

Evolutionary Operation EVOP

Projeto de Experimentos II
Prof. Flávio Fogliatto

1

O que é EVOP?

- Método de monitoramento e melhoria contínua de processos:
 - Objetivo é mover as condições operacionais na direção do ótimo
- Não requer grandes mudanças nas condições operacionais:
 - Idéia é alterar o processo sem afetar sua qualidade
- Proposto como método rotineiro de operação da planta, conduzido por pessoal operacional com mínimo envolvimento da engenharia

2

Origens do EVOP

- Método originalmente proposto por George Box:
 - Box & Draper (1969). *Evolutionary Operation: A Statistical Method for Process Improvement*. New York: John Wiley.
- Versão simplificada do método encontrada em:
 - Hicks (1993). *Fundamental Concepts in the Design of Experiments*. 4ª Ed., Saunders College Publ.
 - Myers & Montgomery (1995). *Response Surface Methodology*. John Wiley.
- Disponível no *software* EVOPtimizer:
 - Versão demo disponível *on line*.

3

Panorâmica do método

- Procedimento consiste na realização de um experimento fatorial 2^k , dentro da amplitude de operabilidade do processo em sua condição atual:
 - Pressuposto é que fatores de controle são mensuráveis e ajustáveis a uma pequena distância dos ajustes atuais sem geração de não-conformes
- Idéia é coletar dados sobre variável de resposta (usualmente rendimento) em vários pontos do design experimental:
 - Quando um conjunto de dados foi obtido para todos os pontos, diz-se que um ciclo foi concluído

4

Panorâmica do método (Cont.)

- Um ciclo normalmente não é suficiente para detectar mudanças no processo:
 - Número de ciclos é variável
 - Efeitos não detectados entre 5 e 8 ciclos costumam não estar presentes no processo na faixa investigada
- Ciclos continuam até que alteração fatores principais, interações ou alterações na média resultem significativos:
 - Uma vez obtido aumento significativo no rendimento, diz-se que uma fase do EVOP está concluída
 - Neste ponto decide-se sobre alterações nos ajustes do processo

5

Panorâmica do método (Cont.)

- Para testar a significância de fatores controláveis e interações, uma estimativa do erro experimental é necessária:
 - Box propôs estimativa baseada no cálculo de amplitudes dos dados do ciclo
- Projeto 2^k é centrado nas condições operacionais ótimas atuais do processo:
 - Utilização de ponto central é prevista no método:
 - Comparando resposta neste ponto com demais 2^k pontos checa-se curvatura ou mudança na média (CIM – *change in mean*)
 - Se processo estiver de fato centrado no ótimo, resposta no centro deve ser significativamente maior que nos pontos periféricos

6

Panorâmica do método (Cont.)

- EVOP pode, em teoria, ser aplicado na investigação de k variáveis de processo:
 - Na prática, duas ou três variáveis são investigadas
- Método foi originalmente concebido para operacionalização através de planilhas:
 - Análise pode ser simplificada com o uso de aplicativos

7

EVOP vs Superfícies de Resposta

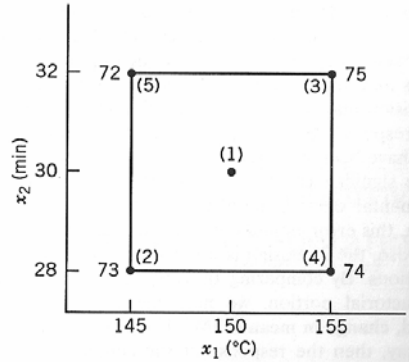
- Metodologia de Superfície de Resposta (MSR) é normalmente utilizada em plantas piloto:
 - Experimentos tendem a ser mais complexos e caros
- EVOP é uma adaptação da MSR para plantas industriais:
 - Método pode ser usado para ajustar processos após *scale up*:
 - Uma vez determinados ajustes operacionais que geram produtos conformes, utiliza-se EVOP para otimizar rendimento

8

EVOP através de um exemplo

- Exemplo ilustra o método originalmente concebido por Box:

