



UNIVERSIDADE
AbERTA
www.uab.pt

Universidade Aberta
Departamento de Ciências e Tecnologia
Mestrado em Estatística, Matemática e
Computação
Área de Especialização: Estatística
Computacional

**"Metodologias de Superfície de Resposta: uma
investigação no âmbito da indústria cerâmica e do vidro"**

Elaborada por:
António Francisco da Silva Mondim

Ano de 2014



UNIVERSIDADE
AbERTA
www.uab.pt

Universidade Aberta
Mestrado em Estatística, Matemática e
Computação
Área de Especialização: Estatística
Computacional

"Metodologias de Superfície de Resposta: uma
investigação no âmbito da indústria cerâmica e do vidro"

Mestrando

António Francisco da Silva Mondim, 904411

Orientadora:

Professora Doutora Teresa Paula Costa Azinheira Oliveira

Dissertação apresentada para obtenção do grau de Mestre em
MEMC

Agradecimentos

Agradeço de modo particular à minha Orientadora Doutora Teresa Paula Costa Azinheira Oliveira toda a paciência que demonstrou, o empenho e carinho que colocou na motivação para este trabalho.

Agradeço a todos os meus Professores das várias áreas disciplinares deste Mestrado.

Agradeço de modo especial ao Dr. Carlos Gonçalves e companheira Júlia Alves, ao seu Pai e irmãos, proprietários da empresa MAGROU, que gentil e graciosamente, me proporcionaram a oportunidade de realização da experiência, que autorizaram, acompanharam e facilitaram.

Agradeço ainda aos meus amigos que sempre me acompanharam e não me deixaram esmorecer ao longo deste tempo de trabalho árduo e empenhado, nomeadamente à minha amiga Biatriz Neves e à colega Lúcia Teves, que pacientemente me foram acompanhando e sugerindo melhorias estéticas de apresentação do trabalho.

Por fim, mas não por último, agradeço a colaboração, sempre prestimosa, da minha filha Anabela Ramos da Fonseca Mondim.

“Deus quer, o Homem sonha e a Obra nasce.”

Fernando Pessoa

Dedico este trabalho à Ciência em geral e à Indústria Cerâmica em particular, desejando ver revertida a situação catastrófica em que grande parte das fábricas cerâmicas de Barcelos sobrevive.

Resumo

A metodologia de superfícies de resposta MSR é um conjunto de técnicas estatístico-matemáticas que visa a otimização. No nosso tipo de sociedade, orientada para um mercado globalizado onde impera a lei do mais forte, qualquer ganho de produtividade, qualquer diminuição dos custos de produção, qualquer inovação, podem marcar a diferença entre o sucesso e a falência. Falar em MSR é um sonho, implementar a metodologia é antecipar o sucesso.

O processo inicia-se com a interpretação do problema, definição clara do objectivo e identificação dos factores que o influenciam. Num primeiro momento, um modelo de primeira ordem é, normalmente, suficiente para modelar o processo, uma regressão múltipla estabelece a relação entre o objectivo e os factores identificados como relevantes. Verificados os pressupostos de aplicação do modelo e a linearidade do processo, urge caminhar no sentido da superfície de resposta que presumivelmente contém o ótimo (máximo ou mínimo). Para isso procuramos o máximo incremento na resposta: pelo método da máxima inclinação ascendente (*steepest ascent*) se procuramos um ponto de máximo, ou pelo método da máxima inclinação descendente (*steepest descent*) se procuramos um mínimo.

Chegados à superfície de resposta ou ajustamos um novo modelo de primeira ordem ou realizamos experiências adicionais, particularmente ajustando um polinómio de ordem dois, para obter uma estimativa mais precisa do ótimo.

A MSR é um processo sequencial cujo algoritmo iremos descrever através de um caso prático no âmbito da indústria cerâmica, mais concretamente no fabrico de “chávenas de porcelana”. A aplicação da metodologia de resposta tem vindo a crescer exponencialmente como o atestam o elevado número de artigos científicos publicados e a diversidade dos campos onde é implementada.

Neste estudo, centramos a nossa atenção em modelos de primeira e segunda ordem, contudo a ciência não tem limite e analistas, apoiados na capacidade numérica e gráfica dos computadores, publicam já muitos trabalhos com modelos de ordem superior.

Palavras-chave: pasta cerâmica, regressão múltipla, regressão polinomial, superfícies de resposta, pontos de estacionaridade.

Abstract

The response surface methodology, MSR is a set of statistical and mathematical techniques aimed at optimization. In our kind of society, oriented towards a global market where the law of the strongest prevails, any productivity gains, any decrease in production costs, or any innovation, can make the difference between success and failure. Speaking MSR is a dream, implementing the methodology is to anticipate success.

The process begins with the interpretation of the problem, clear definition of the subject and identification the factors that influence it. At first, a model of the first order is usually sufficient to model the process, a multiple regression establishes the relationship between the subject and the factors, identified as relevant. Satisfied the conditions of application of the model and the linearity of the process, it is urgent to move towards the surface response that presumably contains the optimum (maximum or minimum). For this, we look for the maximum increment in the response: at the maximum upward slope (steepest ascent) method, if we seek a point of maximum, or by the method of maximum descending slope (steepest descent) if we seek a minimum.

Arriving at the response surface, or we adjust a new model of first order or we perform additional experiments , particularly adjusting a polynomial of order two, to obtain more accurate estimate of the optimal .

The MSR is a sequential process by which algorithm we will describe a study case under the ceramics industry, specifically in the manufacture of "porcelain cups." The response of this methodology has been growing exponentially as a large number of scientific articles have been published and the diversity of the fields where it is implemented.

In this study, we focus our attention on models of first and second order, yet science has no limit and analysts, supported by numerical and graphical capabilities of computers, many papers have been published with models of higher order.

Keywords: ceramic paste, multiple regression, polynomial regression, response surfaces , points of stationarity.

Índice

Resumo	V
Abstract.....	VI
1. Introdução	11
1.1 Conceitos e Noções Introdutórias	11
2. Resenha histórica e revisão da literatura.....	17
3. DOE – Delineamento de experiências	25
4. Análise de Variância.....	33
4.1 Terminologia.....	34
4.2 Pressupostos.....	34
4.3 Análise de Variância a um factor	41
5. Metodologia de Superfícies de Resposta (MSR).....	44
5.1 MSR em R usando “RSM”	46
5.2 Regressão.....	55
5.3 Optimização de uma única Resposta.....	59
6. Enquadramento do Planeamento Experimental.....	62
6.1 Introdução à Cerâmica.....	64
6.2 Controlo das pastas cerâmicas	69
6.3 Desenho Experimental	73
6.3.1 Variáveis dependentes	74
6.3.2 Variáveis independentes	75
7. Superfícies de Resposta: Método e Planeamento	79
7.1 Codificação dos dados e análise	84
7.2 Direção da máxima inclinação descendente	88
7.3 Conclusões do modelo de primeira ordem	91
8. Análise de uma superfície de resposta de segunda ordem.....	93

8.1	Caso de estudo: Modelo de 2ª ordem	94
8.2	Determinação do ponto de estacionaridade	98
8.2.1	Densidade	99
8.2.2	Fluidez	101
8.3	Conclusão do Modelo de Segunda Ordem	102
9.	Conclusões	104
	Bibliografia	106

Lista de gráficos

Gráfico 1- Superfície de Resposta 3D	12
Gráfico 2 - Gráfico de Contornos – Curvas de nível.....	13
Gráfico 3 - Exemplo de Superfície de Resposta 3D.....	14
Gráfico 4 - Exemplo de gráfico de contornos	14
Gráfico 5 – Os resíduos da regressão	56
Gráfico 6 – Curva 1 desfloculação da barbotina 1	78
Gráfico 7 – Curva 2 desfloculação da barbotina 2	78
Gráfico 8 – Superfície de Resposta Densidade	79
Gráfico 9 - Curvas de nível densidade.....	80
Gráfico 10 - Superfície de Resposta Fluidez	80
Gráfico 11 - Curvas de nível fluidez	80
Gráfico 12 - Planeamento fatorial 2^2 com pontos centrais	82
Gráfico 13 - Resíduos da regressão linear “densidade”	87
Gráfico 14 – Superfície de Resposta variável densidade.....	89
Gráfico 15 - Curvas de nível variável densidade	89
Gráfico 16- Superfície de Resposta variável Fluidez	90
Gráfico 17 - Curvas de nível variável fluidez	90
Gráfico 18 - Planeamento fatorial 2^2 com pontos centrais e axiais	93
Gráfico 19 - Planeamento fatorial Modelo	95
Gráfico 20- Resíduos da regressão quadrática “densidade”	97
Gráfico 21 - Resíduos da regressão quadrática “Fluidez”	97

Índice de Ilustrações

Ilustração 1 - Arranjos CCD	15
Ilustração 2 - Arranjo de Box-Behnken para 3 factores	16
Ilustração 3 - Fábrica Magrou	62
Ilustração 4 - Artigos em produção	63
Ilustração 5 - Fornos de cozedura.....	68
Ilustração 6 - Diluidor e Agitador de materiais não plásticos.....	71

Índice de Quadros

Quadro 1 - Organização de dados – distribuição das amostras.....	38
Quadro 2 - Organização de dados – descritivas das amostras	38
Quadro 3 - Tabela Anova (teórica).....	43
Quadro 4 - Lançamento de dados em R.....	47
Quadro 5 - Codificação de dados em R	48
Quadro 6 - Codificação/Descodificação de dados em R	48
Quadro 7 - Modelo de primeira ordem em R.....	49
Quadro 8 - Modelo de primeira ordem com interações em R.....	50
Quadro 9 - Modelo de segunda ordem em R	52
Quadro 10 - Grafico de contornos em R.....	52
Quadro 11 - Método do vetor gradiente	53
Quadro 12 - steepest ascent.....	53
Quadro 13 - Valores estimados da Fluidez, em R.....	54
Quadro 14 - Defeito de Enchimento/ características reológicas	75
Quadro 15 - soluções para os defeitos de enchimento.....	75
Quadro 16 - matérias-primas.....	76
Quadro 17 - cálculo da percentagem de humidade das matérias-primas.....	76
Quadro 18 - Registo de observações reais	82
Quadro 19 - Estatísticas descritivas.....	83
Quadro 20 - Testes de normalidade	83
Quadro 21 – Observações Codificadas (1)	85
Quadro 22 - Cálculo dos coeficientes de regressão linear para “densidade”	86
Quadro 23 - Cálculo dos coeficientes de regressão linear para “fluidez”	86
Quadro 24 - Resíduos da regressão linear múltipla “densidade”	88
Quadro 25- Cálculo dos coeficientes de regressão quadrática para “densidade”	96
Quadro 26 - Cálculo dos coeficientes de regressão quadrática para “fluidez” ...	96
Quadro 27 - Valores previstos para a variável “densidade”	99
Quadro 28 - Valores previstos para a variável “Fluidez”	99

1. Introdução

Com este trabalho foi desenvolvido um estudo aprofundado acerca da Metodologia de Superfícies de Resposta (MSR) no âmbito da Estatística Experimental e do Planeamento de Experiências. A MSR consiste num potente aglomerado de técnicas da área da Estatística e da Matemática que visam a optimização, sendo portanto de crucial importância explorar a importância da sua aplicabilidade na análise e desenvolvimento de Processos Industriais e nas mais variadas áreas. Num mundo onde impera a competitividade e o recurso às mais altas tecnologias é imperativo recorrer a técnicas sofisticadas mas de interpretação fácil, nomeadamente as técnicas de visualização gráfica. Recorrendo a um processo iterativo e sequencial a MSR foi explorada no âmbito de uma aplicação real, um caso prático no âmbito da Indústria Cerâmica, onde foi dada ênfase aos modelos de primeira e segunda ordem.

O trabalho começa por apresentar alguns conceitos e noções introdutórias, seguindo-se uma breve resenha histórica acerca da MSR, acompanhada de uma revisão de literatura. O desenrolar desta abordagem e análise conduz à necessidade de explorar temáticas intrinsecamente relacionadas com a MSR no âmbito do Planeamento de Experiências, que entendemos por bem explorar nos capítulos seguintes. MSR inclui a aplicação da regressão: as suas técnicas são empregues antes, durante e após a análise de regressão - antes, quando a experiência tem de ser desenhada e os valores de entrada têm de ser seleccionados; durante, quando obtemos os respetivos valores experimentais; após a análise de regressão, quando são aplicados certos modelos e procedimentos de teste, nomeadamente a Análise de Variância, testando se a mudança de níveis de um fator ou interação é significativa, e outras técnicas de otimização.

1.1 Conceitos e Noções Introdutórias

O delineamento experimental foi concebido como um método de pesquisa explicativa em que o pesquisador intervém na amostra, impondo deliberadamente os níveis de uma ou mais características explanatórias com o propósito de derivar inferências referentes aos efeitos causais dessas características sobre as respostas da

MSR: Uma Investigação no âmbito da Indústria Cerâmica

unidade experimental: as características explanatórias são denominadas factores e os seus níveis, tratamentos. A Metodologia de Superfícies de Resposta (MSR) é um conjunto de técnicas matemáticas e estatísticas, útil para a modelação e análise de aplicações onde a resposta de interesse é influenciada por muitas variáveis (factores) e o objectivo é otimizar essa resposta (Montgomery 2005).

O “*screening experiment*” é um processo de triagem que visa descobrir quais os factores mais potentes no processo de fabrico de um determinado produto. O objetivo principal da experiência é seleccionar ou *filtrar* alguns dos efeitos principais importantes versus os menos importantes. Importa construir um modelo empírico que relacione as variáveis do projeto com os parâmetros de qualidade do produto e verificar se os factores podem ser convenientemente medidos na monitorização e controlo do produto final. As técnicas de triagem mais comuns são as Técnicas de Taguchi e Plackett-Burman. A MSR, após o “*screening experiment*” na identificação das variáveis mais importantes – factores - e depois de refinada a informação, visa a optimização como objectivo final, determinando os níveis óptimos das variáveis críticas. A representação gráfica das superfícies de resposta pode ser feita por meio de gráficos de superfície 3D e gráficos de contornos (curvas de nível).

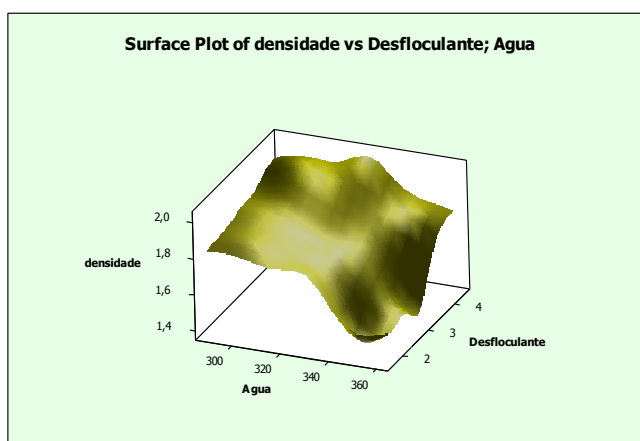


Gráfico 1- Superfície de Resposta 3D

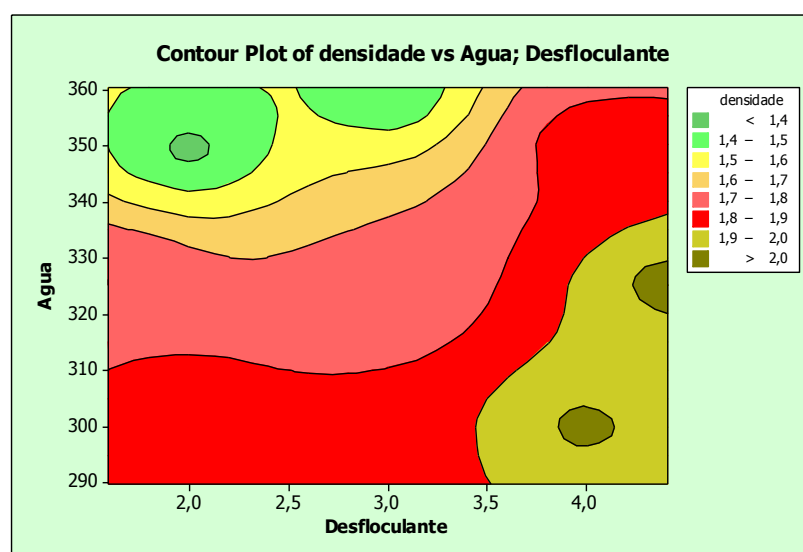


Gráfico 2 - Gráfico de Contornos – Curvas de nível

Normalmente um experimentador não salta diretamente para um problema de otimização: após a triagem, importa localizar a área potencialmente mais significativa da região experimental. Por exemplo, na produção de uma determinada pasta cerâmica, decidimos, após triagem, que a quantidade de água associada à moagem e a adição de determinada quantidade de desfloculantes são os fatores relevantes, suscetíveis de experimentação e controlo, que permitem o objetivo de satisfazer necessidades específicas de produção de um determinado artigo cerâmico. Quando as características reológicas da pasta – densidade, fluidez e tixotropia - derivam de uma série de tratamentos, valores contínuos, então a Metodologia de Superfície de Resposta é útil para o desenvolvimento, melhoria e otimização destas variáveis resposta. Na maioria dos problemas de MSR, a forma da relação entre a resposta e as variáveis independentes é desconhecida; assim, a primeira etapa é encontrar uma aproximação adequada para a relação entre elas. Neste caso y , (y_1) densidade, (y_2) fluidez são as variáveis resposta que são funções da (x_1) quantidade de água, (x_2) quantidade de desfloculantes. Esta relação pode ser expressa como: $y = f(x_1, x_2) + \varepsilon$.

As variáveis x_1 , x_2 são variáveis independentes e a resposta y depende delas. A variável dependente y é função de x_1 , x_2 e de ε , o termo do erro experimental. O erro ε representa os erros aleatórios observados na resposta y , desvios na resposta que não resultam da variação dos factores; é aleatório e segue uma distribuição normal de média zero e desvio padrão σ .

O valor esperado da resposta, $E(y)$, é dado por:

$E(y) = f(x_1, x_2) = \pi$. A superfície representada por $\pi = f(x_1, x_2)$ é a chamada **superfície de resposta**.

