

CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA DE MINAS GERAIS

Engenharia de Automação Industrial

OTIMIZAÇÃO DE PROCESSOS COM AQUISIÇÃO DE DADOS VIA

PADRÃO OPC:

Uma possibilidade a ser explorada

Izabela Caroline Araujo

Araxá-MG

2012

Izabela Caroline Araujo

OTIMIZAÇÃO DE PROCESSOS COM AQUISIÇÃO DE DADOS VIA

OPC:

Uma possibilidade a ser explorada

Trabalho de conclusão de curso apresentado ao curso de Engenharia de Automação Industrial do Centro Federal de Educação Tecnológica de Minas Gerais (CEFET-MG), campus Araxá, pela aluna Izabela Caroline Araujo.

Orientador (a): Prof. Natal Junio Pires

Araxá-MG

2012

Dedico este trabalho aos meus queridos pais Marcio Donizette de Araújo e Sueli da Silva Faria Araujo por todo o amor, esforço e dedicação.

AGRADECIMENTO

Agradeço primeiramente ao meu Deus Jeová, por seu amor e generosidade. Sem Suas bênçãos e proteção meu caminhar não seria bem sucedido.

Aos meus amados pais Marcio e Sueli, que me apoiaram, incentivaram, tornaram essa conquista possível e tantas vezes foram altruístas em meu favor. Por todo carinho, dedicação e amizade o meu muito obrigada e o meu pedido de desculpas por muitas vezes estar ausente durante esse caminho.

Aos meus queridos irmãos Thiago e Poliana agradeço pela amizade e apoio. A presença de vocês alegra minha vida.

Aos meus amigos e colegas de sala agradeço pelo companheirismo, incentivo, convivência harmoniosa e apoio nos momentos mais difíceis. Vocês com certeza tornaram essa jornada mais divertida.

Ao colega Luis Paulo Fagundes por ser tão generoso, pela força, colaboração e confiança que deposita em mim, pela felicidade de ter um amigo tão dedicado.

Ao meu orientador Prof. Natal Junio Pires por ser um orientador tão amigo, por ser um professor tão dedicado e acessível e por inspirar tanta confiança.

Aos demais professores por serem generosos e ensinarem com boa vontade.

Enfim, a todos que sempre estiveram presentes na minha vida. Sem vocês essa caminhada não seria tão feliz.

Muito Obrigada!

RESUMO

Otimização de processos é um assunto de interesse dos profissionais de processo, uma vez que estes enfrentam grande desafio quanto à sintonização de parâmetros e obtenção da melhor resposta devido ao fato de muitos processos industriais dependerem de variáveis que podem ser operadas numa ampla faixa de medição. Uma outra deficiência é percebida quando as tentativas de se otimizar são realizadas de acordo com dados colhidos em laboratórios e plantas piloto, uma vez que tais parâmetros diferem dos processos reais, exigindo-se assim uma readequação dos mesmos para utilização em processos reais. Essa utilização de plantas piloto e laboratórios se deve ao fato de que a aquisição de dados para otimização do processo pode consistir em uma atividade demorada e até inviável devido à difícil coleta dos dados de processo que podem ser numerosos e possivelmente estarem separados por longas distâncias. Acreditamos que seja possível obter parâmetros de uma planta já existente de forma ágil com aquisição de dados online via padrão OPC e manipulá-los a fim de garantir a melhor resposta do processo utilizando planejamento fatorial e metodologia de superfície de resposta. Para isso, faremos uso de pesquisa bibliográfica e experimentação computacional.

Palavras-chave: Otimização. Metodologia de superfície de resposta. Planejamento fatorial. Aquisição de dados. OPC.

ABSTRACT

Process optimization is an interesting subject to process's professionals, once they face a big challenge with the tuning parameters and obtaining the best answer due to the fact many industrial processes depend on variables that can be operated in a wide measuring range. Another deficiency is perceived when the attempts to optimize are made according to data collected in laboratories and pilot plants, since these parameters differ from the actual processes, thus requiring readjustment for use in real processes. This use of pilot plants and laboratories is due to the fact that the data acquisition for process optimization can consist of a time consuming activity and even infeasible due to the difficult process of data collection that can be numerous and possibly be separated by long distances. We believe it is possible to obtain parameters of an existing plant in an agile way with online data acquisition via OPC and manipulate them to ensure the best response process using response surface methodology and experimental design. In order to do so, we are going to use literature and computational experimentation.

Keywords: Optimization. Response surface methodology. Experimental design. Data acquisition. OPC.

LISTA DE FIGURAS

Figura 1 - Desvio dos dados em relação ao modelo estimado de regressão.....	20
Figura 2 - Espaço experimental para misturas de (a) dois, (b) três e (c) quatro componentes.....	25
Figura 3 - Rendimento contra tempo de reação, com temperatura constante em 155°F.....	32
Figura 4 - Rendimento contra temperatura, com o tempo de reação constante em 1,7 hora.....	33
Figura 5 - Experimento de otimização variando um fator de cada vez.....	33
Figura 6 - Um sistema em estudo com respostas influenciadas por fatores de entrada.....	34
Figura 7 - Experimento fatorial sem interação.....	37
Figura 8 - Experimento fatorial com interação com dados referentes ao exemplo citado.....	38
Figura 9 - Gráfico de probabilidade normal de resíduos.....	44
Figura 10 - Gráfico de resíduos contra fator A.....	44
Figura 11 - Gráfico dos resíduos contra fator B.....	45
Figura 12 - Planejamento fatorial de dois níveis com ponto central.....	48
Figura 13 - Plano descrito pela equação $y=68,00-5,25x_1+4,25x_2$	52
Figura 14 - Curvas de nível descritas pela equação $y=68,00-5,25x_1+4,25x_2$ e trajetória de máxima inclinação.....	53
Figura 15 - Planejamento em estrela para duas variáveis codificadas correspondente à Tabela 9.....	58
Figura 16 - Superfície quadrática descrita pela equação $y=89+1,51x_1-2,36x_2-2,81x_{12}-2,81x_{22}+1,75x_1x_2$	61
Figura 17 - Curvas de nível descritas pela equação $y=89+1,51x_1-2,36x_2-2,81x_{12}-2,81x_{22}+1,75x_1x_2$	62
Figura 18 - YSI 2300 Analisador de Glicose e Lactose.....	65
Figura 19 - Arquitetura cliente-servidor OPC.....	68
Figura 20 - Os nove drivers utilizados na solução proprietária.....	69
Figura 21 - Requisições aos diferentes dispositivos na solução OPC.....	70

Figura 22 - Inserção de novo nó como fonte de dados no servidor OPC.....	74
Figura 23 - Seleção do padrão de comunicação.....	74
Figura 24 - Configuração do nó modbus ethernet.....	75
Figura 25 - Tela de configuração do OPCWareXL.....	75
Figura 26 - Configuração do servidor OPC com o software OPCWareXL.....	76
Figura 27 - Configuração de grupo no software OPCWareXL.....	77
Figura 28 - Tela de localização dos itens do servidor OPC.....	78
Figura 29 - OPCWareXL com os itens já configurados.....	79
Figura 30 - Simulador do CLP, usando o software Concept.....	79
Figura 31 - Tela inicial da aplicação para otimização de processos.....	81
Figura 32 - Modelo linear com base no planejamento fatorial de dois níveis com ponto central.....	82
Figura 33 - Análise de variância do modelo linear com base no planejamento fatorial de dois níveis com ponto central.....	82
Figura 34 - Teste F para determinar a validade do modelo.....	83
Figura 35 - Planejamento seguindo o caminho de máxima inclinação.....	83
Figura 36 - Incompatibilidade entre o modelo linear e os dados provenientes do planejamento fatorial.....	84
Figura 37 - Tela de ajuste do planejamento estrela.....	85
Figura 38 - Resultados do modelo quadrático proveniente do planejamento estrela.....	86
Figura 39 - Análise de variância para o modelo quadrático.....	86
Figura 40 - Teste F e valor ótimo das variáveis.....	87

LISTA DE TABELAS

Tabela 1 - Dados para um experimento fatorial.....	35
Tabela 2 - Dados para um planejamento fatorial de dois fatores.....	39
Tabela 3 - Tabela de Análise de variância para um modelo de um fatorial com dois fatores e efeitos fixos.....	42
Tabela 4 - Resultado de um planejamento fatorial 2^2 com ponto central.....	48
Tabela 5 - Análise de variância para o ajuste do modelo $y=68,00-5,25x_1+4,25x_2$	51
Tabela 6 - Caminho com máxima inclinação para o modelo estimado.....	54
Tabela 7 - Resultado do novo planejamento fatorial 2^2 com ponto central.....	55
Tabela 8 - Análise de variância para o ajuste do modelo $y=85,71+1,25x_1-2,25x_2$	56
Tabela 9 - Resultados do planejamento em estrela obtido com ampliação do planejamento da tabela 7.....	57
Tabela 10 - Análise de variância para o ajuste do modelo $y=89+1,51x_1-2,36x_2-2,81x_{12}-2,81x_{22}+1,75x_1x_2$ aos dados da Tabela 9.....	61

SUMÁRIO

INTRODUÇÃO.....	11
1 REVISÃO BIBLIOGRÁFICA.....	15
1.1 Introduzindo o assunto otimização.....	15
1.2 Métodos de Otimização.....	18
1.2.1 Mínimos Quadrados.....	18
1.2.2 Análise de Misturas.....	23
1.2.3 Simplex.....	28
1.3 Planejamento fatorial e metodologia de superfície de resposta.....	30
1.3.1 Planejamento Fatorial.....	31
1.3.2 Metodologia de Superfície de Resposta.....	45
1.3.3 Aplicações.....	63
1.3.4 Considerações sobre os métodos.....	64
2 PROCEDIMENTOS EXPERIMENTAIS.....	65
2.1 Aquisição de dados via OPC.....	65
3 RESULTADOS E DISCUSSÃO.....	72
3.1 Aplicação.....	72
3.1.1 Configuração de OPC para aquisição de dados e integração com Excel.....	73
3.1.2 Descrição da aplicação de otimização usando planejamento fatorial e metodologia de superfície de resposta.....	80
CONCLUSÃO.....	88
REFERÊNCIAS BIBLIOGRÁFICAS.....	89

INTRODUÇÃO

A otimização de processos industriais têm assumido uma importância crucial frente às novas demandas de racionalização de recursos, os mais diversos possíveis. Em um planeta onde a maioria dos recursos explorados são não renováveis é trivial dizer da importância que assume a otimização (racionalização) no uso e exploração destes bens findáveis, os quais têm importantes influências no desenvolvimento tecnológico e econômico do país. O surgimento dos métodos de otimização ocorreu em função da necessidade de se aperfeiçoar o desempenho dos mais variados sistemas. Otimizar implica em tornar algo “tão perfeito, efetivo ou funcional quanto possível”. Assim, pode-se entender a otimização como sendo um processo baseado em instruções que viabilizam a obtenção do melhor resultado de uma determinada situação (EIRAS, 2000).

Muitos processos na indústria apresentam produtos que dependem de variáveis que podem ser operadas em ampla faixa de medição e/ou possibilidades. Assim sendo, os profissionais de processo enfrentam grande desafio para sintonização dos parâmetros do processo e obtenção da melhor resposta. A aquisição de dados para otimização do processo pode consistir em uma atividade demorada e até inviável devido à difícil coleta de dados que podem ser numerosos e possivelmente estarem separados por longas distâncias. Dessa forma, é desejável empreender novas estratégias e interfaces de obtenção de dados que possam ser tratados com as técnicas disponíveis, e assim garantir que possam ser tomadas decisões rápidas e acertadas acerca dos processos que se desejam otimizar.

Quanto mais rápido e facilitado forem os processos de obtenção dos valores das variáveis que supostamente influem um dado processo, tanto mais rápido e eficiente serão as tomadas de decisões para se determinar quais parâmetros deverão ser ajustados para novos valores visando a racionalização do processo.

Nesta perspectiva, torna-se primordial e parece-nos ser bastante produtivo pensar em formas de integrar técnicas de otimização com formas já conhecidas de obtenção de dados acerca de processos industriais.

Os procedimentos de otimização usualmente consistem em ajustes e coleta de dados do processo. Tais procedimentos podem demandar muito tempo devido ao possível grande número de variáveis e elevada distância dos pontos de coleta e da parametrização das mesmas. Essas dificuldades têm levado os profissionais de otimização a realizarem testes em plantas piloto e laboratórios, o que faz com que os parâmetros de otimização obtidos não sejam compatíveis com os do processo, tornando necessária uma reparametrização.

De acordo com a literatura, como no artigo *Modelling tempering behaviour of dark chocolates from varying particle size distribution and fat content using response surface methodology* (AFOAKWA; et al., 2008), onde foram relatados estudos sobre a têmpera do chocolate utilizando metodologia de superfície de resposta, os experimentos geralmente são realizados com dados obtidos em laboratório, com equipamentos específicos e registrados manualmente.

A utilização de parâmetros obtidos da própria planta industrial garante a fidelidade dos resultados obtidos com a otimização, diferentemente dos dados obtidos em plantas piloto e laboratórios, que requerem reparametrização para aplicação em sistemas reais.

Outro aspecto observado foi o uso de softwares proprietários para tratamento dos dados obtidos como, por exemplo, o Minitab (BINGOL; et al, 2011) ; Statgraphics (AFOAKWA; et al, 2008), *Design Expert* (CHEN; et al,2011) e Statistica (LI; LIN; TRAN, 2011).

Esses softwares possuem custos elevados, pois apresentam muitas funcionalidades que nem sempre são utilizadas nos processos de otimização.

Nesse sentido, faremos uso do padrão OPC como meio para aquisição de dados, por apresentar uma boa intercambiabilidade entre os sistemas de controle mais

difundidos no mercado, e trataremos os dados por meio de um algoritmo criado em Excel®.

O uso de sistemas de controle já consolidados juntamente com aquisição de dados via OPC permitem também maior agilidade na coleta de dados e parametrização do sistema, uma vez que não é necessário o deslocamento de pessoal de dentro da sala de controle para o campo.

Dentre as possibilidades de otimização de processos para aplicações industriais nos concentraremos no planejamento fatorial e metodologia de superfície de resposta. Portanto, com este trabalho, mesmo sendo ele de pequeno fôlego, esperamos contribuir de certa forma para fortalecer a necessidade primeira de otimizar a exploração de nossos recursos e segundo incentivar o uso de ferramentas que já existem nos processos industriais e interfaciá-las com técnicas de otimização, as quais vem encontrando um campo promissor nesse sentido.

Desta forma, temos os seguintes objetivos:

Objetivo Geral: Investigar a possibilidade de obtenção da melhor resposta do processo por meio de planejamento fatorial e metodologia de superfície de resposta com aquisição de dados via OPC.

Objetivos Específicos: Averiguar métodos de otimização, bem como os métodos de aquisição de dados. Estudar e aplicar os métodos de planejamento fatorial e superfície de resposta. Realizar a aquisição de dados via padrão OPC.

Este trabalho está dividido da seguinte maneira:

Capítulo 1: Trata alguns conceitos sobre otimização e apresenta algumas técnicas de otimização, bem como algumas aplicações.

Capítulo 2: Aborda a aquisição de dados via OPC com suas principais características, bem como as vantagens trazidas por este padrão.

Capítulo 3: Descreve uma aplicação de Excel® que manipula dados através de planejamentos fatoriais e metodologia de superfície de resposta e que faz a aquisição de dados via OPC.

Por fim, apresentamos as conclusões do trabalho e perspectivas de trabalhos futuros.

1 REVISÃO BIBLIOGRÁFICA

1.1 Introduzindo o assunto otimização

O uso da palavra otimizar com o significado de busca do ótimo é relativamente novo no Brasil. As atividades de busca dos melhores resultados recebiam a denominação de “Acompanhamento Operacional” e em muitas plantas ainda são assim designadas (Intech, 2009).

O que se busca com a otimização em processos industriais não é o melhor absoluto ou o ótimo a qualquer custo, mas busca-se o melhor de acordo com as condições apresentadas pela planta. Essas condições podem ser a campanha que a planta trabalha, a disponibilidade de matéria prima, as limitações logísticas, demandas por produtos, demandas comerciais, dentre outras.

O sentido matemático de otimização – manipular variáveis livres de um sistema, com a finalidade de se obter o valor máximo ou mínimo de uma função com várias variáveis sujeita a um conjunto de restrições – é bastante apropriado quando tratamos de processos industriais, uma vez que nestes, o processo de otimização tenciona levar a planta ao melhor ponto para a utilização das instalações, dadas as restrições existentes e, para isso, testa a sensibilidade do processo à manipulação das variáveis com a finalidade de se descobrirem as condições de melhor resultado.

A função da qual se busca o ponto de máximo ou mínimo é, muitas vezes, chamada de “função objetivo”; as variáveis livres são aquelas que permitem manipulação para modificar o valor da função objetivo e as restrições geralmente são variáveis dependentes, resultantes das ações exercidas nas variáveis livres (GUIMARÃES NETO; NASCIMENTO; NUNES, 2009).

Os métodos de otimização são procedimentos que buscam a solução mais adequada naquele momento, levando em conta as restrições. Por meio de seu uso, é possível alcançar diversos objetivos específicos. A implementação das soluções obtidas através das técnicas de otimização possibilitam a fabricação de produtos com melhores características, a diminuição do tempo de desenvolvimento, o aumento da produtividade dos processos, a minimização da sensibilidade dos produtos às variações das condições ambientais, melhoria no aproveitamento de insumos e assim por diante.

As técnicas de otimização podem ser aplicadas nas mais diversas áreas de conhecimento, como gestão de cadeias de suprimento (produção, distribuição); transporte (itinerários, gestão de frotas e tripulação); indústria petroquímica (aquisição de materiais, distribuição, projeto e operação de refinarias); finanças (operação de carteiras, gestão de recursos financeiros); aplicações militares (logística, alocação de pessoal e operações de guerra). Ainda pode ser citada a otimização do projeto de estruturas de veículos (terrestres, aeroespaciais e marítimos) e otimização de equipamentos para diversas indústrias de processamento em geral.

Para a engenharia o estudo e emprego das técnicas de otimização também são importantes, pois a procura por melhores resultados é um dos seus grandes objetivos. Assim, ela tem investido na área de otimização para que a aplicação desta ferramenta em problemas de grande porte seja possível. Há um grande interesse em se desenvolver programas de otimização eficientes e robustos, de modo a atender a uma gama crescente de necessidades humanas (GOULART, 2005).

É bastante comum, especialmente em indústrias químicas, surgirem problemas em que é necessário estudar ao mesmo tempo várias propriedades que sejam afetadas por um grande número de fatores experimentais. É atribuição da otimização responder questões como por exemplo, quais fatores experimentais devem ser controlados para assegurar a qualidade do produto, como investigar os efeitos desses fatores sobre as propriedades, visando minimizar o trabalho necessário e o custo dos experimentos, melhorando a qualidade do produto (BARROS NETO; SCARMINIO; BRUNS, 2007).

No presente trabalho pretendemos mostrar que o emprego dos métodos de otimização pode ajudar a responder a essas questões de forma racional e econômica. Utilizando planejamentos experimentais baseados em princípios estatísticos, é possível extrair do sistema em estudo o máximo de informação útil, realizando um número mínimo de experimentos.

Mais importante do que analisar os dados obtidos, é planejar os experimentos, projetá-los de forma que eles sejam capazes de fornecer a informação que se procura. Se o experimento for cuidadosamente planejado, é possível tanto minimizar os custos operacionais, como garantir que os resultados do experimento contenham informações úteis para a solução do problema de partida (BARROS NETO; SCARMÍNIO; BRUNS, 2007).

Como já supracitado, a otimização trata do problema da busca da melhor utilização de um conjunto limitado de recursos, escolhendo a alternativa que maximize ou minimize um determinado fator, dentre todas aquelas que satisfazem certas restrições. Contudo, para resolver um problema, é necessário primeiro formulá-lo.

A utilização de modelos facilita o processo de análise e tomada de decisão. Eles permitem que a decisão seja mais bem avaliada e testada, pois possibilitam a experimentação da solução proposta antes de efetivamente implementá-la. A economia obtida justifica a utilização dos métodos de otimização.

Para formular o modelo, é preciso definir o objetivo básico do problema, ou seja, a otimização a ser alcançada. Por exemplo, maximização de lucros, de desempenhos, de bem-estar social; minimização de custos, de perdas, de tempo. Tal objetivo será representado por uma função objetivo, a ser maximizada ou minimizada. Para que esta função objetivo seja matematicamente determinada, devem ser definidas as variáveis de decisão, ou variáveis livres, envolvidas. Essas variáveis são as incógnitas a serem determinadas pela solução do modelo. Podem ser, por exemplo, número de máquinas, as classes de investimento à disposição, a área a ser explorada etc. Normalmente, assume-se que estas variáveis só possam assumir valores positivos. Estas variáveis normalmente estão sujeitas a um conjunto de

restrições, determinadas pelas limitações do sistema, as quais podem ser, por exemplo, quantidade de equipamento disponível, tamanho da área a ser explorada, capacidade de um reservatório, exigências nutricionais para determinada dieta, entre outros tantos exemplos.

A seguir vamos discorrer sobre alguns métodos de otimização que são de ajuda para o planejamento do experimento e obtenção do modelo.

1.2 Métodos de Otimização

1.2.1 Mínimos Quadrados

Uma técnica para criar modelos de fenômenos a partir de tabelas de dados experimentais é a minimização quadrática. Quando se cria um modelo, busca-se uma função que expresse a melhor relação entre os dados.

Esse tipo de aproximação de funções tem uma perspectiva diferente da interpolação. Para exemplificar, podemos pensar numa situação em que temos apenas os valores da função em certos pontos. Para a aproximação por regressão não se exige que a função aproximadora interpole dados. Exige-se apenas que essa função tome valores que minimizem o somatório dos resíduos.

O método de mínimos quadrados é uma técnica de otimização matemática que tem por objetivo determinar uma função que aproxima um conjunto de pontos a partir de combinação linear de funções simples. Esse critério para estimar os coeficientes de regressão, além de buscar o melhor ajuste para o conjunto de dados, também procura minimizar a soma dos quadrados das diferenças entre o valor estimado e os dados observados, sendo que essas diferenças são chamadas de resíduos.

Então, quando temos uma distribuição de pontos aplicamos este método para ajustar a melhor curva a este conjunto de dados.

Para o caso em que a função de ajuste é uma função linear temos que:

$$\hat{y}_i = a + bx_i$$

onde i indica cada uma das n observações da base de dados. Assim como em algumas referências, adicionamos o símbolo “^” como notação para as estimativas.

Para que esta seja a reta que melhor se ajusta aos dados, é preciso minimizar a soma das diferenças entre os valores de $f(x)$ tabelados y_i e os valores da curva de ajuste $a + bx_i$.