- Considere um processo químico onde *rendimento* é função da *temperatura* (x_1) e do *tempo de reação* (x_2)
- Condições operacionais atuais são:
 - $x_1 = 150^\circ\text{C}$
 - $x_2 = 30 \text{ min}$
- EVOP utilizará um fatorial 2^2 acrescido de um ponto central

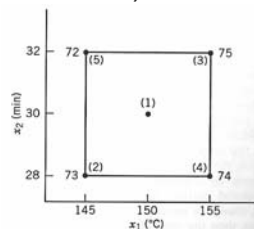


9

Projeto experimental

- Rodadas são numeradas e rodadas na ordem proposta:
 - Isso facilita a memorização das rodadas pelo pessoal operacional
 - Se efeitos de fatores de ruído (como turno) estiverem presentes, eles serão confundidos com os blocos (i.e., ciclos adicionais a serem rodados)

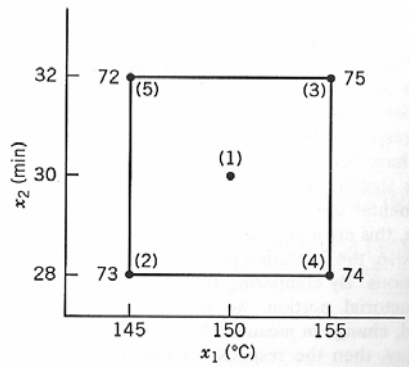
Valores de rendimento estão indicados nas rodadas



10

Efeitos principais e interações calculados da maneira usual para um 2²

- Efeitos principais:
 - Temperatura:
 - $[(3)+(4)-(2)-(5)]/2$
 - Tempo:
 - $[(3)+(5)-(4)-(2)]/2$
- Interação:
 - $[(2)+(3)-(4)-(5)]/2$
- Curvatura (mudança na média):
 - $[(2)+(3)+(4)+(5) - 4(1)]/5$



11

5		3		Cycle: n = 1		Phase: 1		
.1				Response: Yield		Date: 3/27/94		
2		4		Calculation of Averages				
Operating Conditions		(1)	(2)	(3)	(4)	(5)	Calculation of Standard Deviation	
(i) Previous cycle sum							Previous sum S =	
(ii) Previous cycle average							Previous average $\bar{S} =$	
(iii) New observations		74	73	75	74	74	New sum S = range \times	
(iv) Differences [(iii) - (ii)]							$\frac{S - \bar{S}}{\sqrt{n}}$	
(v) New sums [(i) + (iii)]		74	73	75	74	72	New sum S =	
(vi) New averages [$\bar{y}_i = (v)/n$]		74	73	75	74	72	New average $\bar{S} =$	
							$\frac{\text{New sum } S}{n - 1}$	
Calculation of Effects						Calculation of Error Limits		
Temperature effect								
$= \frac{1}{2}(\bar{y}_3 + \bar{y}_4 - \bar{y}_2 - \bar{y}_5) = 2.00$						For new average $\frac{2}{\sqrt{n}} \bar{S} =$		
Time Effect								
$= \frac{1}{2}(\bar{y}_3 + \bar{y}_5 - \bar{y}_2 - \bar{y}_4) = 0.00$						For new effects $\frac{2}{\sqrt{n}} \bar{S} =$		
Interaction effect								
$= \frac{1}{2}(\bar{y}_2 + \bar{y}_3 - \bar{y}_4 - \bar{y}_5) = 1.00$								
Change-in-mean effect								
$\frac{1}{3}(\bar{y}_2 + \bar{y}_3 + \bar{y}_4 + \bar{y}_5 - 4\bar{y}_1) = -0.40$						For change in mean $\frac{1.78}{\sqrt{n}} \bar{S} =$		

