Os totais que interessam serão substituídos por $\mu_1, \mu_2, \mu_3, \ldots$

e os restantes por zero, resolvendo-se o sistema assim obtido. Supondo que os valores dados pela solução são $r_1' \; r_2' \; r_3' \ldots$, a variância da comparação será

$$(\mu_1 r_1' + \mu_2 r_2' + \mu_3 r_3' + \ldots) \, \sigma^2$$

7 - DISCUSSÃO DOS RESULTADOS

Êste trabalho não dá uma base adequada para a discussão dos resultados do ponto de vista agrícola. Para êste fim, deveriam êles ser examinados juntamente com os resultados de outras experiências semelhantes sôbre o mesmo assunto. Os pontos seguintes devem, no entanto, ser considerados :

ALGODÃO

a) O algodão em plantio contínuo mostrou uma deterioração significativa, relativamente às demais rotações.

b) O algodão nas rotações 4 e 6, quando segue imediatamente a uma leguminosa (feijão de porco), apresenta melhores resultados do que naquelas, (3, 9 e 11), em que segue o milho ou feijão comum.

c) Os melhores resultados foram os obtidos na rotação 8, em que o algodão segue uma plantação consociada de milho e mucuna ; no entanto, a diferença entre essa rotação e as rotações 4 e 6 apenas atinge o nível de significância.

d) Quando a plantação de leguminosa antecede de dois anos à cultura do algodão, esta não parece ter sido beneficiada (comparar as rotações 3 e 7 — o êrro da comparação é grande).

MILHO

a) Relativamente às outras rotações, o milho contínuo apresenta (proporcionalmente) uma deterioração mais notável do que a encontrada para a cultura contínua de algodão.

b) A produção de milho na consociação milho e mucuna é aproximadamente a mesma que a do milho nas rotações milho-algodão e milho-feijão. Assim, o benefício ao algodão foi obtido nesta rotação, sem perda apreciável, quanto à produção de milho. (No entanto, no ano preliminar 1936/7, essa consociação teve um mau comportamento ; ver quadro 3.).

RESUMO

Êste artigo descreve os métodos que deveriam ser seguidos na análise de um experimento de rotação contendo rotações de duração diversa e na qual não foram incluídas tôdas as fases dessas rotações, de forma que as produções médias de uma determinada cultura, nas diferentes rotações, podem estar afetadas pela variação anual na produção dessa cultura. Foram ajustadas as constantes dadas pelo método de quadrados mínimos, usando-se, para êsse fim, o processo de aproximação sucessiva, descrito pela primeira vez por Stevens (2). A estimativa dos erros, que afetam as estimativas acima mencionadas, foi considerada detalhadamente.

LITERATURA CITADA

1. **Stevens, W. L.** Experiências de rotação. Bragantia **11** : 317-330. 1951.

2. **Stevens, W. L.** Statistical analysis of a non-orthogonal tri-factorial experiment. Biometrika **35** : Parts III and IV. 346-367. 1948.

3. **Fisher, R. A.** *Em* Statistical methods for research workers pag. 151-169, 9.ª ed., Oliver and Boyd Ltd., Edinburgh, 1944.

Publicamos, a seguir, o trabalho original, em inglês.

THE ANALYSIS OF A ROTATION EXPERIMENT

INTRODUCTION

In a rotation experiment involving rotations of different length, in which all phases of each rotation are not represented, the comparisons of the mean yields of a particular crop in the different rotations are liable to be affected by annual variation in the yield of that crop. Consider, for example, an experiment designed to test the relative performance of continuous cotton, continuous maize, and a two-year rotation of maize and cotton. In each block a plot of continuous cotton, one of continuous maize and one of maize-cotton is included according to the scheme.