Essa diferença pode ser tanto positiva quanto negativa e a soma dos desvios é nula, isto é:

$$\sum_{i=1}^n e_i = 0$$

onde e_i representa cada erro, ou resíduo da regressão, e seu valor é dado pela diferença entre a resposta observada e a resposta estimada.

Para evitar o cancelamento, minimizamos o quadrado da diferença, ou seja, minimizamos:

$$\sum_{i=1}^n e_i^2$$

Daí muitas vezes fazemos referência a esse método como sendo método dos mínimos quadrados. Um requisito para este método é que o desvio seja distribuído aleatoriamente, com distribuição normal e independente.

O método dos mínimos quadrados ajuda a encontrar valores de a e de b que minimizem $S(a,b)$. Esses parâmetros são estimadores que minimizam a soma dos

quadrados dos resíduos da regressão, de forma que o grau de ajuste do modelo aos dados observados seja maximizado.

Com os erros, resíduo ou desvios acrescentados, o que buscamos estimar é:

$$\hat{y}_i = a + bx_i + e_i$$

A figura a seguir (FIG.1) mostra um diagrama típico de dispersão dos dados observados e uma candidata para linha estimada de regressão.

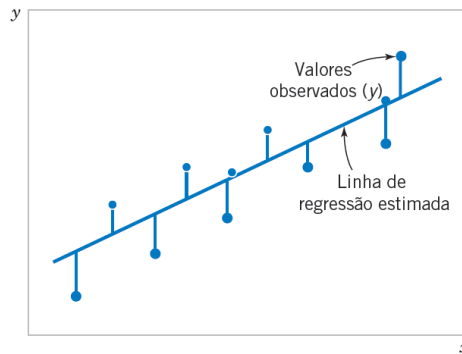


Figura 1 - Desvio dos dados em relação ao modelo estimado de regressão
Fonte: MONTGOMERY; RUNTER, 2009.

Minimizando a soma do quadrado dos resíduos, encontramos a e b que fornecem a menor diferença entre a previsão de \hat{y} e o y realmente observado.

Substituindo e_i por $y_i - a - bx_i$, temos:

$$S(a, b) = \sum_{i=1}^n e_i^2$$

Para que encontrar os valores de a e b que minimizem o somatório, é necessário derivar $\sum e_i^2$ em relação a e b igualar a zero (Box, Hunter e Hunter, 1978, capítulo 14; Pimentel e Neto, 1996 apud NETO; SCARMINIO; BRUNS, 2007):

$$\frac{\partial S}{\partial a} = -2 \sum_{i=1}^n (y_i - a - bx_i) = 0 \quad (I)$$

$$\frac{\partial S}{\partial b} = -2 \sum_{i=1}^n x_i (y_i - a - bx_i) = 0 \quad (II)$$

Se distribuirmos e dividirmos a expressão (I) por $2n$, temos:

$$\frac{-2 \sum_{i=1}^n y_i}{2n} + \frac{2 \sum_{i=1}^n a}{2n} + \frac{2 \sum_{i=1}^n bx_i}{2n} = \frac{0}{2n}$$

$$\frac{-\sum_{i=1}^n y_i}{n} + \frac{\sum_{i=1}^n a}{n} + \frac{\sum_{i=1}^n bx_i}{n} = 0$$

$$-\bar{y} + a + b\bar{x} = 0$$

$$a = \bar{y} - b\bar{x},$$

onde \bar{y} e \bar{x} são as médias amostrais de y e x respectivamente.

Se substituirmos esse resultado na expressão (II) temos:

$$-2 \sum_{i=1}^n x_i (y_i - \bar{y} + b\bar{x} - bx_i) = 0$$

$$\sum_{i=1}^n$$

$$\sum_{i=1}^n x_i (y_i - \bar{y}) + b \sum_{i=1}^n x_i (\bar{x} - x_i) = 0$$

$$b = \frac{\sum_{i=1}^n x_i (y_i - \bar{y})}{\sum_{i=1}^n x_i (x_i - \bar{x})}$$

Alguns livros usam uma fórmula que tem resultado equivalente a esse:

$$b = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

Esse método possibilita tanto criar o modelo quando se tem um conjunto finito de valores dados, o que representa uma perspectiva discreta, como também criar modelos para casos contínuos, em que temos interesse em aproximar a função para todo um intervalo.

Outra forma de calcular os valores de a e b é possível por meio da resolução de equações matriciais, em que:

$$X^t X = \begin{bmatrix} n & \sum X_i \\ \sum X_i & \sum X_i^2 \end{bmatrix} \quad e \quad X^t y = \begin{bmatrix} \sum y_i \\ \sum X_i y_i \end{bmatrix}$$

e

$$X^t X b = X^t y$$

Para resolver essa equação matricial, é preciso multiplicá-la à esquerda pela inversa de $X^t X$, e então isolar o vetor b , que contém os elementos estimadores que estamos procurando.

$$(X^t X)^{-1} (X^t X) b = (X^t X)^{-1} X^t y$$

$$I b = (X^t X)^{-1} X^t y$$

$$b=(X^t X)^{-1} X^t y$$

Essa equação é muito importante, pois se ampliarmos as matrizes X e y adequadamente, chegaremos a solução geral para o ajuste de um modelo pelo método dos mínimos quadrados, para qualquer que seja o número de observações ou de parâmetros necessários para caracterizar o modelo. No entanto, para que a solução exista, é necessário que a matriz $(X^t X)^{-1}$ não seja singular e que os modelos sejam lineares nos parâmetros, ou seja, não deve haver termos como b_0^2 ou $b_0 b_1$ (BARROS NETO; SCARMINIO; BRUNS, 2007).

No texto sobre metodologia de superfície de resposta usaremos o método dos mínimos quadrados em um exemplo e então seu uso ficará mais claro.

1.2.2 Análise de Misturas

Um experimento com mistura é aquele em que duas ou mais substâncias são misturadas para que se forme um produto final, e a resposta a ser medida não consiste em valores absolutos, e sim em é uma função das proporções dos ingredientes presentes.

Os planejamentos experimentais para o estudo de misturas têm como objetivo geral a modelagem e análise da superfície de resposta da mistura, partindo de um pequeno número de observações, que são realizadas combinando os ingredientes de acordo com uma pré-seleção, o que resulta em misturas de diferentes composições.

Esse método permite obter predições da resposta para qualquer que seja a composição da mistura, desde que esteja dentro da região abrangida pelo experimento, permite quantificar a influência de cada componente na resposta, estando ele separado ou em combinação com outros, além de permitir estimar o

ponto, ou região do espaço experimental, que gere a melhor resposta, o ajuste ótimo (NUNES, 1998).

Como já dito anteriormente numa mistura os componentes não representam quantidades, e sim proporções de uma mistura, sua composição. Logicamente, essas proporções devem ser não negativas e sua soma deve ser sempre igual a 1, ou em termos percentuais igual a 100%.

Por causa da restrição de que o somatório deve ser igual a 100%, as outras metodologias de modelagem devem ser modificadas para se ajustarem aos problemas de mistura. Esses métodos modificados são largamente empregados em aplicações na ciência, na engenharia e na indústria (BARROS NETO; SCARMINIO; BRUNS, 2007).

Dessa forma, o espaço de experimentação de uma mistura com q componentes, na qual a proporção do i -ésimo componente é representada por x_i , fica reduzido a uma região simplex limitada pelas condições:

$$x_i \geq 0, \leq i \leq q$$

Para uma mistura de q componentes, podemos escrever:

$$\sum_{i=0}^q x_i = 100\% = 1$$

As variáveis x_i são linearmente dependentes, pois a alteração da proporção de um componente na mistura provoca uma mudança na proporção de pelo menos um dos demais.

A forma do espaço experimental varia de acordo com o número de elementos da mistura, como exemplificado na figura 2:

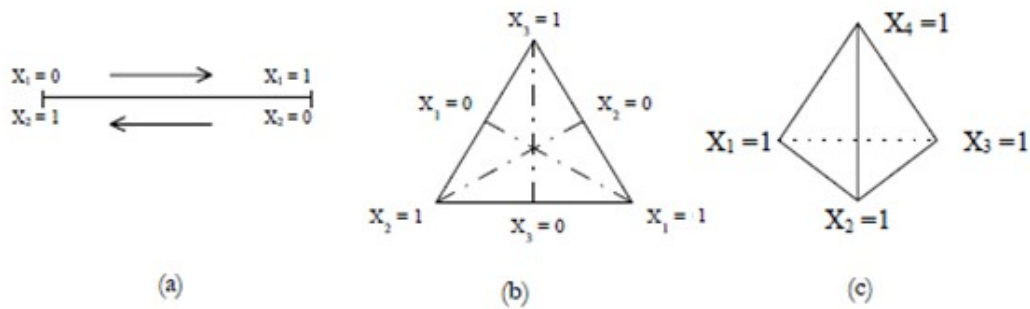


Figura 2 - Espaço experimental para misturas de (a) dois, (b) três e (c) quatro componentes.
Fonte: NUNES, 1998.

De forma geral, a investigação das propriedades de uma mistura se dá de forma bem semelhante aos sistemas com variáveis independentes. Primeiro postulamos um modelo para descrever a forma como as propriedades de interesse variam de acordo com a composição da mistura. Depois, realizamos um planejamento para determinar quais as composições da mistura devem ser estudadas. E então, ajustamos o modelo aos resultados experimentais e o avaliamos.

A modelagem mais simples para uma mistura de dois componentes é o modelo aditivo, ou linear:

$$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon_i$$

onde y_i é a resposta de interesse, β_0 , β_1 e β_2 são os parâmetros do modelo e ε_i representa o erro aleatório. Ajustando o modelo de acordo com as observações feitas, temos:

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2,$$

que nos possibilita estimar a resposta média em um ponto qualquer de composição (x_1, x_2).

Para obter modelos mais fáceis de interpretar, ao invés de usar diretamente a restrição $x_1 + x_2 = 1$, para eliminar x_1 ou x_2 da equação, vamos usar uma estratégia

diferente. Já que a soma x_1+x_2 é sempre igual a 1, se a introduzirmos como coeficiente de b_0 na equação anterior, a igualdade não será alterada:

$$\hat{y} = b_0(x_1+x_2) + b_1x_1 + b_2x_2$$

Fazendo um pequeno rearranjo temos:

$$\hat{y} = (b_0+b_1)x_1 + (b_0+b_2)x_2 = b_1^{\hat{}}x_1 + b_2^{\hat{}}x_2$$

onde $b_i^{\hat{}} = b_0+b_i$. Dessa forma, é possível termos apenas dois coeficientes para serem determinados ($b_1^{\hat{}}$ e $b_2^{\hat{}}$) em vez dos três que apareceram na equação original.

Para determinar esses parâmetros, consideramos as situações em que só há o componente 1, e a equação se reduziria a $\hat{y} = b_1^{\hat{}}y_1$, onde y_1 é a resposta observada para o componente 1 puro e a situação em que só há o componente 2, e a equação se reduziria a $\hat{y} = b_2^{\hat{}}y_2$, onde y_2 é a resposta observada para o componente 2 puro. Assim, os dois coeficientes no modelo aditivo são as próprias respostas dos respectivos componentes puros. Se o modelo for válido, podemos prever as propriedades de uma mistura sem ter sido necessário fazer nenhuma mistura sequer. Isso pode ser observado na figura 2(a) em que percebemos que a superfície de resposta, que neste caso é unidimensional, é uma reta que une y_1 e y_2 . Para qualquer que seja a mistura, a resposta será a média ponderada das respostas dos componentes puros.

É possível aumentar a precisão do modelo por repetir os ensaios com os componentes puros. Assim, teremos $b_1^{\hat{}} = \bar{y}_1$ e $b_2^{\hat{}} = \bar{y}_2$, em que \bar{y}_1 e \bar{y}_2 são as médias das respostas. O erro padrão para $b_1^{\hat{}}$ e $b_2^{\hat{}}$ pode ser obtido através da expressão:

$$\Delta b_i^{\hat{}} = \frac{s}{\sqrt{n_i}}$$

em que s é uma estimativa do erro padrão de uma resposta, obtida através de repetidas observações, n_i representa o número de observações usadas para calcular o valor médio \bar{y}_i e $i=1, \dots, q$.

Apesar de os resultados com componentes puros determinarem completamente o modelo linear, é preciso realizar experimentos com misturas binárias a fim de verificar se o modelo é realmente adequado. No caso de a composição ter efeitos mais complicados sobre a resposta, é necessário um modelo mais elaborado.

O modelo quadrático é a forma mais simples de ampliação do modelo linear, que é definido pela expressão:

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + b_{11} x_1^2 + b_{22} x_2^2 + b_{12} x_1 x_2$$

Esta expressão contém seis parâmetros, mas é possível reduzir esse número para misturas binárias em função da soma constante, resultante em 1, das proporções dos dois elementos. Se substituirmos as identidades $x_1 + x_2 = 1$, $x_1^2 = x_1(1 - x_2)$, $x_2^2 = x_2(1 - x_1)$ e reunirmos os termos em x_1 e x_2 , teremos:

$$\hat{y} = (b_0 + b_1 + b_{11})x_1 + (b_0 + b_2 + b_{22})x_2 + (b_{12} - b_{11} - b_{22})x_1 x_2$$

e finalmente:

$$\hat{y} = b_1' x_1 + b_2' x_2 + b_{12}' x_1 x_2,$$

em que $b_i' = b_0 + b_i + b_{ii}$ e $b_{12}' = b_{12} - b_{11} - b_{22}$. Dessa forma, temos apenas um coeficiente a mais que no modelo linear. Então, só precisamos acrescentar aos dois valores utilizados para determinar o modelo linear outra medida feita numa mistura binária de composição qualquer¹ (BARROS NETO; SCARMINIO; BRUNS, 2007).

¹ É possível obter modelos de misturas com mais componentes ampliando os modelos utilizados para misturas binárias, como pode-se verificar no capítulo 7 do livro *Como fazer experimentos* (NETO; SCARMINIO; BRUNS, 2007).

1.2.3 Simplex

O simplex é uma técnica popular utilizada para determinar, numericamente, a solução de problemas de otimização. Baseia-se em algoritmos simples que utilizam conceitos de álgebra linear e busca encontrar o vértice ótimo pesquisando o subconjunto dos vértices da região factível.

O simplex pode ser definido como uma configuração espacial determinada por $N+1$ vértices num espaço de dimensão igual ou maior que N . Nos métodos de otimização simplex essa configuração é um polígono (ou um equivalente multidimensional) de $p+1$ vértices, onde p é o número de variáveis independentes que queremos ajustar. Caso sejam duas variáveis, o simplex é um triângulo. Caso sejam três, um tetraedro. Se forem quatro ou mais, um hiperpoliedro. O número de fatores define as dimensões em que o simplex se move (BARROS NETO; SCARMINIO; BRUNS, 2007).

Alguns procedimentos deste método são a construção do modelo de acordo com as restrições e a função objetivo impostas e, a partir daí, inicia-se o processo sequencial de resolver um sistema de equações lineares para encontrar as soluções básicas viáveis, maximizando ou minimizando a função objetivo, sempre verificando se a solução atual é ótima e, caso não seja, é necessário determinar uma solução melhor. Verifica-se nos demais vértices qual fornece melhor resposta para a função objetivo que a atual. Então, após fazer essa mudança de vértice na direção que mais contribua para ascensão à melhor resposta, verifica-se novamente se este novo vértice é ótimo. O processo termina quando estando num ponto extremo, todos os outros pontos extremos adjacentes fornecem valores piores para a função objetivo. Essa técnica de pesquisa operacional é utilizada para otimizar:

- Planejamento de produção: por desenvolver planos com menor custo para produção, baseado na previsão de vendas, capacidade disponível da fábrica e fatores reais de custo.

- Distribuição e transporte: em função da demanda de um produto distribuído em diversos locais, da disponibilidade e da quantidade e disposição dos depósitos. De posse dessas informações, é possível determinar as demandas e a quantidade ótima a ser transportada de forma que se minimizem os custos.
- Logística e movimentação de materiais: por aumentar a capacidade de utilização do transporte e manejo de materiais.
- Siderurgia: por determinar quais minérios devem ser carregados no alto forno de modo a produzir, ao menor custo, um determinado aço que siga determinadas especificações de elementos químicos.
- Agricultura: por selecionar os alimentos que devem ser plantados para que o lucro seja máximo, respeitando as características do solo, dos equipamentos disponíveis e do mercado comprador.
- Cálculos de dosagens: para determinar como misturar ingredientes para obter melhor resposta atingindo-se um custo mínimo do produto final.

Os métodos simplex permitem otimizar sistemas controlados por um grande número de variáveis independentes e funcionam bem na presença de erros experimentais. Além disso, apresentam uma vantagem que atrai bastante aos pesquisadores que não têm afinidade com cálculos estatísticos, que é o fato de não exigirem o emprego de testes de significância.

Porém, os métodos simplex apresentam limitações que devem ser levadas em conta quando da decisão de que método de otimização utilizar. Uma primeira limitação a ser destacada é o fato de que este método só pode ser empregado com variáveis quantitativas. Fatores qualitativos, como presença/ausência de determinado elemento ou tipo de catalisador a ser utilizado, não podem fazer parte de uma otimização simplex, em função das características do algoritmo. Outra limitação é que caso o simplex indique condições nas quais não seja possível, ou seja muito difícil, controlar os fatores experimentais, o sistema não poderá ser otimizado por

esse método. Além dessas, outra restrição é que só é possível otimizar uma resposta de cada vez (BARROS NETO; SCARMINIO; BRUNS, 2007).

Numa otimização simplex só é possível passar para o próximo passo depois de se conhecer a resposta da etapa imediatamente anterior. Diferentemente, veremos mais adiante que a metodologia de superfície de resposta permite realizar diversos experimentos simultaneamente, nos métodos simplex só se pode fazer um experimento de cada vez, por isso o termo simplex *sequencial*. Essa particularidade faz com que o emprego do simplex seja mais conveniente em instrumentos de resposta rápida (BARROS NETO; SCARMINIO; BRUNS, 2007).

1.3 Planejamento fatorial e metodologia de superfície de resposta.

Um método de otimização que tem sido amplamente utilizado na modelagem de diversos processos industriais, é a metodologia de superfície de resposta, que consiste em uma técnica baseada em planejamentos fatoriais que possui duas etapas: modelagem e deslocamento. Tais etapas são repetidas até se atingir uma região ótima da superfície de resposta (BARROS NETO; SCARMINIO; BRUNS, 2007).

Contudo, um problema notado no uso do planejamento fatorial e metodologia de superfície de resposta é que o número de ensaios aumenta significativamente de acordo com o número de fatores investigados. Por exemplo, se analisarmos dois níveis de sete variáveis, o número de experimentos para um planejamento fatorial completo seria 2^k com $k = 7$, o que seria equivalente a 128 experimentos. Para solucionar esse problema, é interessante saber que quando o número de fatores analisados aumenta, aumentam também as chances de que um ou mais desses fatores não afetem significativamente a resposta. Então se torna desnecessário realizar todos os ensaios do fatorial completo. Pode-se usar a estratégia de um planejamento fatorial fracionado, que envolve menos do que o conjunto completo dos 2^k experimentos para se obter informação sobre os efeitos principais e as interações de ordem inferior.

1.3.1 Planejamento Fatorial

A validade das conclusões obtidas em um experimento depende bastante da forma como este foi conduzido. Por conseguinte, o planejamento do experimento exerce um papel fundamental na solução futura do problema que motivou o experimento. Por isso, engenheiros e cientistas que desejam melhorar o desempenho de um processo de fabricação consideram o planejamento do experimento uma ferramenta extremamente importante (MONTGOMERY; RUNGER, 2009).

O uso de planejamentos experimentais baseados em conhecimentos estatísticos pode ajudar a solucionar problemas de forma racional e econômica. Permite aos pesquisadores extrair do sistema em estudo o máximo de informação relevante, fazendo para isso um número mínimo de experimentos.

Determinar a influência de um ou mais fatores sobre um outro fator de interesse - a resposta - é um problema bastante comum para quem faz experimentos. O planejamento fatorial é utilizado na análise de experimentos quando se tem vários fatores e é necessário determinar o efeito desses fatores sobre a resposta. A possibilidade de tratar tanto fatores quantitativos quanto qualitativos também torna essa técnica muito atraente.

Os fatores, geralmente, são as variáveis que o experimentador pode manipular. As respostas de interesse serão – ou não – afetadas em função das manipulações provocadas nos fatores. Depois de identificar os fatores e respostas, o próximo passo é definir o objetivo que pretendemos alcançar com esses experimentos, para, então, definir o planejamento mais adequado.

Quando analisamos experimentos com vários fatores, temos os chamados efeitos principais, que são efeitos sobre a resposta causados por apenas um fator, e temos também os efeitos de interação que são causados pela interação de dois ou mais

fatores. Experimentos fatoriais são a única maneira de descobrir interações entre as variáveis (MONTGOMERY; RUNGER; HUBELLE, 2004).

Uma opção ao planejamento fatorial, que infelizmente é usada na prática, é mudar um fator de cada vez em vez de variá-los ao mesmo tempo. Porém, essa não é a maneira mais eficaz de se fazer um experimento. O ótimo só seria alcançado por meio dessa prática em circunstâncias muito particulares, pois são raras as situações em que dois fatores atuem de forma realmente independente (BARROS NETO; SCARMINIO; BRUNS, 2007).

Em Montgomery & Runger, (2009), tem-se um exemplo que ilustra esse procedimento. A situação se trata de um engenheiro interessado em definir os valores de temperatura e tempo que maximizem o rendimento de um determinado processo químico. Supondo que ele fixe a temperatura em 155°F – correspondente ao nível atual de operação – e faça cinco tentativas experimentais em diferentes níveis de tempo, sendo esses: 0,5h; 1h; 1,5h; 2h e 2,5h.

Os resultados desses experimentos são mostrados na figura a seguir (FIG. 3).

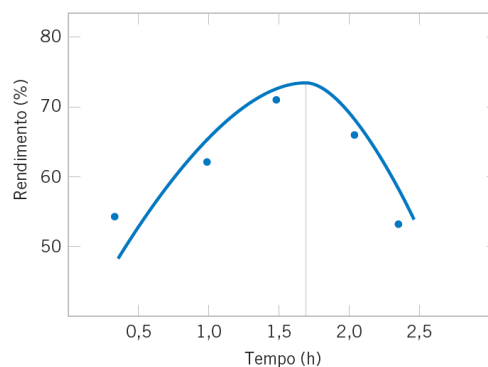


Figura 3 - Rendimento contra tempo de reação, com temperatura constante em 155°F.
Fonte: MONTGOMERY; RUNGER, 2009, p. 333.