Tabela do EV/OP para o 1º ciclo

Primeiro ciclo não gera estimativa do desvio-padrão

5 3
 .1
 2 4

Cycle: $n = 2$
 Response Yield

Phase: 1
 Date: 3/27/94

Operation Conditions	Calculation of Averages					Calculation of Standard Deviation
	(1)	(2)	(3)	(4)	(5)	
(i) Previous cycle sum	74	73	75	74	72	Previous sum $S =$
(ii) Previous cycle average	74	73	75	74	72	Previous average $\bar{S} =$
(iii) New observations	72	71	76	75	73	New $S = \text{range} \times$ $f_{S,n} = (3.6)(0.3) = 0.90$
(iv) Differences [(iii) - (ii)]	-2	-2	-1	1	1	Range of (iv) = 3.0
(v) New sums [(i) + (iii)]	146	144	151	149	145	New sum $S = 0.90$
(vi) New Averages $[\bar{y}_i = (v)/n]$	73	72	75.5	74.5	72.5	New average $\bar{S} =$ $\frac{\text{New sum } S}{n - 1} = 0.90$
Calculation of Effects						Calculation of Error Limits
Temperature effect $= \frac{1}{2}(\bar{y}_3 + \bar{y}_4 - \bar{y}_2 - \bar{y}_5) = 2.75$						For new average $\frac{2}{\sqrt{n}} \bar{S} = 1.27$
Time effect $= \frac{1}{2}(\bar{y}_3 + \bar{y}_5 - \bar{y}_2 - \bar{y}_4) = 0.75$						For new effects $\frac{2}{\sqrt{n}} \bar{S} = 1.27$
Interaction effect $= \frac{1}{2}(\bar{y}_2 + \bar{y}_3 - \bar{y}_4 - \bar{y}_5) = 0.25$						
Change-in-mean effect $= \frac{1}{3}(\bar{y}_2 + \bar{y}_3 + \bar{y}_4 + \bar{y}_5 - 4\bar{y}_1) = 0.50$						For change in mean $\frac{1.78}{\sqrt{n}} \bar{S} = 1.13$

Segundo ciclo do EVOP

13

2º ciclo em detalhes

- Ao final do 2º ciclo, erro experimental pode ser estimado e estimativas dos efeitos podem ser comparados com limites 2-Sigma (aprox. 95%)
- Estimativa do desvio-padrão:

– Calcula-se a amplitude das diferenças:

- $+1,0 - (-2,0) = 3,0$

– Converte-se amplitude em s :

- $s =$ estimativa do desvio-padrão
- Multiplica-se amplitude por fator $f_{5,n}$
- Média da estimativa do desv-pad em cada ciclo é obtida através do cálculo

Calculation of Standard Deviation	
Previous sum $S =$	
Previous average $\bar{S} =$	
New $S = \text{range} \times$	
$f_{S,n} = (3.6)(0.3) = 0.90$	
Range of (iv) = 3.0	
New sum $S = 0.90$	
New average $\bar{S} =$	
$\frac{\text{New sum } S}{n - 1} = 0.90$	

14

Tabela de valores de $f_{k,n}$

Table V. A table of values of $f_{k,n}$

No. of Cycles, n	Number of Runs in the Block, k								
	2	3	4	5	6	7	8	9	10
2	0.63	0.42	0.34	0.30	0.28	0.26	0.25	0.24	0.23
3	0.72	0.48	0.40	0.35	0.32	0.30	0.29	0.27	0.26
4	0.77	0.51	0.42	0.37	0.34	0.32	0.30	0.29	0.28
5	0.79	0.53	0.43	0.38	0.35	0.33	0.31	0.30	0.29
6	0.81	0.54	0.44	0.39	0.36	0.34	0.32	0.31	0.30
7	0.82	0.55	0.45	0.40	0.37	0.34	0.33	0.31	0.30
8	0.83	0.55	0.45	0.40	0.37	0.35	0.33	0.31	0.30
9	0.84	0.56	0.46	0.40	0.37	0.35	0.33	0.32	0.31
10	0.84	0.56	0.46	0.41	0.37	0.35	0.33	0.32	0.31
15	0.86	0.57	0.47	0.42	0.38	0.36	0.34	0.33	0.31
20	0.86	0.58	0.47	0.42	0.38	0.36	0.34	0.33	0.32

15

2º ciclo em detalhes (Cont.)