Year		*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Crop Sequence	1	A	A	A	A	A	A	A	A	A	A
	2	M	M	M	M	M	M	M	M	M	M
	3	M	A	M	A	M	A	M	A	M	A

where: A = cotton; M = maize

If the advantage of the rotation is assessed by comparing the mean yield of the cotton in sequence 3 with that in sequence 1, and the mean yield of the maize in sequence 3 with that in sequence 2, the comparisons will be affected by annual variation in the yield of the crop. Thus if the odd years 1, 3 . . . are, on balance, bad years for both cotton and maize, the maize in the rotation will appear to do worse relative to the continuous maize than it should, and the cotton in the rotation will appear to do better than it should. If the odd years are bad for cotton, and the even years for maize, the rotation will appear to do better than it should for both crops.

In an experiment of this type the difficulty can be simply overcome by comparing the mean yield of the maize of sequence 3 with the mean yield for the odd years of sequence 2, and similarly for cotton, using the even years of sequence 1. This in fact extracts all the information the experiment is capable of giving, if only comparisons of plots bearing the same crop in the same year are to be considered. It should be noted, however, that the experiment would have been much improved if the other phase sequence 3, namely A M A M . . ., had been included, since this would have enabled comparisons to be made with the yields of both sets of continuous plots in all years. Not only may we expect this to result in a reduction of the experimental error, but the performance of both crops will be tested in all years, and consequently variations in the advantage of the rotation relative to continuous cropping in different years (due to seasonal variations) will be more quickly eliminated.

It is in fact now a recognised principle of the design of rotation experiments, that all phases of each rotation should be included. The design of such experiments is discussed by Stevens (1). This present paper gives an account of the analysis of an experiment in which all phases were not represented, and to which the simple method of overcoming the difficulty given in the last paragraph is not applicable. Nothing in the paper should be taken as justifying the use of a design of this type in future experiments. Not only is the amount of arithmetical work required in the analysis greatly increased ; more important, the efficiency of the experiment and the accuracy, clarity and comprehensiveness of the conclusions, can be much improved by the the use of a good design.

Since, however, rotation experiments take a long time to carry out, it is important to extract as much information as possible from experiments which are already in progress or have been completed, even though their design could in fact be greatly improved.

DESCRIPTION OF THE EXPERIMENT

The experiment considered in the present paper was designed to investigate the effect of various rotations of crops on the maintenance of fertility. It was carried out at the Central Experiment Station, Campinas, over the crop years 1936/7 to 1943/4. In all, ten different rotations, involving cotton, maize, beans and leguminous cover crops, together with continuous cotton and continuous maize, were included. Only a single phase of each rotation was used. The rotations are shown in Table 1. * There were four replicates of each rotation. The arrangement of the plots within each replicate was systematic, but in the subsequent analysis the design has been treated as if it were one in randomized blocks.

The yields of cotton and maize in each year (totals of 4 plots) are given in Tables 2 and 3. Inspection of these Tables shows immediately that for both cotton and maize the means of the different rotations over all years are affected by variations in yield from year to year. 1939/40, for example, was a bad year for both cotton and maize, and this depresses the mean yield of cotton in rotations 1, 3, 4, 8 and 11, and of maize in rotations 2 and 10, relative to the remaining rotations.

The standard errors shown in Tables 2 and 3 are calculated from the yields for each year separately. They therefore provide valid estimates of error (apart from the lack of randomization) for comparisons of the rotations in a single year. Since, however, the same rotation occupies the same plots during the whole of the experiment the errors affecting the same rotation in the different years are likely to be correlated. Consequently the differences between the same pair of rotations in the different years are to a certain extent affected by the same errors, and therefore cannot be regarded as independent.

The estimation of the various components of error will be discussed below. At this stage it is sufficient to note that some at least of the differences of the yields shown in Tables 2 and 3 represent real effects. The difference between the yield of cotton in the maize-cotton rotation and that of the continuous cotton in 1943/4, for example, is 15 kg, and this is 2.1 times its standard error ($5.1\sqrt{2} = 7.2$). Similarly the difference in the corresponding yields of maize in 1942/3 is 27 ± 11 kg. The advantage, towards the end of the experiment, of the maize-cotton rotation relative to the continuous cropping by either crop is therefore apparent.