O tempo correspondente ao rendimento máximo encontrado é então fixado para que sejam realizadas cinco corridas em temperaturas diferentes, como 140°F, 150° F, 160°F, 170°F e 180°F (FIG. 4).

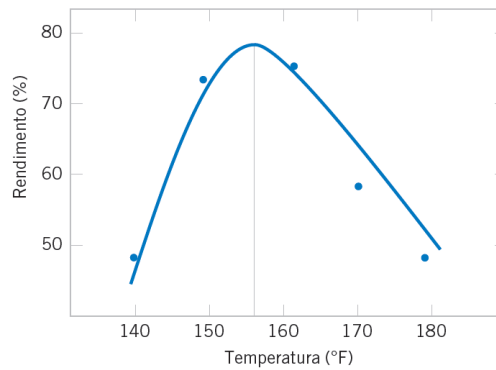


Figura 4 - Rendimento contra temperatura, com o tempo de reação constante em 1,7 hora.
Fonte: MONTGOMERY; RUNTER, 2009, p. 334.

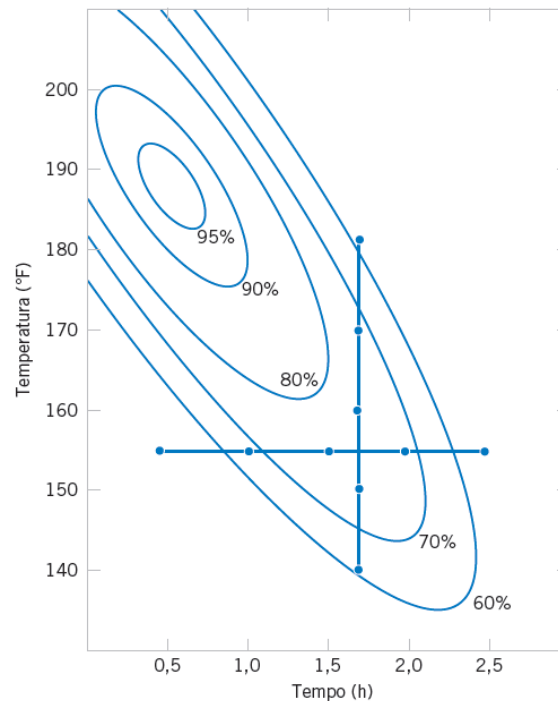


Figura 5 - Experimento de otimização variando um fator de cada vez.
Fonte: MONTGOMERY; RUNTER, 2009, p. 334.

Na figura 5 há curvas de nível do rendimento como uma função da temperatura e do tempo. Podemos facilmente perceber que essa técnica de variar um fator de cada vez falhou drasticamente, pois o rendimento ótimo verdadeiro acontece em tempos bem menores e em temperaturas maiores. Descobrir que se chegaria ao ótimo em tempos menores é interessante, pois isso pode impactar significativamente a capacidade de produção, no custo de fabricação e na produtividade.

Esse tipo de falha acontece quando se utiliza como abordagem a variação de um fator de cada vez, porque esta técnica não permite identificar a interação entre os fatores temperatura e tempo. Além disso, este é um método ineficiente, pois necessita de mais experimentos do que um planejamento fatorial e, como vimos, não há garantia de se conseguir o ótimo real por meio dele (MONTGOMERY; RUNGER, 2009).

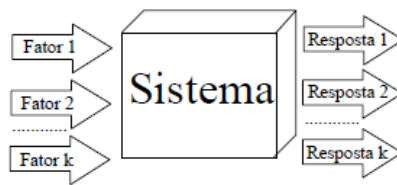


Figura 6 - Um sistema em estudo com respostas influenciadas por fatores de entrada.
Fonte: GONTIJO, 2006.

Na figura 6 podemos ver certo número de fatores atuando sobre um sistema e produzindo assim um conjunto de respostas. Este sistema pode ser representado por uma função que atua sobre as variáveis de entrada e produz as respostas observadas como saída. O objetivo é determinar essa função ou pelo menos chegar a uma aproximação aceitável para ela. Assim, é possível entender melhor o sistema em estudo e como ele reage às manipulações dos fatores, para enfim escolher suas melhores condições de operação.

Para exemplificar um planejamento fatorial, pensemos na situação em que se deseja saber como o rendimento de um processo é afetado com a variação de temperatura e do reagente. Para isso foi feita a análise do efeito para dois níveis de temperatura e o uso de dois reagentes distintos, caracterizando um planejamento 2^2 . Como este é um exemplo para fins didáticos, o sistema é fictício e a escolha de níveis arbitrária. Os resultados para esse planejamento podem ser verificados na tabela a seguir (TAB. 1).

Tabela 1 - Dados para um experimento fatorial.

Ensaio	Temperatura (°C)	Reagente	Rendimento (%)				Média
1	50	A	52	55	54	51	53
2	70	A	95	97	91	93	94
3	50	B	55	51	57	53	54
4	70	B	66	63	65	62	64

Fonte: Elaboração Própria.

Nesta tabela há o registro das respostas observadas, ou rendimentos, em todas as 4 possíveis combinações dos níveis escolhidos. Como se pode notar, todos os ensaios foram feitos quatro vezes, produzindo dezesseis respostas no total. Assim, será possível estimar o erro experimental de uma determinada resposta. É importante conhecer a extensão desse erro para decidirmos se existem ou não efeitos significativos na ação dos fatores.

O efeito principal é assim denominado porque ele se refere a fatores primários no estudo. Ele é a variação que a mudança de nível do fator produz na resposta. Por definição o efeito principal da temperatura é obtido através da média dos efeitos da temperatura nos dois níveis do reagente. Usando a letra T para representar a temperatura e \bar{y}_i para representar a resposta média observada no i -ésimo ensaio, temos:

$$T = \frac{(\bar{y}_2 - \bar{y}_1) + (\bar{y}_4 - \bar{y}_3)}{2}$$

$$T = \frac{(94 - 53) + (64 - 54)}{2}$$

$$T = \frac{(41) + (10)}{2}$$

$$T = 25,5\%$$

Esse valor indica que o rendimento da reação aumenta em média 25,5% quando a temperatura passa do nível inferior (50°C) para o nível superior (70°). Porém, essa

conclusão é incompleta, pois ainda é necessário interpretar os efeitos dos dois fatores conjuntamente, os efeitos de interação.

Nos planejamentos fatoriais de dois níveis é comum identificar os níveis superior e inferior com os sinais (+) e (-) respectivamente. Essa atribuição de sinais também pode ser feita para os níveis dos fatores qualitativos. Dessa forma, a equação acima poderia ser escrita da seguinte forma:

$$T = \bar{t}_i$$

onde podemos ver que o efeito principal é a diferença entre as médias de nível alto e baixo.

Para o efeito principal do reagente os cálculos são:

$$R = \bar{r}_i$$

$$R = \frac{(54+64) - (94+53)}{2}$$

$$R = -14,5\%$$

Quando trocamos o reagente A pelo reagente B, o rendimento diminui 14,5% em média.

O efeito de interação existe quando as variáveis se influenciam mutuamente, de forma que o valor ideal para um fator depende do valor do outro.

Caso não houvesse interação, o efeito da temperatura seria o mesmo para qualquer reagente. Porém, como vimos o efeito da temperatura no reagente A é de 41% e de 10% quando usamos o reagente B, ou seja, o efeito do fator temperatura depende dos níveis do fator reagente.

O efeito de interação é, por definição, metade da diferença dos efeitos principais. Então, temos:

$$T \times R = TR = \frac{10 - 41}{2} = -15,5 \%$$

Para esse cálculo fizemos a diferença subtraindo o valor do efeito T correspondente ao nível inferior do reagente do valor correspondente ao nível superior, levando em conta que convencionamos que B seria nível + e A seria nível -.

O conceito de interação pode ser ilustrado graficamente de diversas maneiras.

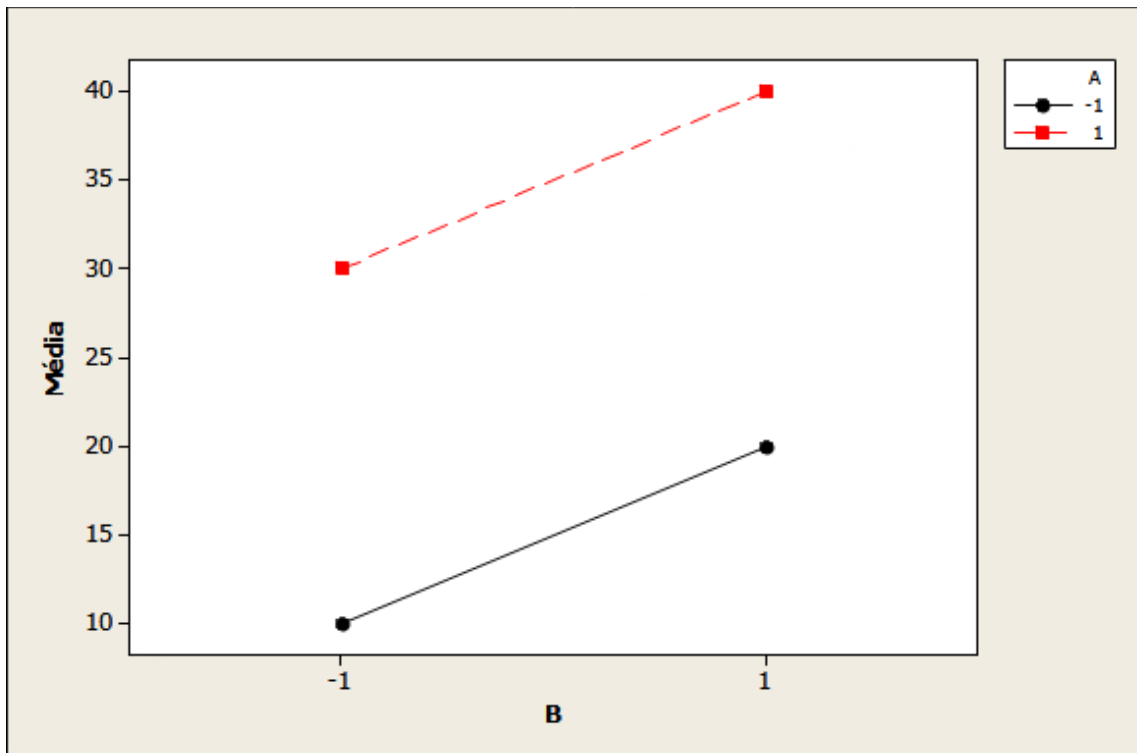


Figura 7 - Experimento fatorial sem interação.
Fonte: Elaboração própria.

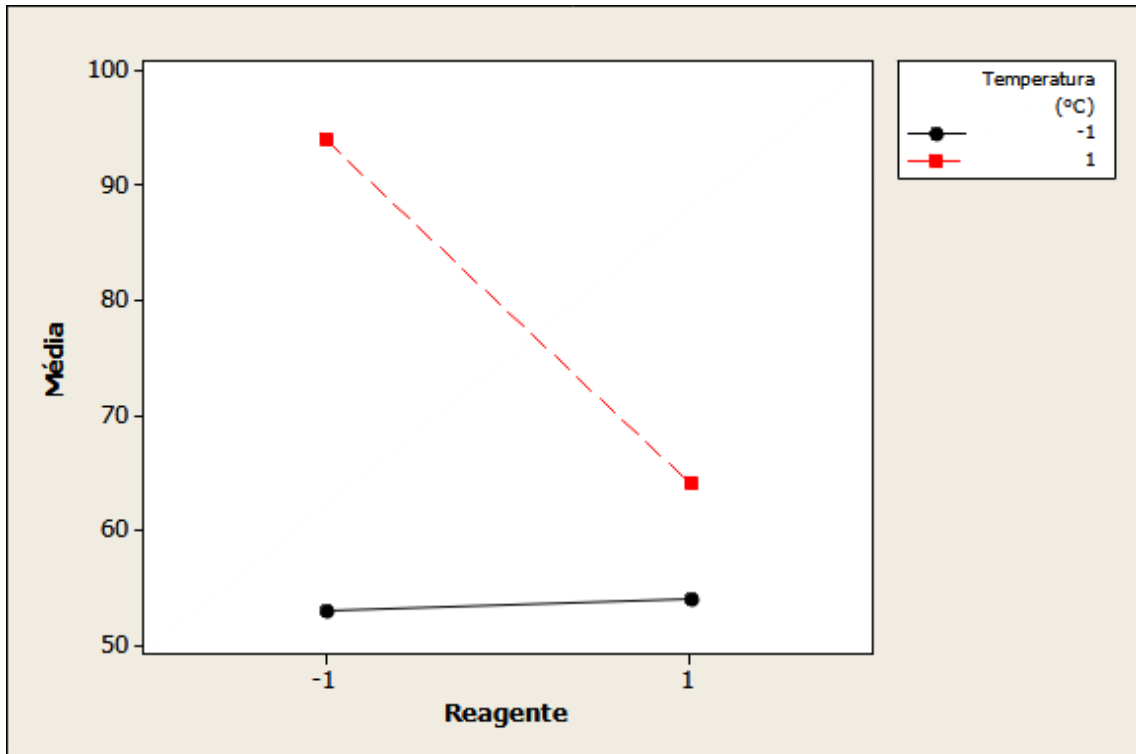


Figura 8 - Experimento fatorial com interação com dados referentes ao exemplo citado.
Fonte: Elaboração própria.

No gráfico da figura 7 podemos ver que as linhas são paralelas, o que indica que os fatores A e B não têm interação significativa. Já no gráfico da figura 8, que plota os dados da tabela 1, as linhas não são paralelas, o que indica interação entre os fatores temperatura e reagente. Esses gráficos são chamados de gráfico de interação entre dois fatores e foram plotados no software Minitab, porém podem ser construídos por diversos outros programas computacionais.

Como vimos, para qualquer efeito que queiramos calcular é necessário usar todas as respostas observadas. Esta é uma importante particularidade do planejamento fatorial de dois níveis: nenhuma das respostas obtidas fica ociosa.

Os exemplos citados trataram de planejamentos fatoriais com dois fatores, pois esse é o tipo mais simples de experimento fatorial. Para tratar desse planejamento de uma forma mais genérica, digamos que temos os fatores A e B, com a níveis do fator A e b níveis do fator B. O experimento tem n réplicas, sendo que cada réplica contém todas as combinações de ab . A tabela a seguir (TAB. 2) mostra o experimento fatorial de dois fatores. A observação na ij -ésima célula na k -ésima

replicação é representada por y_{ijk} . Esse é um planejamento completamente aleatorizado, então as abn observações devem ser corridas em uma ordem aleatória.

Tabela 2 - Dados para um planejamento fatorial de dois fatores.

Fator A	Fator B			
	1	2	...	b
1	$y_{111}, y_{112}, \dots, y_{11n}$	$y_{121}, y_{122}, \dots, y_{12n}$		$y_{1b1}, y_{1b2}, \dots, y_{1bn}$
2	$y_{211}, y_{212}, \dots, y_{21n}$	$y_{221}, y_{222}, \dots, y_{22n}$		$y_{2b1}, y_{2b2}, \dots, y_{2bn}$
...				
a	$y_{a11}, y_{a12}, \dots, y_{a1n}$	$y_{a21}, y_{a22}, \dots, y_{a2n}$		$y_{ab1}, y_{ab2}, \dots, y_{abn}$

Fonte: HINES et al., 2006.

As observações podem ser descritas por meio do modelo estatístico linear a seguir:

$$Y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk},$$

onde μ representa o efeito médio global, τ_i representa o efeito do i -ésimo nível do fator A, β_j é o efeito do j -ésimo nível do fator B, $\tau\beta$ é o efeito de interação entre A e B e ε_{ijk} é um componente do erro aleatório, que tem distribuição normal com média zero e variância σ^2 . Nosso interesse é testar as hipóteses de que não haja efeito principal para A, nenhum efeito principal para B e nenhum efeito de interação AB. Para testar essas hipóteses usaremos análise de variância (HINES et al., 2006).

Vamos supor que os a níveis do fator A e os b níveis do fator B são escolhidos pelo experimentador e as inferências são restritas apenas a esses níveis. Para esse modelo, é costume definir τ_i , β_j e $(\tau\beta)_{ij}$ como desvios em relação a média, de forma

$$\text{que } \sum_{i=1}^a \tau_i = 0, \sum_{j=1}^b \beta_j = 0, \sum_{i=1}^a (\tau\beta)_{ij} = 0 \text{ e } \sum_{j=1}^b (\tau\beta)_{ij} = 0.$$

Seja $y_{i..}$ o total de observações feitas no i -ésimo nível da variável A, $y_{.j}$ o total de observações feitas no j -ésimo nível da variável B, y_{ij} o total de observações feitas no ij -ésima célula da tabela 2 e $y_{...}$ o total global de todas observações feitas (HINES et al., 2006).

Defina $\bar{y}_{i..}$, $\bar{y}_{.j.}$, $\bar{y}_{ij.}$, $\bar{y}_{...}$ como médias totais de linhas, colunas e células correspondentes. Ou seja,

$$y_{i..} = \sum_{j=1}^b \sum_{k=1}^n y_{ijk} \quad \bar{y}_{i..} = \frac{y_{i..}}{bn} \quad i=1,2,\dots,a$$

$$y_{.j.} = \sum_{i=1}^a \sum_{k=1}^n y_{ijk} \quad \bar{y}_{.j.} = \frac{y_{.j.}}{an} \quad j=1,2,\dots,a$$

$$y_{ij.} = \sum_{k=1}^n y_{ijk} \quad \bar{y}_{ij.} = \frac{y_{ij.}}{n} \quad i=1,2,\dots,a$$

$$j=1,2,\dots,a$$

$$y_{...} = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n y_{ijk} \quad \bar{y}_{...} = \frac{y_{...}}{abn}$$

A soma de quadrados total pode ser escrita como:

$$\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (y_{ijk} - \bar{y}_{...})^2$$

$$= \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (y_{ijk} - \bar{y}_{ij.} + \bar{y}_{ij.} - \bar{y}_{i..} + \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{.j.} - \bar{y}_{...} + \bar{y}_{...})^2$$

$$= \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (y_{ijk} - \bar{y}_{ij.})^2 + \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (\bar{y}_{ij.} - \bar{y}_{i..} + \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{.j.} - \bar{y}_{...})^2$$

$$= \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (y_{ijk} - \bar{y}_{ij.})^2 + \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (\bar{y}_{ij.} - \bar{y}_{i..})^2 + \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (\bar{y}_{i..} - \bar{y}_{.j.})^2 + \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (\bar{y}_{.j.} - \bar{y}_{...})^2$$

ou, simbolicamente,

$$SQ_T = SQ_A + SQ_B + SQ_{AB} + SQ_E$$

A soma total dos quadrados, SQ_T , é particionada em uma soma de quadrados devida ao fator A (SQ_A), uma soma de quadrados devida ao fator B (SQ_B), uma soma de quadrados devida à interação entre os fatores A e B (SQ_{AB}) e uma soma de quadrados devida ao erro (SQ_E). Para que a soma de quadrados do erro seja diferente de zero, é necessário que se faça no mínimo duas réplicas de cada experimento.

O total de graus de liberdade é igual a $abn - 1$. Os efeitos principais A e B têm $(a - 1)$ e $(b - 1)$ graus de liberdade, respectivamente. Já o efeito de interação AB tem $(a - 1)(b - 1)$ graus de liberdade. Para cada uma das células ab na tabela 2, há $n - 1$ entre as n réplicas, podendo as observações dentro de cada célula diferirem unicamente por causa do erro aleatório. Assim sendo, há $ab(n - 1)$ graus de liberdade para o erro (HINES et al., 2006).

Se dividirmos cada soma quadrática pelo número correspondente de graus de liberdade obteremos médias quadráticas. Considerando que os fatores A e B estejam fixados, os valores esperados para as médias quadráticas são:

$$E(MQ_A) = E\left(\frac{SQ_A}{a-1}\right) = \sigma^2 + \frac{bn \sum_{i=1}^a \tau_i^2}{a-1},$$

$$E(MQ_B) = E\left(\frac{SQ_B}{b-1}\right) = \sigma^2 + \frac{an \sum_{j=1}^b \beta_j^2}{b-1},$$

$$E(MQ_{AB}) = E\left(\frac{SQ_{AB}}{(a-1)(b-1)}\right) = \sigma^2 + \frac{n \sum_{i=1}^a \sum_{j=1}^b (\tau\beta)_{ij}^2}{(a-1)(b-1)},$$

e

$$E(MQ_A) = E\left(\frac{SQ_E}{ab(n-1)}\right) = \sigma^2.$$

Desta forma, para testarmos $H_0: \tau_i = 0$ (nenhum efeito do fator A), $H_0: \beta_j = 0$ (nenhum efeito do fator B) e $H_0: \delta$ (nenhum efeito da interação) é preciso dividir a média quadrática correspondente pela média quadrática do erro. Se essas hipóteses forem todas verdadeiras, todas as quatro médias quadráticas serão estimativas não tendenciosas de σ^2 (HINES et al., 2006).

Essas razões terão uma distribuição F, com o número de graus de liberdade do numerador igual ao número de graus de liberdade da média quadrática do numerador e $ab(n-1)$ graus de liberdade para o denominador.

O teste F permite verificar se os efeitos são significativos ou não. Para determinar o nível α de significância do fator A, comparamos se $f_0 > f_{\alpha, a-1, ab(n-1)}$. Para determinar o nível α de significância do fator B, comparamos se $f_0 > f_{\alpha, b-1, ab(n-1)}$. Finalmente, para determinar a significância α dos efeitos de interação, comparamos se $f_0 > f_{\alpha, (a-1)(b-1), ab(n-1)}$. Na tabela de análise de variância a seguir (TAB. 3) é possível ver o procedimento de teste, bem como encontrar os cálculos de F_0 .

Tabela 3 - Tabela de Análise de variância para um modelo de um fatorial com dois fatores e efeitos fixos.

Fonte de Variação	Soma de Quadrados	Graus de Liberdade	Média Quadrática	F_0
Tratamentos A	SQ_A	$a-1$	$MQ_A = \frac{SQ_A}{a-1}$	$\frac{MQ_A}{MQ_E}$
Tratamentos B	SQ_B	$b-1$	$MQ_B = \frac{SQ_B}{b-1}$	$\frac{MQ_B}{MQ_E}$
Interação	SQ_{AB}	$(a-1)(b-1)$	$MQ_{AB} = \frac{SQ_{AB}}{(a-1)(b-1)}$	$\frac{MQ_{AB}}{MQ_E}$
Erro	SQ_E	$ab(n-1)$	$MQ_E = \frac{SQ_E}{ab(n-1)}$	
Total	SQ_T	$abn-1$		

Fonte: HINES et al., 2006.