- Média \bar{s} é usada para calcular limites do erro
- Ao final do 2º ciclo, efeito da *Temperatura* excede seu limite de erro:
 - Ou seja, estimativa do efeito $\neq 0$ por pelo menos 2σ
 - Mudança nas condições operacionais é garantida

Calculation of Effects	Calculation of Error Limits
Temperature effect $= \frac{1}{2}(\bar{y}_3 + \bar{y}_4 - \bar{y}_2 - \bar{y}_5) = 2.75$	For new average $\frac{2}{\sqrt{n}}\bar{s} = 1.27$
Time effect $=$	For new effects
Interaction effect $=$	
Characterization effect $= \frac{1}{5}(\bar{y}_2 + \bar{y}_3 + \bar{y}_4 + \bar{y}_5 - 4\bar{y}_1) = 0.50$	$\frac{1.78}{\sqrt{n}}\bar{s} = 1.13$

- Como efeito da *Temp* é positivo, deve-se aumentar a *Temp* para aumentar o rendimento
- Estratégia é iniciar nova fase do EVOP centrada no ponto $x_1 = 155^\circ\text{C}$ e $x_2 = 30$ min.

16

Estimativa do desvio-padrão em detalhes

- Desvio-padrão σ é estimado pelo método da amplitude:
 - Seja $y_i(n)$ a observação derivada do ponto i no ciclo n
 - Seja $\bar{y}_i(n)$ a média dos $y_i(j)$ após o ciclo n ($j=1,2,\dots,n$)
- Quantidades na linha (iv) da planilha do EVOP correspondem às diferenças $y_i(n) - \bar{y}_i(n-1)$:
 - A variância dessas diferenças é:

$$\text{Var}[y_i(n) - \bar{y}_i(n-1)] \equiv \sigma_D^2 = \sigma^2 \left(1 + \frac{1}{n-1} \right) = \sigma^2 \frac{n}{n-1}$$

17

Continuando...

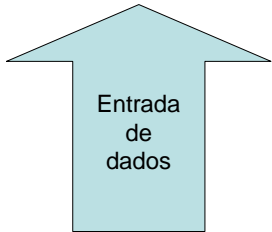
- A amplitude das diferenças, R_D , está relacionada à estimativa do desv-pad das diferenças por $\sigma_D = R_D / 2$
- O fator d_2 (muito usado no CEP) depende do nº de observações usadas no cálculo do R_D
- Como $R_D/d_2 = \sigma\sqrt{n/(n-1)}$, então:

$$\hat{\sigma} = \sqrt{\frac{n-1}{n}} R_D / d_2 = f_{k,n} R_D = s$$

18

Mesmos dados podem ser analisados no Statgraphics

	x1	x2	Rendimento
	-1	-1	73
	+1	-1	74
	-1	+1	72
	+1	+1	75
	0	0	74



Analyse de régression multiple					
Variable à expliquer: Rendimento					
Paramètre	Estimation	Erreur type	T	Proba.	
CONSTANTE	73,6	0,2	368,0	0,0017	
x1	1,0	0,223607	4,47214	0,1400	
x2	0,0	0,223607	0,0	1,0000	
x1*x2	0,5	0,223607	2,23607	0,2677	

Analyse de variance					
Source	Somme des carrés	Ddl	Carré moyen	F	Proba.
Modèle	5,0	3	1,66667	8,33	0,2447
Résidu	0,2	1	0,2		
Total (Corr.)	5,2	4	Nenhum fator significativo		

R-carré = 96,1538 %
R-carré (ajusté pour les ddl) = 84,6154 %
Estimation de l'écart-type du résidu = 0,447214
Erreur absolue moyenne = 0,16
Test de Durbin-Watson = 1,25
Autocorrélation résiduelle d'ordre 1 = -0,05