Comparisons of this type can be made for any pair of rotations which carry the same crop in some years at least. If, however, they carry different crops in some years there will generally be some additional information arising from the yields in those years. Rotations which never carry the same crop in the same year cannot be compared by this method.

A comprehensive set of estimates for all the rotations can be obtained by fitting constants by the method of least squares. The procedure to be followed is described in the next section. The adjusted means resulting from the analysis are shown for comparisons in the last columns of Table 2 and 3. The adjustments have made substantial difference to the mean yields, particularly in the case of cotton.

THE LEAST SQUARES ANALYSIS

Suppose that each of the yields of the different rotations in the different years given in Table 2 and 3 is made up of a pair of additive constants, one representing the effect of the rotation (taken as the same in all years) and the other the effect of the year (taken as the same for all rotations), together with a residual representing

* These tables will be found in the Portuguese version.

the difference of the actual yield from the sum of the constants. The least squares analysis gives the values of these constants which make the sum of the squares of the residuals a minimum, and this provides estimates of the differences between the rotations. Provided that the errors of the yields of the different rotations in the different years are normally and independently distributed, these estimates are the most accurate possible.

An alternative way of looking at the matter is to regard the least squares analysis as providing adjustments to the mean yields shown in Tables 2 and 3 so as to compensate for the fact that these means are made up of yields from different years. Looked at in this way, it is easy to see how to obtain the values of the constants by sucessive aproximation (Stevens, **2**). This method is not only simpler than the construction of the mathematical equations, but also usually requires less arithmetical work with data of the type we are considering.

In the present instance, inspection of the data shows that the yields of both the continuous cotton and continuous maize appear to be deteriorating relatively to the yields of the rotations. We will therefore include an additional constant b representing the deterioration per year in each of the two analyses (for cotton and maize), to allow for, and estimate, the magnitude of this deterioration. If, for cotton, the constants for the different rotations are r_1, $r_3, \ldots$ and those for the years are s_1, $s_2, \ldots$ the scheme of constants for cotton is that shown in Table 4. (Only the first two lines are shown — the extension to the remainder is obvious, but is not actually required in the computations.)

The actual computations are shown in Table 5. One decimal place appears adequate for the adjustments. By inspection it is seen that the biggest differences are those due to years. We therefore start by calculating values for the year constants, ignoring the other constants. These are simply the year means, and are shown, **to one decimal place,** in line (1). They must be carefully checked. Column (2) is then calculated by adding the means from line (1) corresponding to the occupied cells of each row and subtracting the total of the row. Thus

$$+\ 37.6 = \text{sum of all line (1)} - 435$$
$$+\ 3.9 = 89.2 + 39.8 + 66.2 + 60.7 - 252$$

etc. It will be noted that yields for rotations 3, 4, 8, and 11 occur in the same years, and therefore have the same total of means. The total of column (2) should be zero, except for rounding off errors in the means. Their total amounts to

$$6(89.2) + 3(75) + \ldots - 2080 = +\ 0.7$$

and this provides an exact check.

The means in column (3) are then formed from column (2) by dividing the number of cells in each line, **and reversing the signs.** Column (3) must be carefully checked. The means of columns (3) represent first approximations to the differences between rotations, allowing for differences between years. The deterioration **b** of rotation 1 is calculated in a similar manner, using the formula

$$b = -\frac{Sy(x - \bar{x})}{S(x - \bar{x})^2}$$

appropriate to a linear regression, and taking the difference of the regression on line (1) (+) and that on the actual yields of rotation 1 (—). This gives the sum of products

$$\begin{aligned} &+ 3(89.2) + 2(75) + 1(39.8) - 1(66.2) - 2(65) - 3(60.7) \\ &\quad - 3(92) - 2(74) - 1(24) + 1(67) + 2(58) + 3(45) = -50.9 \end{aligned}$$

This is entered in column 2. The divisor is $3^2 + 2^2 + 1^2 + 1^2 + 2^2 + 3^2 = 28$. Dividing and changing signs gives +1.8, which is entered in column (3).