Na maioria das vezes é melhor avaliar primeiro os efeitos de interação, pois o conhecimento destes geralmente é mais importante que o conhecimento sobre efeitos principais. Caso a interação não seja significativa, a interpretação dos efeitos principais é direta. Contudo, os efeitos principais podem não ter muito valor prático se a interação for significativa.

A seguir estão as fórmulas de cálculo para somas quadráticas em uma análise de variância com dois fatores:

$$SQ_T = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n y_{ijk}^2 - \frac{y_{...}^2}{abn}$$

$$SQ_A = \sum_{i=1}^a \frac{y_{i..}^2}{bn} - \frac{y_{...}^2}{abn}$$

$$SQ_B = \sum_{j=1}^b \frac{y_{.j.}^2}{an} - \frac{y_{...}^2}{abn}$$

$$SQ_{AB} = \sum_{i=1}^a \sum_{j=1}^b \frac{y_{ij.}^2}{n} - \frac{y_{...}^2}{abn} - SQ_A - SQ_B$$

$$SQ_E = SQ_T - SQ_{AB} - SQ_{AB} - SQ_{AB}$$

É interessante ressaltar que as soluções para esses cálculos podem ser facilmente obtidas por meios computacionais.

Os resíduos de um experimento fatorial cumprem um importante papel para a verificação da adequação do modelo. Para um experimento fatorial de dois fatores os resíduos são:

$$e_{ijk} = y_{ijk} - \bar{y}_{ij.}$$

Nessa equação podemos ver que os resíduos são iguais à diferença entre as observações e as médias de célula correspondentes.

A seguir (FIG. 9) temos um gráfico de probabilidade normal de resíduos para um determinado planejamento fatorial. No gráfico podemos ver que as extremidades não estão localizadas exatamente ao longo de uma linha reta que passaria pelo centro do gráfico, o que indica alguns problemas quanto à suposição de normalidade. Contudo, esse desvio de normalidade não parece ser severo. As figuras 10 e 11 plotam os resíduos contra o fator A e o fator B, nos quais podemos observar que os resíduos estão distribuídos de forma bastante simétrica com relação ao eixo 0 o que indica aleatoriedade.

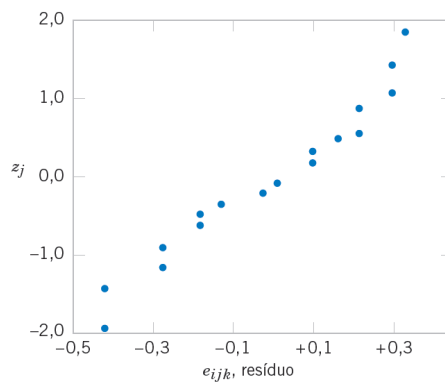


Figura 9 - Gráfico de probabilidade normal de resíduos.
Fonte: MONTGOMERY; RUNTER, 2009.

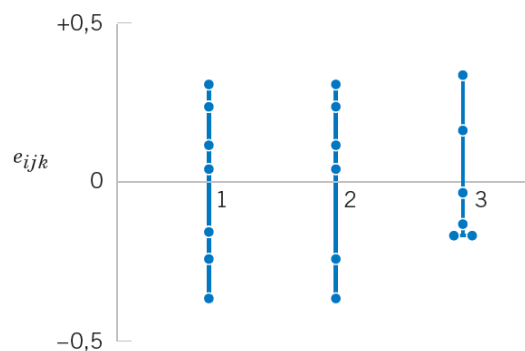


Figura 10 - Gráfico de resíduos contra fator A.
Fonte: MONTGOMERY; RUNGER, 2009.

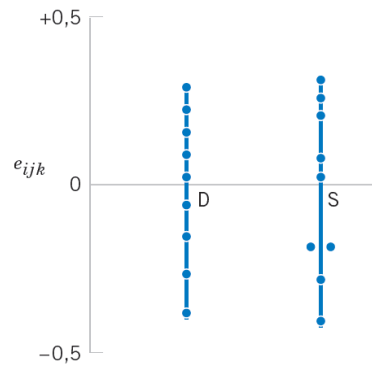


Figura 11 - Gráfico dos resíduos contra fator B.
Fonte: MONTGOMERY; RUNGER, 2009.

1.3.2 Metodologia de Superfície de Resposta

A metodologia de superfícies de resposta (ou MSR) consiste em uma coleção de técnicas matemáticas e estatísticas, eficaz para otimização de processos. Essa metodologia, fundamentada em planejamentos fatoriais, foi introduzida nos anos 1950 por G.E.P. Box e a partir de então, tem sido usada largamente e com grande sucesso em processos industriais (BARROS NETO; SCARMINIO; BRUNS, 2007).

Aplicações importantes em planejamentos, desenvolvimento de novos produtos, assim como melhorias de projetos e produtos já existentes, utilizam a MSR. Todavia, a mais ampla aplicação dessa metodologia é na área industrial, especialmente em circunstâncias em que há várias variáveis que podem influenciar no desempenho ou na qualidade de um determinado produto ou processo. Ou seja, seu emprego é bastante útil em aplicações em que a resposta de interesse sofre influência de diversas variáveis e se tem o objetivo de otimizar essa resposta (MYERS & MONTGOMERY, 1995 apud LOUZADA, et. al).

A técnica tem duas etapas – modelagem e deslocamento – que se repetem até cumprir o objetivo de alcançar uma região ótima da superfície examinada. A modelagem normalmente é feita por meio de ajustes de modelos simples (geralmente lineares ou quadráticos) às respostas conseguidas através de planejamentos fatoriais. O deslocamento é feito ao longo do caminho de maior

inclinação para um determinado modelo, que é onde a resposta varia de forma mais acentuada (BARROS NETO; SCARMINIO; BRUNS, 2007).

Um primeiro passo na MSR é encontrar o modelo adequado que se aproxime da relação verdadeira entre Y e as variáveis independentes. Em geral, utilizamos um polinômio de baixo grau em alguma região das variáveis independentes. O modelo de primeira ordem se dá pela função de aproximação:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon$$

Caso haja curvatura no sistema, então é necessário usar um polinômio de maior grau, como um modelo de segunda ordem:

$$Y = \beta_0 + \sum_{i=1}^k \beta_{1i} x_i + \sum_{i=1}^k \beta_{2i} x_i^2 + \sum_{i=1}^{k-1} \sum_{j>i}^k \beta_{ij} x_i x_j + \varepsilon$$

em que $\sum_{i=1}^k \beta_{1i} x_i$ representam o efeito linear, $\sum_{i=1}^k \beta_{2i} x_i^2$ representam o efeito quadrático e $\sum_{i=1}^{k-1} \sum_{j>i}^k \beta_{ij} x_i x_j + \varepsilon$ representam o efeito de interação.

Esses modelos geralmente funcionam bem para uma região relativamente pequena do espaço das variáveis independentes (MONTGOMERY; RUNGER, 2007).

Para exemplificar uma situação em que o emprego da MSR seria útil, podemos citar a situação suposta por Barros Neto et al. (2009) em que um pesquisador deseja definir os níveis de concentração de um reagente (x_1) e da velocidade de agitação (x_2) que maximizem o rendimento y de um processo. Ele conhece as condições atuais de funcionamento do processo, os valores desses fatores estão fixados em 50% para concentração e 100 rpm para velocidade de agitação, para os quais o rendimento tem ficado em torno de 68%.

O rendimento do processo é uma função dos níveis de concentração de reagente e velocidade de agitação, como

$$Y=f(x_1, x_2)+\varepsilon,$$

onde ε representa o erro observado na resposta Y . Se a resposta esperada for denotada por $E(Y)=f(x_1, x_2)$, então a superfície representada por $f(x_1, x_2)$ é denominada de superfície de resposta (MONTGOMERY; RUNGER; HUBELE, 2004).

A MSR é um procedimento sequencial. Quando se está num ponto da superfície de resposta longe do ótimo, não existe muita curvatura, portanto o modelo de primeira ordem será o mais adequado. O objetivo é que o pesquisador chegue, de forma rápida e eficiente, à região de ótimo. Tendo encontrado a melhor região, um modelo mais elaborado pode ser empregado; por exemplo, um modelo de segunda ordem, e pode-se também fazer uma análise para localizar o ponto ótimo, de máximo ou mínimo.

A finalidade da MSR é determinar as condições de operações ótimas para o sistema ou determinar uma região do espaço fatorial na qual os requerimentos de operação são satisfeitos.

O autor começa analisando a superfície de resposta em torno das condições operacionais atuais. Isso é feito por meio do planejamento fatorial (FIG. 12). Como podemos notar, esse planejamento contém um ponto central, e, desta forma, ao invés de varrer apenas dois níveis dos fatores, varre três. Este fato torna possível verificar se há falta de ajuste para um modelo linear, o que seria impossível caso fossem usados apenas dois níveis. Na tabela a seguir (TAB. 4), podemos ver a matriz de planejamentos e os rendimentos observados. Foram realizados, ao todo, sete ensaios, e três desses foram repetições no ponto central.

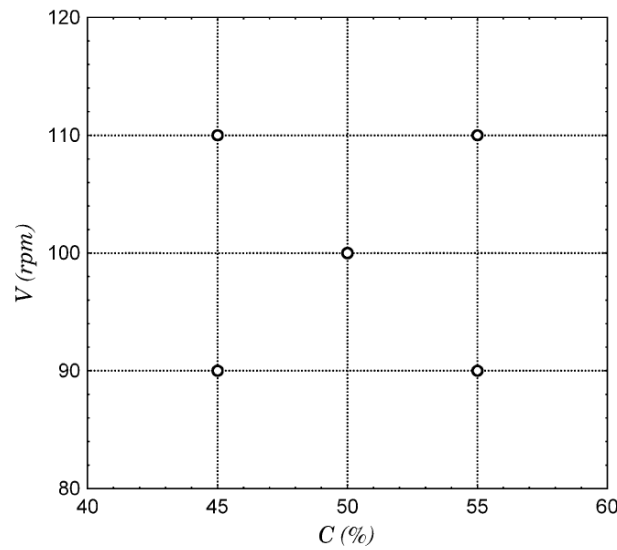


Figura 12 - Planejamento fatorial de dois níveis com ponto central.
Fonte: NETO; SCARMINIO; BRUNS, 2007, p.307.

Tabela 4 - Resultado de um planejamento fatorial 2^2 com ponto central.

Ensaio	C(%)	V(rpm)	x_1	x_2	$y(\%)$
1	45	90	-1	-1	69
2	55	90	1	-1	59
3	45	110	-1	1	78
4	55	110	1	1	67
5	50	100	0	0	68
6	50	100	0	0	66
7	50	100	0	0	69

Fonte: BARROS NETO; SCARMINIO; BRUNS, 2007.

Os valores x_1 e x_2 (TAB. 4) representam os valores dos dois fatores (concentração do reagente e velocidade de agitação), codificados pelas equações

$$x_1 = \frac{C-50}{5} \quad \text{e} \quad x_2 = \frac{v-100}{10}$$

Para começar a análise, o autor admite que a superfície de resposta na região analisada seja uma função linear dos fatores e, desta forma, o rendimento pode ser estimado pela equação:

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2,$$

em que b_0 , b_1 e b_2 são estimadores dos parâmetros do modelo e x_1 e x_2 representam os fatores codificados.

Os estimadores dos parâmetros podem ser obtidos pelo método dos mínimos quadrados. Como descrito no tópico 1.2.1 Mínimos Quadrados, é possível lançar mão de várias equações para executar esses cálculos, no entanto, uma das melhores formas de calcular os estimadores é por meio da utilização de equações matriciais. Contudo, se o número de fatores for grande, estes cálculos seriam demasiado trabalhosos para serem feitos manualmente, uma vez que envolveriam uma grande quantidade de manipulações numéricas, e, desta forma, é interessante utilizar softwares para este fim, como, por exemplo, o *Scilab* (PIRES, 2004).

Nesse caso a matriz X seria:

$$X = \begin{bmatrix} 1 & -1 & -1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

em que a primeira coluna corresponde ao termo b_0 , enquanto a segunda e terceira colunas são formadas pelos valores codificados dos fatores x_1 e x_2 .

A última coluna na (TAB. 4) contém os rendimentos observados em cada ensaio. Esses valores são as respostas e constituem o que se denomina de vetor y . Assim temos que:

$$y = \begin{bmatrix} 69 \\ 59 \\ 78 \\ 67 \\ 68 \\ 66 \\ 69 \end{bmatrix}$$

Então, é realizado o cálculo $X^t X$ e $X^t y$ (onde X^t refere-se à matriz transposta de X):

$$X^t X = \begin{bmatrix} 7 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix} \text{ e } X^t y = \begin{bmatrix} 476 \\ -21 \\ 17 \end{bmatrix},$$

em seguida

$$b = (X^t X)^{-1} X^t y = \begin{bmatrix} 1/7 & 0 & 0 \\ 0 & 1/4 & 0 \\ 0 & 0 & 1/4 \end{bmatrix} \times \begin{bmatrix} 476 \\ -21 \\ 17 \end{bmatrix} = \begin{bmatrix} 68,00 \\ -5,25 \\ 4,25 \end{bmatrix}.$$

Dos três ensaios repetidos no ponto central, pode-se calcular s^2 como uma estimativa da variância das observações.

$$s^2 = \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{n-1} = \sum_{i=1}^3 \frac{(x_i - 67,66)^2}{3-1} = 2,33$$

A variância estimada dos elementos do vetor b , é dada pela equação:

$$\hat{V}(b) = (X^t X)^{-1} s^2 = \begin{bmatrix} 1/7 & 0 & 0 \\ 0 & 1/4 & 0 \\ 0 & 0 & 1/4 \end{bmatrix} \times 2,33 = \begin{bmatrix} 0,33 & 0 & 0 \\ 0 & 0,58 & 0 \\ 0 & 0 & 0,58 \end{bmatrix}$$

Tirando as raízes quadradas o autor encontra os desvios padrões de b_0 , b_1 e b_2 . Com eles e com os valores estimados para b , foi encontrado o modelo ajustado e seu erro padrão:

$$\hat{y} = 68,00 - 5,25 x_1 + 4,25 x_2$$

$$(\pm 0,58) \quad (\pm 0,76) \quad (\pm 0,76)$$

Se a superfície ajustada for uma aproximação apropriada da função verdadeira de resposta, então a análise desta superfície será aproximadamente equivalente à análise do sistema real (MONTGOMERY; RUNGER; HUBELE, 2004).

Como o erro é relativamente pequeno, conclui-se que esse modelo é significativo. Contudo, no texto de onde tiramos o exemplo, ainda são feitas outras análises de variância para verificar se o modelo é adequado. A seguir temos essa análise de variância (TAB. 5).

Tabela 5 - Análise de variância para o ajuste do modelo $\hat{y} = 68,00 - 5,25 x_1 + 4,25 x_2$.

Fonte de variação	Soma quadrática	Nº Graus de Liberdade	Média quadrática
Regressão	182,50	2	91,25
Resíduos	5,50	4	1,38
Falta de ajuste	0,83	2	0,42
Erro puro	4,67	2	2,34
Total	188,0	6	
% de variação explicada: 97,07			
% máxima de variação explicada: 97,52			

Fonte: BARROS NETO; SCARMINIO; BRUNS, 2007.

A conclusão é a mesma para esta análise, pois como a razão entre as médias quadráticas de falta de ajuste e de erro é igual à de $MQ_{faj}/MQ_{ep} = 0,42/2,34 = 0,18$, que não representa um valor estatisticamente significativo, não existe evidência de falta de ajuste para este modelo. A teoria que justifica esta análise pode ser verificada no capítulo 5 do livro *Como fazer experimentos* (BARROS NETO; SCARMINIO; BRUNS, 2007).

A superfície de resposta descrita pelo modelo é definida pelo plano representado a seguir (FIG. 13).

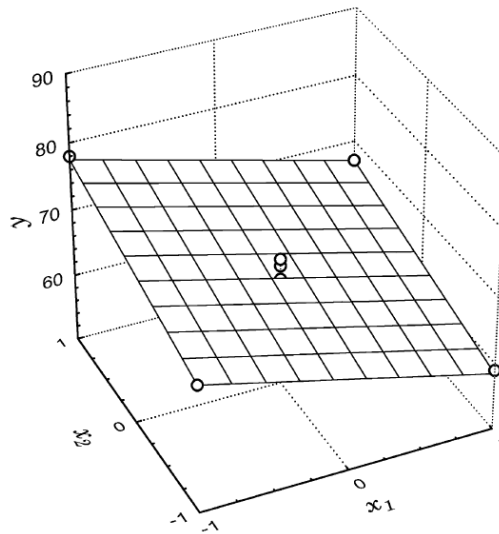


Figura 13 - Plano descrito pela equação $\hat{y} = 68,00 - 5,25x_1 + 4,25x_2$
 Fonte: BARROS NETO; SCARMINIO; BRUNS, 2007.

É possível obter uma representação bidimensional da superfície modelada por desenhar curvas de nível, que são linhas em que a resposta é constante. Na figura a seguir (FIG. 14), podemos ver que, para sistemas de primeira ordem, as curvas de nível de um plano são representadas por segmentos de retas paralelas.

Para que o pesquisador se mova rapidamente para a vizinhança do ótimo, de forma simples e eficiente economicamente, ele pode seguir o método da ascendente de maior inclinação, que é uma técnica para se deslocar sequencialmente ao longo do trajeto ascendente com maior inclinação, ou seja, deslocar na direção de aumento máximo na resposta, que é uma trajetória perpendicular às curvas de nível. Obviamente, se o objetivo for uma minimização, então estaríamos falando sobre o método da descendente de maior inclinação. Na situação em questão, podemos ver tanto pela superfície de resposta e curvas de nível descritas pelo modelo ajustado, quanto pela própria equação do modelo, que teremos maiores rendimentos com menores valores de x_1 e maiores valores de x_2 .

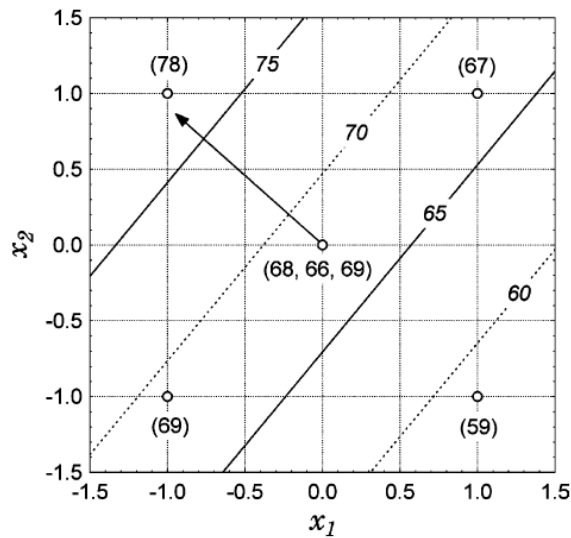


Figura 14 - Curvas de nível descritas pela equação $\hat{y} = 68,00 - 5,25x_1 + 4,25x_2$ e trajetória de máxima inclinação.

Fonte: BARROS NETO; SCARMÍNIO; BRUNS, 2007.

A determinação do caminho de maior inclinação pode ser feita algebricamente a partir dos coeficientes do modelo. Para conseguir a inclinação máxima, os deslocamentos devem ser feitos ao longo dos eixos x_2 e x_1 , segundo a proporção b_2/b_1 .

Numa situação genérica em que temos uma superfície de resposta definida por p fatores, o caminho de maior inclinação é proporcional aos módulos e aos sinais dos coeficientes do modelo. Tendo como orientação o procedimento descrito a seguir, o caminho pode ser traçado facilmente:

a) Primeiro escolhemos um dos fatores, digamos i , para servir de base, e então mudamos seu nível, para mais ou para menos, de acordo com o objetivo do experimento. Alguns autores recomendam que o fator escolhido seja o de maior coeficiente, em módulo. Comumente, o deslocamento inicial é de uma unidade na escala codificada.

b) Após isto, podemos determinar os deslocamentos dos demais fatores $j \neq i$, em unidades codificadas, por meio da equação:

$$\Delta x_j = \frac{b_j}{b_i} \Delta x_i$$

c) Então, convertamos os valores codificados de volta às unidades originais e determinamos os novos níveis dos fatores.

Para a situação tratada neste exemplo, temos:

$$\Delta x_2 = \frac{b_2}{b_1} = \frac{4,25}{(-5,25)} \Delta x_1 = -0,81 \Delta x_1,$$

que indica que para cada unidade recuada no eixo x_1 deve-se avançar 0,81 ao longo do eixo x_2 . A seguir, podemos ver as coordenadas de pontos ao longo da trajetória de ascensão ao máximo, nela há tanto as variáveis codificadas como as unidades reais de concentração e velocidade de agitação (TAB. 6).

Tabela 6 - Caminho com máxima inclinação para o modelo estimado.

Etapa	x_1	x_2	$C(\%)$	$v(rpm)$	$y(\%)$
<i>Centro</i>	0	0,00	50	100,0	68; 66;69
<i>Centro</i> + Δ	-1	0,81	45	108,1	77
<i>Centro</i> +2 Δ	-2	1,62	40	116,2	86
<i>Centro</i> +3 Δ	-3	2,43	35	124,3	88
<i>Centro</i> +4 Δ	-4	3,24	30	132,4	80
<i>Centro</i> +5 Δ	-5	4,05	25	140,5	70

Fonte: BARROS NETO; SCARMINIO; NETO, 2007.

Podemos notar que, em princípio, esse deslocamento melhora a resposta, o rendimento aumenta. Porém, depois do terceiro ensaio a resposta começa a diminuir. Para interpretar esses resultados podemos comparar a superfície de resposta a um “morro”, onde o topo do morro representaria o ponto de resposta máxima. Caso o objetivo fosse encontrar um ponto de resposta mínima, a superfície seria comparada a um “vale” para o qual procuraríamos o caminho da descendente de maior inclinação.

Concluimos, então, que até o terceiro ensaio o pesquisador estava deslocando os experimentos ladeira acima, mas após este ensaio já começa a descer o “morro” pelo outro lado. Isso nos indica que o pesquisador está próximo ao ótimo, e ao invés de continuar se deslocando, é necessário analisar essa região que apresentou as melhores respostas.

Para isso ele faz um novo planejamento fatorial, desta vez com centro em torno do melhor ensaio, que é o terceiro com concentração em 35% e velocidade em 125 rpm (TAB. 7).

Tabela 7 - Resultado do novo planejamento fatorial 2^2 com ponto central.

Ensaio	C(%)	V(rpm)	x_1	x_2	$y(\%)$
1	30	115	-1	-1	86
2	40	115	1	-1	85
3	30	135	-1	1	78
4	40	135	1	1	84
5	35	125	0	0	90
6	35	125	0	0	88
7	35	125	0	0	89

Fonte: BARROS NETO; SCARMINIO; BRUNS, 2007.