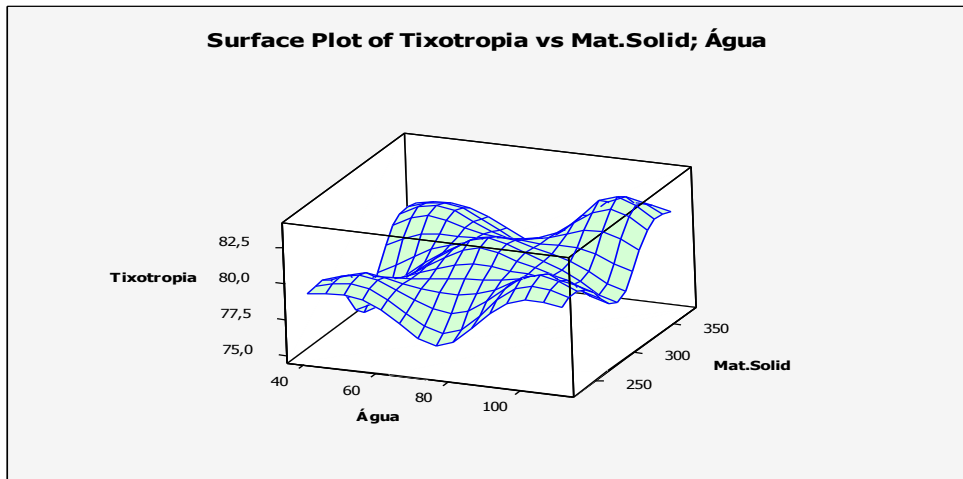


Gráfico 3 - Exemplo de Superfície de Resposta 3D

Na imagem do gráfico 3 observamos a superfície (π) de resposta, representada no gráfico de Superfície de Resposta 3D pelo quadriculado azul.

No gráfico de contornos, gráfico 4, a superfície (π) de resposta encontra-se representada pelas áreas coloridas. A cada par experimental (x_1, x_2), corresponde uma determinada solução, o valor observado da resposta $f(x_1, x_2)$.

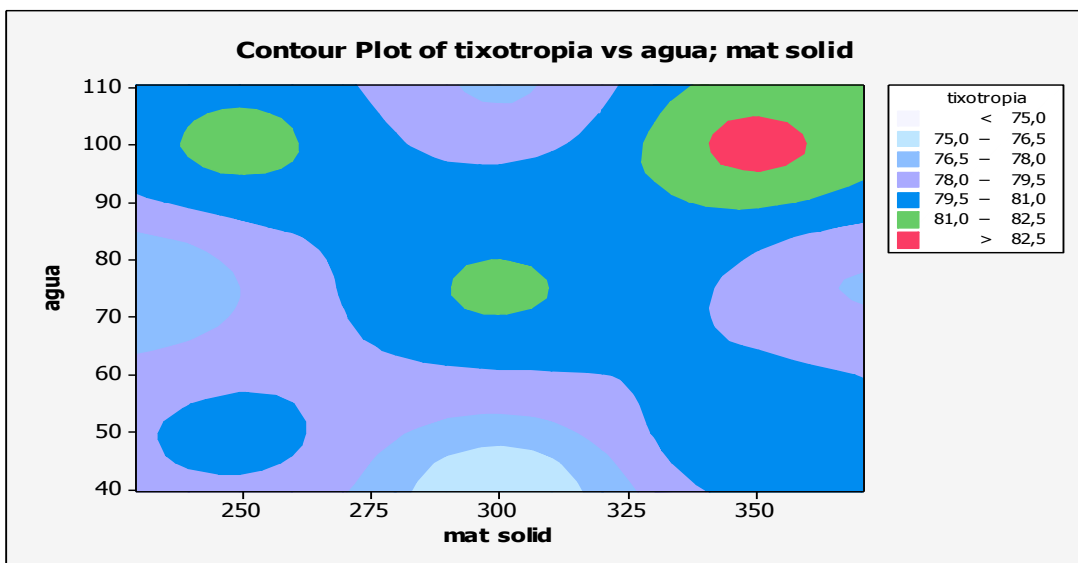


Gráfico 4 - Exemplo de gráfico de contornos

Existem dois tipos principais de arranjos experimentais para a organização e análise de dados em MSR: CCD – *Central Composite Design (Box-Wilson)* e BBD – *Box Behnken Design*. O arranjo central composto (*Box-Wilson Central Composite Design - CCD*) é uma matriz formada por três grupos distintos de elementos experimentais: um factorial completo ou fracionado; um conjunto de pontos centrais e, adicionalmente, um grupo de níveis extras denominados pontos axiais. Se a distância entre o centro do arranjo e ponto factorial (-1, +1) for aproximadamente 1 (em módulo), a distância do centro ao ponto axial será maior que a unidade. Esta distância, normalmente representada por α , depende de certas propriedades desejadas para a experiência e do número de fatores envolvidos. O número de pontos axiais em CCD é igual ao dobro do número de fatores e estes pontos representam os valores extremos. Em função da sua localização, podem ser circunscritos, inscritos ou de face centrada

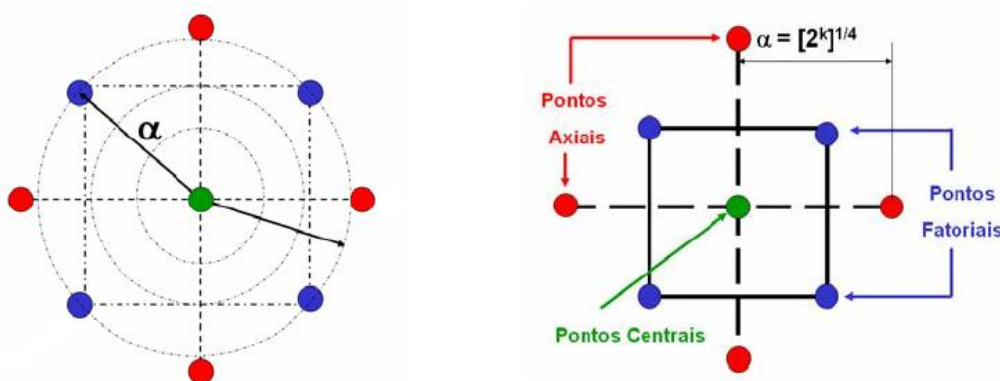


Ilustração 1 - Arranjos CCD

O arranjo circunscrito é o original CCD e requer 5 níveis para cada factor. Os pontos axiais estão a uma distância α do centro e cumprem as propriedades de rotacionalidade e ortogonalidade.

O arranjo de face centrada, à direita, requer 3 níveis para cada variável e é o eleito para o nosso caso de estudo.

Um arranjo do tipo *Box-Behnken* (Box e Behnken, 1960) é um modelo quadrático independente que não contem fatoriais completos ou fracionários embutidos. Neste tipo de arranjo, as combinações de fatores são os pontos médios das arestas das faces do espaço experimental, Ω , além dos pontos centrais, como se pode ver na figura

MSR: Uma Investigação no âmbito da Indústria Cerâmica

seguinte. Estes arranjos têm a propriedade da rotacionalidade em 3 níveis para cada fator.

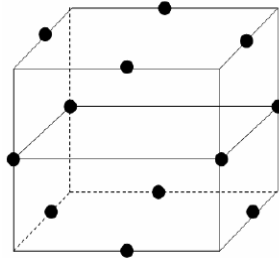


Ilustração 2 - Arranjo de Box-Behnken para 3 factores

A internet veio facilitar e acelerar o intercâmbio cultural e científico: o conhecimento e aplicação da metodologia (MSR) sofreu um incremento vertiginoso. Efetivamente o ótimo, máximo ou mínimo, é o paradigma da sociedade atual em toda e qualquer área.

2. Resenha histórica e revisão da literatura

Na incessante busca por desvendar os segredos do universo, o homem começa por aderir ao culto das forças da natureza como forma de conhecimento. Na passagem da época primitiva para a antiguidade, o homem alarga os horizontes, passando das explicações mitológicas para explicações de natureza religiosa. Segue-se a etapa pelo caminho da razão: com o aparecimento da filosofia, a explicação da natureza passa a ser eminentemente racional. A eterna busca da verdade conduz à passagem da filosofia para a ciência através da racionalização da abordagem empírica, da explicação das relações causais entre fenómenos por rigorosos processos analíticos e racionais.

O uso do método científico na implementação dos delineamentos experimentais surge associado aos fundamentos da moderna estatística, teórica e experimental, iniciada a partir do sec. XX e deve-se em muito a Ronald Fisher. A “Rothamstead *Experimental Station*”, fundada em 1843 em Londres, foi o Instituto pioneiro, em termos mundiais, da pesquisa agro-pecuária: por volta de 1920 Fisher assume a chefia do Departamento de Estatística deste Instituto e com o suporte dos dados acumulados das “experiências clássicas”, cria novas técnicas e conceitos que se tornam incontornáveis na moderna Estatística. Fisher lançou as bases da inferência estatística e delineou muitos métodos originais para a resolução dos vários problemas encontrados em Rothamsted e em outras instituições de pesquisa. Introduziu diversas técnicas de análise de dados, como a análise da variação, que passou a ser amplamente utilizada na análise estatística de dados de experiências, e a técnica de polinómios ortogonais para o uso de características ambientais.

As pesquisas científicas de natureza experimental exigiam tratamento adequado às pequenas amostras, com objetivo inferencial, como revelaram os trabalhos pioneiros de Gosset, conhecido pelo pseudónimo de Student. Esses trabalhos foram continuados no mais alto nível teórico por R. A. Fisher, a figura mais representativa da Fase da Experimentação. As contribuições de Fisher para a estatística começaram com a resolução do problema da estimação pontual. Coube-lhe o mérito de ter criado o conceito de verossimilhança, diferente de probabilidade, e ter desenvolvido o Método de Máxima Verossimilhança, aplicando-o na estimação de parâmetros, inclusive na Genética, onde encontrou campo fértil de aplicação.

MSR: Uma Investigação no âmbito da Indústria Cerâmica

Statistical Methods for Research Workers (1925) e *The Design of Experiments* (1935) são duas obras de Fisher, onde este explica as suas ideias sobre os métodos estatísticos que perspectiva como um todo: Análise de Variância e Delineamentos Experimentais. Sir Ronald A. Fisher foi inovador no uso dos métodos estatísticos e da análise de dados; após o trabalho na Estação Experimental de Rothamsted, onde era feita pesquisa agrícola nomeadamente ensaios com fertilizantes químicos, em 1933 leva o seu trabalho para a Universidade de Londres, onde o DOE (do inglês *Design of Experiments*) foi aplicado e obteve enorme sucesso. Esta metodologia, desenvolvida entre 1920 e 1930 por Fisher, foi posteriormente aperfeiçoada por outros importantes pesquisadores como Box, Hunter e Taguchi, dentre outros. A primeira aplicação prática do DOE data de 1930, na British Textile.

Em 1925 desenvolve a análise de variância ANOVA, incidindo na estimação das componentes da variância e nos delineamentos experimentais, realçando o papel da repetição, da casualização e do controlo local na eficiência das experiências. Criou os delineamentos em blocos completos casualizados e, mais tarde e para melhor controlar a heterogeneidade experimental, criou os delineamentos em blocos incompletos (ou látice). As principais características da abordagem introduzida por Fisher são as seguintes: - Cada experiência deve fornecer uma estimativa da variabilidade atribuível a características estranhas a que estão sujeitas as respostas a quaisquer tratamentos: erro aleatório (ϵ); - Uso da repetição para validar a estimativa da variabilidade e de casualização; - Controle local com o propósito de controlar e reduzir a fonte de variação estranha; - Princípio de que a análise estatística dos resultados é determinada pelo modo como a experiência é conduzida (casualização); - Conceito de experiência fatorial, ou seja, da pesquisa dos efeitos de dois ou mais fatores numa única unidade experimental, em alternativa a experiências separadas para cada característica.

A partir das contribuições relevantes de Fisher, muitos delineamentos experimentais foram desenvolvidos para as mais variadas situações. Esses delineamentos têm sido extensivamente utilizados e tornaram-se tradicionais em muitas áreas de pesquisa. Muito se evoluiu desde então devido essencialmente à facilidade de cálculo, proporcionada pela computação. Estes recursos tornaram viável a utilização do delineamento experimental mais apropriado para cada situação particular e, conseqüentemente, um planeamento de experiências mais eficiente.

MSR: Uma Investigação no âmbito da Indústria Cerâmica

A Metodologia da Superfície de Resposta, enquadrada nos delineamentos experimentais, revela-se um método importante na conceção, formulação, desenvolvimento e análise de produtos científicos; um método igualmente eficaz na melhoria de estudos e produtos já existentes, sendo usada em muitos sectores industriais onde há espaço para o desenvolvimento e otimização de processos de fabricação, mesmo no sector primário onde acontece a experimentação agrícola e na indústria extrativa onde se busca a racionalização dos processos. As aplicações mais comuns da MSR centram-se na produção industrial nomeadamente no campo das Ciências Químicas, Biológicas e Clínicas, Ciências Sociais, Ciência dos Alimentos, Física e Ciências da Engenharia. O método tem uma extensa aplicação no mundo real, sendo por isso importante saber como e onde a Metodologia de Superfícies de Resposta começou e o seu historial. Segundo Hill e Hunter (Wikipedia, 2006) a Metodologia das Superfícies de Resposta foi introduzida por GEP Box e KB Wilson em 1951 quando propuseram o uso de polinómios do primeiro grau para aproximar a variável resposta. A ideia principal da MSR é usar uma sequência de experiências projetadas para obter uma resposta ótima: Box e Wilson sugeriram ainda o uso de um modelo polinomial do segundo grau para o efeito. Eles reconhecem que este modelo é apenas uma aproximação e não uma certeza, mas usam-no porque tal modelo é fácil de estimar e aplicar, mesmo quando pouco se sabe sobre o processo; mas, de acordo com Meal e Pike(1989) a origem da MSR remonta a 1930 com o uso das *Curvas de Resposta*.

A metodologia RSM surgiu a partir de estudos relativos a problemas levantados por Hotelling (1941) e Friedman e Savage (1947), que buscavam, nas suas pesquisas, a realização de experiências para a obtenção de condições ótimas. Nesta pesquisa é realçada a contribuição importante advinda de Hartley (1959), que prestou um enorme contributo à eficiência e economicidade do modelo e o de muitos outros cujos testemunhos se encontram disseminados na literatura e no terreno.

Destaca-se ainda o papel determinante da Internet nos atuais desenvolvimentos dos modelos que usam polinómios de nível três, de que ressaltam os estudos sobre os arranjos fatoriais fracionados de Connor e Zelen (1959). Quanto aos modelos de segunda ordem, o seu uso advém dos conhecimentos práticos dos arranjos centrais compostos (CCDs) de *Box-Wilson (1951)* e dos arranjos de três níveis (BBD) de *Box-Behnken (1960)*.

MSR: Uma Investigação no âmbito da Indústria Cerâmica

Os CCD são, de longe os desenhos mais populares entre os utentes de MSR. Durante os anos 70, num período em que o interesse pelo uso prático de MSR não era grande, houve um interesse considerável no desenho D-Optimal. Para além deste, há muitas outras classes de desenhos MSR que foram propostos durante os anos 70 e 80, contudo muitos destes foram pouco usados entre os entusiastas de MSR, facto que se deve, sobretudo, a uma generalizada falta de conhecimento da sua existência. Exemplos de algumas destas classes de desenhos são o desenho Box-Draper (1974), desenhos Hoke (1974), desenhos atribuídos a Roquemore (1976) e os atribuídos a Notz (1982). Estes últimos, também chamados desenhos híbridos, são usados numa extensa panóplia de indústrias em problemas de superfície de resposta.

Inúmeros trabalhos relacionados com a metodologia MSR têm aparecido nos tempos mais recentes, grande parte produzidos por Box, juntamente com seus colaboradores. Dentre os trabalhos que abordam esse tema estão os de Hinkelmann e Kempthorne (1994), Grosh e Rao (1996), Mason, Gunst e Hess (2003), Montgomery e Myers (2005), Montgomery e Anderson-Cook (2009), cada um com abordagens específicas mostrando a viabilidade do emprego da metodologia

Nos anos 80, o sucesso dos engenheiros e cientistas Japoneses alcançado pelo uso prático do desenho experimental e pela determinação das condições de optimalidade, vieram reanimar o interesse por MSR. O trabalho de Genichi Taguchi (1980) foi tanto polémico como determinante para este avanço.

Os últimos avanços da MSR aparecem ligados à arte e ciência da planificação que engloba quer a capacidade de integrar todo o conhecimento não estatístico que possa ser útil à definição dos objetivos e a uma escolha experimental dos fatores e das variáveis resposta, quer os aspetos logísticos do decurso duma experiência, etc. Alguns artigos da literatura que versam este tema devem-se a Hann (1977,1984), Bishop, Petersen e Trayser (1982), Hoadley e Kettenring (1990), Coleman e Montgomery (1993).

Nesta última década, embora não se tenha assistido a desenvolvimentos teóricos significativos, constatamos uma enorme massificação da aplicação da metodologia. Hoje a MSR é comumente usada na generalidade da indústria e nos centros de investigação. As exigências deste século não se compadecem com o bom, exigem o ótimo em cada circunstância; indústrias como a nanotecnologia exigem o rigor e

MSR: Uma Investigação no âmbito da Indústria Cerâmica

precisão que esta metodologia permite. A MSR é encarada como um processo completo, global, de planificação de experiências produtivas nos vários setores da atividade económica; um meio de controlo contínuo de garantia de conformidade aos objetivos de qualidade fixados. As inovações pontuais surgem ligadas à evolução matemático-estatística da otimização em geral, por um lado, e ao rigoroso controlo da variância dos fatores, por outro. Novos desenhos amostrais subordinados a novos e mais expeditos critérios, incluindo a rotacionalidade para experiências com respostas múltiplas, geram o sucesso da metodologia MSR.

A literatura é rica em metodologia para determinar níveis do desenho das variáveis que maximizam (minimizam) **uma** resposta. Diversos critérios têm sido propostos para determinar condições de compromisso no caso de **várias respostas** como podemos observar através duma resenha histórica por Myers, Khuri e Carter (1989).

Até aos anos recentes, foram muito raros os packages que tratam estes procedimentos de análise MSR (procedimentos que usualmente não constavam dos pacotes de software estatístico standardizados). O aparecimento de alguns pacotes específicos de DOE e MSR vieram permitir a agilização dos processos e facilitar a simulação.

No que concerne à localização espacial do uso da MRS, inicialmente aplicada na indústria química com um considerável sucesso, observamos que, durante os últimos 20 anos, esta encontrou uma aplicação extensiva numa enorme variedade de indústrias, bem longe das suas origens no processo químico, destacando-se a sua importância em áreas que incluem os semicondutores e produtos eletrónicos, maquinaria, corte de ferro, processos de fusão, nanotecnologia, otimização para análise numérica e na indústria automóvel, entre outros. Alguns dos desenvolvimentos mais recentes podem ser lidos em Myers e Montgomery.