Inserindo dados do 2º ciclo

	x1	x2	Rendimento
	-1	-1	73
	+1	-1	74
	-1	+1	72
	+1	+1	75
	0	0	74
	-1	-1	71
	+1	-1	75
	-1	+1	73
	+1	+1	76
	0	0	72

Analyse de régression multiple					
Variable à expliquer: Rendimento					
Paramètre	Estimation	Erreur type	T	Proba.	
CONSTANTE	73,5	0,319505	230,043	0,0000	
x1	1,375	0,357217	3,8492	0,0085	
x2	0,375	0,357217	1,04978	0,3343	
x1*x2	0,125	0,357217	0,349927	0,7383	

Analyse de variance					
Source	Somme des carrés	Ddl	Carré moyen	F	Proba.
Modèle	16,375	3	5,45833	5,35	0,0393
Résidu	6,125	6	1,02083		
Total (Corr.)	22,5	9			

R-carré = 72,7778 %
R-carré (ajusté pour les ddl) = 59,1667 %
Estimation de l'écart-type du résidu = 1,01036
Erreur absolue moyenne = 0,7
Test de Durbin-Watson = 1,90561 (P=0,4389)
Autocorrélation résiduelle d'ordre 1 = -0,239796

x₁ significativo com coeficiente positivo

EVOP Simplex

- Simplex é uma alternativa aos projetos fatoriais:
 - São projetos ortogonais saturados de 1ª ordem:
 - Não permitem verificação do efeito de interações e falta-de-ajuste
 - Demandam $N = k + 1$ observações p/ ajustar um modelo de 1ª ordem com k variáveis:
 - Linhas da matriz de projeto são pontos em k dimensões tal que ângulo formado entre pares de pontos e a origem é θ , onde:

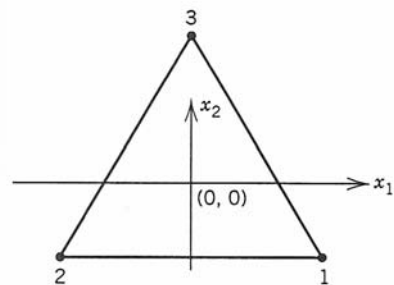
$$\cos(\theta) = -\frac{1}{(N-1)} = -\frac{1}{k}$$

21

Projeto Simplex para $k = 2$

- Pontos são coordenadas de um triângulo equilátero:
 - $k = 2, N = 3, \cos(\theta) = -1/2$
- Matriz de projeto é dada por:

$$\mathbf{D} = \begin{bmatrix} x_1 & x_2 \\ \sqrt{3}/2 & -1/\sqrt{2} \\ -\sqrt{3}/2 & -1/\sqrt{2} \\ 0 & 2/\sqrt{2} \end{bmatrix}$$

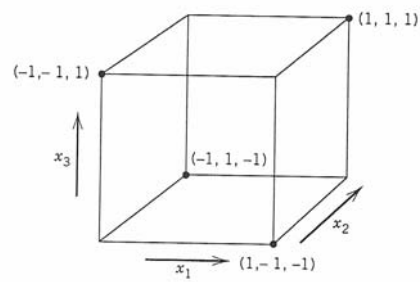


22

Projeto Simplex para $k = 3$

- Pontos são vértices de um tetraedro:
 - $k = 3, N = 4, \cos(\theta) = -1/3$
- Matriz de projeto é dada por:

$$\mathbf{D} = \begin{matrix} & x_1 & x_2 & x_3 \\ \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \end{matrix}$$



23

Exemplo: Simplex com $k = 3$

- Objetivo é obter uma relação de 1ª ordem entre crescimento (y) de um organismo e:
 - % de glicose (x_1),
 - concentração de extrato de fermento (x_2) e
 - tempo em horas (x_3) disponibilizados para o crescimento do organismo
- Fatores de controle codificados são:

$$x_1 = \frac{\% \text{ glic} - 30}{1,0} \quad x_2 = \frac{\% \text{ ferm} - 0,5}{0,1} \quad x_3 = \frac{\text{horas} - 45}{15}$$