Os valores x_1 e x_2 (TAB. 7) representam os valores dos dois fatores (concentração do reagente e velocidade de agitação), codificados pelas equações:

$$x_1 = \frac{C-35}{5} \quad \text{e} \quad x_2 = \frac{v-125}{10}$$

Seguindo o mesmo procedimento que no primeiro planejamento, é encontrado o modelo:

$$\hat{y} = 85,71 + 1,25 x_1 - 2,25 x_2$$

$(\pm 0,49)$ $(\pm 0,65)$ $(\pm 0,65)$

Os erros padrões foram obtidos por meio de uma estimativa conjunta da variância, combinando os ensaios repetidos dos dois planejamentos. Os erros deste modelo,

em comparação com os coeficientes, são mais importantes do que os erros do primeiro modelo. A tabela 8 traz a análise de variância e nos mostra que a percentagem da variação explicada é apenas 27,20% e o valor de $MQ_{faj}/MQ_{ep}=34,36$, que é maior que o valor de $F_{2,2}$ no nível 95% de confiança, que é igual a 19, confirmando que o modelo ajustado não é adequado para a região para onde o caminho de máxima inclinação levou, pois já não descreve satisfatoriamente a superfície de resposta.

Tabela 8 - Análise de variância para o ajuste do modelo $\hat{y} = 85,71 + 1,25x_1 - 2,25x_2$.

Fonte de variação	Soma quadrática	Nº Graus de Liberdade	Média quadrática
Regressão	26,50	2	13,25
Resíduos	70,93	4	17,73
Falta de ajuste	68,93	2	34,46
Erro puro	2,00	2	1,00
Total	97,42	6	

% de variação explicada: 27,20
 % máxima de variação explicada: 97,95

Fonte: BARROS NETO; SCARMINIO; BRUNS, 2007.

Quando o experimentalista chega à vizinhança do ótimo, geralmente é necessário que se use o modelo de segunda ordem para aproximar a resposta, em função da curvatura na verdadeira superfície de resposta (MONTGOMERY; RUNGER; HUBELE, 2004), essa teoria se confirma no exemplo estudado, e como o modelo linear não atende mais, vamos então usar um modelo quadrático que tem como função geral para duas variáveis a expressão:

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_{11}x_1^2 + b_{22}x_2^2 + b_{12}x_1x_2$$

O planejamento possui apenas cinco diferentes combinações de valores de concentração e velocidade, e este novo modelo tem seis parâmetros. É necessário então, ampliar o planejamento, pois não é possível determinar as estimativas se há mais parâmetros do que níveis. Há diferentes formas de ampliação. Nesse caso foi usada a mais comum, chamada de planejamento em estrela.

Essa ampliação é feita acrescentando ao planejamento original, um outro planejamento, porém, desta vez, girado de 45 graus em relação a orientação de partida. Caso a ampliação tenha sido feita dessa forma rotacionada, o desvio padrão da resposta estimada \hat{y} será o mesmo para todos os pontos que estiverem equidistantes do centro do planejamento. Para rotabilidade, devemos escolher corretamente o espaçamento axial α , para isso fazemos $\alpha=(F)^{1/4}$, sendo que F é o número de pontos na porção fatorial do planejamento, geralmente $F=2^k$. Para o caso em que há dois fatores, então $k=2$, $\alpha=(2^2)^{1/4}=1,414$. O planejamento em estrela com ponto central é também chamado de planejamento composto central rotacionado (MONTGOMERY; RUNGER; HUBELE, 2004).

A seguir temos os resultados para o novo planejamento (TAB. 9).

Tabela 9 - Resultados do planejamento em estrela obtido com ampliação do planejamento da tabela 7.

Ensaio	C(%)	v(rpm)	x_1	x_2	y(%)
1	30	115	-1	-1	86
2	40	115	1	-1	85
3	30	135	-1	1	78
4	40	135	1	1	84
5	35	125	0	0	90
6	35	125	0	0	88
7	35	125	0	0	89
8	28	125	$-\sqrt{2}$	0	81
9	35	139	0	$\sqrt{2}$	80
10	42	125	$\sqrt{2}$	0	86
11	35	111	0	$-\sqrt{2}$	87

Fonte: BARROS NETO; SCARMINIO; BRUNS, 2007.

Na tabela 9, x_1 e x_2 representam os valores das variáveis codificadas, como na tabela 7.

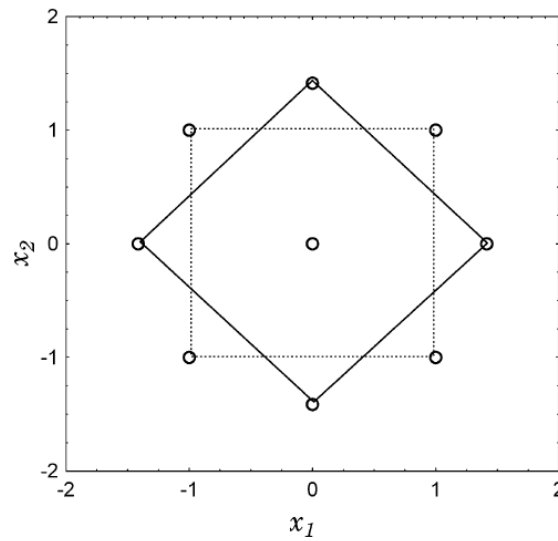


Figura 15 - Planejamento em estrela para duas variáveis codificadas correspondente à Tabela 9.
Fonte: BARROS NETO; SCARMINIO; BRUNOS, 2007.

Agora a matriz X tem dimensão 11×6 , com suas seis colunas correspondentes aos seis termos do modelo quadrático. Sua primeira coluna corresponde ao termo independente b_0 , enquanto que a segunda e terceira colunas correspondem aos valores codificados dos fatores x_1 e x_2 . As três novas próximas colunas referem-se a x_1^2 , x_2^2 e x_1x_2 , nesta ordem. Para obtê-las elevamos ao quadrado ou multiplicamos as colunas referentes aos fatores x_1 e x_2 (TAB. 9). O vetor y é formado pelas respostas na última coluna da tabela.

$$X = \begin{bmatrix} 1 & -1 & -1 & 1 & 1 & 1 \\ 1 & 1 & -1 & 1 & 1 & -1 \\ 1 & -1 & 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & -\sqrt{2} & 0 & 2 & 0 & 0 \\ 1 & 0 & \sqrt{2} & 0 & 2 & 0 \\ 1 & \sqrt{2} & 0 & 2 & 0 & 0 \\ 1 & 0 & -\sqrt{2} & 0 & 2 & 0 \end{bmatrix} \text{ e } y = \begin{bmatrix} 86 \\ 85 \\ 78 \\ 84 \\ 90 \\ 88 \\ 89 \\ 81 \\ 80 \\ 86 \\ 87 \end{bmatrix}$$

E assim como no planejamento anterior, foi seguido o método dos mínimos quadrados para calcular os estimadores. A seguir, apresentamos os cálculos de $X^t X$

e $X^t y$, que desta vez foram realizados no *Matlab* 7.6.0, em função das maiores dimensões das matrizes.

$$X^t X = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & 1 & -1 & 0 & 0 & -\sqrt{2} & 0 & \sqrt{2} & 0 \\ -1 & -1 & 1 & 0 & 0 & 0 & \sqrt{2} & 0 & -\sqrt{2} \\ 1 & 1 & 1 & 0 & 0 & 2 & 0 & 2 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 2 & 0 & 2 \\ 1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\times \begin{bmatrix} 1 & -1 & -1 & 1 & 1 & 1 \\ 1 & 1 & -1 & 1 & 1 & -1 \\ 1 & -1 & 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & -\sqrt{2} & 0 & 2 & 0 & 0 \\ 1 & 0 & \sqrt{2} & 0 & 2 & 0 \\ 1 & \sqrt{2} & 0 & 2 & 0 & 0 \\ 1 & 0 & -\sqrt{2} & 0 & 2 & 0 \end{bmatrix}$$

$$X^t X = \begin{bmatrix} 11 & 0 & 0 & 8 & 8 & 0 \\ 0 & 7.99 & 0 & 0 & 0 & 0 \\ 0 & 0 & 7.99 & 0 & 0 & 0 \\ 8 & 0 & 0 & 12 & 4 & 0 \\ 8 & 0 & 0 & 4 & 12 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4 \end{bmatrix}$$

$$X^t y = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & 1 & -1 & 0 & 0 & -\sqrt{2} & 0 & \sqrt{2} & 0 \\ -1 & -1 & 1 & 0 & 0 & 0 & \sqrt{2} & 0 & -\sqrt{2} \\ 1 & 1 & 1 & 0 & 0 & 2 & 0 & 2 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 2 & 0 & 2 \\ 1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} 86 \\ 85 \\ 78 \\ 84 \\ 90 \\ 88 \\ 89 \\ 81 \\ 80 \\ 86 \\ 87 \end{bmatrix}$$

$$X^t y = \begin{bmatrix} 934 \\ 12 \\ -18,9 \\ 667 \\ 667 \\ 7 \end{bmatrix}$$

Em seguida, mostramos o cálculo da matriz inversa de $X^t X$, obtendo:

$$(X^t X)^{-1} = \begin{bmatrix} 0,33 & 0 & 0 & -0,16 & -0,16 & 0 \\ 0 & 0,12 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0,12 & 0 & 0 & 0 \\ -0,16 & 0 & 0 & 0,17 & 0,05 & 0 \\ -0,16 & 0 & 0 & 0,05 & 0,17 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0,25 \end{bmatrix}$$

E a partir dessas matrizes é possível encontrar o vetor b que contém os valores numéricos dos estimadores dos parâmetros do modelo quadrático, conforme podemos ver a seguir:

$$b = (X^t X)^{-1} X^t y$$

$$b = \begin{bmatrix} 0,33 & 0 & 0 & -0,16 & -0,16 & 0 \\ 0 & 0,12 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0,12 & 0 & 0 & 0 \\ -0,16 & 0 & 0 & 0,17 & 0,05 & 0 \\ -0,16 & 0 & 0 & 0,05 & 0,17 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0,25 \end{bmatrix} \times \begin{bmatrix} 934 \\ 12 \\ -18,9 \\ 667 \\ 667 \\ 7 \end{bmatrix}$$

$$b = \begin{bmatrix} 89 \\ 1,51 \\ -2,36 \\ -2,81 \\ -2,81 \\ 1,75 \end{bmatrix}$$

E dessa forma o modelo quadrático ajustado é:

$$\hat{y} = 89 + 1,51 x_1 - 2,36 x_2 - 2,81 x_1^2 - 2,81 x_2^2 + 1,75 x_1 x_2$$

$$(\pm 0,75) \quad (\pm 0,46) \quad (\pm 0,46) \quad (\pm 0,54) \quad (\pm 0,54) \quad (\pm 0,65)$$

Novamente, os erros foram calculados a partir de uma estimativa conjunta da variância, obtida de todos os ensaios repetidos, inclusive os da tabela 4. A análise de variância para este modelo é apresentada a seguir (TAB. 10). Agora o valor de MQ_{aj}/MQ_{ep} é apenas 0,25, não havendo então evidência de falta de ajuste para o modelo.

Tabela 10 - Análise de variância para o ajuste do modelo

$\hat{y} = 89 + 1,51 x_1 - 2,36 x_2 - 2,81 x_1^2 - 2,81 x_2^2 + 1,75 x_1 x_2$ aos dados da Tabela 9.

Fonte de variação	Soma quadrática	Nº Graus de Liberdade	Média quadrática
Regressão	144,15	5	28,83
Resíduos	2,76	5	0,55
Falta de ajuste	0,76	3	0,25
Erro puro	2,00	2	1,00
Total	146,91	10	

% de variação explicada: 98,12
 % máxima de variação explicada: 98,64

Fonte: BARROS NETO; SCARMINIO; BRUNS, 2007.

Nas figuras 16 e 17 podemos ver a superfície de resposta e curvas de nível correspondentes ao modelo ajustado. O ponto de máximo dessa região está localizado aproximadamente em $x_1=0,15$ e $x_2=-0,37$, isto é, o ótimo dessa região se forem empregadas uma concentração de 36% e uma velocidade de agitação de 121 rpm. Aplicando esses valores ao modelo obtido, o rendimento da reação deve ficar em torno 89,6%, que é 32% melhor que o rendimento inicial, que era 68%.

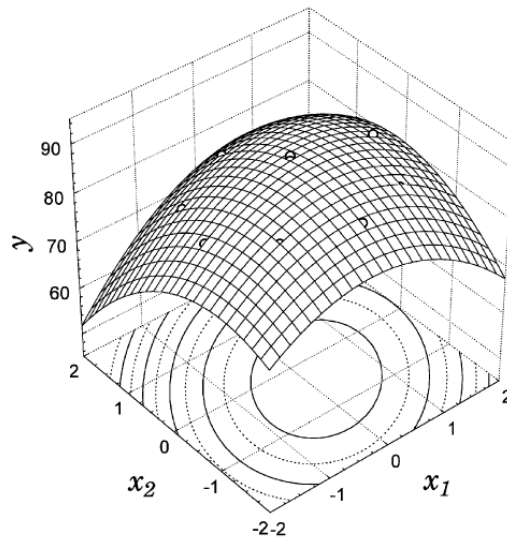


Figura 16 - Superfície quadrática descrita pela equação $\hat{y} = 89 + 1,51x_1 - 2,36x_2 - 2,81x_1^2 - 2,81x_2^2 + 1,75x_1x_2$.
Fonte: BARROS NETO; SCARMÍNIO; BRUNS, 2007.

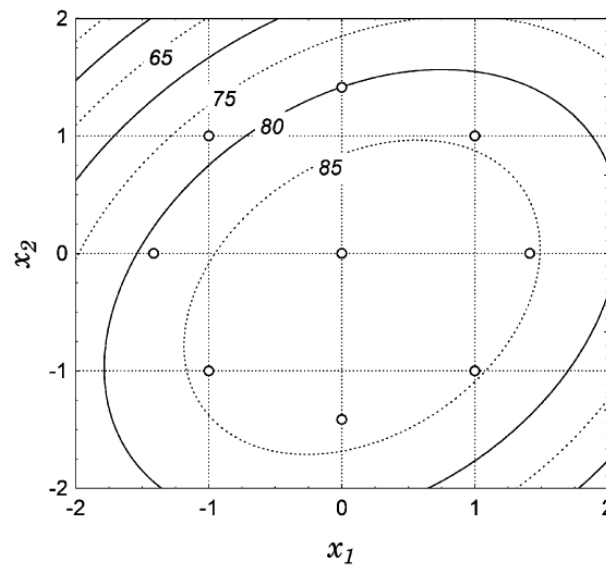


Figura 17 - Curvas de nível descritas pela equação $\hat{y} = 89 + 1,51x_1 - 2,36x_2 - 2,81x_1^2 - 2,81x_2^2 + 1,75x_1x_2$.
Fonte: BARROS NETO; SCARMÍNIO; BRUNS, 2007.

Na MSR a escolha da faixa inicial de variação dos fatores é uma questão muito importante, pois isso é o que vai determinar o tamanho do planejamento e a velocidade com que os experimentos subsequentes se moverão em direção ao ótimo ao longo da superfície.

Para determinar a melhor escala de variação dos fatores, o pesquisador deve apoiar-se em todo conhecimento disponível sobre o sistema estudado, também em

sua experiência, e assim escolher deslocamentos nem tão pequenos que não forneçam efeitos significativos e nem tão grandes que saiam da região linear, ou que estejam “do outro lado do morro”. O pesquisador deve ter em mente a importância de se fazer os experimentos de forma sequencial e interativa (BARROS NETO; SCARMINIO; BRUNS, 2007).

Na MSR o número de fatores não representa uma restrição, bem como o número de respostas também não representa, muito embora a análise possa ficar um tanto complicada. Essa característica é interessante, pois muitas vezes um processo ou produto deve atender a mais de um requisito, como ter custo mínimo, e, no entanto, deve também manter os parâmetros de qualidade dentro das especificações.

À medida que o número de fatores vai aumentando, aumenta também a complexidade da análise. Para tratar problemas com muitas variáveis podemos recorrer a outros métodos como o de programação linear, ou mesmo não linear, ou ainda podemos usar a metodologia de otimização simultânea proposta por G. C. Derringer e R. Suich que se baseia na definição de uma função de desajabilidade².

1.3.3 Aplicações

Como já dito, a MSR tem sido amplamente empregada para otimizar produtos e processos. Exemplos dessas aplicações podem ser verificados na vasta literatura sobre o assunto. Vejamos alguns exemplos a seguir.

Num estudo sobre otimização do processo de coagulação - floculação para o tratamento de águas residuais de reciclagem de papel foram avaliados os efeitos e interações de três fatores: dosagem de coagulante, dosagem de floculante e do pH sobre a eficiência do tratamento, um planejamento composto central, que é a versão

² Este estudo não faz parte do escopo deste trabalho. Essa teoria pode ser verificada no capítulo 6 do livro “Como fazer experimentos” (BARROS NETO; SCARMINIO; BRUNS, 2007, p.332-338).

padrão da metodologia de superfície de resposta, foi utilizado para se otimizar o rendimento de forma rápida e eficaz (WANG, et al, 2007).

Já no artigo de Chen et. al. a metodologia de superfície de resposta e o planejamento com ponto central foram utilizados para otimizar o processo de síntese do rutilo sintético a partir de escória de titânia para teor máximo de TiO_2 . Os efeitos de diversos parâmetros sobre o conteúdo TiO_2 foram sistematicamente investigados para identificar as condições experimentais ótimas (CHEN, G.; et al., 2011).

No trabalho de PIRES, T. S. é apresentado o emprego de uma função de aproximação, obtida por meio da MSR, no lugar de uma função objetivo a ser minimizada na otimização de sistemas de cogeração modelados com auxílio de um simulador de processos. O estudo conclui que para um problema de otimização termoeconômica com um grande número de variáveis de decisão, ou com uma função objetivo onerosa o emprego da função de aproximação implica na redução de cálculos sem prejuízo quanto a obtenção do ótimo, podendo assim ser mais eficiente do que o modelo que emprega a função objetivo inicial (PIRES, T. S, 2010).

1.3.4 Considerações sobre os métodos

Em todas estas técnicas de otimização, observamos que é de grande importância a obtenção de dados para que possam ser analisados, e tão imediato ocorram as análises toma-se decisões importantes sobre quais parâmetros influenciam ou não um dado processo, bem como quais deverão ser os valores de operação desses parâmetros. Assim, o desenvolvimento de estratégias de obtenção de dados de forma rápida e eficiente para serem trabalhados é de suma importância para o profissional cuja perspectiva seja a de otimizar um processo.

Neste trabalho que desenvolvemos, procuramos simular a obtenção de dados via OPC que possam ser analisados usando a técnica de planejamento fatorial e metodologia de superfície de resposta de modo a apresentar

possibilidades reais que podem ser implementadas nas mais diferentes indústrias e que com certeza trazem grandes ganhos em termo de operacionalidade.

2 PROCEDIMENTOS EXPERIMENTAIS

2.1 Aquisição de dados via OPC

Quando tratamos de ambientes laboratoriais os equipamentos de aquisição de dados, isto é, os analisadores e medidores de grandezas, apresentam uma interface de comunicação com computadores já bem consolidada. Tais interfaces são em sua maior parte definidos por um padrão serial, sendo em sua maioria compostos por padrões como RS232, RS485 e USB.

Exemplos que comprovam esta evidência são equipamentos como o analisador de glicose e lactose YSI 2300 que apresenta padrão de comunicação RS232 e RS485, a balança determinadora de umidade SHI-MOC-63U que utiliza padrão de comunicação USB além do padrão RS232 e também o espectrofotômetro CM-5 que possui comunicação USB 1.1., todos comercializados no Brasil pela TECNAL equipamentos para laboratórios (www.tecnallab.com.br).



Figura 18 - YSI 2300 Analisador de Glicose e Lactose.
Fonte: www.tecnallab.com.br

Nas indústrias o cenário é um pouco diferente. Elas possuem fontes de dados com padrões diferentes como CLPs, medidores, sensores, acionadores, válvulas, relés de proteção, dentre outros dispositivos, sendo que muitos desses são integrados a sistemas de gerenciamento de dados e sistemas de supervisão, automação e controle como sistemas SCADA. Por meio de variados protocolos como Modbus, Fieldbus, Profibus, Device Net, TCP/IP, é realizada a comunicação de equipamentos que podem, eventualmente, estar separados por longas distâncias. Além disso, há uma infinidade de tipos de aplicações de controle de processo, onde é possível manipular os dados, como sistemas MES (Sistema de Execução Industrial), PIMS (Sistema de Gerenciamento de Informação da Planta Industrial), ERP (Planejamento de Recurso de Empreendimentos) e aplicações Windows (Excel, Access, etc.). Toda essa diversidade de tecnologias de fontes de dados, interfaces com o usuário, redes de comunicação, plataformas de aplicações, sistemas operacionais cria uma barreira à intercambiabilidade de dados entre equipamentos de fabricantes que utilizam protocolos de comunicação diferentes (RIBEIRO; ANJOS; LEANDRO, 2008).

Com o intuito de possibilitar a integração dos dados de toda a empresa, sem esbarrar nos inconvenientes da incompatibilidade de protocolos proprietários e equipamentos, o padrão OPC (OLE for Process Control) foi desenvolvido pela OPC Foundation, no ano de 1995.

Esta organização sem fins lucrativos surgiu da reunião de empresas que tinham o objetivo de criar um mecanismo padronizado de troca de informações entre dispositivos dos mais diversos fabricantes. Entre essas empresas estão a Fisher-Rosemount (hoje Emerson Process Management), Rockwell Software, Opto 22, Intellution (hoje parte da GE Fanuc) e Intuitive Technology (hoje parte da Axeda). Contudo, hoje a fundação conta com mais de 300 membros, (CÂNDIDO, 2004).

O padrão OPC, que hoje é amplamente empregado em diversas aplicações, como pode ser verificado em <http://www.matrikonopc.com/resources/case-studies.aspx>, consiste, basicamente, num protocolo aberto, ou seja, não desenvolvido exclusivamente pelos fabricantes dos equipamentos. É composto por diversas especificações, em constante desenvolvimento, que padronizam a comunicação das arquiteturas de aquisição de dados online, para isto ele apresenta regras para o

desenvolvimento de sistemas com interfaces padrões para comunicação dos dispositivos de campo, como CLPs, sensores, balanças, com sistemas de supervisão, monitoração e gerenciamento como SCADA, MES, ERP, etc (FONSECA, 2002).

O padrão cumpre seu objetivo primário de possibilitar a interconectividade entre plataformas que usam diferentes protocolos e precisam se comunicar, permitindo a troca transparente de dados entre diversos tipos de aplicações, tanto gerenciais quanto de chão de fábrica, para que a informação se mova verticalmente.