Uma busca de MSR na Internet, revela a pujança de publicações, jornais e revistas científicas, e da enorme quantidade de artigos científicos publicados. Algumas áreas laboratoriais e industriais onde a metodologia MSR é aplicada, mereceram uma atenção especial e uma leitura mais atenta na medida em que revelam quer a renovada atualidade do método, quer a pertinência do uso, quer ainda a heterogeneidade da sua aplicação. Destas destacamos:

MSR: Uma Investigação no âmbito da Indústria Cerâmica

- A “*modelagem e otimização de uma variedade de água e processos de tratamento de águas residuais*”, uma aplicação onde são revistos os princípios e as etapas teóricas da MSR na coagulação - floculação , adsorção, processos oxidativos avançados, processos de electro -químicas e desinfecção (IWA – Publishing Online:2013);

- “*A otimização de Cr (VI) a redução e remoção por eletrocoagulação utilizando a metodologia de superfície de resposta*”, onde se investigam os efeitos de diferentes condições de operação na remoção de cromo hexavalente (Cr (VI)), pela eletrocoagulação com elétrodos de aço inoxidável (Journal of Hazardous Materials:2009);

- “*Otimização de processo de coagulação-floculação para efluentes Palm Oil Mill utilizando Metodologia de Superfície de Resposta*”, onde é proposta uma nova abordagem para a fábrica de óleo de palma de efluentes (POME) tratamento, bem como a recuperação e reutilização da água (Environmental Science & technology:2005);

- “*Um método numérico para a construção da frente de Pareto de problemas multi-objetivos de otimização*”, o método proposto é aplicado a quatro problemas de teste que ilustram as dificuldades específicas encontradas em problemas de otimização multi-objetivos (Jornal de Matemática Aplicada e Computacional:2014);

- “*Otimização do projeto robusto de amortecedores de fricção para o controle de resposta estrutural*”, onde se propõe a otimização do projeto robusto de amortecedores de fricção para controlar a resposta estrutural contra terremotos (Wiley Online Library:2014);

- “*Otimização da produção heterogénea de biodiesel a partir de óleo de palma de uso culinário através da metodologia de superfície de resposta*” a metodologia foi usada para estudar as relações de metanol e o rácio de óleo molar, o tempo de reação e temperatura de reação no rendimento do éster metílico e na conversão de ácidos gordos livres, uma experiência CCD onde foi aplicado o fatorial completo 2^4 e usados dois pontos centrais (Biomassa e Bioenergia, Elsevier:2010);

- “*Otimização da extração reativa de mamona para produção de biodiesel utilizando a metodologia de superfície de resposta*”, um projeto composto central (CCD) para estudar os efeitos da concentração catalisadora de metanol para o rácio de óleo molar, a temperatura de reação e a intensidade da mistura no rendimento de

FAME. O biodiesel foi produzido por extracção reactiva (RE) de sementes de rícino para reduzir o custo e o tempo associado com o método convencional (Fuel, Elsevier:2012);

- *“Otimização de um Secador de gasóleo de vácuo usando Metodologia de Superfície de Resposta”*. A eficiência da desidratação e perda de gasóleo neste tipo de secador de vácuo são afetados significativamente pelo tempo de uso, temperatura e pressão, como variáveis independentes. Para concretizar a experiência foi delineado um desenho CCD com três níveis destas variáveis, Narimani e Kadijani (2014);

- *“Otimização da succinilação de amido de milho, utilizando a metodologia de superfície de resposta”*. Foi escolhido um método de regressão multi-linear para ajustar a relação entre a variável dependente (y , o grau de substituição (DS)) e os quatro regressores (x_i), ou seja, a concentração de bicarbonato de sódio (x_1), o conteúdo da reacção humidade sistema (x_2), a temperatura de reacção (x_3) e tempo (x_4), em Wu et al (2013).

- *“Aplicações de Desenho Experimental e Método de Superfície de Resposta em Probabilidade: Avaliação de Recursos Geotérmicos”*. Havendo limitadas informações publicadas sobre avaliações probabilísticas de perspectivas de recursos geotérmicos, este estudo pretende estimar a quantidade da capacidade de produção de eletricidade a partir de um sistema sintético de energia geotérmica. A variável resposta (capacidade de geração eléctrica) é função de parâmetros principais incertos, Quinao e Zarrouk (2014).

- *“Lixiviação de poligalacturonases obtidas pela fermentação semissólida da casca e albedo do maracujá-amarelo”*, Sousa et al (2012) artigo publicado na revista brasileira de engenharia agrícola e ambiental em 2012, da autoria de Carlos Sousa e outros, onde se pretendeu otimizar a etapa de recuperação das enzimas, sendo os parâmetros da lixiviação-agitação, tempo de contato e relação solvente/substrato, investigados num planeamento experimental fatorial 2^3 ;

- *“Efeito da Força e da Velocidade de Compressão sobre as Propriedades de Comprimidos contendo alta Concentração de Extrato Seco Vegetal”*, Soares et al (2013)um trabalho original de Luís Soares, Peter Schmidt, George Ortega e Pedro Petrovick, publicado na ata farmacêutica Bonaerense em 2003 que estuda a influência da força e da velocidade de compressão sobre as propriedades de dureza, tempo de desintegração e friabilidade;

MSR: Uma Investigação no âmbito da Indústria Cerâmica

- *“Desenvolvimento e aplicação de um sensor voltamétrico para análise simultânea de dopamina, ácido ascórbico e ácido úrico empregando um eletrodo modificado com naftoquinona adsorvida sobre nanotubos de carbono”* Luz et al (2013) de Rita Luz e outros, publicado em 2013, onde se descreve o desenvolvimento de um sensor voltamétrico seletivo e sensível para dopamina, ácido ascórbico e ácido úrico, tendo os processos de oxidação destes elementos sido investigados por voltametria cíclica e voltametria de pulso diferencial, concluindo que o eletrodo modificado apresentava excelente atividade catalítica para oxidação e fornecia picos bem definidos que levaram à otimização dos parâmetros experimentais.

- *“Determinação Experimental e Predição dos Limites de Inflamabilidade do Etanol Anidro e Hidratado para Uso na Indústria Aeronáutica”*, Cartagena, J.(2013) dissertação de Juan Cartagena em 2013, onde importa a determinação dos limites de inflamabilidade para o etanol hidratado e para o etanol anidro, os quais podem ser um substituto total ou parcial do combustível tradicional na aviação comercial.

3. DOE – Delineamento de experiências

O planeamento de experiências e a experiência em si, foram concebidos como um método de pesquisa explicativa em que o pesquisador intervém na amostra, impondo deliberadamente os níveis de uma ou mais características explanatórias com o propósito de derivar inferências referentes aos efeitos causais dessas características sobre as características respostas da unidade experimental. As características explanatórias são denominadas factores e os seus níveis, tratamentos. Exemplos comuns de tratamentos são os diferentes estímulos ministrados a animais, plantas ou materiais, tais como diferentes dietas alimentares ou diferentes fungicidas, diferentes quantidades de água e de desfloculante para o fabrico de uma determinada pasta cerâmica.

Num estágio inicial, o objetivo de uma experiência pode ser, simplesmente, descobrir se os estímulos produzem qualquer resposta mensurável, enquanto num estágio de pesquisa posterior pode ser o de confirmar ou refutar certas conjeturas que tenham sido formuladas sobre as respostas aos tratamentos. Uma pesquisa científica completa procede através de estágios:

- Identificação e estabelecimento do problema científico ou problema de pesquisa.
- Formulação da hipótese científica ou hipótese de pesquisa.
- Revisão da literatura.
- Construção do plano de pesquisa.
- Colheita dos dados.
- Análise e interpretação dos dados.
- Conclusões, que pode conduzir à confirmação ou rejeição da hipótese original, e confirmação ou questionamento de resultados de outras pesquisas.
- Apresentação dos resultados através de relatório, e difusão desses resultados.

A identificação e estabelecimento do problema e a formulação da hipótese definem os objetivos da experiência. Estas decisões exigem uma clara e inequívoca caracterização quer do propósito da investigação quer da população alvo e das suas unidades.

MSR: Uma Investigação no âmbito da Indústria Cerâmica

A planificação da experiência deve ser consistente com os objectivos da pesquisa, garantindo o planeamento do desenho amostral, das acções, procedimentos e timings e também a definição do modelo estatístico adequado.

Na etapa da colheita de dados há que assegurar obediência à planificação estabelecida e salvaguardar a exatidão dos dados recolhidos. A análise dos dados é usualmente feita com recurso a packages estatísticos – R, SPSS, Minitab, entre outros; os pressupostos de aplicação dos testes a promover devem ser cuidadosamente verificados e há que ter o cuidado de inferir somente no âmbito do estudo. Finalmente, a análise de dados origina resultados donde devem ser extraídas conclusões e recomendações.

O objetivo, o delineamento e a análise estatística estão intrinsecamente ligados. Embora o planeamento para a consecução dos objetivos da experiência possa levar a mais que um delineamento possível, há que seleccionar o que associa a maior eficiência à melhor prática.

Um fator experimental é uma característica explanatória cujos níveis, na amostra, têm as seguintes propriedades: são escolhidos e definidos no plano experimental para cada unidade da amostra; compreendem um pequeno conjunto de níveis que se repetem nessas unidades; classificam as unidades em grupos que constituem uma partição da amostra com a propriedade de cada unidade pertencer a um grupo e nenhuma unidade pertencer a mais de um grupo; as relações com os níveis dos demais fatores experimentais constituem uma estrutura significativa que determina a estrutura da experiência ou delineamento experimental.

Cada nível específico de um fator de tratamento definido para a amostra é um tratamento; no caso da experiência com dois ou mais fatores experimentais de tratamento, cada combinação distinta dos níveis desses fatores, definida para a amostra, também é um tratamento. Os níveis de um fator experimental e as combinações distintas dos níveis de dois ou mais fatores experimentais definidos para a amostra são denominados condições experimentais. Unidades experimentais distintas com uma mesma condição experimental constituem **repetições** dessa condição experimental.

Para estabelecer a existência de efeito causal de um fator experimental sobre uma variável resposta são requeridas três condições: a) o fator experimental e a característica resposta variam conjuntamente; b) o fator experimental precede a variável

resposta; c) não há explicação plausível para a variação da característica resposta que não inclua efeito do fator experimental.

A primeira condição é estabelecida por procedimentos de inferência estatística. A segunda ocorre por definição: níveis do factor experimental são aplicados ou manifestam-se nas unidades experimentais previamente à mensuração das características respostas. A questão crítica é a terceira condição, ou seja, a garantia de que a variação dos valores observados da variável resposta em unidades com níveis diferentes de fatores experimentais, não possa ser atribuível exclusivamente a características estranhas.

O erro experimental (global) é constituído pelo conjunto dos valores da variável resposta, nas unidades de observação, que é atribuível a características estranhas. Esse erro experimental compreende diversas fontes de variação estranha nomeadamente a variação inerente ao processo de medida e ao registro dos dados; a variação devida à reprodução não uniforme das condições experimentais; as interações das condições experimentais e as características estranhas com outras fontes de variação relevantes, fora do controle do pesquisador. Importa controlar o erro experimental, diminuindo a confusão entre os efeitos dos fatores sobre as variáveis resposta e os efeitos de características estranhas já que as diferenças reveladas pelas variáveis resposta devem ser atribuídas aos fatores experimentais de modo tão inequívoco quanto possível. Este controlo deve ser exercido a vários níveis: controlo das técnicas experimentais, controlo local, controlo estatístico e casualização.

O controlo local consiste na classificação das unidades de observação segundo os níveis de uma ou mais características estranhas relevantes e na afetação dos tratamentos às unidades experimentais de modo que os efeitos dessas características estranhas não fiquem misturados com efeitos importantes dos fatores experimentais e sejam separados do erro experimental que afeta esses efeitos.

A casualização na atribuição dos tratamentos é o processo de afetação aleatória dos níveis de fatores de tratamento às respetivas unidades experimentais através dum procedimento objetivo de sorteio que atribua a todas as unidades experimentais, para cada um desses fatores de tratamento, a mesma chance de receber qualquer dos correspondentes tratamentos.

MSR: Uma Investigação no âmbito da Indústria Cerâmica

A sensibilidade da experiência para detetar efeitos de fatores experimentais que se manifestem na amostra está relacionada com a precisão. Esta depende da grandeza do erro que afeta os efeitos de fatores experimentais e é medida pela variância do erro experimental, que afeta os efeitos dos fatores experimentais. Normalmente a variância populacional σ^2 não é conhecida e no seu lugar é utilizada uma estimativa s^2 , um valor apropriado para aproximá-la, determinado a partir dos valores observados na amostra. Essa estimativa da variância do erro experimental por unidade é a variação dos valores observados da variável resposta devida a características estranhas dividida pelo correspondente número de graus de liberdade. Fisher define a precisão ou quantidade de informação fornecida pela experiência como o inverso da variância do erro experimental.

Para além da repetição, do controlo local, da casualização, do confundimento e da precisão, o delineamento experimental subordina-se a outros princípios básicos: ortogonalidade ou balanceamento e eficiência. Um delineamento experimental é ortogonal quando todos os pares de fatores são ortogonais: a condição necessária e suficiente para que dois fatores sejam ortogonais é que a proporção do número de unidades de observação, para as combinações de níveis correspondentes a cada nível particular de um dos fatores, seja constante para todos os níveis do outro fator. Os delineamentos equilibrados são apropriados para situações em que se torna inconveniente ou impossível o uso de delineamentos ortogonais: um delineamento equilibrado se provê igual precisão para todas as comparações simples de tratamentos, ou seja, para todas as comparações individuais de tratamentos.

Um delineamento é mais eficiente do que outro delineamento se a quantidade de informação que ele provê é maior do que a provida por este delineamento alternativo. Regra geral, o pesquisador deve escolher o delineamento experimental que faculte o máximo de informação relevante aos objetivos da experiência, com o custo mínimo.

O procedimento para inferências referentes a um efeito de fatores experimentais consiste, basicamente, em comparar a componente da variação dos valores observados da variável resposta (que inclua esse efeito em sua origem) com um outro componente que expresse os mesmos efeitos, mas não subordinado a fatores experimentais. Se o primeiro componente da variação se revelar maior do que o segundo por diferença que não possa ser originada apenas de variação aleatória, confirma-se a existência real do

efeito de fatores experimentais. Uma experiência é válida quando as inferências derivadas da amostra para a população alvo são válidas, ou seja, não enviesadas nem tendenciosas.

A partir das contribuições relevantes de Fisher, muitos delineamentos experimentais foram desenvolvidos para as mais variadas situações. Esses delineamentos têm sido extensivamente utilizados e tornaram-se tradicionais em muitas áreas de pesquisa. Muito se evoluiu desde então devido essencialmente à facilidade de cálculo, proporcionada pela computação. Estes recursos tornaram viável a utilização do delineamento experimental mais apropriado para cada situação particular e, conseqüentemente, um planeamento de experiências mais eficiente.

Efetivamente o século XX é o século da especialização. Segundo Douglas Montgomery, podem observar-se quatro importantes fases no desenvolvimento da estatística que têm o planeamento de experiências como denominador comum.

A *primeira fase*, já citada, está relacionada com testes e experiências feitas, sobretudo na agricultura, por **Sir Ronald S. Fisher** (1890-1962), quando cria os princípios do planeamento experimental: aleatoriedade, replicação e partição em blocos. A enumeração destes princípios foi realizada a partir de um importante conceito, também desenvolvido por Fisher, denominado “Análise de Variância” (ANOVA).

A *segunda fase* é denominada de “Industrial”. Nesta altura, o planeamento de testes e experiências incidia, sobretudo, no desenvolvimento industrial do pós-guerra. Ao contrário das experiências realizadas no âmbito agrícola, as experiências de carácter industrial eram caracterizadas, essencialmente pela observação imediata de qualquer variável responsável e pela possibilidade, de o investigador poder retirar importantes informações após um pequeno número de testes. Um dos grandes investigadores desta época é Abraham Wald (1902-1950) para quem as experiências são efetuadas para servir de base à acção – teoria da decisão. Outro grande nome é C. R. Rao, que desenvolveu a aplicabilidade de matrizes ortogonais no planeamento de experiências, conceito amplamente aplicado na fase seguinte, do desenvolvimento da estatística moderna.

Montgomery considera que a *terceira fase* é indissociável do nome Taguchi. Genichi Taguchi é um engenheiro e estatístico japonês que criou um método para a optimização do processo de produção industrial. O final da 2ª Guerra Mundial deixou o

Japão com a sua indústria praticamente inoperante. Com o desenvolvimento industrial japonês, as décadas de 50 e 60, surge a necessidade de assegurar competitividade. E uma das formas de o conseguir foi através da redução e controlo da variabilidade do processo produtivo. No essencial, da metodologia de Taguchi consiste em tomar algumas variáveis (ou factores) do processo e realizar uma experiência com o processo de produção para que sejam determinados novos níveis de operação dos factores controláveis envolvidos, e em que a variabilidade do processo seja a menor possível. A experiência permite, também, definir quais os factores influentes no processo e ajustá-los de tal forma, que a sua performance seja a melhor possível dentro das especificidades do processo. A consequência desta forma de otimizar todo o processo produtivo, desde o planeamento até à produção, foi o de criar um enorme interesse (ou interesse renovado) no planeamento estatístico em investigadores e industriais.

O ensino superior da estatística só começou a ser lecionado de forma generalizada pelas importantes Universidades mundiais após o uso mais abrangente do método de Taguchi. Esta era é considerada a quarta fase do desenvolvimento da estatística moderna.

Pensar de maneira sistemática sobre um problema contribui significativamente para a sua solução. Segundo Montgomery (2001), uma maneira sistemática de avaliar a grandeza de várias fontes de variação que influenciam um processo deve iniciar-se com a identificação e seleção dos fatores que possam contribuir para a variação. Uma contribuição para a escolha apropriada destes fatores passa pela construção de fluxogramas e diagramas de causa-efeito, realizados por especialistas no processo. Em seguida, deve selecionar-se um modelo que inclua os fatores escolhidos e planejar experiências eficientes para estimar os seus efeitos. A novidade introduzida por Fisher no planeamento experimental foi o princípio da casualização, da aleatoriedade, que garantia a validade da estimativa do erro e possibilitava a aplicação dos testes de significância para verificar o efeito dos tratamentos; pela aleatoriedade (um mecanismo objectivo de sorteio) nenhum tratamento seria continuamente favorecido ou desfavorecido nas sucessivas repetições por qualquer fonte estranha de variação. Era absolutamente necessário garantir a aleatoriedade para que as variações que contribuem para o erro experimental fossem convertidas em variáveis aleatórias. Realizar as experiências de acordo com o planeado, detetar, documentar e analisar as anormalidades

que ocorrerem na condução experimental, é essencial para garantir o sucesso do delineamento. Realizadas as experiências, o passo seguinte consiste em estimar os efeitos dos fatores incluídos no delineamento, usando métodos estatísticos adequados e, no final, importa interpretar e discutir os resultados, recomendando as melhorias possíveis.