The adjustments to the means of line (1) are now calculated from the figures in column (3). The appropriate multiple of **b** is included for each year. Thus the total for the first year is

$$(+ 3) (+ 1.8) - 5.4 - 1.0 + 4.5 + 4.5 + 11.8 - 8.7 = + 11.1$$

and that for the second year is

$$(+ 2) (+ 1.8) - 5.4 - 3.1 + 1.1 = -3.8$$

These totals are entered in line (4). Their total is equal to

$$(7) (-5.4) + (4) (-1.0). + \ldots\ldots = -0.8$$

Dividing by number of occupied cells in each column, and changing signs, as before, we obtain the adjustments of line (5). These are the corrections to the year means necessary on account of the regression and of differences between rotations. Thus a second approximation to the constant s_1 of the first year is $89.2 - 1.8 = 87.4$.

The process is now repeated, using the adjustments of line (5) to give the totals and means of columns (6) and (7). The values in column (7) are adjustments to the values for the rotation differences in column (3). A further repetitition of this whole cycle gives lines (8) and (9) and columns (10) and (11).

The values in column (11) are sufficiently small to show that there is no need for further approximations. The adjustments of columns (3), (7) and (11) are therefore added to give the totals in the last columns. These totals give the final estimates of the relative yields of the different rotations. They may conveniently be increased by the mean yield, 67.1 over the whole experiment, for purposes of presentation of the results.

Similarly the totals of lines (1), (5) and (9) give estimates of the mean yearly yields, freed from rotation effects.

The rules for signs followed above may be summarised as follows.

Apart from the initial means, which form the first approximations to the first set of constants, the totals from which the first approximations and subsequent adjustments are calculated, for convenience carry the opposite sign to the approximations and adjustments themselves i.e. they are **plus** the appropriate function of the first approximations or subsequent adjustments to the other sets of constants **minus** (where necessary) the corresponding function of the actual yields. When taking the means to obtain the actual approximations and adjustments, therefore, **the signs are changed.**

It may be noted that the computations can be slightly abbreviated by amalgamating the lines for rotations 3, 4, 8 and 11 for which yields occur in the same years, and calculating a constant for the average of these rotations. The operative table showing the multiples of the approximations and adjustments to be taken (which may replace the actual yields of Table 5) would then have 4's in this line and 1's elsewhere where yields occurred. The sum of the totals would be entered in the total column, and the constant obtained would be the mean of the constants for the four rotations. These four constants will have the same deviations from their mean as do the original rotation means from their mean. If there were no regression to be fitted a similar amalgamation could be made for the years 1937/8 and 1943/4. In extensive tables with a certain amount of internal symmetry amalgamations of this kind may substantially shorten the work.

CALCULATION OF THE RESIDUAL ERROR

The sum of squares of the residuals after fitting the constants in a least squares analysis gives an estimate of the errors to which the data are subject. The meaning of this estimate will be discussed in the next section. Here we are concerned with the method of calculation.

In any least squares analysis the sum of squares of the residuals after fitting constants is given by the general formula

$$S(y^2) - b_1\ S(x_1y) - b_2\ S(x_2y) \text{----------} \qquad (A)$$

where y represents the original values, b_1, b_2, ... the fitted constants, and x_1, x_2, ... the coefficients of these constants in the expressions for the expected y in terms of the constants. Thus Table 4 shows that constant r_1 has coefficients of +1 for all the y's of row 1 and coefficients of zero elsewhere. Consequently for r_1 we have $S(x_1y)$ equal to the total of row 1, namely 435. Also for s_1, $S(x_1y) = 535$, etc.. Similarly the coefficients of the deterioration constant **b** are +3, +2, +1, 0, —1, —2, —3 for the first line and 0 elsewhere. Hence $S(x_by)$ equals +130. Consequently the reduction in the sum of squares, which includes the correction for the mean of the ordinary analysis of variance, is given by

$$(-6.2)\ (435) - (0.4)\ (252) + (5.1)\ (274) + \ldots\ldots$$
$$+ (86.9)\ (535) + (77.1)\ (225) + \ldots + (2.3)\ (+130) = 148208.3$$