24

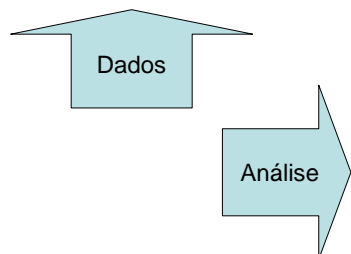
Dados do experimento

Run Number	Percent Glucose, x_1		Percent Yeast, x_2		Time, x_3 (hr)		Growth, (g/ liter)
	Uncoded	Coded	Uncoded	Coded	Uncoded	Coded	
	1	3.0	0	0.641	$\sqrt{2}$	30	
2	1.586	$-\sqrt{2}$	0.500	0	60	1	9.76
3	3.0	0	0.359	$-\sqrt{2}$	30	-1	7.38
4	4.414	$\sqrt{2}$	0.500	0	60	1	12.50

25

Gerando modelo no Statgraphics

	x1	x2	x3	Crescimento
0	1,41421	-1		8,52
-1,41421	0	1		9,76
0	-1,41421	-1		7,38
1,41421	0	1		12,5



Régression multiple - Crescimento

Libellé: _____

Analyse de régression multiple

Variable à expliquer: Crescimento

Paramètre	Estimation	Erreur type	T	Proba.
CONSTANTE	9,54			
x1	0,968739			
x2	0,403052			
x3	1,59			

Modelo (ajuste não pode ser verificado)...

Analyse de variance

Source	Somme des carrés	Ddl	Carré moyen	F	Proba.
Modèle	14,516	3	4,83867		
Résidu	0,0	0	0,0		
Total (Corr.)	14,516	3			

R-carré = 100,0 %
R-carré (ajusté pour les ddl) = 0,0 %
Estimation de l'écart-type du résidu = 0,0
Erreur absolue moyenne = 0,0
Test de Durbin-Watson =
Autocorrélation résiduelle d'ordre 1 =

EVOP Simplex

- Considere um EVOP com k variáveis e matriz de projeto \mathbf{D} :
 - j -ésima linha de \mathbf{D} será designada por \mathbf{d}'_j
 - vantagem desse design com relação ao fatorial é o menor número de rodadas
- Suponha que o objetivo seja maximizar a resposta:
 - Procedimento do EVOP Simplex ajustaria as variáveis de processo de acordo com três regras, descritas a seguir

27

EVOP Simplex Regra 1

- Seja y_i a resposta medida no i -ésimo ponto experimental ($i = 1, 2, \dots, n$):
 - Considere que o menor valor de y_i ocorra no ponto \mathbf{d}'_j
- Um novo simplex será formado deletando a linha \mathbf{d}'_j da matriz \mathbf{D} e substituindo-a pelo novo ponto experimental:

$$\mathbf{d}'_j^* = 2n^{-1}(\mathbf{d}'_1 + \mathbf{d}'_2 + \dots + \mathbf{d}'_{j-1} + \mathbf{d}'_{j+1} + \dots + \mathbf{d}'_n) - \mathbf{d}'_j$$

- Rode o processo no próximo ciclo usando níveis para x_1, \dots, x_n dados por \mathbf{d}'_j^*

28

EVOP Simplex Regra 2

- Aplique a Regra 1 novamente, a menos que um dado ponto experimental tenha ocorrido em n simpleces sucessivos sem ser eliminado:
 - Se esta situação ocorrer para o ponto i , descarte y_i e rode o processo no próximo ciclo usando os níveis dos fatores dados em \mathbf{d}'_i
 - Então aplique a Regra 1