Um bom planeamento de experiências deve contemplar a utilização de réplicas, de aleatorização e partição em blocos. Ao repetir-se várias vezes um teste em unidades experimentais diferentes, cria-se uma variação para a variável de resposta que pode ser utilizada para se avaliar a significância estatística do incremento experimental. Quando se aleatoriza uma experiência, aumenta-se sua validade já que se aumenta a chance dos efeitos desconhecidos serem distribuídos através dos níveis dos fatores. A partição em blocos usa-se quando não for possível manter a homogeneidade das condições experimentais. Esta técnica estatística permite avaliar se a falta de homogeneidade interfere nos resultados. Um bloco consubstancia uma parte da experiência completa, presumindo-se que é mais uniforme do que o todo.

A ordem-padrão, utilizada pelos arranjos DOE, gera condições experimentais balanceadas e ortogonais de modo a que os fatores investigados sejam experimentados uniformemente em cada um dos seus níveis.

O arranjo experimental mais comum é o fatorial completo, para o qual o número de experiências é igual ao número de níveis experimentais, elevado ao número de fatores. No caso típico de fatoriais em dois níveis, o número de experiências (N) para avaliar os k fatores é dado por $N = 2^k$. Os arranjos fatoriais completos podem ser gerados para qualquer quantidade de fatores e os níveis de cada fator alternam-se nas colunas segundo uma mesma ordem tal que para a primeira coluna, os níveis se alteram a cada experiência 2^0 ; para a segunda coluna, os níveis alteram-se a cada 2^1 ; para a terceira coluna, os níveis alteram-se a cada 2^2 e assim por diante. O procedimento repete-se para tantas colunas quantos os k fatores, até à k-ésima coluna. Os Fatoriais Completos cobrem todo o espaço experimental; porém, enquanto o número de fatores cresce linearmente, o número de experiências cresce exponencialmente, tornando inviável, a partir de dado limite, o espaço experimental.

“Havendo pouco interesse nas interações, estas podem ser negligenciadas” Montgomery e Runger (2003). Neste caso seria suficiente um menor número de

experiências para avaliar os efeitos principais ou as interações de baixa ordem e, desta forma, bastariam frações da experiência completa para detetar a presença de fatores influentes. Por exemplo, um arranjo 2^{3-1} possui um fatorial completo 2^2 e a coluna representativa do fator remanescente iguala-se aos fatores que fazem parte do fatorial completo 2^2 (Box et al., 1978); assim para três fatores originais A, B e C, toma-se A e B para compor um fatorial completo e admite-se $C=AB$. Esta relação é chamada de gerador de confundimento (Paiva, 2004). Em experiências exploratórias onde muitos fatores são examinados numa quantidade relativamente pequena de experiências, usam-se estes arranjos com o objetivo de se identificarem os (poucos) fatores realmente influentes.

Para testarmos se a mudança de níveis de um fator ou interação é significativa empregamos um teste de hipótese para a média; no caso do planeamento de experiências, este teste é a Análise de Variância (ANOVA). (Montgomery e Runger, 2003).

4. Análise de Variância

A análise de variância é um procedimento estatístico usado para estudar a relação entre uma variável dependente e uma ou mais variáveis independentes. Para a sua utilização correta não é necessário admitir pressupostos sobre a natureza da relação estatística, nem é necessário que as variáveis independentes sejam quantitativas.

As variáveis independentes na análise de variância podem ser qualitativas – sexo, localização geográfica, tratamento, dieta, etc. Pode até dizer-se que são geralmente medidas numa escala nominal ou ordinal. Se as variáveis independentes forem quantitativas, não assumimos qualquer pressuposto quanto à relação entre as variáveis independentes e a variável dependente: o problema da especificação do tipo de função que descreve a relação que se encontra na regressão não se coloca neste caso.

Com base nos elementos observados numa amostra podemos tirar conclusões para um domínio mais vasto, a população, donde a amostra foi recolhida: estamos no âmbito da estatística indutiva, processo que se subdivide em duas áreas: a teoria da estimação e teoria da decisão.

A teoria da estimação tem como objectivo estimar o valor dos parâmetros da população teórica a partir das estimativas amostrais. Na prática o que fazemos é testar se as observações da amostra se ajustam a uma determinada distribuição teórica, usando testes de ajustamento de que destacamos o teste de *Kolmogorov-Smirnov*. Os parâmetros de uma população podem ser estimados pontualmente, estimação pontual (ex. a média amostral); a estimativa por intervalos associa um determinado grau de confiança ao estimador pontual.

A teoria da decisão, através dos testes de hipóteses, tem como objectivo confirmar ou infirmar uma determinada hipótese acerca de um ou mais parâmetros da população a partir de uma ou mais estimativas obtidas nas amostras.

Se a técnica ANOVA geralmente envolve comparações de médias parece inadequado a denominação de análise de variância. Porém a designação de ANOVA é perfeitamente justificada: embora frequentemente se comparem médias, estas comparações são feitas usando estimativas da variância. A ideia de comparar a variação entre grupos com a variação dentro de grupos, como testes de homogeneidade, foi primeiramente estudada por Wilhelm Lexis (1837 – 1914) que criou um critério de

dispersão (o termo variância não era conhecido na época), também chamado razão de Lexis. Contudo, somente graças aos trabalhos de Fisher, essas ideias tiveram maior avanço e pleno desenvolvimento.

Se da análise de variância concluirmos que existe diferença, então poderemos fazer comparações entre pares de médias. Para isso usaremos métodos chamados de comparações múltiplas, que permitem fazer qualquer tipo de comparação das médias envolvidas.

4.1 Terminologia

A análise de variância tem a sua própria terminologia: **Factor** – utiliza-se para designar cada uma das variáveis independentes que estudamos; **Níveis** – os diferentes grupos definidos por um factor; **Modelo unifactorial** – modelo em que estudamos apenas um factor; **Modelos multifactoriais** - modelos onde estudamos o efeito de dois ou mais factores simultaneamente; **Modelo de efeitos fixos** – modelos que envolvem inferências apenas para os níveis específicos do factor incluído no estudo; **Modelo de efeitos casualizados** – modelos em que os níveis selecionados são escolhidos aleatoriamente entre todos os níveis possíveis do factor. Podemos tirar conclusões para outros níveis do factor não representados no estudo. Ambos os modelos são usados para testar hipóteses sobre a igualdade de efeitos dos níveis do factor. Contudo, no modelo de efeitos randomizados, podemos também testar hipóteses e construir intervalos de confiança para as variâncias dos níveis, e por isso é também referido como um modelo de componentes de variância.

4.2 Pressupostos

Os pressupostos subjacentes à construção do modelo de análise de variância:

1. Para cada nível do factor, os valores na população seguem uma distribuição normal; a verificação da normalidade é feita através da utilização dos testes de ajustamento (*Kolmogorov*) ou por recurso ao uso de papel de probabilidade.
2. Estas distribuições têm igual variância (desvio padrão); a homocedasticidade é calculada através da utilização de vários testes, nomeadamente os testes de

Hartley quando os n_i , isto é, os tamanhos amostrais são iguais, de Bartlett, que se aplica em situações de igualdade ou desigualdade dos tamanhos amostrais e Levene.

3. As observações para cada nível do factor são observações aleatórias independentes das observações para os outros níveis do factor. A independência resulta, normalmente, da aleatoriedade do processo.

A verificação dos pressupostos assenta em testes de hipóteses cuja finalidade é desenvolver regras que permitam, uma vez formulada a hipótese, decidir, correndo um certo risco de errar, se essa hipótese é aceitável, face à informação contida nos dados. Seja X uma variável aleatória com função distribuição $F_X(x|\theta)$, sendo θ o parâmetro desconhecido e Θ o espaço do parâmetro; sejam Θ_0 e Θ_1 dois subconjuntos de Θ . Hipótese estatística é qualquer conjectura acerca do parâmetro desconhecido. Os elementos básicos de um teste estatístico são genericamente: Hipótese nula, H_0 ; Hipótese alternativa, H_1 ; Nível de significância, α ; Estatística de teste; Região de rejeição; decisão.

Existem, basicamente, duas metodologias para realizar testes estatísticos: testes paramétricos, que exigem que a forma da distribuição amostral seja conhecida e testes não paramétricos que não exigem à partida o conhecimento da distribuição amostral. Seria assim racional a opção por testes não paramétricos, contudo de um modo geral a potência dos testes paramétricos é superior à dos testes não paramétricos. Efetivamente os testes não paramétricos só devem usar-se quando não é possível validar as condições de aplicação dos testes paramétricos.

Os testes paramétricos exigem a verificação simultânea de duas condições: que a variável dependente tenha distribuição normal; e que as variâncias populacionais sejam homogêneas, quando se comparam duas ou mais populações. Para testar a normalidade e a homogeneidade das variâncias usam-se habitualmente os testes de *Kolmogorov-Smirnov* e o teste de *Levene*, respetivamente.

Os testes não paramétricos, vulgarmente apelidados de “*distribution-free tests*” (testes adistribucionais) não exigem que a variável em estudo tenha distribuição conhecida. Embora se aceite que os testes não paramétricos são menos potentes que os correspondentes paramétricos, a afirmação só colhe quando as amostras são de dimensão grande e igual; para amostras de pequena e diferente dimensão e onde as

variáveis não verificam os pressupostos dos métodos paramétricos, os testes não paramétricos podem revelar-se mais potentes.

Sempre que alguma das condições de aplicabilidade dos testes paramétricos falhe, o estatístico pode e deve recorrer a transformações matemáticas das variáveis de modo a cumprir as exigências dos métodos paramétricos, antes de recorrer aos métodos não paramétricos. A generalidade das revistas científicas só aceita estudos com base em testes paramétricos.

Sendo a distribuição normal, uma distribuição teórica necessária em análise de variância, importa fazer o estudo da normalidade da distribuição dos valores da população em cada nível do factor. A compatibilidade de um conjunto de valores observados com a distribuição normal ou outra é analisada através de testes de ajustamento. Os mais usados: teste do Qui-quadrado e teste de *Kolmogorov*. Quando a nossa amostra é retirada de uma população contínua, recorreremos ao teste de aderência à normal *Kolmogorov-Smirnov*, usado para decidir se a distribuição da variável $F(X)$ provém de uma distribuição específica $F(X_0)$, testando se é ou não normal, com os parâmetros μ e σ : $H_0: X \rightarrow N(\mu, \sigma)$ vs $H_1: X \rightarrow N(\mu, \sigma)$.

Para calcular a estatística de teste ordenamos, por ordem crescente as observações da variável; calculamos em seguida a frequência acumulada de cada observação; calculamos a diferença entre esta e a frequência acumulada que a observação teria se a distribuição de probabilidade fosse normal assim como a mesma diferença relativamente à observação anterior; a estatística de teste é a maior destas duas diferenças – $D = \max\{\text{Max}(|F(x_i) - F_0(x_i)|), \text{Max}(|F(x_{i-1}) - F_0(x_i)|)\}$. O valor crítico está tabelado. Qualquer hipótese está, como vimos, associada a um dado nível de significância. O erro do tipo I, α , usualmente estabelecido em $\alpha = 0,05$; associado a este erro (rejeitar H_0 sendo H_0 verdadeira) está sempre o erro do tipo II, β , de Aceitar H_0 sendo H_0 falsa.

Em ANOVA, se admitimos que as distribuições são normais e têm uma variabilidade constante, as diferenças nas médias vão refletir o efeito dos níveis no factor, e é por essa razão que a análise de variância se refere às médias das observações para cada nível do factor.

O procedimento da análise de variância resume-se a duas etapas: - verificar se as médias dos níveis do factor são iguais, considerando um determinado nível de

significância; - Se as médias forem diferentes, examinarmos o modo como diferem e quais as implicações destas diferenças.

Precisamos de um teste global para averiguar da existência de diferenças nas médias, antes de podermos testar várias combinações de médias para determinar diferenças entre populações específicas. Se não fizermos um teste global, os múltiplos testes para a comparação de pares de médias alterarão a probabilidade de cometer um erro do Tipo I (α), não para cada comparação, mas para a experiência global. Se cada uma destas comparações fosse feita com base em $\alpha = 0,05$, teríamos uma probabilidade de 0,05 de falsamente considerarmos significativa cada comparação, isto é, um erro do Tipo I pode ocorrer 5 vezes de modo diferente. Globalmente, podemos afirmar que teríamos uma probabilidade de aproximadamente 0.25 ($5 \times 0,05$) de declarar uma das comparações incorretamente significativa. O procedimento por ANOVA (*Analysis of Variance*) impede a inflação mencionada do erro Tipo I, porque nos indica se existe realmente alguma diferença entre as médias das populações correspondentes.

Quando a hipótese nula de igualdade das variâncias é rejeitada, torna-se necessário proceder à transformação prévia das observações. As transformações das observações têm em vista a redução ou eliminação da heterocedasticidade. Assim quando as amplitudes amostrais são significativamente diversas é recomendável proceder, entre outras, a uma das seguintes transformações: $\ln Y_{ij}$, $\sqrt{Y_{ij}}$, $\frac{1}{Y_{ij}}$. Estas transformações permitem muitas vezes atingir mais que um objetivo: reposição da normalidade, verificação da homocedasticidade, tornar os efeitos aditivos.

Formulação do modelo:

Consideremos que o factor em estudo tem k níveis e notemos estes níveis pelo índice j ($j = 1, 2, \dots, k$).

N_j – número de observações para o nível j do factor.

Y_{ij} – observação i para o nível j do factor.

Com esta notação, o modelo pode então ser formulado do seguinte modo:

$$Y_{ij} = \mu_j + \varepsilon_{ij}$$

Em que μ_j são os parâmetros e ε_{ij} são variáveis independentes $N(0, \sigma^2)$.

O valor observado no ensaio i do nível j do factor Y_{ij} é a soma de duas componentes: um termo constante μ_j e um termo de erro aleatório ε_{ij} ; como o valor esperado de $E(\varepsilon_{ij})=0$, então $E(Y_{ij}) = \mu_j$; como os μ_j são constantes e a variância de uma

MSR: Uma Investigação no âmbito da Indústria Cerâmica

constante é zero, todas as observações têm a mesma variância: $\sigma^2 (Y_{ij}) = \sigma^2 (\varepsilon_{ij}) = \sigma^2$; como os ε_{ij} se distribuem normalmente, o mesmo acontece aos Y_{ij} – porque são funções lineares de ε_{ij} ; assumimos que os termos de erro são independentes, logo o valor do termo de erro num ensaio não tem qualquer efeito sobre o termo de erro num outro ensaio para o mesmo nível do factor ou para níveis diferentes. Como os ε_{ij} são independentes, também os Y_{ij} o são. Genericamente temos:

Amostra	Dados	Distribuição
1	$Y_{11}, Y_{12}, \dots, Y_{1...}$	$N(\mu_1, \sigma^2)$
2	$Y_{21}, Y_{22}, \dots, Y_{2...}$	$N(\mu_2, \sigma^2)$
3	$Y_{31}, Y_{32}, \dots, Y_{3...}$	$N(\mu_3, \sigma^2)$

Quadro 1 - Organização de dados – distribuição das amostras

Amostra	Média	Variância
1	\bar{Y}_1	S_1^2
2	\bar{Y}_2	S_2^2
3	\bar{Y}_3	S_3^2

Quadro 2 - Organização de dados – descritivas das amostras

Se notarmos por μ_1 , μ_2 e μ_3 as médias das variáveis, podemos escrever as hipóteses, nula e alternativa:

$$H_0 : \mu_1 = \mu_2 = \mu_3 \text{ (as médias são iguais)}$$

$$H_1 : \mu_1, \mu_2, \mu_3 \text{ não são todas iguais.}$$

Em relação a estas duas hipóteses podemos afirmar que pequenas diferenças entre as médias, \bar{Y}_s , podem ser atribuídas ao acaso; sendo assim, estas médias são muito semelhantes e nesse caso não rejeitamos a hipótese nula das médias das populações serem iguais. Por outro lado, se as diferenças entre estas médias são demasiado grandes para as podermos atribuir meramente ao acaso, iremos rejeitar a hipótese nula de as médias das populações serem todas iguais. Para um teste destas hipóteses precisaremos de uma medida de discrepância existente entre as médias, \bar{Y}_s , e

MSR: Uma Investigação no âmbito da Indústria Cerâmica

correspondentemente um critério bem definido para testar a hipótese nula em relação aos μ_s . Este critério baseia-se na estimação e comparação das variâncias.

O teste H_0 é baseado na estimativa de σ^2 . Vamos então construir duas estimativas da variância:

A 1ª estimativa, baseada na variação entre médias das amostras, só é válida se $\mu_1 = \mu_2 = \mu_3$.

A 2ª estimativa, baseada na variação nas amostras, é sempre válida assumindo que os pressupostos iniciais são corretos – as populações das quais selecionamos as amostras seguem aproximadamente uma distribuição normal com a mesma variância σ^2 .

- Se os μ_i são todos iguais temos duas estimativas de σ^2 , portanto estas devem ser iguais.
- Se os μ_i não são todos iguais, a 1ª estimativa tende a ser maior que σ^2 e portanto deve ser maior que a 2ª estimativa.

Concluimos então que:

- Os dados são consistentes com a hipótese nula quando as duas estimativas são aproximadamente iguais.
- Os μ_i não são todos iguais, quando as duas estimativas diferem substancialmente.

Uma medida óbvia da diferença entre os \bar{Y}_s é a sua variância, e podemos calculá-la para o exemplo:

$$s_{\bar{Y}}^2 = \frac{(\bar{Y}_1 - \bar{Y})^2 + (\bar{Y}_2 - \bar{Y})^2 + (\bar{Y}_3 - \bar{Y})^2}{3 - 1}$$

Com os pressupostos de que: - os valores nas populações das quais selecionamos as amostras, seguem aproximadamente uma distribuição normal com a mesma variância σ^2 ; - a hipótese nula é verdadeira.

Podemos então considerar que as três amostras foram selecionadas da mesma e única população que segue a distribuição $N(\mu, \sigma^2)$. E neste caso sabemos que: - o erro

padrão da distribuição destas médias é $\sigma_{\bar{Y}} = \frac{\sigma}{\sqrt{n}}$. E a variância é $\sigma_{\bar{Y}}^2 = \frac{\sigma^2}{n}$.

Daqui obtemos a variância da população (da qual selecionamos as 3 amostras):

MSR: Uma Investigação no âmbito da Indústria Cerâmica

$\sigma^2 = n\sigma_{\bar{Y}}^2$. Podemos considerar que $s_{\bar{Y}}^2$ é um estimador de $\sigma_{\bar{Y}}^2$; e $ns_{\bar{Y}}^2$ é um estimador de σ^2 .

A 2ª estimativa de σ^2 , no pressuposto de que as amostras são provenientes de populações com a mesma variância, permite usar cada uma das variâncias amostrais s_1^2 , s_2^2 , s_3^2 como estimativas de σ^2 .