The formula (A) from which this quantity is calculated is, however, only correct if the **b**'s are the solutions of the least squares equations. In a solution by successive approximation there are small residual errors, which though they do not materially affect the estimates of the constants, introduce considerable error into the sum of squares calculated from (A). In so far as the errors consist of constant errors in all the s's or in all the r's, an adjustment may be made. If δs is the difference of the weighted mean of the s's (+) from the mean of the original yields (—), the weights being equal to the number of yields in each column, the amount to be added to the reduction in the sum of squares is $-\delta s\ T$, where T is the total of all the yields.

$$\text{Here } \delta s = \left[6\ (86.9) + 3\ (77.1) + \ldots\right]/31 - 67.09677 = +\ 0.12581$$

and hence (—0.12581) (2080) = — 261.7 must be added the total given above. For the r's the weighted mean is — 0.18064, and the correction is therefore +375.7.

The ordinary correction for the mean is 139,561.3, and consequently the reduction in the sum of squares due to the fitting of the constants other than the mean is 148208.3 — 261.7 + 375.7 — 139561.3 = 8761.0.

A part of this sum of squares is accounted for by the variation in the year means. The sum of squares attributable to this is calculated in the ordinary manner. (If year means are used more decimal places must be included than are shown in Table 5).

The full analysis of variance is therefore as shown in Table 6. For convenience the mean squares for maize are also shown in this Table. The sums of squares have been divided by 4 to bring them to a single plot basis. The degrees of freedom for the aggregate of the constants fitted to cotton is 14, i.e., 7 for years less 1 for the general mean, plus 8 for rotations less 1 for a redundant constant (there are only 7 independent contrasts), plus 1 for the deterioration.

The residual sum of squares calculated by this method is 169.8. This is not in fact very accurate owing to the errors mentioned above. An accurate value may be obtained by calculating the actual residuals and summing their squares, and although this is not really necessary for the estimation of error it provides a useful check on the whole analysis. In addition an examination of the residuals is sometimes advisable to see if there are additional effects for which constants have not been included in the first analysis.

Table 7 shows the actual residuals. That for rotation 1 in 1937/8, for example, is given by

$$92 - 86.9 + 6.2 - 3 \times 2.3 = +4.4$$

It will be seen that the sums of the rows and of the columns are all nearly zero, as they should be. The sum of squares divided by 4 is 161.7 which checks reasonably with that obtained by the ordinary method.

Inspection of the residuals does not reveal any additional effects not included in the first analysis.

COMPONENTS OF ERROR

The errors to which the plot yields are subject can be divided into three components :

(1) A component which is independent from year to year and plot to plot.

(2) A component which is constant for year to year but independent from plot to plot.

(3) A real variation from year to year in the differences between the rotations owing to seasonal influences on the effects of the various rotations.

This sub-division of the errors is equivalent to regarding the errors of the plot yields as made up of three additive parts

$$e_1 + e_2 + e_3$$

independently distributed with variances v_1, v_2, v_3, e_1 assuming a new value for each plot each year, e_2 having the same value in all years for a particular plot, and e_3 having the same value for all plots of a particular rotation in a given year.

Components (1) and (2) can theoretically be supplemented by further components representing differences from plot to plot in the linear trends and higher components of regression on years. In practice it is found that after removal of the constant component, (1) the remainder of the plot to plot variation can be regarded as virtually independent from year to year except possibly for experiments of very long duration.

If yields of a single crop are available for all plots in all years the various components can be isolated in a comprehensive analysis of variance. This is not possible here, but a partial analysis can be carried out on the cotton crop of the rotations 1, 3, 4, 8 and 11 which are represented in the four years 3 1937/8, 1939/40, 1941/2 and 1943/4. A similar analysis can be carried out on maize for the rotations 2, 3, 5, 8 and 12 in the three years 1938/9, 1940/1 and 1942/3. The results of these analyses are shown in Table 8.