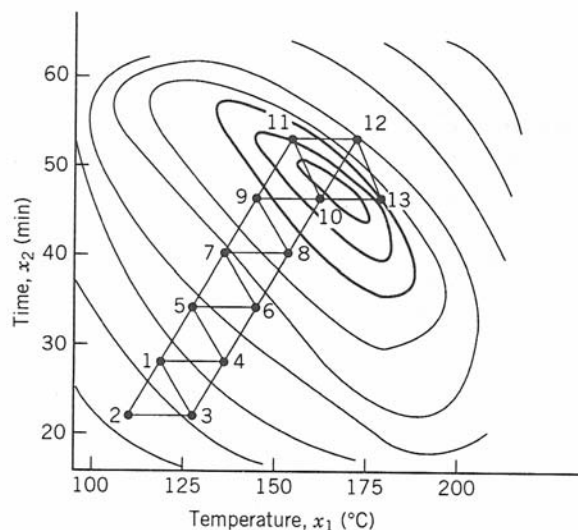
29

EVOP Simplex Regra 3

- Evite ciclos:
 - Se y_i for a menor resposta no m -ésimo simplex e y_i^* a menor resposta no simplex $m+1$, não retorne ao m -ésimo simplex:
 - Saia do simplex $m+1$ descartando o segundo maior erro absoluto

30

Esquema gráfico do EVOP Simplex $p/k = 2$



31

Problemas com o EVOP Simplex

- Não fornece informações sobre interações:
 - Por se tratarem de projetos de resolução III, efeitos principais e interações de 1ª ordem estão vinculadas
- No caso de presença de erro experimental moderado associado à resposta, o procedimento tende a se comportar de maneira errática:
 - Replicações dos pontos eliminaria esse problema, mas agregaria mais rodadas experimentais ao projeto:
 - Por que não rodar um fatorial, nesses casos?

32

Conselhos práticos sobre o EVOP

- Quanto à lista de variáveis:
 - Comece com as duas ou três mais importantes segundo a equipe de processo:
 - Se ao cabo de 5 – 8 ciclos nenhum efeito significativo emergir, troque as variáveis ou seus níveis de investigação
 - É relativamente freqüente que uma variável não seja considerada como importante pelo fato de nunca ter sido variada em busca de um melhor ajuste

33

Conselhos práticos sobre o EVOP (Cont.)

- Use o desvio-padrão da resposta medida em vários ciclos como segunda variável de resposta:
 - Processo será otimizado com relação à rendimento e dispersão
- Treinamentos em EVOP para pessoal de fábrica pode ser feito em ½ dia:
 - Não aplique EVOP sem treinamento, já que trata-se de um procedimento orientado ao pessoal operacional

34

Resumo do EVOP

- **Variação do planejamento fatorial:**
 - não se trata de modificação do método básico, mas sim de uma forma simples de utilização da técnica para processos contínuos.
- **Características mais importantes:**
 - Efetuado durante um processo contínuo, que não pode ser parado para efetuar o estudo com o projeto fatorial comum
 - Para evitar grandes mudanças no processo, somente pequenos intervalos nos níveis das variáveis são estudados
 - Para determinar o efeito desta mudança, diversas medidas são feitas nestas condições (número de vezes maior que o normalmente utilizado na escala de laboratório) e faz-se uma média das observações.

35

Resumo do EVOP (Cont.)

- **Características mais importantes:**
 - Realizado por operadores de processos e no próprio processo
 - Efetuado enquanto o processo se realiza, logo não podem ser estudadas muitas variáveis simultaneamente, 2 ou 3 no máximo
 - Custo da aplicação do EVOP é baixo
 - Costuma-se adotar o fatorial 2^k com um ponto central
 - Após analisar-se os resultados, efetuam-se novos fatoriais deslocando as variáveis no sentido que o fatorial anterior indicou:
 - Repete-se o processo até encontrar as condições ótimas

36

Exercício

Yield during four cycles of a chemical process is shown in the following table. The variables are percent concentration (x_1) at levels 50, 55, and 60 g/liter and temperature (x_2) at 240°C, 242°C, and 244°C. Analyze by EVOP worksheet methods.

Cycle	Conditions				
	(1)	(2)	(3)	(4)	(5)
1	80.7	79.8	80.2	84.2	77.5
2	79.1	82.8	82.5	84.6	78.3
3	76.6	79.1	79.0	82.3	81.1
4	80.5	79.8	84.5	81.0	80.1