Então a média será também uma estimativa de σ^2 e é dada por: $\frac{s_1^2 + s_2^2 + s_3^2}{3}$

Temos assim duas estimativas para σ^2 : a 1ª baseada na variação entre as médias das amostras (between groups) $ns_{\bar{Y}}^2$; a 2ª baseada na variação nas amostras (*within groups*) $\frac{s_1^2 + s_2^2 + s_3^2}{3}$. Se as compararmos podemos dizer que: se a primeira, baseada na variação entre as médias das amostras, for muito maior que a segunda, baseada na variação nas amostras, será razoável rejeitar a hipótese nula.

Se a hipótese nula for rejeitada, a primeira estimativa deverá ser maior que a segunda, porque refletirá as diferenças que possam existir entre as médias das populações e também a variação casual, enquanto a segunda reflete apenas a variação casual. Para compararmos rigorosamente estas duas estimativas de σ^2 usamos a estatística dada pela razão das variâncias:

$$F = \frac{\text{Estimativa de } \sigma^2 \text{ baseada na variação entre amostras } (\bar{Y}_s)}{\text{Estimativa de } \sigma^2 \text{ baseada na variação nas amostras}}$$

Se os vários pressupostos já mencionados são verdadeiros – a hipótese nula ser verdadeira e as amostras serem provenientes de populações normais com igual variância – a distribuição amostral desta estatística é a distribuição F. Se o valor de F exceder F_α , a hipótese nula será rejeitada.

Os graus de liberdade: numerador k-1 (considerando genericamente k amostras), porque se baseia em k-1 desvios independentes da média; denominador k(n-1), porque se baseia em n-1 desvios independentes para cada uma das k amostras.

No exemplo, para k-1 = 2 e k(n-1) = 3 x = Graus de liberdade, se considerarmos o nível de significância $\alpha = 0,05$, temos que (recorrendo às tabelas da distribuição F) $F_{0,05; 2; \dots} = \dots$ e o quociente: $F = \dots > F_\alpha$, logo podemos rejeitar a hipótese nula para este nível de significância.

Há uma forte evidência para rejeitar a hipótese de igualdade de todas as médias. Concluimos que as diferenças entre as médias amostrais não podem ser atribuídas simplesmente ao acaso.

4.3 Análise de Variância a um factor

Do mesmo modo que a análise de variância para o modelo de regressão decompõe a soma total dos quadrados na soma dos quadrados de regressão e do erro, também no caso do modelo de análise de variância existe uma decomposição correspondente.

Se usarmos qualquer informação sobre os níveis do factor, a variabilidade total dos Y_s é medida pelos desvios das observações Y_{ij} em torno da média global \bar{Y} : $Y_{ij} - \bar{Y}$.

A medida convencional da variabilidade total é a soma dos quadrados destes desvios, que notamos por Soma dos Quadrados Totais (SQT).

$$SQT = \sum_{j=1}^k \sum_{i=1}^{n_j} (Y_{ij} - \bar{Y})^2$$

Em que Y_{ij} é a observação i no nível j , para $j = 1, 2, \dots, k$ e $i = 1, 2, \dots, n_j$
 \bar{Y} é a média global das observações.

Quando usamos a informação referente aos níveis do factor, os desvios que refletem a restante incerteza dos dados são os desvios entre cada observação Y_{ij} e a média do respetivo nível do factor \bar{Y}_j : $Y_{ij} - \bar{Y}_j$.

A diferença entre os desvios para a média global e os desvios para a média em cada nível, reflete a diferença entre as médias dos níveis do factor e a média global:

$$(Y_{ij} - \bar{Y}) - (Y_{ij} - \bar{Y}_j) = \bar{Y}_j - \bar{Y}$$

Podemos então decompor o desvio global, $Y_{ij} - \bar{Y}$, em duas componentes:

1. O desvio da média do nível do factor em torno da média global;
2. O desvio de Y_{ij} em torno da média do nível do factor.

$$\underbrace{Y_{ij} - \bar{Y}}_{\text{Desvio total}} = \underbrace{\bar{Y}_j - \bar{Y}}_{\text{Desvio do nível do factor para a média global}} + \underbrace{Y_{ij} - \bar{Y}_j}_{\text{Desvio em torno do nível do factor}}$$

Quando elevamos ao quadrado ambos os membros da igualdade, os produtos cruzados no termo da direita simplificam e obtém-se:

$$\underbrace{\sum_j \sum_i (Y_{ij} - \bar{Y})^2}_{SQT} = \underbrace{\sum_j n_j (\bar{Y}_j - \bar{Y})^2}_{SQ_t} + \underbrace{\sum_j (\sum_i (Y_{ij} - \bar{Y}_j)^2)}_{SQE}$$

O termo da esquerda é a soma total dos quadrados, notada por SQT. Esta soma, no modelo de análise de variância engloba duas componentes:

1. A soma dos quadrados para o tratamento, notada SQ_t, é uma medida que reflete as diferenças entre as médias dos níveis do factor, baseada nos desvios das médias das amostras para os diversos níveis \bar{Y}_j em torno da média global \bar{Y} . Se todas as médias amostrais para os diferentes níveis forem iguais, então SQ_t = 0. Quanto maior for a diferença entre estas médias, maior será SQ_t.
2. A soma dos quadrados do Erro, notada por SQE, é uma medida da variação aleatória das observações em torno das respectivas médias dos níveis. Quanto menor for a variação entre as observações dentro de um mesmo nível do factor, mais pequena será SQE. Se SQE = 0, todas as observações para esse nível são iguais, mantendo-se isto para todos os níveis. Quanto maior for a diferença entre as observações num nível e a média desse nível, maior será SQE.

Decomposição dos graus de liberdade. Para obtermos estimativas da variância (média dos quadrados já calculados), devemos decompor os graus de liberdade a exemplo do que foi feito para a soma total dos quadrados: - a SQT estão associados n – 1 graus de liberdade, porque temos um total de n observações mas como usamos uma estimativa da média, temos uma restrição $\sum_j \sum_i (Y_{ij} - \bar{Y}) = 0$, logo menos um grau de liberdade; a SQ_t estão associados k – 1 graus de liberdade, temos um total de k médias dos níveis, e como usamos a média global, $\sum_j n_j (\bar{Y}_j - \bar{Y}) = 0$, logo menos um grau de liberdade; a SQE estão associados n – k graus de liberdade porque temos um total de n observações e usamos as k médias dos níveis, logo menos k graus de liberdade.

MSR: Uma Investigação no âmbito da Indústria Cerâmica

As médias dos quadrados são obtidas dividindo cada soma dos quadrados pelos graus de liberdade que lhe estão associados. Deste modo:

$$QMt = \sum_{j=1}^k n_j \frac{(\bar{Y}_j - \bar{Y})^2}{k - 1} \quad \text{Média dos quadrados dos tratamentos}$$

$$QME = \sum_{j=1}^k \sum_{i=1}^{n_j} \frac{(Y_{ij} - \bar{Y}_j)^2}{n - k} \quad \text{Média dos quadrados dos erros}$$

A média dos quadrados dos tratamentos (mean square between – entre tratamentos) e média dos quadrados dos erros (mean square within – intra tratamentos).

Podemos construir a tabela de ANOVA considerando as diversas fontes de variação. Esta tabela é um sumário das contribuições para a variabilidade total e é muito útil quando estudamos o efeito de vários factores. A tabela foi obtida com os dados do exemplo.

Fonte da variação	Soma dos quadrados	Graus de liberdade	Média dos quadrados	Teste F
Entre tratamentos	$SQt = \sum n_j (\bar{Y}_j - \bar{Y})^2$	k-1	$QMt = \frac{SQt}{k-1}$	$\frac{QMt}{QME}$
Erro (nos tratamentos)	$SQE = \sum (Y_{ij} - \bar{Y}_j)^2$	n-k	$QME = \frac{SQE}{n-k}$	
Total	$SQT = \sum (Y_{ij} - \bar{Y})^2$	n-1		

Quadro 3 - Tabela Anova (teórica)

A análise de variância deve ser entendida como um conjunto de técnicas usadas na análise de dados que estão dependentes de vários efeitos operando simultaneamente. Na vida real dificilmente ocorrem factos isolados pelo que o campo de aplicabilidade da ANOVA é muito vasto e diversificado. O uso de métodos estatísticos cresce vigorosamente em pesquisas das variadas áreas de conhecimento, social, económica, política, científica, etc. A atração pelo quantitativo, a necessidade de medir e comparar para validar resultados, é uma necessidade da sociedade actual e neste contexto percebe-se a frequência com que se recorre à estatística, nomeadamente à análise de variância.

5. Metodologia de Superfícies de Resposta (MSR)

De acordo com Myers e Montgomery (1995), a Metodologia de Superfícies de Resposta (MSR) é uma coleção de técnicas matemáticas e estatísticas que são utilizadas para modelar e analisar problemas nos quais a resposta de interesse é influenciada por muitas variáveis; a resposta deve alcançar um valor ótimo; a forma de relacionamento entre a variável de resposta e as variáveis independentes é desconhecida. Portanto, o primeiro passo dentro da metodologia MSR é encontrar uma aproximação razoável para o verdadeiro relacionamento entre y e o conjunto de variáveis. Usualmente, emprega-se um polinómio de baixa ordem.

Supondo que a resposta esperada $E(Y)$ seja função de k variáveis preditoras codificadas x_1, x_2, \dots, x_k , o relacionamento entre y e as variáveis preditoras pode ser expresso segundo uma expansão em Série de Taylor (Box e Draper, 1987), tal que:

$$E(Y) = \pi = \pi_0 + \sum_{i=1}^k \left[\frac{\delta \eta}{\delta x_i} \right]_0 x_i + \frac{1}{2} \sum_{i=1}^k \sum_{j=1}^k \left[\frac{\delta^2 \eta}{\delta x_i \delta x_j} \right] + \dots$$

Onde 0 indica a avaliação na origem.

Se os termos de maior ordem forem ignorados, a expansão produzirá

$$\pi = \beta_0 + \sum_{i=1}^k \beta_i x_i$$

Se forem mantidos os termos de segunda ordem

$$y = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i < j} \beta_{ij} x_i x_j + \epsilon$$

Se houver curvatura no processo, então um polinómio de ordem superior deve ser utilizado, tal como o modelo de segunda ordem,

$$y = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_{ii} x_i^2 + \sum_{i < j} \beta_{ij} x_i x_j + \epsilon$$

A generalidade dos modelos de superfície de resposta utiliza um destes modelos. Muito provavelmente um modelo polinomial não será uma aproximação razoável do modelo real em todo o espaço experimental, mas pelo menos para uma determinada região ele funcionará muito bem (Montgomery, 2001). Para estimar os parâmetros (β) do modelo polinomial emprega-se o método dos mínimos quadrados, que também pode ser escrito em forma matricial.

Quando o ajustamento é adequado, a superfície de resposta encontrada é passível de uma análise aproximadamente equivalente à análise da superfície real, sendo para isso necessário definir um arranjo experimental adequado onde possamos colher os

dados de y e, assim, ficarmos capacitados para estimar eficazmente os parâmetros do modelo.

A metodologia de superfícies de resposta é um procedimento sequencial: quando estamos num ponto da superfície de resposta que está longe do ótimo, há pouca curvatura no sistema e o modelo de 1ª ordem é apropriado, urge encontrar, de forma rápida e eficiente, a melhor região de estudo, a região de ótimo. Encontrada esta região, então podemos usar um modelo mais elaborado, de segunda ordem por exemplo, para localizar o ponto ótimo de máximo ou mínimo.

Na maioria das experiências, assume-se inicialmente que o modelo linear é adequado; entretanto, para se confirmar se existe ou não falta de ajuste, deve-se utilizar pontos experimentais extras denominados de Pontos Centrais. De acordo com Box et al (1978), tal procedimento consiste em adicionar ao fatorial completo, pontos que sejam intermediários aos níveis dos fatores. Serão adicionados tantos pontos centrais quantos forem os K fatores da experiência. Os pontos centrais são utilizados para avaliar a existência de curvatura. Se a resposta média obtida com os nf pontos fatoriais \bar{y}_f for significativamente diferente da média formada pelos nc pontos centrais \bar{y}_c , então, pode-se caracterizar uma região de curvatura significativa. Segundo Montgomery (2001), a soma de quadrados para curvatura é dada por:

$$SS_{QP} = \frac{(nf \times nc)(\bar{y}_f - \bar{y}_c)^2}{nf + nc}$$

O Erro quadrático médio e o teste de hipóteses são respetivamente:

$$MS_E = \frac{SSE}{nc-1} = \frac{\sum_{i=1}^{nc} (y_i - \bar{y})^2}{nc-1}$$

$$H_0: \sum_{i=1}^k \beta_{ii} = 0 \quad \text{vs} \quad H_1: \beta_{ii} \neq 0$$

Se SS_{QP} for muito maior do que MS_E , haverá evidência suficiente para rejeitar a hipótese nula de que os coeficientes quadráticos são nulos.

O teste para a significância do modelo é realizado como um procedimento de ANOVA: calculando-se a razão entre a média quadrática dos termos de regressão e a média quadrática do erro, encontra-se a estatística F ; se F for maior que $F_{\text{crítico}}$, então o modelo é adequado (Montgomery, 2001).

5.1 MSR em R usando “RSM”

A MSR insere-se no conjunto dos métodos de otimização que, devido à especificidade e à necessidade contínua e intensiva de recursos computacionais, vem sendo alvo, em R, de packages específicos como *lm* e mais recentemente *rsm* que apresenta soluções padronizadas para modelos de primeira e segunda ordem, adotando os modelos CCD- *Central Composite Design* e BBD- *Box-Behnken Design*. O “R”, para além do package *lm*, dispõe, a partir de 2010, de um package específico, **rsm**, criado por V. Lenth Russel, para o estudo das superfícies de resposta. A metodologia das superfícies de resposta recorre a funções lineares ou quadráticas e neste sentido está próxima dos usuais modelos de regressão contudo apresenta especificidades próprias: a codificação das variáveis, a avaliação do ajuste, os diferentes tipos de análise em função do modelo ajustado, os resultados da análise e a visualização gráfica da superfície de resposta. Atualmente o package “rsm” cobre apenas os modelos *standard* de primeira e segunda ordem e o seu método prevê apenas uma única variável de resposta, não sendo a otimização de respostas múltiplas aqui contemplada.

Para tratar o problema das respostas múltiplas, Derringer e Swich (1980) criaram uma estratégia eficiente de otimização simultânea. O algoritmo criado originalmente por Harrington (1965) e, posteriormente aprimorado pelos pesquisadores, foi denominado de *Desirability*. O princípio é bastante simples: primeiro estabelece-se um relacionamento entre as respostas e as variáveis independentes. Em geral, isto pode ser feito utilizando MSR e OLS. Então, usando a formulação de Harrington (unilateral ou bilateral), cada uma das respostas do conjunto original é transformada, tal que d_i pertença ao intervalo $0 \leq d_i \leq 1$. O valor de d_i aumenta quando a i -ésima resposta se aproxima dos limites impostos. Combinando as respostas transformadas através de uma média geométrica, encontra-se o índice global D , tal como descreve a equação

$$D = (d_1 (Y_1) \cdot d_2 (Y_2) \dots d_k (Y_k))^{1/k}$$

O índice global D também pertence ao intervalo $[0, 1]$ e será maximizado quando todas as respostas se aproximarem o máximo possível de suas especificações. Quanto mais próximo de 1 estiver D , mais próximas as respostas originais estarão dos seus respetivos limites de especificação. O ponto de ótimo geral do sistema é o ponto de ótimo alcançado pela maximização da média geométrica calculada a partir das funções

desirability individuais. Em R, o package *desirability* de (Kuhn, 2009) é específico para o tratamento destes casos, embora manifeste algumas limitações quando as variáveis em estudo apresentam forte correlação.

Seja um exemplo simples no âmbito da cerâmica. Na produção de uma pasta cerâmica: são necessárias determinadas quantidades de água e desfloculante que permitam a obtenção de uma pasta específica. A água (x_1) e o desfloculante (x_2) são as variáveis independentes e a fluidez a variável dependente (y). O ponto central, correspondente à produção normalizado, correspondeu a 300g de água e 75g de desfloculante. Definimos os níveis baixo e superior da nossa experiência com base no ponto central, decidindo fazer variar a água entre 250 e 350g e o desfloculante entre 50 e 100g. Adotamos o modelo fatorial 2^2 aumentado com 3 pontos centrais.

A codificação dos dados é o primeiro passo e constitui um elemento importante na análise das superfícies de resposta. A forma como os dados são codificados afecta o resultado da análise canónica e da *steepest- ascent path*. Usando um método de codificação que faça com que todas as variáveis codificadas da experiência variem de acordo com o mesmo padrão, impomos a cada preditor a mesma capacidade de determinar o percurso até à superfície de resposta.

O package *rsm* dispõe de uma extensão das *data-frame*, a classe *coded.data* para este efeito: as funções *coded.data*, *as.coded.data*, *decode.data*, *code2val* e *val2code* criam ou revertem o valor destes objetos.

Um exemplo com base num caso de estudo:

```
> BD<-read.table('clipboard', header=T)
> BD
```

	Água	Desfloculante	Bloco	Fluidez
1	250.0	50.0	B1	79.9
2	350.0	50.0	B1	80.8
3	250.0	100.0	B1	81.4
4	350.0	100.0	B1	82.9
5	300.0	75.0	B1	83.3
6	300.0	75.0	B1	83.4
7	300.0	75.0	B1	83.7
8	300.0	75.0	B2	79.1
9	300.0	75.0	B2	79.0
10	300.0	75.0	B2	79.2
11	300.0	110.4	B2	77.8
12	300.0	39.6	B2	75.0
13	370.7	75.0	B2	77.9
14	229.3	75.0	B2	76.5

Quadro 4 - Lançamento de dados em R

MSR: Uma Investigação no âmbito da Indústria Cerâmica

```
> library("rsm")
> pasta<-coded.data(BD, x1 ~ (Água - 300)/50, x2 ~ (Desfloculante - 75)/25)
> Fluidez[1:7,]
```

	x ₁	x ₂	Bloco	Fluidez
1	-1.000	-1.000	B1	79.9
2	1.000	-1.000	B1	80.8
3	-1.000	1.000	B1	81.4
4	1.000	1.000	B1	82.9
5	0.000	0.000	B1	83.3
6	0.000	0.000	B1	83.4
7	0.000	0.000	B1	83.7
8	0.000	0.000	B2	79.1
9	0.000	0.000	B2	79.0
10	0.000	0.000	B2	79.2
11	0.000	1.416	B2	77.8
12	0.000	-1.416	B2	75.0
13	1.414	0.000	B2	77.9
14	-1.414	0.000	B2	76.5

Variable codings ...
x1 ~ (Água - 300)/50
x2 ~ (Desfloculante - 75)/25

Quadro 5 - Codificação de dados em R

Nesta experiência os dados do bloco B1 foram os primeiros a ser colhidos e analisados, após o que foi acrescentado o Bloco B2 para ser feita uma nova análise. Podemos facilmente codificar ou decodificar matrizes ou *data frames* com novos valores arbitrários, por exemplo:

```
> code2val(data.frame(x1 = c(-0.25, 0.75), x2 = c(-1.5, -0.25)),
+ attr(Fluidez, "codings"))
```

	Água	Desfloculante
1	287.5	37.50
2	337.5	68.75

Quadro 6 - Codificação/Decodificação de dados em R

MSR: Uma Investigação no âmbito da Indústria Cerâmica

A função `rsm` é, como já foi dito, uma extensão da função `lm` e funciona de forma similar; contudo usa funções especiais para modelos FO – *first order*, TWI – *Two-way interaction*, PQ – *pure quadratic* e SO – *second order*.