In order to allow for the deterioration of the continuous cropping relative to the rotations, the appropriate degree of freedom has been separated from the years $\times$ rotations component. The mean square is obtained from the years $\times$ rotations table by using the multipliers given in Table 9 to calculate a quantity Q. The contribution from year 1937/8 for cotton is equivalent to +12 times the yield of rotation 1, less 3 times the yields of the other four rotations. Each yield in the table is the total of four plot yields. Hence for cotton the sum of squares of the coefficients of the individual plot yields in Q is

$$(4)\ (2)\left[12^2 + 4(3^2) + 4^2 + 4(1^2)\right] = 1600$$

The value of Q is found to be 167.9, and the sum of squares is therefore $167.9^2/1600 = 17.62$. For maize the divisor is

$$(4)\ (2)\left[4^2 + 4(1^2)\right] = 160$$

The mean square for cotton, for rotations $\times$ blocks $\times$ years, is an estimate of v_1, the mean square for rotations $\times$ blocks is an estimate of $v_1 + 4v_2$ (since there are four years included in the analysis) and the mean square for rotations $\times$ years (excluding the deterioration component) is an estimate of $v_1 + 4v_3$ (since there are four plots of each rotation). The same holds for maize except that the mean square for rotations $\times$ blocks is an estimate of $v_1 + 3v_2$.

The relation between the partial analysis of variance given in Table 8, and the least squares analysis of variance of Table 6 will now be clear. The 11 and 7 degrees of freedom for the years $\times$ rotations remainder components of Table 8 form part of the 16 and 13 degrees of freedom of the residual components of Table 6. It will be best to use the values of Table 6 (from residuals) for the estimation of component (3) of the variance. The estimates of the three components of variance are therefore.

Cotton	Maize
$v_1 = 6.15$	$v_1 = 10.79$
$v_2 = 1.15$	$v_2 = 6.77$
$v_3 = 0.99$	$v_3 = 15.66$

Relative to (1) components (2) and (3) are both substantially greater for maize than for cotton.

ERRORS OF THE COMPARISONS BETWEEN ROTATIONS

If a rotation has a given crop in p years and there are q replicates per year, the same plots being used in all years, the error variance applicable to the mean over all p years, appropriate to comparisons with other rotations which have the same crop as the first in all p years, is

$$\frac{1}{pq} v_1 + \frac{1}{q} v_2 + \frac{1}{p} v_3 = \frac{1}{pq} (v_1 + qv_3) + \frac{1}{q} v_2$$

The remainder mean square of the least squares analysis of variance has an expectation of $v_1 + qv_3$. The variance of each rotation mean must therefore be increased by $v_2/4$, regardless of the number of years entering into the mean.

The standard errors calculated in this manner are entered in Table 10, which gives the adjusted mean yields of cotton and maize from the least squares analyses converted to 1000 Kg per Hectare.

For comparisons of rotations in which the crop does not always occur in the same years in both rotations these standard errors provide lower limits only to the true standard errors. In the special case in which the rotation with the crop in the greater number of years has the crop in all the years in which it occurs in the other rotation, the standard error of the comparisons will be $\sqrt{2}$ times the standard error of the second rotation. In other cases no quick method of obtaining the exact standard error is available. The standard errors of all comparisons are given by inverting the matrix of coefficients in the least squares equations (see, for example, Fisher, Statistical Methods for Research Workers, Sect. 29) but the labour involved would be prohibitive in work of this kind. Instead upper limits can usually be calculated for comparisons of special interest by the method of the next paragraph.