Para desenvolver o padrão, a OPC Foundation se baseou em padrões já existentes da indústria de computação e na direção das companhias membro que conhecem as necessidades da indústria, contando com a colaboração da Microsoft, também membro da organização, que atua como consultora de tecnologia.

Enquanto houver necessidade a Fundação continuará a criar novas especificações e gerenciar as especificações existentes para utilizar nova tecnologia, assegurando a interoperabilidade e garantindo mais funcionalidade ao padrão OPC (CÂNDIDO, 2004).

O funcionamento do padrão OPC se baseia na tradicional arquitetura cliente-servidor. Os computadores responsáveis pelo reconhecimento de drivers e compartilhamento dos dados constituem os servidores OPC. Os clientes OPC são as aplicações que “consomem” estes dados e podem estar em quaisquer computadores conectados à rede do servidor OPC (RIBEIRO; ANJOS; LEANDRO, 2008). Uma aplicação cliente pode solicitar dados a um ou mais servidores OPC, e um servidor OPC pode fornecer dados para um ou mais clientes OPC. Isso aumenta bastante a possibilidade de comunicações, sendo necessário apenas que os aplicativos sejam compatíveis com OPC.



Figura 19 - Arquitetura cliente-servidor OPC
Fonte: PUDA, 2008.

Como o OPC não elimina os protocolos proprietários nativos dos equipamentos, apenas os “traduz” para o padrão OPC, algumas empresas, como a Matrikon OPC, se especializaram no desenvolvimento de servidores OPC específicos para cada protocolo de comunicação existente e possuem soluções para a maioria dos sistemas de automação conhecidos em seu portfólio (MATHEUS; ABRAHÃO, 2011).

O OPC é baseado no OLE (Object Linking and Embedding) de componentes orientados a objeto, por meio das tecnologias COM (Component Object Model) e DCOM (Decentralized Component Object Model) da Microsoft, que permitem que aplicações troquem dados. Graças à arquitetura cliente-servidor, esses dados podem ser acessados por um ou mais computadores, mesmo que as aplicações trabalhem sobre sistemas que utilizem protocolos diferentes.

No trabalho *Padrão OPC: Uma Alternativa de Substituição dos Drivers Proprietários para Acessar Dados de PLCs* (CÂNDIDO, 2004), é mostrada uma comparação entre a conectividade proprietária versus a conectividade OPC.

O autor supõe que uma grande planta petroquímica quer monitorar um de seus equipamentos por meio de uma interface homem-máquina (HMI), para visualização; por meio de um historiador de processo (Process Historian), para gravação dos

dados e por meio de uma aplicação que monitore a máquina (Machine Condition Monitor).

Os dados virão de três diferentes fontes: um PLC, um sistema de monitoramento de vibração (Vibration Monitoring System) e uma máquina para calcular (Calculation Engine), como o Excel.

Na solução proprietária, cada aplicação precisa de três drivers para se comunicar com cada uma das fontes de dados. Dessa forma, para esta solução seria necessários um total de nove drivers.

Cada fonte de dados teria que fornecer o mesmo dado três vezes, uma para cada aplicação. Isso geraria uma lentidão na comunicação, pois os muitos drivers criariam requisições demais para cada fonte. A figura 20 ilustra como seriam as requisições das aplicações às fontes e suas respostas, dadas através dos drivers específicos de cada aplicação.

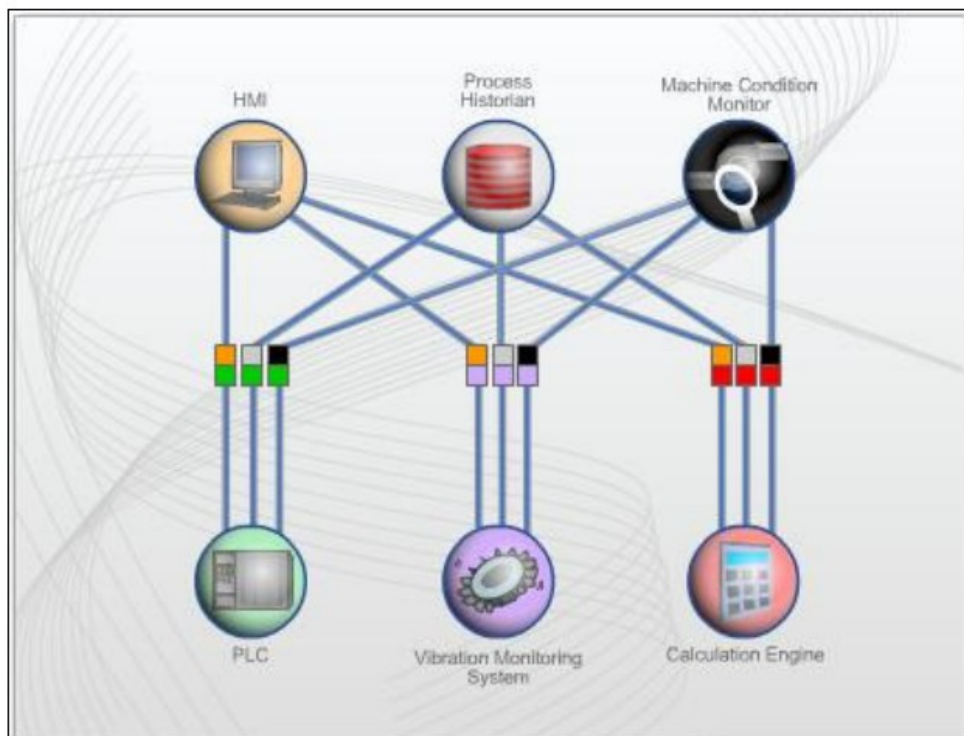


Figura 20 - Os nove drivers utilizados na solução proprietária.
Fonte: CÂNDIDO, 2004.

A implementação dessa solução, incluindo instalação e comunicação dos drivers, demandaria aproximadamente dez dias e teria um custo total de cerca de cinquenta mil dólares incluindo software e trabalhos relacionados (CÂNDIDO, 2004).

Já com a solução OPC, seriam necessários três servidores OPC e três interfaces, sendo um servidor e uma interface para cada fonte de dados (FIG. 21).

Por se tratar de um padrão aberto, as interfaces OPC estão disponíveis, não adicionando custos com software.

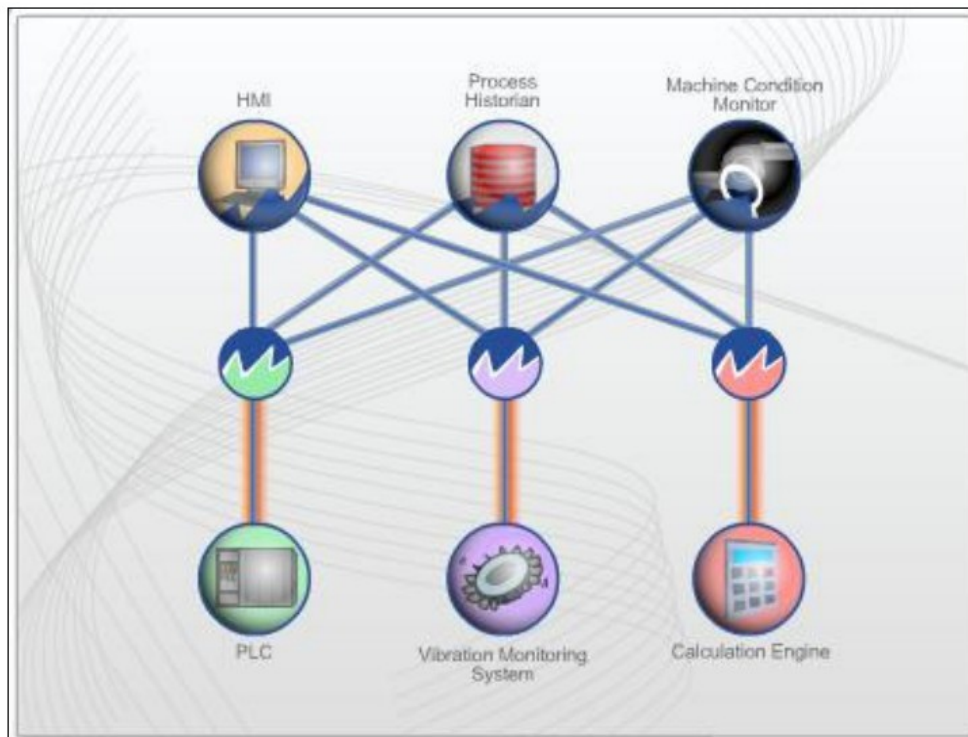


Figura 21 - Requisições aos diferentes dispositivos na solução OPC.
Fonte: CÂNDIDO, 2004.

Se um servidor OPC receber três requisições leitura, ele enviará apenas uma requisição de leitura à fonte de dados. Isso diminui significativamente o número de requisições de dados em cada fonte de dados, além de melhorar significativamente o desempenho de protocolos seriais.

De acordo com Cândido (2004), a implementação desta solução demandaria apenas dois dias e os custos com software e serviços relacionados seriam inferiores a dez mil dólares.

Essa minimização do tempo de implementação e redução dos custos de aquisição de software, com certeza tem contribuído para a popularidade do padrão OPC.

Tendo em vista esse exemplo e as outras características do OPC já supracitadas, entendemos que ele resolveu o desafio da indústria moderna quanto ao acesso e compartilhamento de dados entre controladores, dispositivos e aplicações, independentemente de qual seja o fornecedor ou o protocolo proprietário nativo usado pelos equipamentos (DAREK, 2009). Desta forma o OPC dá aos usuários liberdade para escolha das melhores soluções, que levam em conta funcionalidade, tempo e custo, fazendo com que não exista mais o domínio de algumas poucas empresas no segmento de automação industrial (CÂNDIDO, 2004).

Além disso, a crescente utilização de sistemas abertos de comunicação tem possibilitado a redução de cabos, graças a sua instalação mais simples, e tem também facilitado a manutenção (NOGUEIRA, 2009).

Outro aspecto interessante é que o conhecimento profundo do funcionamento interno do OPC não é um requisito para usá-lo, mesmo uma breve familiaridade com seus principais conceitos é útil (DAREK, 2009).

Em função dos benefícios advindos deste padrão, ele teve uma rápida ascensão na popularidade, tornando-se o mais popular do mundo de tecnologia de conectividade de dados. Por ser tão versátil, o OPC é usado em todos os níveis da empresa e em todos os setores industriais (DAREK, 2009).

Em razão dos benefícios que o padrão trouxe e, principalmente, a popularidade decorrente destes benefícios, decidimos abordar esse tema descrevendo o que é OPC e explicando a função que clientes e servidores OPC têm na comunicação OPC, tratando do assunto de forma conceitual.

3 RESULTADOS E DISCUSSÃO

3.1 Aplicação

Como apresentado, a metodologia de superfície de resposta necessita de uma grande quantidade de dados coletados em campo, devido ao relativamente grande número de ensaios para modelamento.

Se utilizarmos parâmetros obtidos da própria planta, ao invés de dados obtidos em plantas piloto ou laboratório, garantimos maior fidelidade aos resultados obtidos com a otimização.

Como vimos, o padrão OPC, já bastante consolidado, possibilita a obtenção desses dados de forma rápida, sem necessidade de deslocamento de pessoal, o que facilita e agiliza as tomadas de decisões para se determinar quais parâmetros deverão ser ajustados.

Portanto, nos parece proveitoso pensar em formas de integrar técnicas de otimização, como a metodologia de superfície de respostas, com formas já conhecidas de obtenção de dados acerca de processos industriais, como a aquisição de dados via OPC.

Dessa forma, fizemos uso do padrão OPC como meio para aquisição de dados, por apresentar uma boa intercambiabilidade entre os sistemas de controle mais difundidos no mercado, e trataremos os dados por meio de um algoritmo criado em Excel, para exemplificar a integração de técnicas de otimização com aplicações já existentes.

A aplicação desenvolvida tem por objetivo realizar a otimização de um processo utilizando a metodologia de superfície de resposta.

O servidor utilizado foi o Modbus OPC da Matrikon®. Ele permite a conexão com os CLP's por meio dos padrões ethernet ou serial, de forma que foi adotado o padrão ethernet. A configuração deste servidor será tratada posteriormente.

Os dados provenientes do servidor OPC foram tratados por meio da linguagem de programação Visual Basic utilizando o software Excel da Microsoft. O Excel foi escolhido como ferramenta para programação por se tratar de uma ferramenta comumente encontrada nas indústrias, por apresentar um fácil tratamento dos dados dispensando uso de banco de dados e por possuir uma forma fácil de gerar gráficos. A integração do servidor OPC com a aplicação feita no Excel foi feita utilizando uma ferramenta chamada OPCWareXL. Esta ferramenta permite transformar o Excel em um cliente OPC possibilitando a leitura e escrita de variáveis dentro de um servidor OPC.

Por fim, para realizar os testes simulando a conexão de um CLP com a aplicação, foi utilizado o programa Concept da Schneider Electric™ que consiste em um software para programação de CLP's.

3.1.1 Configuração de OPC para aquisição de dados e integração com Excel

A configuração do servidor OPC é simples e rápida, conforme descreveremos a seguir.

O primeiro passo é inserir um nó que corresponde ao dispositivo que será a fonte de dados para o servidor OPC. Para fazer isso clicamos no botão "New Node", como na figura a seguir (FIG. 22).

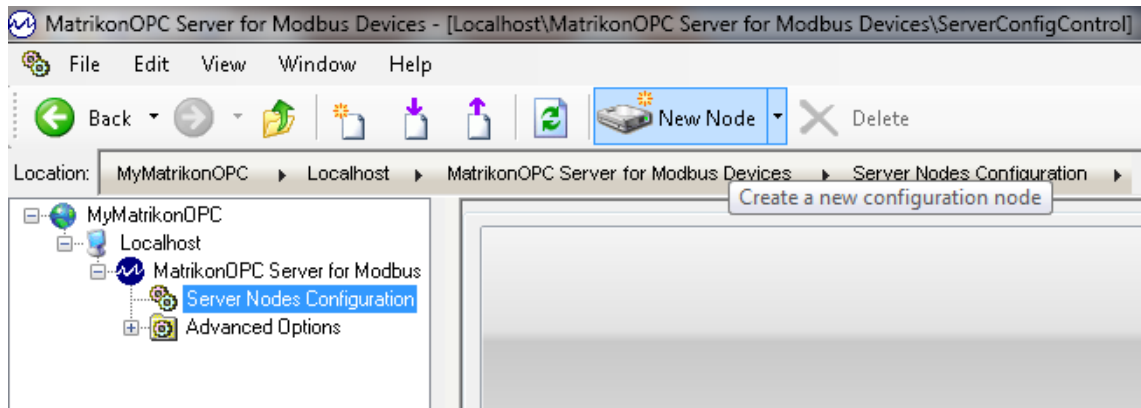


Figura 22 - Inserção de novo nó como fonte de dados no servidor OPC.
Fonte: Elaboração própria.

Em seguida selecionamos o padrão de comunicação do dispositivo. Como para esta aplicação utilizamos o software Concept para simular um CLP em rede ethernet, este deve ser o padrão escolhido (FIG. 23).

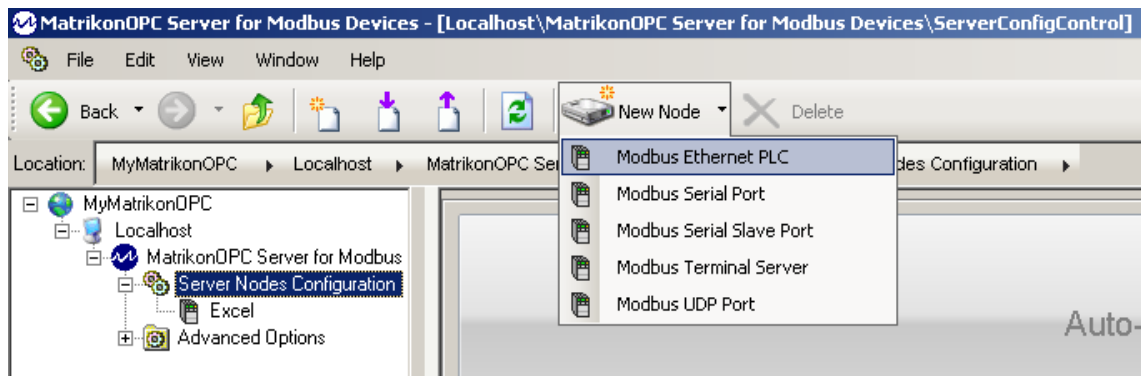


Figura 23 - Seleção do padrão de comunicação.
Fonte: Elaboração própria.

Ao selecionar o padrão ethernet, é exibida a tela (FIG. 24) que mostra os parâmetros a serem configurados.

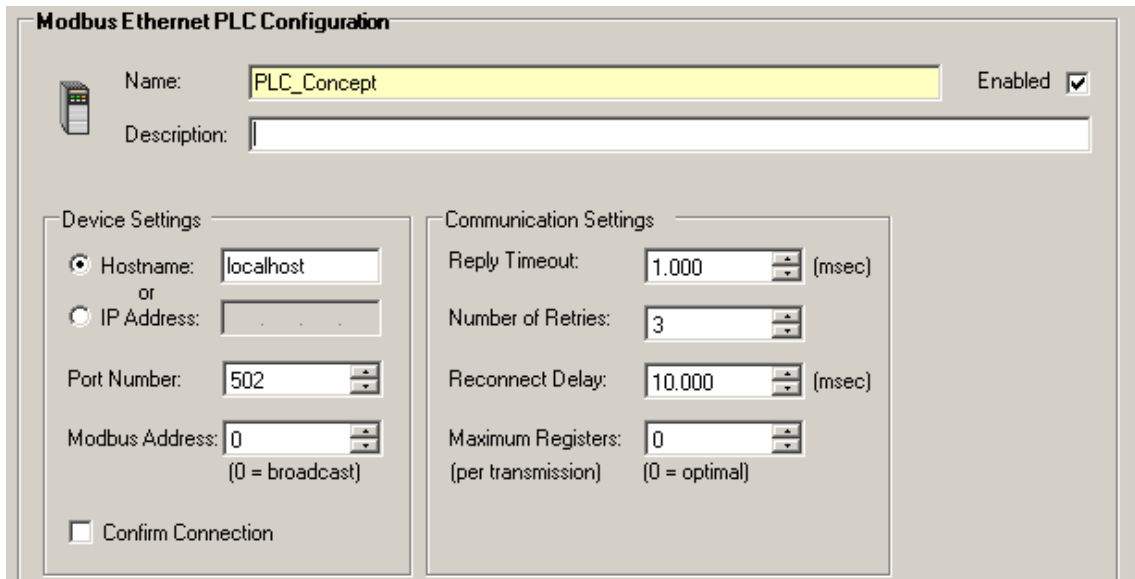


Figura 24 - Configuração do nó modbus ethernet.
Fonte: MatrikonOPC for Modbus.

O parâmetro configurado foi o “Hostname” que foi definido como “localhost”, pois a fonte de dados, no caso o CLP simulado, se encontra na própria máquina onde foi instalado o servidor OPC. Os outros parâmetros não foram alterados e suas configurações foram deixadas nos valores padrões.

Tendo sido feitas estas configurações, clicamos no botão “Create” e concluímos a configuração do servidor OPC.

A integração com o Excel foi feita por meio do software OPCWareXL que permite a leitura e escrita de valores no servidor OPC por meio da instalação de um suplemento no Excel (FIG. 25).

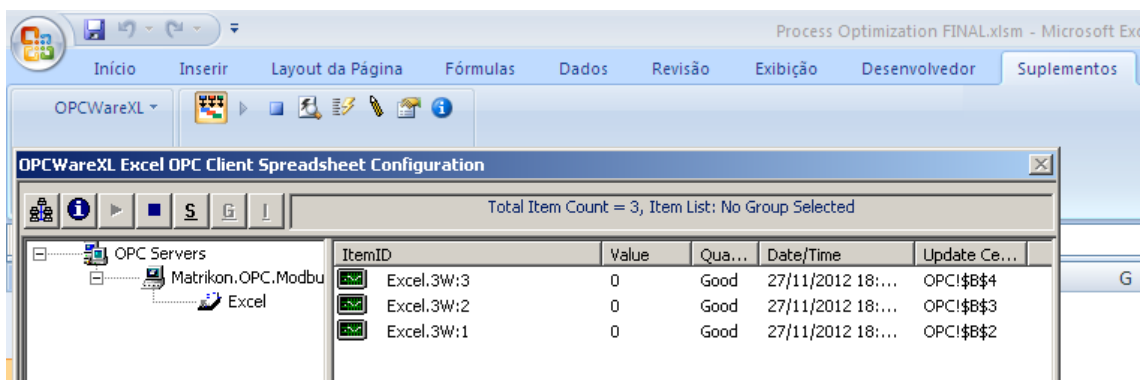


Figura 25 - Tela de configuração do OPCWareXL.
Fonte: OPCWareXL.

Para configurar a conexão com o servidor OPC devemos clicar em “OPC Servers” e em seguida em “Add Server”. Na tela que se abre (FIG. 26), devemos determinar o computador onde está instalado o servidor OPC e em seguida selecionar o servidor Modbus OPC.

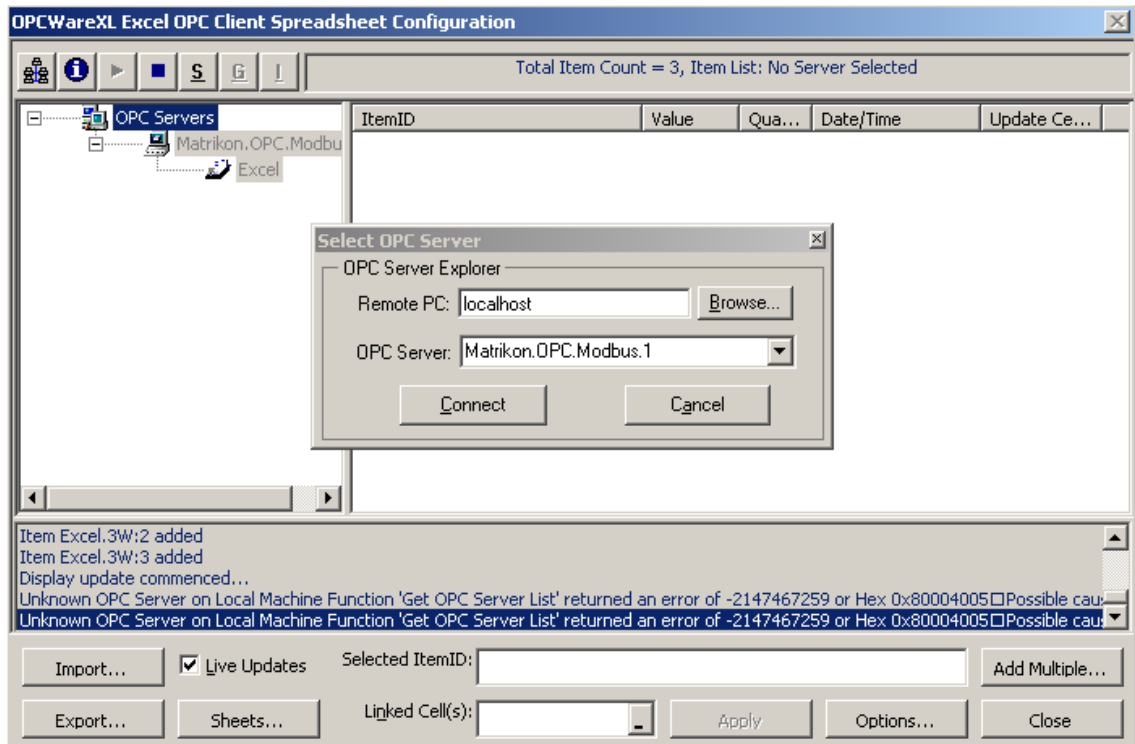


Figura 26 - Configuração do servidor OPC com o software OPCWareXL.
Fonte: OPCWareXL.