No caso de estudo, temos uma variável resposta “Fluidez” e dois preditores x_1 e x_2 , bem como o factor Bloco. No primeiro estágio foram recolhidos os elementos constantes da tabela e pertencentes ao bloco B1 com base nos quais construímos um modelo de superfície de resposta de primeira ordem:

```
> pasta.1<- rsm(Fluidez ~FO(x1,x2), data =pasta, subset = (Bloco == "B1"))
> summary(pasta.1)
```

Call:

```
rsm(formula = Fluidez ~ FO(x1, x2), data = pasta, subset = (Bloco ==
  "B1"))
```

Residuals:

```
 1  2  3  4  5  6  7
-0.8 -1.1 -1.1 -0.8  1.1  1.2  1.5
```

Coefficients:

	Estimate	Std. Error	t	value Pr(> t)
(Intercept)	82.2000	0.5542	148.321	1.24e-08 ***
x1	0.6000	0.7331	0.818	0.459
x2	0.9000	0.7331	1.228	0.287

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.466 on 4 degrees of freedom

Multiple R-squared: 0.3524, Adjusted R-squared: 0.02861

F-statistic: 1.088 on 2 and 4 DF, p-value: 0.4194

Analysis of Variance Table

Response: Fluidez

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
FO(x1, x2)	2	4.6800	2.3400	1.0884	0.41937
Residuals	4	8.6000	2.1500		
Lack of fit	2	8.5133	4.2567	98.2308	0.01008
Pure error	2	0.0867	0.0433		

Direction of steepest ascent (at radius 1):

x1	x2
0.5547002	0.8320503

Corresponding increment in original units:

Água	Desfloculante
27.73501	20.80126

Quadro 7 - Modelo de primeira ordem em R

MSR: Uma Investigação no âmbito da Indústria Cerâmica

Observamos o sumário de rsm. Constatamos que o modelo de primeira ordem ajustado não é significativo, nem ajustado: R^2 ajustado = 0,03 e lack of fit com p-value = 0,01. É-nos fornecida a direção da próxima etapa (steepest ascent), pouco relevante já que o modelo não é ajustado. Devemos assim procurar ajustar um modelo de ordem superior.

```
> pasta.2<- update(bolo.1, . ~. + TWI(x1,x2))
> summary (bolo.2)

Call:
rsm(formula = Fluidez ~ FO(x1, x2) + TWI(x1, x2), data = pasta,
     subset = (Bloco == "B1"))

Residuals:
 1  2  3  4  5  6  7
-0.95 -0.95 -0.95 -0.95  1.10  1.20  1.50

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  82.2000    0.6366  129.127  1.02e-06 ***
x1             0.6000    0.8421   0.712  0.528
x2             0.9000    0.8421   1.069  0.364
x1:x2          0.1500    0.8421   0.178  0.870
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.684 on 3 degrees of freedom
Multiple R-squared: 0.3592, Adjusted R-squared: -0.2816
F-statistic: 0.5605 on 3 and 3 DF, p-value: 0.6769

Analysis of Variance Table

Response: Fluidez
      Df Sum Sq Mean Sq F value Pr(>F)
FO(x1, x2)  2  4.6800  2.3400  0.8249  0.518235
TWI(x1, x2)  1  0.0900  0.0900  0.0317  0.869976
Residuals    3  8.5100  2.8367
Lack of fit  1  8.4233  8.4233 194.3846  0.005105
Pure error   2  0.0867  0.0433

Stationary point of response surface:
x1  x2
-6  -4

Stationary point in original units:
      Água      Desfloculante
5.968559e-12 -2.500000e+01

Eigenanalysis:
$values
[1] 0.075 -0.075

$vectors
      [,1] [,2]
[1,] 0.7071068 -0.7071068
[2,] 0.7071068  0.7071068
```

Quadro 8 - Modelo de primeira ordem com interações em R

MSR: Uma Investigação no âmbito da Indústria Cerâmica

Este modelo, onde incluímos as potenciais iterações entre as variáveis, revela-se igualmente desajustado, p -value de lack of fit = 0,005.

Para avançarmos no estudo necessitamos obrigatoriamente de recolher mais dados. Vamos agora aceder aos dados do segundo bloco B2 havendo assim condições para implementar um modelo CCD e ajustarmos um modelo de segunda ordem; para o efeito vamos acrescentar um “PQ” ao modelo com iterações anterior, sem esquecer a interação entre os blocos.

Quando o modelo de primeira ordem mostra uma significativa falta de ajuste, o analista pode utilizar um modelo de segunda ordem para descrever a superfície de resposta. Há vários desenhos disponíveis para realizar um projeto de segunda ordem: vamos continuar com o CCD, o mais vulgar. O analista pode começar com os pontos fatoriais 2k, adicionando em seguida pontos centrais e pontos axiais para obter o Delineamento Central Composto. Ao adicionar os pontos axiais obterá termos quadráticos para incluir no modelo de segunda ordem.

O seguinte quadro apresenta o modelo de segunda ordem SO que abarca os termos de primeira ordem FO, as iterações TWI e os termos quadráticos puros PQ.

```
> pasta.3<- rsm( Fluidez ~ Bloco + SO(x1, x2), data = pasta)
> summary(pasta.3)

Call:
rsm(formula = Fluidez ~ Bloco + SO(x1, x2), data = pasta)

Residuals:
    Min       1Q   Median       3Q      Max
-0.18961 -0.07983  0.01603  0.05414  0.21039

Coefficients:
            Estimate Std. Error    t      value Pr(>|t|)
(Intercept)  83.48961   0.07141 1169.23 < 2e-16 ***
Bloco[T.B2] -4.41253   0.07824  -56.40  1.45e-10 ***
x1           0.54753   0.05175   10.58  1.47e-05 ***
x2           0.94441   0.05172   18.26  3.66e-07 ***
x1:x2        0.15000   0.07319    2.05  0.0796 .
x1^2        -0.93021   0.05388  -17.27  5.37e-07 ***
x2^2        -1.32660   0.05376  -24.68  4.57e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1464 on 7 degrees of freedom
Multiple R-squared:  0.9984,    Adjusted R-squared:  0.9971
F-statistic: 746.5 on 6 and 7 DF,  p-value: 1.853e-09

Analysis of Variance Table

Response: Fluidez
            Df Sum Sq Mean Sq  F value    Pr(>F)
Bloco       1   68.201   68.201  3183.3470  1.441e-10
FO(x1, x2)  2    9.542    4.771  222.6970  4.608e-07
TWI(x1, x2) 1    0.090    0.090    4.2009  0.07958
```

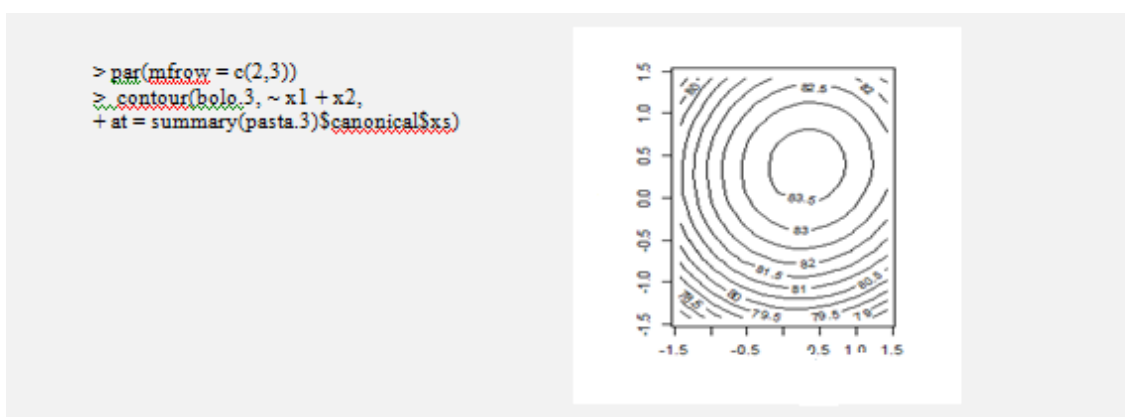
MSR: Uma Investigação no âmbito da Indústria Cerâmica

PQ(x1, x2)	2	18.126	9.063	423.0349	5.005e-08
Residuals	7	0.150	0.021		
Lack of fit	3	0.043	0.014	0.5413	0.67931
Pure error	4	0.107	0.027		
Stationary point of response surface:					
	x1	x2			
	0.3244824	0.3742934			
Stationary point in original units:					
	Água	Desfloculante			
	316.22412	84.35734			
Eigenanalysis:					
Svalues					
[1]	-0.9164988	-1.3403205			
Svectors					
	[,1]	[,2]			
[1,]	-0.9836856	-0.1798961			
[2,]	-0.1798961	0.9836856			

Quadro 9 - Modelo de segunda ordem em R

O lack of fit é agora não significativo ($p=0,679$) – estamos perante um modelo ajustado. O sumário do modelo de segunda ordem fornece resultados duma análise canónica da superfície melhores do que os fornecidos por *Steepest ascent*. A análise indica que o ponto de estacionaridade da superfície ajustada é, em unidades de código, (0.32 ; 0.37) e situa-se no interior da região experimental. Como ambos os valores próprios são negativos ficamos com a certeza de estarmos na presença de um ponto de máximo.

Os gráficos de contorno são úteis para a visualização da superfície de resposta quando os factores não são mais que três, porque a partir daí é quase impossível visualizar a superfície.

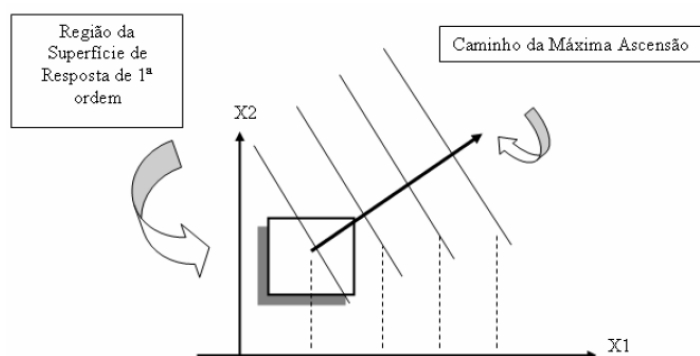


Quadro 10 - Grafico de contornos em R

MSR: Uma Investigação no âmbito da Indústria Cerâmica

Quando o ponto estacionário é localizado, o analista fica em condições de tirar conclusões sobre o resultado que satisfaz a investigação ou então em condições de promover nova pesquisa.

Este caso configura-se como uma situação quase ideal o que, na realidade, poucas vezes acontece. Já vimos que a MSR é uma metodologia que funciona por etapas. Quando o modelo de primeira ordem se manifesta desajustado, devemos estabelecer o percurso que nos leve à superfície de resposta onde se encontre o máximo.



Quadro 11 - Método do vetor gradiente

Aqui constatamos que o modelo de primeira ordem não correspondia às expectativas pelo que deveríamos decidir sobre que caminho, na direção ascendente, deveríamos seguir. O modelo de regressão ajustado (mau) $y = 82,2 + 0,6 x_1 + 0,9 x_2$. A direção do caminho ascendente era dada por $x_1 = 0,555$ e $x_2 = 0,832$

Para nos movermos do centro do delineamento ($x_1=0$ e $x_2=0$) no caminho da inclinação ascendente, deveremos mover 0,555 unidades na direção x_1 para cada 0,832 unidades na direção de x_2 . Assim, a reta da inclinação ascendente contém o ponto central ($x_1=0$ e $x_2=0$) e tem inclinação $\frac{0,832}{0,555} = 1,5$.

Decidimos usar a unidade 10 (g de Água) como tamanho do passo inicial que correspondem na variável codificada a $\Delta x_1 = 0,2$. Assim os passos no caminho da inclinação ascendente são: $\Delta x_1 = 0,2$ $\Delta x_2 = \frac{0,725}{0,275} * \Delta x_1 = 0,528$.

```

> steepest(pasta.1, dist = c(0, 0.2, 0.528))
Linear path of steepest ascent
  dist   x1   x2 | Água  Desfloculante| yhat
1  0.000 0.000 0.000 | 300.00    75.000 | 82.200
2  0.200 0.111 0.166 | 305.55    79.150 | 82.416
3  0.528 0.293 0.439 | 314.65    85.975 | 82.771
    
```

Quadro 12 - steepest ascent

MSR: Uma Investigação no âmbito da Indústria Cerâmica

Reparamos que os valores *yhat* são valores previstos no modelo teórico, pelo que o mais indicado será confrontar estes com os valores observados.

No que concerne ao modelo de segunda ordem muitas vezes acontece que em vez de encontrarmos o ponto de ótimo, o ponto de estacionaridade não passa de um ponto de sela: a verificação pode ser dada pelos sinais dos valores próprios. Nesta situação o melhor é investigar nas proximidades já que estamos próximos do máximo. Supondo que o ponto encontrado no nosso caso de estudo não fosse um máximo, poderíamos investigar, em pontos próximos, qual o comportamento da função encontrada.

```
> canonical.path(pasta.3, dist = seq(-2,2,by = 0.2))
```

dist	x1	x2	Água	Desfloculante	yhat	
1	-2.0	2.292	0.734	414.60	93.350	80.089
2	-1.8	2.095	0.698	404.75	92.450	80.786
3	-1.6	1.898	0.662	394.90	91.550	81.410
4	-1.4	1.702	0.626	385.10	90.650	81.958
5	-1.2	1.505	0.590	375.25	89.750	82.435
6	-1.0	1.308	0.554	365.40	88.850	82.839
7	-0.8	1.111	0.518	355.55	87.950	83.169
8	-0.6	0.915	0.482	345.75	87.050	83.425
9	-0.4	0.718	0.446	335.90	86.150	83.609
10	-0.2	0.521	0.410	326.05	85.250	83.719
11	0.0	0.324	0.374	316.20	84.350	83.755
12	0.2	0.128	0.338	306.40	83.450	83.719
13	0.4	-0.069	0.302	296.55	82.550	83.608
14	0.6	-0.266	0.266	286.70	81.650	83.425
15	0.8	-0.462	0.230	276.90	80.750	83.169
16	1.0	-0.659	0.194	267.05	79.850	82.839
17	1.2	-0.856	0.158	257.20	78.950	82.435
18	1.4	-1.053	0.122	247.35	78.050	81.958
19	1.6	-1.249	0.086	237.55	77.150	81.410
20	1.8	-1.446	0.050	227.70	76.250	80.786
21	2.0	-1.643	0.015	217.85	75.375	80.089

Quadro 13 - Valores estimados da Fluidez, em R

Confirmamos que o valor encontrado é o máximo estimado e, de acordo com o modelo obtivemos estimativas para os valores circundantes da superfície de resposta.

O método consiste em encontrar o caminho certo que leve o investigador a atingir um patamar, uma superfície onde provavelmente se encontra a solução ótima do problema em estudo. Sempre aprendemos que o caminho mais curto entre dois pontos é

o segmento de reta, também aqui a estratégia começa por definir uma orientação que nos leve a direito para a superfície de resposta, usando a regressão múltipla linear. Chegados à superfície de resposta cumpre encontrar o ponto de estacionaridade.

Nem sempre é possível encontrar um ponto ótimo quando há restrições. Nestes casos importa estudar as superfícies e encontrar a melhor solução possível. Este é um tópico que extravasa o âmbito deste trabalho que se subordinou à metodologia.

O exemplo que acima apresentamos resulta das observações registadas ao longo das longas horas dedicadas à observação da produção de pastas cerâmicas. Numa primeira, aquando da identificação do problema de pesquisa e identificação das variáveis mais relevantes, tendo em vista a formulação de hipóteses de pesquisa, fomos construindo computacionalmente planos que integrassem as observações e arbitrando valores que suportassem todas as possibilidades. O arbítrio de valores para a variável resposta permitiu-me um conhecimento e uma prática desta metodologia que me foi extremamente útil na modelação da planificação do caso de estudo desta dissertação.

5.2 Regressão

Em MSR está patente a combinação de técnicas de planeamento de experiências: os testes de hipótese, a ANOVA e a regressão. A realização de inferências acerca de parâmetros desconhecidos da população com base na informação contida nos dados amostrais é feita em dois sentidos: estimando os respetivos parâmetros ou testando hipóteses em relação aos seus valores. A estratégia da MSR é, como vimos, começar por definir uma orientação que nos leve à superfície de resposta usando a regressão múltipla linear, o que implica o cálculo dos parâmetros de regressão, parâmetros desconhecidos que deverão ser estimados a partir dos dados disponíveis. Após a estimação dos parâmetros β_n , a construção dos respetivos intervalos de confiança e testes de significância, devemos ainda verificar se o modelo é correto e, em caso afirmativo, verificar se as variáveis x_n contribuem significativamente com informação para explicar Y. Seja

$H_0: \beta_1 = \beta_2 = \dots = \beta_n = 0$ vs $H_1: \text{Algum } \beta_i (i=1,2,\dots,n) \text{ diferente de zero};$

Se H_0 for verdadeira ter-se-á o modelo $Y = \beta_0 + \epsilon$, modelo não significativo.

O teste para a significância do modelo é realizado como um procedimento de ANOVA: calculando-se a razão entre a média quadrática dos termos de regressão e a média quadrática do erro, encontra-se a estatística F; se F for maior que $F_{\text{crítico}}$, então o modelo é adequado (Montgomery, 2001).

O planeamento de experiências cumpre o objetivo de colher dados que serão posteriormente ajustados, segundo algum algoritmo de modelação. O método dos mínimos quadrados ordinários (OLS) (do inglês Ordinary Least Squares) é o algoritmo vulgarmente usado para estimar os coeficientes de um modelo de regressão linear múltipla, fatoriais completos, fracionados ou superfícies de resposta. O método dos mínimos quadrados para o ajustamento da reta de regressão consiste na minimização da soma dos quadrados dos desvios verticais de cada ponto (observação) relativamente à reta ajustada. Obtido o valor estimado do i-ésimo valor de y, quando $x = x_i$, o desvio do valor observado de y para a reta (\hat{y}_i) é da forma $\epsilon_i = y_i - \hat{y}_i$ que é designado por resíduo ou erro de ajustamento.