Suppose, for example, we require the standard error of the difference of the cotton yields of rotations 6 and 8. Suppose for the moment that each yield of Table 5 is independent and has a residual variance of σ^2. Cotton occurs in both rotations in 1937/8 and 1943/4 and the variance of the difference of the means of these two years is $(\frac{1}{2} + + \frac{1}{2}) \sigma^2 = \sigma^2$. The weight is therefore $1/\sigma^2$. The yield of rotation 6 in 1940/1 can be compared with that of rotation 1 in that year, and the yields of rotation 1 can be compared with the yields of rotation 8 in 1939/40 and 1941/42. None of these yields has entered into the previous comparison and the two comparisons are therefore independent. The variance of the second comparison is $(1+1+1/2+1/2)\,\sigma^2 = 3\,\sigma^2$. The weight a therefore $1/3\sigma^2$, and the weight of the weighted mean of the two comparisons is $4/3\sigma^2$. The variance is therefore $3\sigma^2/4$, or $0.75\sigma^2$. In the present case the independent component of variation is $(v_1 + 4v_3)/4$. There is also the additional component of variance (2) which increases the variance of the difference by $2v_2/4$. An upper limit to the variance is therefore provided by

$$0.1625(v_1 + 4v_3) + 0.5\, v_2 = 2.47$$

In comparisons of special importance, for which the best available estimate of the standard error is required, this may be obtained by a solution by successive approximation similar to that used for the estimates, first replacing the totals of the two rotations to be compared by +1 and —1 and all the other rotation totals, the yearly totals and the regression total by zero. The difference of the final values for the two rotations in question gives a quantity corresponding to the coefficient of σ^2 in the last paragraph. For rotations 6 and 8 the final values for the two rotations are +0.29 and —0.38, giving a variance due to components (1) and (3) of 0.69 σ^2, which may be compared with the upper limit already obtained of 0.75 σ^2 and the lower limit of $(1/4 + 1/3)\,\sigma^2 = 0.58\,\sigma^2$. The additional variance due to v_2 must be introduced in all cases, as explained above, and this will reduce the proportional differences between the various values.

In the more general case in which some comparison of the type

$$\mu_1 r_1 + \mu_2 r_2 + \mu_3 r_3 \ldots\ldots$$

with $\mu_1 + \mu_2 + \mu_3 + \ldots\ldots = 0$, requires to be tested, the following extension of the above method may be noted. Replace the relevant totals by $\mu_1, \mu_2, \mu_3, \ldots$ and the remainder by zero, and solve. Suppose the final values given by the solution are $r_1', r_2', r_3' \ldots$ then the error variance of the comparison will be

$$(\mu_1 r_1' + \mu_2 r_2' + \mu_3 r_3' + \ldots)\,\sigma^2$$

DISCUSSION OF THE RESULTS

This is not the appropriate place to discuss the agricultural implications of the results, which should be examined in conjunction with the results of similar experiments on the same subject.

The following points may, however, be noted.

COTTON :

1. The continuous cotton has shown significant deterioration relative to the rotations.
2. In the rotations 4 and 6, in which cotton immediately follows the legume (Feijão de porco), the cotton is better than in those (3, 9 and 11) in which the cotton follows maize or beans.
3. The cotton has done best in the rotation 8 in which it follows a mixture of maize and mucuna, though the difference between this and rotations 4 and 6 barely attains significance.
4. The legume two years prior to the cotton does not appear to have benefited the cotton (contrast the rotations 3 and 7 — the error of the contrast is high).

MAIZE :

1. The continuous maize has deteriorated relative to the rotations to a greater extent (proportionally) than the continuous cotton.
2. The mixture of maize and mucuna has yielded about the same amount of maize as the maize in the maize-cotton and maize-beans rotations. Thus the benefit to the cotton of this mixture has been obtained without any appreciable loss of maize. (In the preliminary year 1936/7, however the mixture did very badly (see Table 3).

SUMMARY

The paper describes the methods which should be followed when analysing a rotation experiment containing rotations of different length in which all phases of each rotation are not represented, and in which consequently the mean yields for the different rotations of a particular crop are liable to be affected by annual variations in the yield of that crop. Constants are fitted by the method of least squares, the arithmetical method of successive approximation first described by Stevens being used. The estimation of the errors to which the estimates are subject is considered in detail.