Em seguida clicamos com o botão direito em “Matrikon.OPC.Modbus.1” e depois em “Add Group” para adicionar um grupo que irá relacionar com as tags do nó já configurado no servidor OPC. A tela que se abre deve ser configurada como na figura 27.

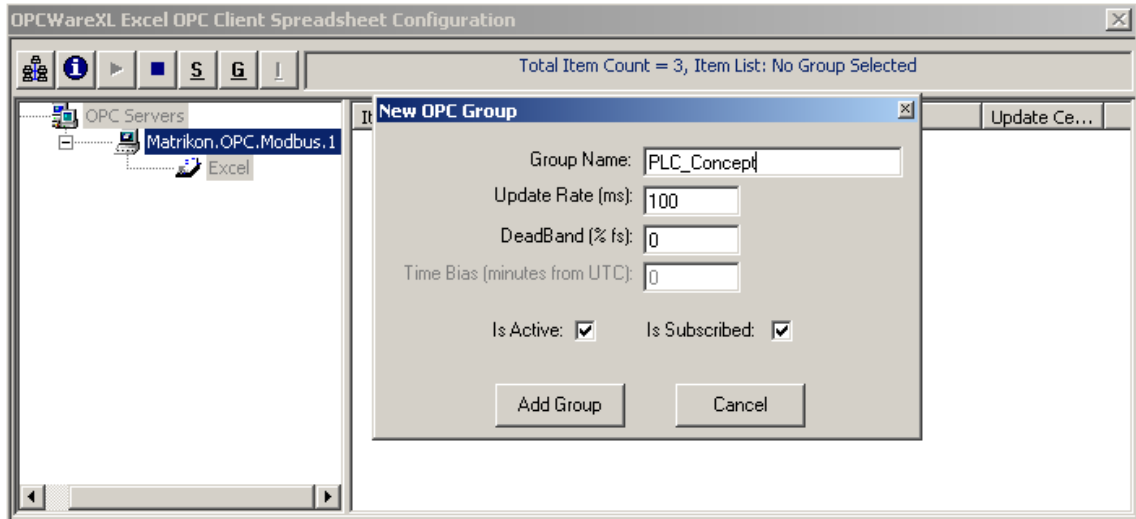


Figura 27 - Configuração de grupo no software OPCWareXL.
Fonte: OPCWareXL.

Para adicionar os itens, clicamos com o botão direito no grupo criado “PLC_Concept” e em “New Item”. Na nova tela, navegamos com o “OPC item Browser”, localizado no canto superior esquerdo da tela, até encontrar a nossa fonte de dados (FIG. 28).

No nosso caso, a variável que se deseja ler é contínua. Variáveis desse tipo são representadas no padrão Modbus pelo grupo “3W”. Assim os itens que queremos configurar são: “Plc_concept.3W:1”, “Plc_concept.3W:2” e “Plc_concept.3W:3”.

Para adicioná-los ao OPCWareXL, digitamos estes endereços no campo “Item ID” e em seguida clicamos em “Validate” e “Add” para adicioná-los. Após configurar as 3 entradas, clicamos em “Done”.

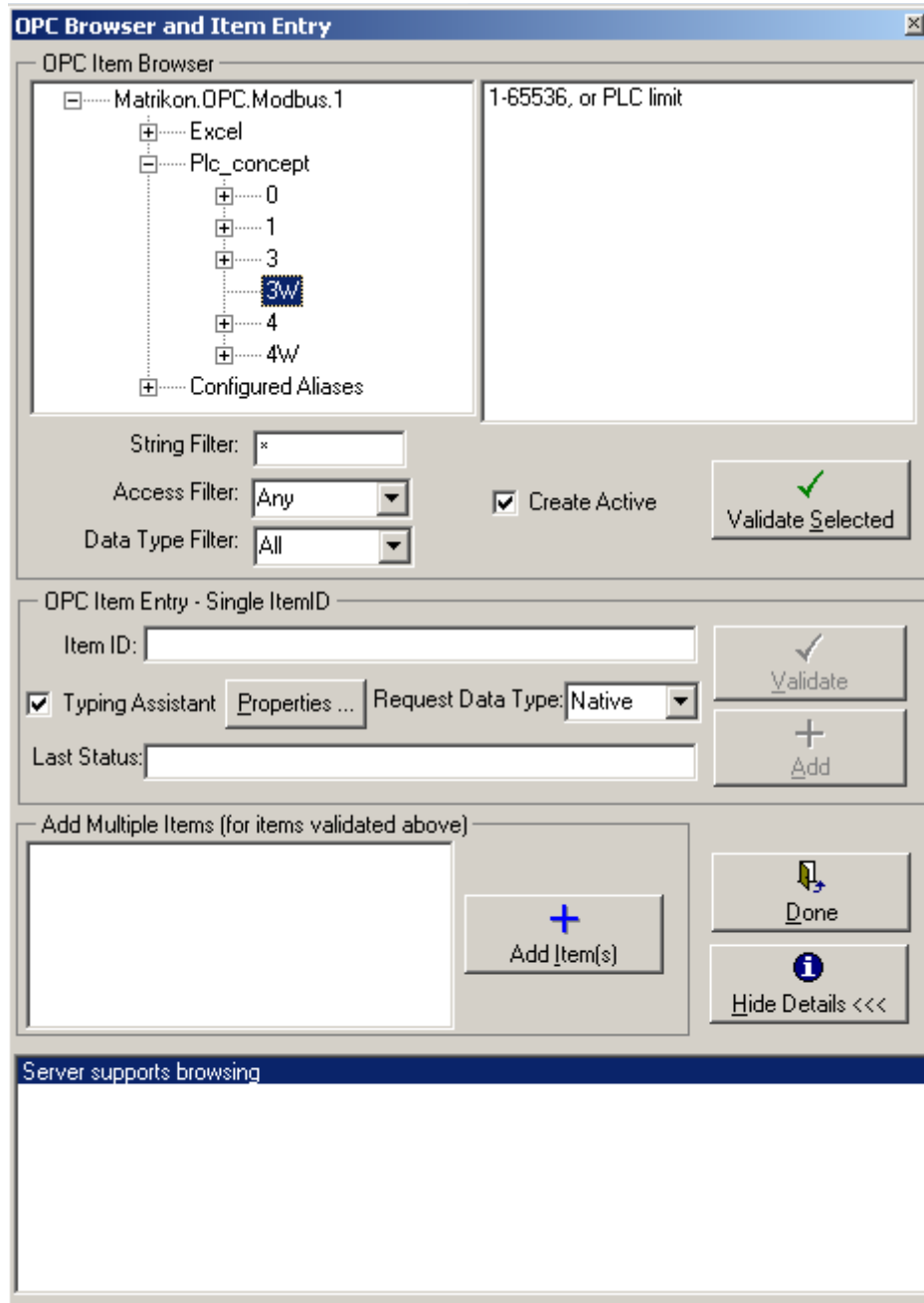


Figura 28 - Tela de localização dos itens do servidor OPC.
Fonte: OPCWareXL.

Com os itens prontos para serem ligados com células, digitamos o nome do item em “Selected ItemID” e o endereço da célula em “Linked Cell” (FIG. 29).

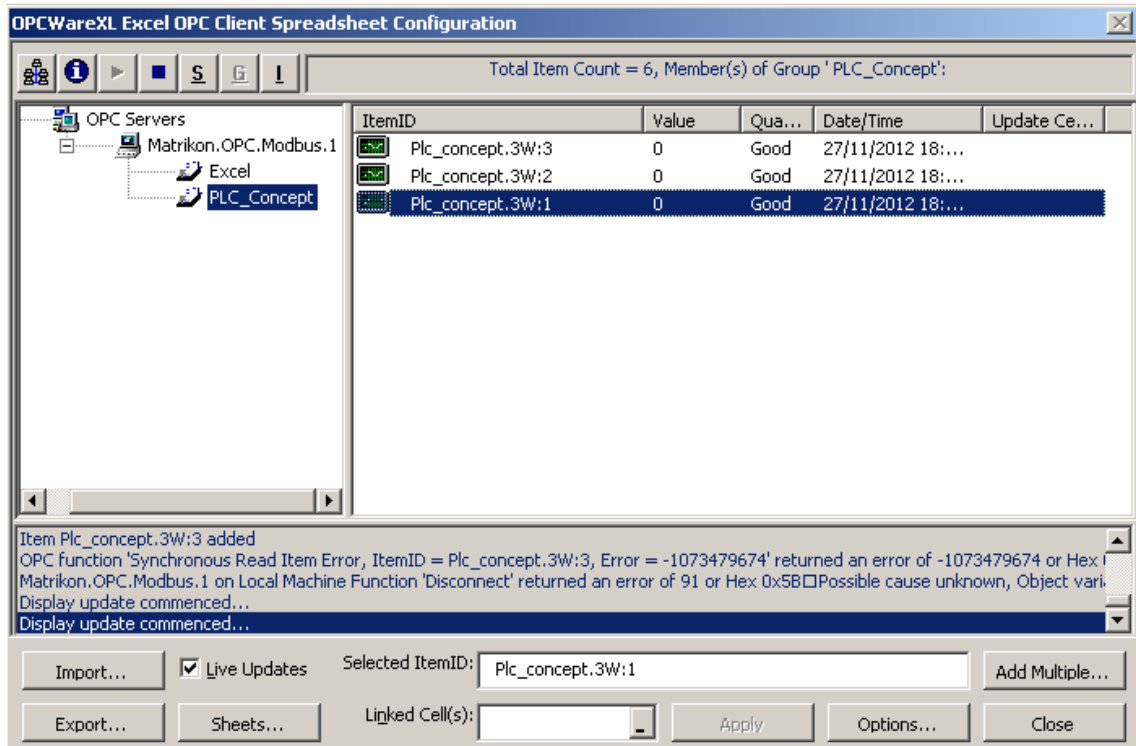


Figura 29 - OPCWareXL com os itens já configurados.
 Fonte: OPCWareXL.

A figura 30 mostra a simulação do CLP, feita utilizando o programa Concept da Schneider Electric.

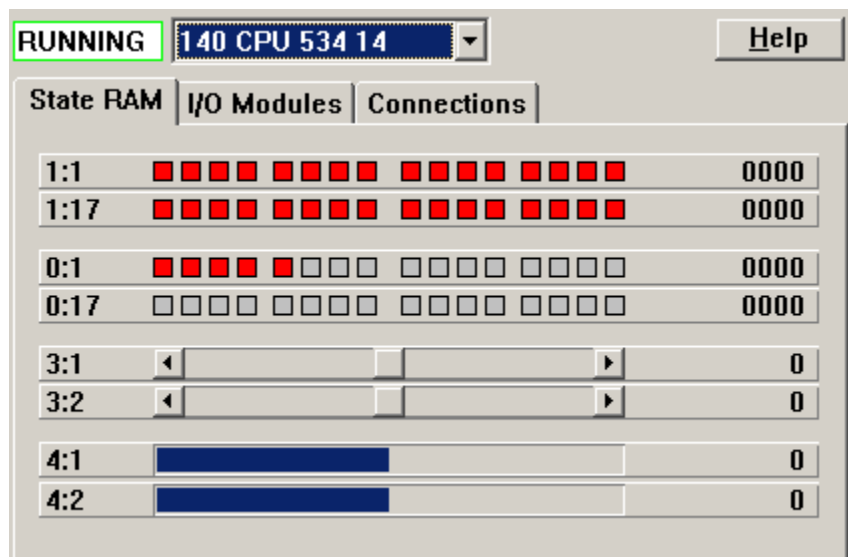


Figura 30 - Simulador do CLP, usando o software Concept.
 Fonte: Concept, Schneider Electric.

Para alterar os valores lidos pelo servidor OPC, modificamos a posição das barras no item 3:1 ou 3:2. Ainda é possível selecionar outro endereço de interesse, clicando nos itens 3:1 ou 3:2 e digitando o endereço requerido (FIG. 30).

3.1.2 Descrição da aplicação de otimização usando planejamento fatorial e metodologia de superfície de resposta.

Por se tratar de um método iterativo a aplicação foi desenvolvida em etapas. A primeira etapa, é apresentada uma tela onde se deve selecionar o objetivo da otimização, maximização ou minimização (FIG. 31). Nos campos “Passo X1” e “Passo X2” definimos a região onde será realizado o planejamento da otimização.

No frame “Leitura das Variáveis do Processo” são realizadas as leituras dos valores das variáveis junto ao servidor OPC, por meio das quais o usuário não mais necessita se deslocar ao campo para recolhimento dos dados de processo.

Em “Ponto Central” configuramos o ponto central do planejamento, com os valores relativos às duas variáveis de interesse.

Após o preenchimento dos campos de passo e de ponto central, o acionamento do botão “Gerar Planejamento” gera automaticamente o planejamento fatorial de dois níveis com ponto central. Após gerar esta matriz de planejamento, é necessário uma comunicação do profissional de otimização com o operador do sistema para que este segundo leve as variáveis x_1 e x_2 para os respectivos valores do planejamento, de forma que a resposta em cada experimento possa ser coletada por meio dos campos de leitura do OPC³ e digitada nos resultados dos respectivos ensaios.

³ Para descrever a aplicação, fizemos o modelamento nos baseando nos dados do exemplo citado no subtítulo 1.3.2 Metodologia de Superfície de Resposta.

Process Optimization

Process Optimization
By Bell e Tipautim

Configurações

Objetivo: Maximizar

Passo X1: 5

Passo X2: 10

Leitura das Variáveis do Processo

Variável X1: 50,01

Variável X2: 100,00

Resultado: 69,01

Experimentos

Variável X1: 50 Variável X2: 100

Matriz de Planejamento

X1	X2	Variáveis codificadas		Resultado
45,00	90,00	-1,00	-1,00	69
55,00	90,00	1,00	-1,00	59
45,00	110,00	-1,00	1,00	78
55,00	110,00	1,00	1,00	67
50,00	100,00	0,00	0,00	68
50,00	100,00	0,00	0,00	66
50,00	100,00	0,00	0,00	69

Gerar Planejamento

Próximo Passo >>

Figura 31 - Tela inicial da aplicação para otimização de processos.
Fonte: Elaboração própria.

Depois de ter coletado as respostas para todos os ensaios, o botão “Próximo Passo” é habilitado. Ao clicar neste botão, são gerados os gráficos da superfície e sua respectiva curva de nível, que são apresentados na tela subsequente (FIG. 32).

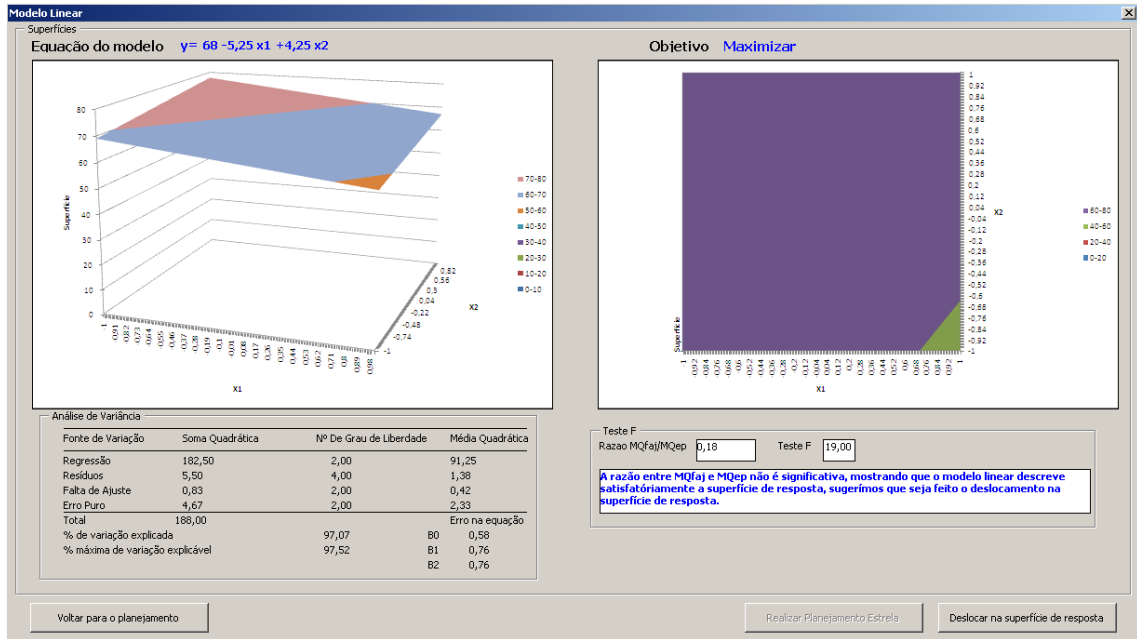


Figura 32 - Modelo linear com base no planejamento fatorial de dois níveis com ponto central.
Fonte: Elaboração própria.

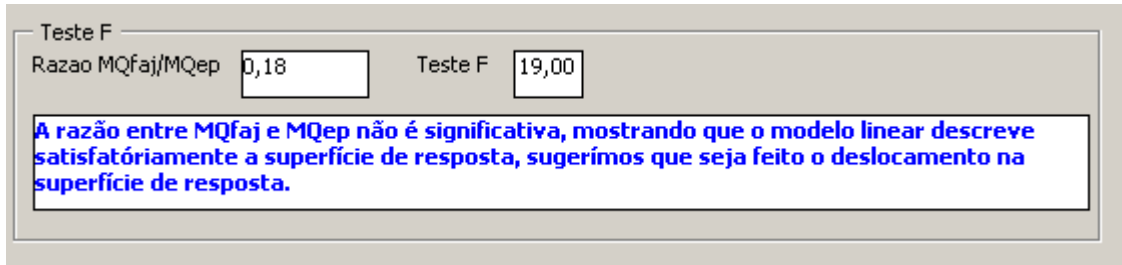
Na figura 32, podemos observar a equação do modelo estimado no canto superior esquerdo e o objetivo da otimização no canto superior direito. Abaixo do gráfico da superfície temos a tabela de análise de variância, que pode ser melhor visualizada na figura 33.

Fonte de Variação	Soma Quadrática	Nº De Grau de Liberdade	Média Quadrática
Regressão	182,50	2,00	91,25
Resíduos	5,50	4,00	1,38
Falta de Ajuste	0,83	2,00	0,42
Erro Puro	4,67	2,00	2,33
Total	188,00		Erro na equação
% de variação explicada		97,07	B0 0,58
% máxima de variação explicável		97,52	B1 0,76
			B2 0,76

Figura 33 - Análise de variância do modelo linear com base no planejamento fatorial de dois níveis com ponto central.
Fonte: Elaboração própria.

No canto inferior direito, melhor visualizado na figura 34, temos o resultado de um teste F, que verifica a compatibilidade do modelo linear com os dados do planejamento e informa se o modelo se adequa aos dados. Com base no resultado, é sugerido o deslocamento na superfície de resposta em direção à máxima

inclinação, ou, caso o modelo não se adéqüe é sugerida a realização de um planejamento estrela.



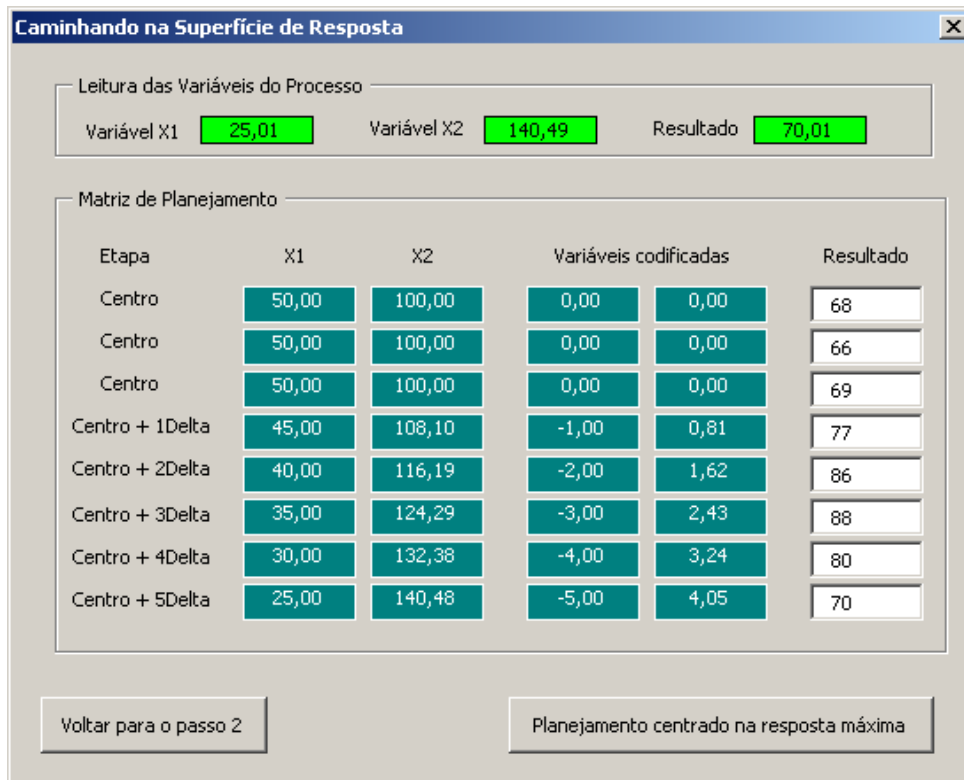
Teste F

Razao MQfaj/MQep Teste F

A razão entre MQfaj e MQep não é significativa, mostrando que o modelo linear descreve satisfatoriamente a superfície de resposta, sugerimos que seja feito o deslocamento na superfície de resposta.