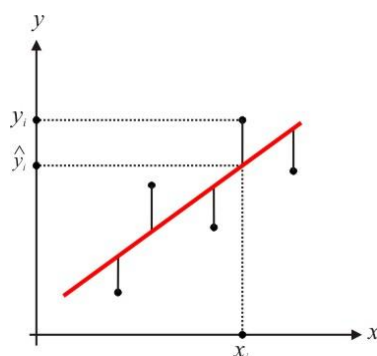


Gráfico 5 – Os resíduos da regressão

Sejam consideradas as $n > k$ observações da variável de resposta disponíveis, como y_1, y_2, \dots, y_n . Assumindo-se que o valor esperado do erro seja zero, podemos escrever que:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \epsilon_i = \beta_0 + \sum_{j=1}^k \beta_j x_{ij} + \epsilon_i$$

O método dos mínimos quadrados baseia-se na escolha de valores para β minimizem a soma dos quadrados dos erros. A função de mínimos quadrados é assim escrita:

$$L = \sum_{i=1}^n \epsilon_i^2 = \sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^k \beta_j x_{ij})^2$$

A função L deve ser minimizada em função de $\beta_0, \beta_1, \dots, \beta_k$.

MSR: Uma Investigação no âmbito da Indústria Cerâmica

Em notação matricial, o sistema de equações resultante da derivação, pode ser escrito como:

$$Y = \beta X + \epsilon, \text{ onde}$$

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad X = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}, \quad \epsilon = \begin{bmatrix} \epsilon_0 \\ \epsilon_1 \\ \vdots \\ \epsilon_k \end{bmatrix}$$

Podemos então reescrever L

$$L = y^t y - \beta^t X^t y - y^t X \beta + \beta^t X^t X \beta = y^t y - 2 \beta^t X^t y + \beta^t X^t X \beta$$

Como os estimadores de mínimo quadrado devem satisfazer a igualdade a zero

$$\frac{\delta L}{\delta \beta} = -2 X^t y + 2 X^t X \hat{\beta} = 0$$

Simplificando a equação anterior: $X^t X \hat{\beta} = X^t y$

E multiplicando ambos os termos pela inversa de $X^t X$

$$\hat{\beta} = (X^t X)^{-1} X^t y$$

Esta equação é a expressão matricial das equações normais de mínimos quadrados, descrita anteriormente (Montgomery, 2001).

Como a verdadeira relação entre x e y normalmente não é conhecida, utilizam-se aproximações de funções, usando-se para o efeito um modelo de regressão linear múltipla que é relativamente adequado dentro de certa faixa de variação das variáveis independentes.

Observemos o modelo quadrático: $y = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_{ii} x_i^2 + \sum_{i < j} \sum \beta_{ij} x_i x_j + \epsilon$

\hat{y} representa um valor ajustado, obtido a partir do modelo de regressão. A diferença entre o valor real observado y e o valor ajustado \hat{y} , denomina-se resíduo.

$$e = y - \hat{y}$$

A soma dos quadrados dos resíduos pode escrever-se

$$SS_E = e^t e = y^t y - \hat{\beta}^t X^t y$$

A soma dos quadrados totais

$$SS_T = y^t y - \frac{1}{n} (\sum_{i=1}^n y_i)^2$$

MSR: Uma Investigação no âmbito da Indústria Cerâmica

A Soma dos quadrados dos tratamentos

$$SS_t = SS_T - SS_e = \hat{\beta} X^t y - \frac{1}{n} (\sum_{i=1}^n y_i)^2$$

O teste para a significância da regressão indica se existe uma relação linear entre a variável de resposta y e os regressores.

$$H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0 \quad \text{vs} \quad H_1: \beta_i \neq 0$$

A rejeição da hipótese nula implica que, no mínimo, um dos regressores é diferente de zero e, por isso, contribui significativamente para a explicação da variação de y . Dado que raramente um modelo explica a totalidade da variação da resposta modelada, é provável que exista uma diferença entre o valor observado e o valor modelado. Sob a hipótese nula, SS_t / σ^2 é uma variável aleatória χ^2 com k graus de liberdade. O número de graus de liberdade para esta variável é igual ao número de regressores. O mesmo acontece com SS_e / σ^2 , com $n-k$ graus de liberdade. As duas expressões anteriores são independentes; dividindo-se as duas variações encontramos uma distribuição F. Portanto, a estatística de teste para a significância do modelo:

$$F_0 = \frac{\frac{SS_t}{k-1}}{\frac{SS_e}{n-k}} = \frac{MSt}{MSe}$$

Rejeitamos a hipótese nula se a estatística do teste for maior que o valor de $F_{(\alpha, k, n-k)}$

O coeficiente de determinação R^2 é a medida mais comum de adequação de um modelo. Este termo representa a percentagem de variação na resposta que é explicada pelo modelo. O valor de R^2 ajustado é o mais apropriado para se comparar modelos com diferentes quantidades de termos, já que a inclusão de muitos termos no modelo de regressão aumenta substancialmente R^2 . O valor de R^2 é calculado:

$$R^2 = \frac{SS_t}{SST} = 1 - \frac{SS_e}{SST}$$

O coeficiente de determinação ajustado (R^2 adj.) considera o número p de variáveis incluídas no modelo, podendo observar-se que diminui quando p aumenta.

$$R^2 \text{ adj} = 1 - \left(\frac{n-1}{n-p}\right) (1 - R^2)$$

A adequação de um modelo pode ainda ser apreciada através dos resíduos. O conjunto de resíduos deve ter distribuição normal com médias zero e variância σ^2 e, para além disto, devem ser não correlacionados quer com outros resíduos, quer com as médias dos tratamentos, quer com as variáveis preditoras quer ao longo do tempo;

devem ter padrão aleatório, não tendencioso e devem preferencialmente assumir a forma padronizada.

5.3 Otimização de uma única Resposta

Encontrar uma condição ótima de operação para um processo, é o objetivo fundamental da metodologia de superfície de resposta: quando obtida a função de aproximação, fica viabilizada a determinação do ponto de ótimo.

Segundo Box e Wilson (1951), a otimização experimental de sistemas com uma única variável de resposta deve ser conduzida em duas fases. A primeira consiste em procurar sequencialmente uma direção linear que leve à máxima otimização. Esta procura é realizada até que haja evidência de que a direção escolhida não resulta em melhorias adicionais para o modelo, ou, enquanto não houver evidências de falta de ajustamento para o modelo de primeira ordem. De acordo com Lin e Chou (2002), a segunda fase só é realizada se for detetada falta de ajustamento no modelo de primeira ordem. Obviamente, nem todas as respostas requererão ajustes dessa natureza, e, em tais casos, o modelo da primeira fase será suficiente quando não houver melhorias adicionais.

Quando uma experiência é planeada para um processo sobre o qual se conhece pouco, existe uma chance de que as condições iniciais estejam muito distantes da real região de ótimo. Já anteriormente referimos que o objetivo é manter a experiência ao longo da direção de máxima ascensão (Path of Steepest Ascent), até que a resposta não denote melhorias adicionais. Neste ponto, uma nova experiência fatorial com pontos centrais deve ser aplicado para se determinar a nova direção. Este processo deve ser repetido até que alguma curvatura seja detetada. A curvatura ou a falta de ajuste considerável do modelo linear indica que os parâmetros da experiência estão próximos da região de máximo. Segundo Lin e Chou (2002) há duas coisas importantes a observar na fase 1: determinar a direção de busca e o comprimento do passo utilizado para conduzir a experiência para a região do ótimo.

O método do vetor gradiente, segundo Forster e Barhe (1990), indica a direção na qual a resposta aumenta mais rapidamente. Esta direção é paralela a uma reta normal traçada sobre as superfícies de resposta ajustadas, passando pelo centro da região de

interesse, a partir do modelo ajustado de primeira ordem. Consequentemente o tamanho dos passos ao longo deste caminho de melhoria é proporcional aos coeficientes de regressão. Uma vez determinado o comprimento do passo para cada variável, as experiências devem ser realizadas alterando-se os níveis dos fatores a partir da origem, de tal forma que se obtenham níveis do tipo $X_0, X_0 + \Delta X, X_0 + 2\Delta X, \dots$, enquanto houver melhorias na resposta y . Uma vez alcançado o ponto a partir do qual y não sofre melhorias adicionais, uma nova experiência deve ser realizado utilizando-se pontos centrais para que a falta de ajuste do modelo seja avaliada. Contudo, se não for detetada falta de ajuste, o novo modelo de primeira ordem deverá fornecer uma direção para a investigação. De outro modo o projeto experimental deve ser aumentado com pontos axiais para que um modelo de segunda ordem seja avaliado.

Segundo Khuri (2003), para se determinar a direção a ser seguida, tal que se alcance um incremento na resposta, é necessário que se maximize $\hat{y}_{(x)}$ sujeita à restrição de que o próximo ponto deva pertencer à região experimental de raio r , tal que:

$$\sum_{i=1}^k x_i^2 = r^2$$

O raio da região experimental coincide com a distância axial α . Usando o método dos multiplicadores de Lagrange e considerando-se esta restrição, temos:

$$S(x) = \hat{\beta}_0 + \sum_{i=1}^k \beta_i x_i - \lambda (\sum_{i=1}^k x_i^2 - r^2)$$

Onde λ é um multiplicador de Lagrange. Calculando as derivadas de $S(x)$ e igualando a zero, obtemos: $x_i = \frac{1}{2\lambda} \hat{\beta}_i$

Para um máximo, λ deve ser positivo, pelo que, usando a função de restrição temos:

$$\lambda = \frac{1}{2r} (\sum_{i=1}^k \beta_i^2)^{1/2}$$

Um máximo local é alcançado no ponto de coordenadas:

$$x_i = r \hat{\beta}_i (\sum_{i=1}^k \beta_i^2)^{-1/2}, \quad i = 1, 2, \dots, k$$

Se $e_i = \hat{\beta}_i (\sum_{i=1}^k \beta_i^2)^{-1/2}$, $i = 1, 2, \dots, k$. Então $(e_1, e_2, \dots, e_k)^T$ é um vetor unitário na direção de $(\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k)$.

A equação do ponto de máximo indica que há uma distância de r unidades da origem, um máximo aumento na resposta ocorre ao longo da direção do vetor $(e_1, e_2, \dots, e_k)^T$. Sendo que este é o único máximo local dentro da hipersfera de raio r , então é

também o máximo absoluto. Se o valor da resposta obtida no ponto $x = re$ exceder o seu valor na origem da experiência, então a movimentação de e é adequada.

Otimizar uma resposta significa encontrar o melhor conjunto de soluções para as variáveis independentes x_1, x_2, \dots, x_k que satisfaçam uma determinada condição de Y . Se uma função possui ponto de máximo ou de mínimo, a primeira derivada dessa função é nula. No caso de a função ser composta de um conjunto de variáveis independentes, as derivadas parciais da função em relação a cada uma das variáveis independentes também devem ser nulas, para que haja um ponto de máximo ou de mínimo. Este é denominado ponto de estacionaridade.

Seja a seguinte função quadrática: $\hat{y} = \frac{1}{2} x^T B x - b^T x$

Onde $x = (x_1, x_2, \dots, x_k)$, b é um vetor de comprimento n e B é uma matriz simétrica $n \times n$.

A função também pode ser escrita: $\hat{y} = \hat{\beta}_0 + x' b + x' B x$

Onde

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_k \end{bmatrix} \quad b = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \cdot \\ \cdot \\ \hat{\beta}_k \end{bmatrix} \quad B = \begin{bmatrix} \hat{\beta}_{11} & \hat{\beta}_{12}/2 & \cdot & \cdot & \hat{\beta}_{1k}/2 \\ & \hat{\beta}_{22} & \cdot & \cdot & \hat{\beta}_{2k}/2 \\ & & \cdot & & \\ & & & \cdot & \\ \text{sim.} & & & & \hat{\beta}_{kk} \end{bmatrix}$$

Na equação $\hat{y} = \hat{\beta}_0 + x' b + x' B x$, b representa o vetor dos coeficientes de regressão de primeira ordem e B é uma matriz simétrica, com a diagonal principal formada pelos termos quadráticos puros e onde os outros coeficientes são representados pela metade do valor dos coeficientes das interações.

O ponto de estacionaridade: $x_s = -\frac{1}{2} B^{-1} b$.

Uma aplicação prática deste desenvolvimento é mostrada nos cálculos a que procedemos no caso de estudo que se segue.

6. Enquadramento do Planeamento Experimental

Em Barcelos, estrada do Prado, fica situada a fábrica de Cerâmica “MAGROU”, empresa industrial criada há cerca de 60 anos e mantida na posse da mesma família desde então.



Ilustração 3 - Fábrica Magrou

Nos velhos tempos, o patriarca da família, Bisavô dos atuais proprietários, iniciou a atividade cerâmica, produzindo louça durante o inverno para, no verão vender nas feiras; o negócio progrediu e, poucos anos depois, a laboração estendeu-se ao ano inteiro. Nos anos 90 tem lugar o apogeu desta pequena indústria que chegou a empregar cerca de noventa trabalhadores. A geração atual, que a administra, preparou-se para a continuidade do negócio, tendo cada um dos três irmãos optado por preparar-se cientificamente: Dr. Carlos Gonçalves, licenciado em direito, na área de relações públicas e vendas para o comércio nacional e internacional; Eng.º Paulo Gonçalves, licenciado em engenharia cerâmica, na área da produção; e Dr. João Nuno Gonçalves, contabilista, na área da administração e contabilidade.

A fábrica tem vindo a decair ao longo dos últimos anos, sendo que o volume de encomendas tem diminuído na ordem dos 20% ao ano, encontrando-se hoje numa fase de pré-desmantelação. As razões para a decadência da empresa e o conseqüente desencanto destes jovens empresários radicam em problemas estruturais que enumeram: o aumento das responsabilidades empresariais, nomeadamente os salários e a legislação laboral rígida; o Euro e a CEE, moeda de valor demasiadamente elevado para permitir

MSR: Uma Investigação no âmbito da Indústria Cerâmica

concorrência com os mercados a quem A CEE abriu as portas; a concorrência avassaladora dos produtos chineses que praticam preços de produto final abaixo dos preços da matéria-prima aqui consumida.

Embora conheçam bem o regime de incentivos à modernização da indústria, qualquer investimento em nova maquinaria estaria destinado ao insucesso, dizem. “Modernizar para quê, maquinaria nova para quê, se quer o preço da pasta cerâmica, quer do vidro – matérias-primas com que laboram – é, a partida, já superior ao do produto acabado chinês? Só mesmo se acabasse este verdadeiro dumping!” A generalidade das outras indústrias cerâmicas da área sofre das mesmas vicissitudes e o panorama é bastante desolador.

Quando solicitamos o planeamento da experiência que seguidamente apresentamos, ficamos a saber que a antiga “casa do segredo” área onde funcionava o laboratório e se fazia a preparação das pastas cerâmicas, já há muito tinha sido desativado. Hoje compram-se as pastas secas e com elas se prepara a barbotina em função do produto específico.

A pasta seca é misturada com água e desfloculantes para a preparação da barbotina com características específicas para a produção de determinado artigo. Abaixo mostramos alguns dos artigos, ainda em produção:



Ilustração 4 - Artigos em produção

Centrámos na produção de barbotinas a nossa intervenção, de que resultou o estudo experimental que se segue.

6.1 Introdução à Cerâmica

Há cerca de dez a quinze mil anos que a cerâmica é produzida pelo homem. Segundo a definição do ceramólogo Giuseppe Liveranni, por cerâmica (da palavra grega “Keramos” – argila) entende-se a terra sobre a qual interveio a obra do homem e a acção consolidadora do fogo, para a transformar em objetos de uso prático e ornamental. Deduz-se daqui a existência de três elementos: a terra, a arte humana, o fogo.

A descoberta fundamental na génese da História da Cerâmica foi a de que o calor transforma a argila num material duro e inalterável pela água. A argila não tem, em si mesma, nenhuma forma definida nem valor intrínseco e, por isso, não impõe restrições à liberdade criativa de quem com ela trabalha; mas, por outro lado, goza da propriedade de poder ser dimensionada simultaneamente como atividade industrial, como atividade artística, como técnica, como reflexo da vida do quotidiano. Com a revolução industrial e as necessidades de produção em massa, novos equipamentos, tecnologicamente avançados, se adaptaram à produção cerâmica; mas é de realçar que estes, atualmente, coexistem com as rodas de oleiro e que, ao mesmo tempo que se desenvolvem grandes fábricas para produção de louça utilitária ou de construção, proliferam os ateliers onde a cerâmica é produzida a nível artístico ou artesanal.

O campo dos produtos cerâmicos é vasto, indo desde o grupo dos cerâmicos tradicionais (faiança, grés, porcelana, sanitário, telhas, tijolos, revestimento e pavimento, etc.) até ao dos cerâmicos técnicos (porcelana dentária, isoladores de alta tensão, próteses, etc.). Enquanto no primeiro grupo se utilizam matérias-primas naturais como as argilas, as areias, os feldspatos e os carbonatos, no segundo grupo recorre-se a óxidos puros, aos nitretos, carbonetos, etc..

Podemos classificar e dividir os produtos cerâmicos tradicionais usando como critério a sua porosidade. Os produtos porosos podem subdividir-se em faiança comum (pasta vermelha), faiança fina (pasta branca), faiança feldspática (pasta branca), refratários, revestimentos e cerâmica de construção. A nossa observação centrou-se numa fábrica de produção de faiança fina de pasta branca, especificamente de louça utilitária e decorativa.

A argila é o principal componente dos produtos cerâmicos: as matérias-primas usadas no fabrico cerâmico são rochas sedimentares e os seus principais elementos são a

sílica e a alumina. As matérias-primas dividem em matérias plásticas e não plásticas: as primeiras englobam os barros e caulinos; as segundas, não plásticas – rijas ou duras – englobam a sílica, o feldspato, a calcite e dolomite, a chamote e o gesso.

Na composição de uma pasta cerâmica podem intervir materiais plásticos (argilas e caulinos), desengordurantes (sílica), fundentes (feldspato, calcite, dolomite, etc.), aditivos tais como branqueadores (sulfato de cobalto), desfloculantes (silicato de sódio, carbonato de sódio, etc.), secantes, etc. A proporção em que cada um dos componentes entra na pasta exige um conhecimento prévio das suas propriedades nomeadamente plasticidade, grau de moagem, cor, composição química aproximada, etc.. A preparação da pasta pode fazer-se por um processo de via seca, via plástica ou via húmida, sendo necessário assegurar uma boa mistura e homogeneização (diluição), alterar o tamanho do grão das matérias-primas rijas (moagem) e eliminar impurezas (peneiração).

Os minerais das argilas constituintes da pasta são formados por partículas de tamanhos muito pequenos (2μ) e por isso não necessitam moagem, mas como se encontram aglomerados é necessário separá-los. Esta dispersão é feita num diluidor em que se adiciona a argila e/ou caulino na quantidade imposta pela formulação e a água. Para facilitar a diluição pode adicionar-se à mistura um desfloculante.