Figura 34 - Teste F para determinar a validade do modelo.
Fonte: Elaboração própria.

Clicando no botão “Deslocar na superfície de resposta”, temos uma nova tela (FIG. 35), com uma matriz de planejamento com novas variáveis codificadas de acordo com o caminho de máxima inclinação. Na parte superior da tela temos novamente a leitura das variáveis do processo coletadas pelo servidor OPC e mais uma vez é necessário uma comunicação com o operador para levar as variáveis aos respectivos pontos do planejamento.



Caminhando na Superfície de Resposta

Leitura das Variáveis do Processo

Variável X1 Variável X2 Resultado

Matriz de Planejamento

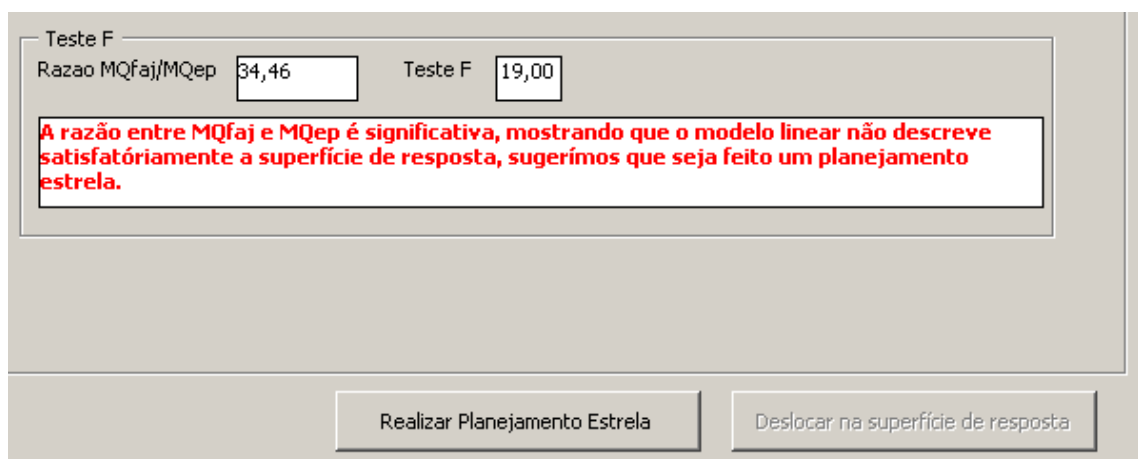
Etapa	X1	X2	Variáveis codificadas		Resultado
Centro	50,00	100,00	0,00	0,00	68
Centro	50,00	100,00	0,00	0,00	66
Centro	50,00	100,00	0,00	0,00	69
Centro + 1Delta	45,00	108,10	-1,00	0,81	77
Centro + 2Delta	40,00	116,19	-2,00	1,62	86
Centro + 3Delta	35,00	124,29	-3,00	2,43	88
Centro + 4Delta	30,00	132,38	-4,00	3,24	80
Centro + 5Delta	25,00	140,48	-5,00	4,05	70

Figura 35 - Planejamento seguindo o caminho de máxima inclinação.
Fonte: Elaboração própria.

Com base no objetivo escolhido, maximizar ou minimizar, haverá um botão denominado “Planejamento centrado na resposta máxima” ou “Planejamento centrado na resposta mínima”, clicando neste botão, será gerado uma nova tela com um planejamento fatorial de dois níveis com ponto central, baseado na resposta de maior ou menor rendimento, dependendo do objetivo da otimização. O programa será levado para a tela inicial onde será feito uma nova análise do planejamento com o ponto central localizado na resposta máxima do planejamento anterior e novamente será necessária a comunicação com o operador para recolher os dados dos experimentos.

Clicando em “Próximo Passo” será exibida uma tela com os gráficos do novo planejamento e sua análise de variância, bem como o teste F.

Caso o novo modelo linear se adéqüe aos dados do planejamento será sugerido fazer um novo deslocamento na superfície de máxima inclinação. Porém, caso o modelo não se ajuste aos dados provenientes do planejamento fatorial, fato este constatado pelo teste F, no canto inferior direito, será sugerido que se faça um planejamento estrela para que se possa estimar um modelo quadrático (FIG. 36).



The screenshot shows a window titled "Teste F". It contains two input fields: "Razao MQfaj/MQep" with the value "34,46" and "Teste F" with the value "19,00". Below these fields is a red text box with the following message: "A razão entre MQfaj e MQep é significativa, mostrando que o modelo linear não descreve satisfatoriamente a superfície de resposta, sugerimos que seja feito um planejamento estrela." At the bottom of the window, there are two buttons: "Realizar Planejamento Estrela" and "Deslocar na superfície de resposta".

Figura 36 - Incompatibilidade entre o modelo linear e os dados provenientes do planejamento fatorial.
Fonte: Elaboração própria.

Clicando no botão “Realizar Planejamento Estrela” (FIG. 36), é exibida uma nova tela (FIG. 37), com o planejamento em estrela, que nada mais é que um planejamento fatorial ampliado.

Planejamento Estrela

Leitura das Variáveis do Processo

Variável X1 **35,02** Variável X2 **110,01** Resultado **87,01**

Experimentos

Variável X1 Variável X2

Ponto Central 35,00 124,29

Matriz de Planejamento

X1	X2	Variáveis codificadas		Resultado
30,00	114,29	-1,00	-1,00	86
40,00	114,29	1,00	-1,00	85
30,00	134,29	-1,00	1,00	78
40,00	134,29	1,00	1,00	84
35,00	124,29	0,00	0,00	90
35,00	124,29	0,00	0,00	88
35,00	124,29	0,00	0,00	89
27,93	124,29	-1,41	0,00	81
35,00	138,43	0,00	1,41	80
42,07	124,29	1,41	0,00	86
35,00	110,15	0,00	-1,41	87

Próximo Passo

Figura 37 - Tela de ajuste do planejamento estrela.
Fonte: Elaboração própria.

Na parte superior da tela podemos observar os campos para leitura das variáveis de interesse, de forma que novamente deve-se fazer uma requisição ao operador do processo, para que coloque as variáveis nos seus respectivos valores de ensaio, e assim sejam coletadas as respostas em cada nível.

Por fim, clicamos no botão “Próximo Passo” para gerar a superfície de resposta correspondente ao novo modelo estimado pelo planejamento estrela (FIG. 38).

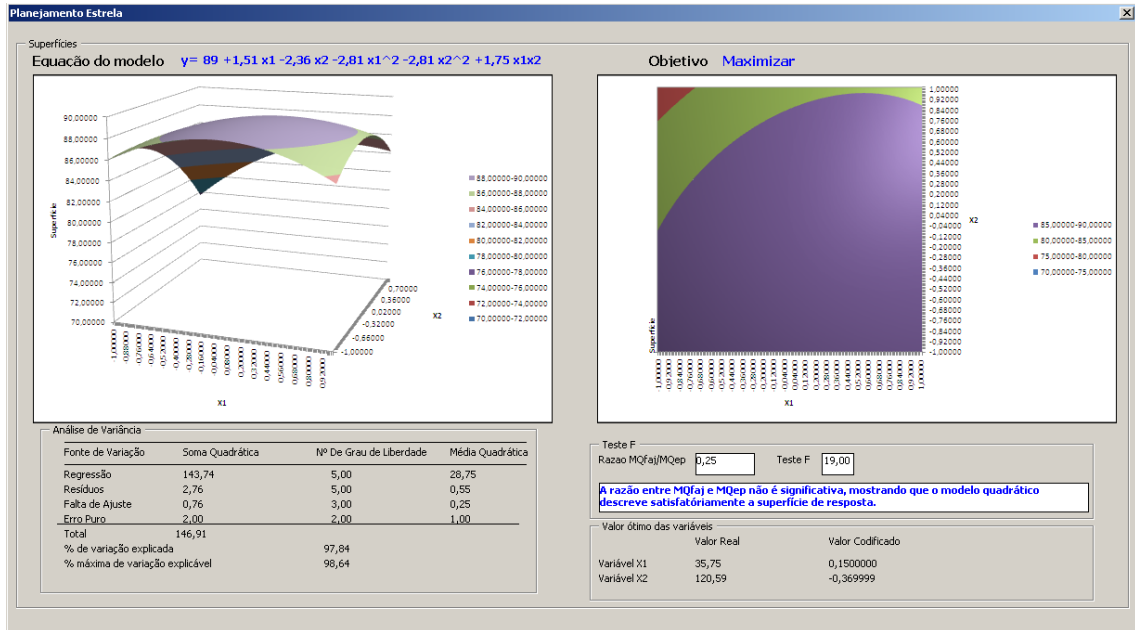


Figura 38 - Resultados do modelo quadrático proveniente do planejamento estrela.
Fonte: Elaboração própria.

Na figura 38, observamos a superfície de resposta descrita pelo modelo quadrático no canto superior esquerdo e suas curvas de nível no canto superior direito.

Na parte de baixo da figura é apresentada a análise de variância, que pode ser melhor visualizada na figura 39, e a análise do teste F e o valor ótimo das variáveis, o que pode ser melhor visualizado na figura 40.

Análise de Variância

Fonte de Variação	Soma Quadrática	Nº De Grau de Liberdade	Média Quadrática
Regressão	143,74	5,00	28,75
Resíduos	2,76	5,00	0,55
Falta de Ajuste	0,76	3,00	0,25
Erro Puro	2,00	2,00	1,00
Total	146,91		
% de variação explicada		97,84	
% máxima de variação explicável		98,64	

Figura 39 - Análise de variância para o modelo quadrático.
Fonte: Elaboração própria.

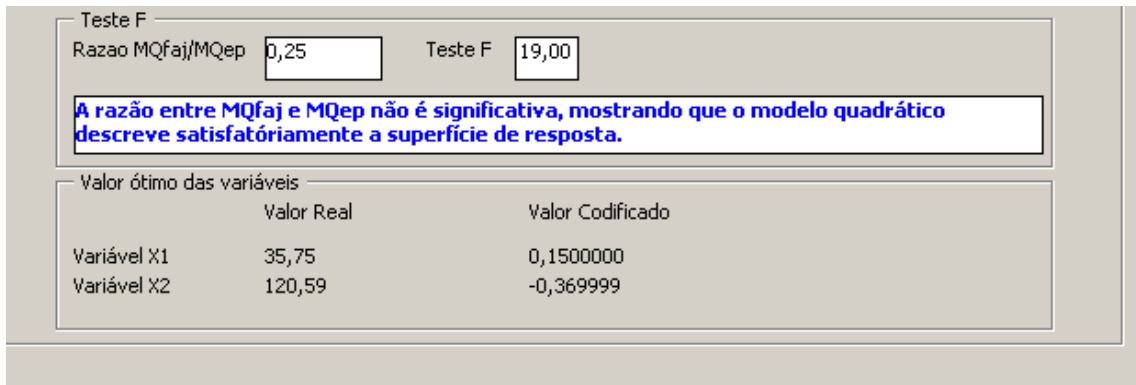


Figura 40 - Teste F e valor ótimo das variáveis.

Fonte: Elaboração própria.

Dáí termina-se o processo de otimização e passaríamos então a verificação real de aderência do modelo ao processo específico a partir da utilização dos valores ótimos encontrados para os fatores em estudo.

CONCLUSÃO

Este trabalho apresentou um estudo sobre otimização de sistemas com ênfase na metodologia de superfície de resposta e apresentação de uma solução para agilizar a coleta de dados de processos industriais, com base em tecnologias já disponíveis.

O uso da tecnologia OPC como meio de aquisição de dados apresentou diversas vantagens tais como: fácil integração entre diversas fontes de dados e a aplicação desenvolvida, elimina a necessidade de drivers específicos e é uma ferramenta presente na maior parte dos processos industriais.

Apresentamos uma aplicação para simular a possibilidade de integrar as técnicas de otimização às tecnologias já consolidadas, como solução para reduzir tempo e custo nos processos de otimização e seu desempenho foi satisfatório.

Como sugestão para trabalhos e estudos futuros, incentivamos a elaboração de um software de otimização que dispense o uso de outras ferramentas, como o OPCWareXL e Excel, já que o cliente OPC estaria intrínseco, bem como o desenvolvimento de uma aplicação que permita analisar mais do que duas variáveis.

REFERÊNCIAS BIBLIOGRÁFICAS

AFOAKWA, E. O.; et al. **Modelling tempering behaviour of dark chocolates from varying particle size distribution and fat content using response surface methodology**. Elsevier, Dec. 2007/Mar. 2008.

AGUIAR, F. L.; MOREIRA JÚNIOR, W. I. **Ajuste de curvas por quadrados mínimos lineares**. Disponível em: <http://www.mat.ufmg.br/gaal/aplicacoes/quadrados_minimos.pdf>. Acesso em: 29 set. 2012.

BARROS NETO, B. B.; SCARMINIO, I. S.; BRUNS, E. R. **Como fazer experimentos**. 3.ed. Campinas, SP: Editora UNICAMP, 2007.

BINGOL, D; et al. **Comparison of the results of response surface methodology and artificial neural network for the biosorption of lead using black cummin**. Elsevier, Jan./Feb. 2011.

CÂNDIDO, R. V. B. **Padrão OPC: Uma alternativa de Substituição dos Drivers Proprietários para acessar dados de PLCs**. 2004. Trabalho de Conclusão de Curso (Bacharel em Ciência da Computação) – FUMEC, Belo Horizonte. 2004.

CHEN, G.; et al. **Application of response surface methodology for optimization of the synthesis of synthetic rutile from titania slag**. Elsevier, Aug. /Nov. 2011.

EIRAS, S. P.; et al. **Métodos de Otimização em Química**. 2000. Disponível em: <www.chemkeys.com> Acesso em: 18 abril 2012.

FONSECA, M. de O. **Comunicação OPC – Uma abordagem prática**. In: SEMINÁRIO DE AUTOMAÇÃO DE PROCESSOS DA ABM, 6., 2002, Vitória, Brasil. Anais... Vitória, 2002.

GOULART, E. da S. **Matrizes Quase-Newton esparsas para problemas de otimização não-linear de grande porte**. 2005. Tese (Doutorado em Ciências em Engenharia Mecânica) - Universidade Federal do Rio de Janeiro, Rio de Janeiro 2005. Disponível em: <<http://www.optimize.ufrj.br/files/EvandroDaSilveiraGoulart.pdf>>. Acesso em 06 out. 2012.

GONTIJO, L. C. **Estudo da formação do 4,5-benzopiazoselenol nas misturas homogêneas dos solventes água - etanol – clorofórmio (ou tolueno)**. 2006. Dissertação (Mestrado em Química) - Instituto de Química, Universidade Federal de Uberlândia, Uberlândia –MG. 2006.

GUIMARÃES NETO, J. F. G.; NASCIMENTO, M. L.; NUNES, M. K. **O uso de simuladores para otimização de processos**. Intech America do Sul, São Paulo, n. 116 p. 48 – 54, 2009.

HINES, W. W. et al. Planejamento de experimentos com vários fatores. . In: _____. **Probabilidade e Estatística na Engenharia**. Tradução Vera Regina Lima de Faria Flores. Revisão técnica Ana Maria Lima de Farias. 4. ed. Rio de Janeiro: LTC, 2006. Cap. 13, p. 315-365.

KOMINEK, D. **OPC: The Ins and Outs to What It's About** : “The Every Man’s Guide to OPC”. 2009, Alberta. Disponível em:
<http://www.automation.com/pdf_articles/Guide_to_OPC.pdf>. Acesso em: 20 out. 2012.

Laboratório de Otimização em Engenharia. **Saiba mais sobre otimização**. Disponível em:<<http://www.optimize.ufrj.br/corpo/lab.htm>>. Acesso em: 17 abril 2012.

LI, A.; LIN, C.; TRAN, D. **Optimizing the response surface for producing ethanol from avicel by Brevibacillus strain AHPC8120**. Elsevier, Nov. 2010/Apr. 2011.

LOPES, F. A. B. et al. **Protocolo OPC: Introdução e Aplicações na Automação Industrial**. 2007. Trabalho de Conclusão de Curso (Especialização em Automação Industrial) – Universidade do Estado do Rio de Janeiro – Universidade Petrobrás, Rio de Janeiro, 2007.

LOUZADA, F. et al. **Metodologia de superfície de resposta: uma introdução nos softwares R e Statistica**. São Carlos. Disponível em:
<http://www2.icmc.usp.br/~biblio/BIBLIOTECA/rel_tec/RT_376.pdf>. Acesso em: 29 set. 2012.

MATHEUS, I. M. ABRAHÃO, L. C. L. **AUTOMAÇÃO INDUSTRIAL: SISTEMAS SUPERVISÓRIOS**. Uberlândia, 2011.

Método dos mínimos quadrados. Disponível em:
<<http://www.math.ist.utl.pt/~calves/cursos/mmq.htm>>. Acesso em: 29 set. 2012.

Método dos mínimos quadrados. Disponível em:
<<http://evunix.uevora.pt/~stadler/FAN-06-07/reglin.pdf>>. Acesso em: 29 set. 2012.

MONTGOMERY, D. C.; RUNGER, G. C. Planejamento de experimentos com vários fatores. In: _____. **Estatística aplicada e probabilidade para engenheiros**. 1943. Tradução e revisão técnica de Verônica Calado. 4. ed. Rio de Janeiro: LTC, 2009. Cap. 14, p. 330-370.

MONTGOMERY, D. C.; RUNGER, G. C.; HUBELE, N. F. Construindo modelos empíricos. In: _____. **Estatística aplicada à Engenharia**. 2. ed. Rio de Janeiro: LTC, 2004. Cap. 6 e 7, p. 177-277.

NEVES NETO; M. G. **Uma análise de alguns métodos de otimização longe do minimizador**. Trabalho de Conclusão de Curso (Engenharia da Computação) – Universidade Federal do Rio Grande de Norte, Natal, 2010. Disponível em:
<<http://www.engcomp.ufrn.br/publicacoes/TCC-2010-1-15.pdf>>. Acesso em 17/04/2012.

NOGUEIRA, T. A. Padrão OPC (OLE for Process Control). In: NOGUEIRA, T. A. **REDES DE COMUNICAÇÃO PARA SISTEMAS DE AUTOMAÇÃO**. Trabalho de Conclusão de Curso (Engenharia de Controle e Automação) – CECAU, Universidade Federal de Ouro Preto, Ouro Preto, 2009. p. 65-69. Disponível em
<<http://www.em.ufop.br/cecau/monografias/2009/THIAGO%20AUGUSTO.pdf>>. Acesso em 17 abril 2012.

NUNES, D. B. **Rotinas para otimização experimental de misturas**. 1998. Dissertação (Mestrado em Engenharia de Produção) – Programa de pós-graduação em Engenharia de Produção, Universidade Federal do Rio Grande do Sul, Porto Alegre. 1998. Disponível em:
<<http://www.lume.ufrgs.br/bitstream/handle/10183/2639/000236213.pdf?sequence=1>>. Acesso em: 29 set. 2012.

NUNES; R. M. **Otimização de matéria prima**. Disponível em:
<<http://www.trabalhosfeitos.com/ensaios/Otimiza%C3%A7%C3%A3o-De-Materia-Prima/39836.html>>. Acesso em: 17 abril 2012.

O método simplex. Disponível em:
<http://wwwp.fc.unesp.br/~arbalbo/Iniciacao_Cientifica/simplex/teoria/3_simplex.pdf>. Acesso em 29 set. 2012.

OPC Foundation. **What is OPC?**. Disponível em:
<http://www.opcfoundation.org/Default.aspx/01_about/01_what_is.asp?MID=AboutOPC>. Acesso em: 18 abril 2012.

PAIVA, A. P. **Metodologia de Superfície de Resposta e Análise de Componentes Principais em Otimização de Processos de Manufatura com Múltiplas Respostas Correlacionadas**. 2006. Tese (Doutorado em Engenharia Mecânica) - Programa de Pós-Graduação em Engenharia Mecânica, Universidade Federal de Itajubá, Itajubá. 2006. Disponível em:
<http://www.portal.unifei.edu.br/files/arquivos/PRPPG/Engenharia_mecanica/Projeto_fabrica%C3%A7%C3%A3o_doutorado/Anderson%20Paulo%20de%20Paiva.pdf>. Acesso em 06 out. 2012.

PEREIRA, A. S. et al. **Pesquisa Operacional: programação linear**. Disponível em:
<<http://www.ebah.com.br/content/ABAAAANOsAG/trabalho-pesquisa-operacional>>. Acesso em 29 set. 2012.

PIRES, N. J. **Determinação catalítica de selênio em cabelo humano empregando a redução de TAN a DHTAN com íon brometo como ativador em sistema de AFCM com detecção espectrofotométrica**. 2004. Dissertação (Mestrado em Química) – Instituto de Química, Universidade Federal de Uberlândia, Uberlândia. 2004.

PIRES, T. S. **Método de superfície de resposta aplicado à otimização termoeconômica de sistemas de cogeração modelados em um simulador de processos**. 2010. Dissertação (Mestrado em Engenharia Mecânica) – COPPE, Universidade Federal do Rio de Janeiro, Rio de Janeiro. 2010. Disponível em:
<http://objdig.ufrj.br/60/teses/coppe_m/ThiagoDaSilvaPires.pdf>. Acesso em 06 out. 2012.

Portal Action. **Estimação dos parâmetros do modelo**. Disponível em:
<<http://www.portalaction.com.br/content/12-estima%C3%A7%C3%A3o-dos-par%C3%A2metros-do-modelo>>. Acesso em: 29 set. 2012.

PUDA, A.P. **Padronização da comunicação através da tecnologia OPC**. 2008. Disponível em:<www.isarj.org.br/artigos/Padronizacao-da-Comunicacao-atraves-da-Tecnologia-OPC.pdf> Acesso em 20 out. 2009.

RIBEIRO, D. F.; ANJOS, R.; LEANDRO. **Utilização do Drive OPC na integração de sistemas de automação industrial**. In: Seminário Nacional de Sistemas Industriais e Automação, 3., 2008, Belo Horizonte.

SOUZA, L. C. A.; SEIXAS FILHO, C.; PENA, R. T. **Padrão de acesso a dados OPC e sua implementação em um driver OPC-Modbus**. Disponível em: <http://dc136.4shared.com/doc/_DLa412y/preview.html>. Acesso em: 21 out. 2012.

Tecnal equipamentos para laboratórios. Disponível em: <www.tecnallab.com.br>. Acesso em: 15 nov. 2012.

WANG, J. et al. **Optimization of coagulation–flocculation process for a paper-recycling wastewater treatment using response surface methodology**. Elsevier, Feb. 2007.