As matérias-primas rijas como a areia, feldspatos, calcite e dolomite ou chamote necessitam ser moídas. A moagem pode ser feita a seco ou a húmido, sendo neste caso adicionada água.

A peneiração é realizada depois da diluição e moagem e tem como objetivo extrair da pasta partículas estranhas de algumas matérias-primas. Faz-se passar a pasta por uma rede metálica com fios de aço ou latão com orifícios quadrados, animada de um movimento vibratório.

A preparação da pasta é normalmente feita por via húmida e como tal utilizam-se grandes quantidades de água que terão de ser removidas antes do processo de conformação. A desumidificação é feita com filtros prensa se a percentagem de humidade for da ordem dos 20% (prensagem ou contramoldagem húmida) ou com atomizadores se a percentagem de humidade não ultrapassar 5% (prensagem a seco).

Depois de preparada a pasta cerâmica é necessário proceder à conformação das peças que se pretende fabricar. Existem atualmente três processos de conformação: via

MSR: Uma Investigação no âmbito da Indústria Cerâmica

seca (5% de humidade), via plástica (20% de humidade) ou via húmida (pasta líquida ou, como normalmente se designa, enchimento por barbotina).

A prensagem a seco utiliza-se na fabricação de peças de revestimento ou pavimento; a pasta seca sob a forma de pó atomizado é introduzida num molde inferior e submetida a pressão elevada (até $\pm 500 \text{ Kgf/cm}^2$) por meio de um molde superior que vai permitir a aglomeração dos grãos.

Plasticidade é a propriedade que permite a alteração da forma de um material, sem ruptura, por aplicação de uma força exterior, e a manutenção da forma adquirida depois de retirada a força. Quanto menor o tamanho dos minerais das argilas e maior a quantidade de matéria orgânica presente, maior a plasticidade de uma argila.

Na conformação por via húmida, a barbotina (pasta líquida) é colocada num molde de gesso com a forma interior igual à forma exterior do objeto que se pretende. Os poros capilares do molde absorvem a água da barbotina deixando uma certa espessura de pasta à sua superfície; assim que se atinge a espessura pretendida, procede-se ao vazamento do excesso de barbotina e deixa-se ficar a peça a secar dentro do molde durante algum tempo até que a peça contraia o suficiente para se descolar e adquira a consistência necessária para que possa ser manuseada. A barbotina deve ter as seguintes características: conter o mínimo de água possível para evitar a saturação dos moldes de gesso, ser suficientemente fluida para permitir encher completamente os moldes e suficientemente viscosa para evitar a sedimentação das partículas.

Noção de densidade. As barbotinas são suspensões de matérias plásticas e não plásticas em água. A densidade permite medir a quantidade de água na barbotina: a densidade da barbotina é o peso de um quilo de barbotina. Quando dizemos que a densidade da barbotina é 1,7 queremos dizer que 1 litro pesa 1700g. Na cerâmica usam-se normalmente barbotinas com densidade entre [1.65, 1.75].

Noção de viscosidade, fluidez e tixotropia. A viscosidade de uma barbotina dá indicação sobre a facilidade com que ela flui. Quanto maior o seu valor mais dificilmente se processa o escoamento do líquido. Esta grandeza é função da quantidade de água e da quantidade de desfloculante (torna mais fluida uma barbotina). Tixotropia é a capacidade de se tornar viscosa uma barbotina após repouso e é o factor determinante na velocidade de formação da parede no processo de conformação.

MSR: Uma Investigação no âmbito da Indústria Cerâmica

Após o processo de conformação, segue-se o processo de secagem das peças. Em cerâmica, a água é um componente fundamental durante as operações de conformação mas tem que ser eliminada na sua totalidade antes das peças entrarem no forno. Os fatores que influenciam a velocidade de secagem são os seguintes: temperatura e humidade do ar, velocidade do ar de secagem, superfície específica das peças e espessura das peças. A secagem pode ser feita ao ar livre ou com a utilização de secadores. Os secadores contínuos são equipamentos nos quais as peças se movem uniformemente através de um túnel onde a temperatura aumenta gradualmente.

Após a conformação, as peças executadas por qualquer dos processos (enchimento, contramoldagem, prensagem, etc.) têm de ser sujeitas a operações de acabamento que consistem fundamentalmente em: cortar e aparar, esponjar para disfarçar costuras e alisar bordos, desbastar, colar asas ou partes da peça.

A cozedura é uma fase muito importante no processo de produção de peças cerâmicas. A primeira etapa da cozedura de uma pasta consiste na eliminação da água não combinada e que não foi expelida durante a secagem. Na segunda etapa vão dar-se transformações acompanhadas por eliminação de gases, vapores e água de constituição dos componentes da pasta (até 550°C). A matéria orgânica – raízes, carvão, compostos químicos de carbono – queima-se entre os 300 e os 400°C, mas alguns compostos podem não arder antes dos 1000°C. Durante este processo é necessário que exista excesso de oxigénio no forno a fim de que a combustão com libertação de CO₂ (dióxido de carbono), seja possível. Os carbonetos de cálcio e magnésio decompõem-se cerca dos 800 – 900°C com libertação de CO₂. Aos 573°C dá-se a transformação do quartzo α para quartzo β com variação brusca do volume da pasta. Numa terceira fase a fusão do componente fundente da pasta permite a sua consolidação que se designa normalmente por vitrificação.

Quase toda a cerâmica fina se coze em atmosfera oxidante, mas há casos que requerem atmosfera redutora, como a cozedura de um vidro vermelho “sangue de boi” (com óxido de cobre) ou a cozedura de porcelana branca. A teoria da cozedura redutora é simples: quando o carbono (C) que contém o combustível (gás, lenha, etc.) se combina com o oxigénio do ar resulta um desenvolvimento de calor e dióxido de carbono: $C + O_2 \rightarrow CO_2 + \text{calor}$. Quando se consome todo o oxigénio do ar, o carbono

MSR: Uma Investigação no âmbito da Indústria Cerâmica

vai reagir agora com o oxigénio existente no corpo cerâmico, por isso se diz que está reduzido e esta redução tem o poder de mudar de cor um vidrado.

O fundente de uma pasta forma um vidro que funciona como elemento de consolidação do conjunto quando o ponto de fusão desta substância for atingido.



Ilustração 5 - Fornos de cozedura

Vidrados. Muitas das peças cerâmicas apresentam porosidade (faiança, grés), pelo que é necessário revesti-las com uma camada impermeável, o vidrado. Um vidrado é uma mistura de óxidos que se aplica sobre um corpo cerâmico e aquecida até à temperatura de fusão 900 – 1400°C (conforme a composição) forma uma camada vítrea que o impermeabiliza. Fontes de óxidos: o quartzo para fornecer a sílica, o feldspato para fornecer o óxido de sódio e o de potássio, a calcite para fornecer o óxido de cálcio, o zarcão para o óxido de chumbo, etc.

Os vidrados preparam-se geralmente misturando-se os elementos constituintes crus e moendo esta mistura com água. Assim dado que se utilizam materiais solúveis em água como o óxido de boro e o óxido de potássio, quando a água se evaporasse essas partículas, dissolvidas na água, evaporar-se-iam também. Para evitar que tal aconteça, os sais solúveis em água são incorporados em silicatos por meio de um processo de fusão prévio, a “fritagem”, que os transforma num composto insolúvel na água. Este processo tem ainda a vantagem de tornar menos tóxicas substâncias contidas no vidrado, como por exemplo o chumbo. Os vidrados fritados têm ainda a vantagem de poderem ser cozidos a temperaturas mais baixas e de permitirem uma maior uniformidade, conduzindo a um menor número de defeitos.

A diferença entre os coeficientes de dilatação da pasta e do vidrado é a responsável pelo aparecimento de defeitos: descasque e craquelê. O descasque – vidrado que se destaca da peça em placas – resulta do facto de o corpo cerâmico ter, durante o arrefecimento, uma contração muito superior à do vidrado que fica demasiado em compressão, devendo diminuir-se o coeficiente de dilatação da pasta reduzindo o teor em areia ou moendo menos os materiais rijos da pasta. Quando o coeficiente de dilatação do corpo é inferior ao do vidrado pode acontecer que este fendilhe (craquelê), o que se pode evitar atuando de forma oposta à anteriormente sugerida, isto é, aumentando a % de areia da pasta ou moendo mais os seus componentes.

Se o vidrado apresenta bolhas é porque iniciou a solidificação antes de ter libertado todos os gases, pelo que terá de cozer durante mais tempo ou a uma temperatura mais elevada. Tanto os vidrados mal cozidos como os que estão cozidos em excesso apresentam defeitos muito comuns: os primeiros não apresentam brilho e podem ainda apresentar crateras provenientes de bolhas que não tiveram tempo de resolver por isso, para remediar a situação pode cozer-se a peça a uma temperatura mais elevada ou acrescentar ao vidrado um fundente para baixar o seu ponto de fusão; os segundos apresentam vidrados intensamente brilhantes, mas pode dar-se a acumulação de vidrado na parte inferior da peça colando-a à placa do forno.

6.2 Controlo das pastas cerâmicas

O termo pasta é usado num sentido geral que inclui pasta seca, semi-húmida e barbotina. Para cada tipo de produto final – faiança, grés, porcelana, etc. – é preparada uma determinada pasta, utilizando determinadas matérias-primas. As matérias-primas são submetidas a diversos processos tecnológicos, nomeadamente moagem, diluição, peneiração, homogeneização, etc., até obtermos uma pasta. Ao longo deste percurso tecnológico há que introduzir operações de controlo para assegurar a perfeita sequência dos processos tecnológicos seguintes, a qualidade dos produtos intermédios e evidentemente do produto final.

Há nos referidos subsectores cerâmicos dois processos principais para a conformação de peças cerâmicas: por via líquida, normalmente chamado enchimento e por via semi-húmida, vulgarmente chamado contramoldagem.

MSR: Uma Investigação no âmbito da Indústria Cerâmica

Vejam os quais os procedimentos tecnológicos desde o início do tratamento das matérias-primas até à pasta estar pronta a ser utilizada. Vamos relatar somente dois processos, o mais simples e o mais complexo:

Processo 1

1. Moagem dos materiais duros: Areia, Calcite, Dolomite, Feldspato, etc..
2. Diluição dos materiais plásticos no próprio moinho: Argilas, Caulinos, etc..
3. Peneiração e desferrização da mistura.
4. Homogeneização e armazenamento da mistura. Utilização da barbotina para enchimento por via líquida.

Processo 2

1. Moagem dos materiais duros: Areia, Calcite, Dolomite, Feldspato.
2. Diluição dos materiais plásticos no diluidor de alta velocidade: Argilas, Caulinos.
3. Peneiração e desferrização dos materiais duros.
4. Mistura e homogeneização dos materiais duros e plásticos no diluidor.
5. Peneiração e desferrização da mistura.
6. Homogeneização e armazenamento da mistura.
7. Homogeneização e desarejamento da pasta. Utilização como pasta semi-húmida em contramoldagem (jauler roller, prensa, etc.).

Havendo diferentes processos tecnológicos de preparação de pastas, há operações comuns. Quais os controlos a efetuar e os pontos do processo tecnológico onde estes devem ser inseridos:

- Controlo da moagem, que deve ser feito imediatamente a seguir à moagem dos materiais duros.

A moagem dos materiais duros destina-se como é evidente, a reduzir as dimensões dos mesmos. Normalmente as matérias-primas são enviadas pelo fornecedor com determinadas dimensões granulométricas geralmente demasiadamente elevadas para permitirem a sua incorporação direta nas pastas cerâmicas. Devem portanto sofrer uma redução das suas dimensões. Todo o futuro comportamento das peças cerâmicas ao longo do percurso fabril está dependente do grau de moagem sofrido pelas matérias-primas. A distribuição granulométrica de uma pasta condiciona as reações no estado

MSR: Uma Investigação no âmbito da Indústria Cerâmica

sólido entre os diferentes componentes dessa pasta: quanto menores forem as partículas sólidas e portanto maior for a superfície de contacto total entre estas, mais facilmente se verificará a reação entre elas e mais completa esta será.



Ilustração 6 - Diluidor e Agitador de materiais não plásticos

Outras propriedades das pastas estão ainda diretamente relacionadas com o tamanho das partículas suas constituintes: o tamanho das partículas influencia, como vimos, a superfície e portanto a relação com a água, condicionando a plasticidade, a viscosidade e a contração. Influencia ainda a resistência em seco e a porosidade depois da cozedura. O tamanho das partículas influencia ainda o tempo de formação de parede – um tamanho de partícula demasiado pequeno conduz a um tempo excessivo de formação de parede.

Este controlo de moagem deve ser feito imediatamente antes da carga de materiais duros ser misturada e homogeneizada com os materiais plásticos. Como é compreensível, após um certo tempo de moagem efetuado numa carga de materiais duros não vamos obter um tamanho de partículas uniforme, mas sim uma certa distribuição granulométrica que pode ser representada graficamente através das “curvas de distribuição granulométrica”.

Há vários processos, mais ou menos sofisticados, de avaliar a eficiência da moagem, mas certamente o que melhor se aplica às pequenas e médias empresas é o processo de peneiração: limitamo-nos a verificar e determinar o resíduo obtido ao passar por uma peneira uma certa quantidade de material moído. Não há valores absolutos para o valor ótimo do resíduo assim obtido, mas considera-se pronta a ser utilizada nas operações seguintes uma carga de materiais duros que origine um resíduo de 1 – 3% referido ao peso seco da amostra ao passar no peneiro de 0,040mm (40 μ m).

MSR: Uma Investigação no âmbito da Indústria Cerâmica

- Controlo do peso específico da mistura resultante da homogeneização da carga de materiais duros e dos materiais plásticos.

Este controlo é necessário em qualquer dos processos tecnológicos seguidos para produzir pastas cerâmicas.

O peso por litro ou peso específico da barbotina, depende das matérias-primas utilizadas e da relação entre o peso de sólidos e o peso de água existentes na barbotina. Quando usamos a pasta cerâmica sob a forma de barbotina num processo de enchimento, o controlo do peso específico é necessário para permitir um tempo de formação de parede constante e também para evitar que as formas utilizadas no enchimento fiquem rapidamente encharcadas. O controlo e possíveis correções, deve ser feito imediatamente antes do processo de homogeneização e armazenamento da barbotina.

Não há valores absolutos para este parâmetro, peso específico, mas usam-se normalmente pesos específicos da ordem dos 1650 – 1800 gr/litro. O controlo do peso específico da barbotina é muito simples e para o efetuar necessitamos apenas de um recipiente de volume conhecido e de uma balança.

- Controlo da viscosidade e tixotropia das barbotinas utilizadas por via líquida (enchimento)

A viscosidade está relacionada com a maior ou menor facilidade com que as partículas argilosas das suspensões cerâmicas deslizam umas sobre as outras. Quanto à tixotropia entende-se como a propriedade que as suspensões têm de aumentar a sua viscosidade ao longo do tempo, quando deixadas em repouso.

A viscosidade está intimamente relacionada com o peso por litro de barbotina, isto é, com a relação peso de sólidos, peso de água da barbotina. Esta deve ser controlada cuidadosamente pois que, se for excessiva, os moldes não ficarão devidamente cheios originando bolhas de ar, etc., e a barbotina não escorrerá perfeitamente, provocando ainda outros defeitos – cordões, grumos, etc.. Por outro lado se a viscosidade for muito reduzida, originará habitualmente um tempo de formação de parede excessivamente longo.

Quanto à tixotropia deve ser também controlada pois se for demasiado elevada origina peças demasiado sensíveis ao manuseamento por serem muito macias. Se a

tixotropia for muito baixa, pode provocar gretas, rachaduras e dureza excessiva das peças dificultando o acabamento das mesmas.

Não há valores absolutos para viscosidade e tixotropia de cada barbotina cerâmica. Há sim valores ótimos de trabalho que convém ao responsável pela preparação de barbotinas reproduzir dia a dia mantendo-os dentro de certos limites. Estes valores dependem de vários fatores tais como composições das barbotinas, tipo de enchimento a que se destinam, etc.; devem ser otimizados e, tanto quanto possível, reproduzidos. A título meramente informativo, podemos referir que uma barbotina para faiança se considera bem desfloculada se, no viscosímetro de torção, apresentar uma viscosidade de 290° - 310° e uma tixotropia ao fim de 3 minutos de 80° - 110°.

O viscosímetro de torção. Um aparelho que se compõe essencialmente de um cilindro metálico suspenso na suspensão amostra por um arame de torção. Este liga-se no seu extremo superior a uma cabeça de torção ajustável e situada na parte superior do instrumento. No extremo inferior do arame há um disco volante provido com uma escala 0° – 360°. A Amostra coloca-se no vaso debaixo do cilindro de modo que este fique submerso. A viscosidade determina-se rodando o disco volante de uma volta completa e soltando-o. Uma vez solto, este retrocede até à posição original, ultrapassando-a de uma certa distância. Esta distância, que pode medir-se em graus, depende então da viscosidade da suspensão e é tanto menor quanto maior for a viscosidade. Após um período de repouso determinado, o processo é repetido sendo a diferença entre as duas leituras o valor atribuído à tixotropia.

Conclusão. Há 3 parâmetros que devem ser controlados antes da utilização das pastas cerâmicas:

- Controlo granulométrico ou da eficiência da moagem
- Controlo do peso específico das barbotinas
- Controlo da viscosidade e tixotropia das barbotinas

6.3 Desenho Experimental

Poderíamos, como é evidente, focalizar a nossa experiência em vários pontos do processo da produção cerâmica mas decidimos centrar o nosso planeamento experimental na produção de uma pasta cerâmica líquida – barbotina - que satisfaça os

requisitos de qualidade e eficiência necessários à continuação do processo de conformação, secagem, acabamento, cozedura e vidrado de um determinado produto de “Louça de Porcelana”.

6.3.1 Variáveis dependentes

Definimos como função objectivo, o controlo das características reológicas, subdividido nas vertentes densidade, viscosidade e tixotropia das barbotinas a utilizar por via líquida (enchimento) e registar os valores de: **Densidade, Fluidiez.**

Densidade. As barbotinas são suspensões de matérias plásticas e não plásticas em água. A densidade permite medir a quantidade de água na barbotina: a densidade da barbotina é o peso de um quilo de barbotina. Quando dizemos que a densidade da barbotina é 1,7 queremos dizer que 1 litro pesa 1700g. Na cerâmica usam-se normalmente barbotinas com densidade entre [1.65, 1.75]. A densidade influencia o tempo de formação de parede e o potencial encharcamento dos moldes.

Fluidiez. A resistência que um fluido oferece ao movimento, provocada pela fricção interna entre camadas adjacentes chama-se viscosidade. A fluidez define-se como o inverso da viscosidade. Para um líquido verdadeiro, se a temperatura se mantiver constante, a viscosidade também é constante; mas quando se trata de uma suspensão, a viscosidade já não é constante, pois é afetada por perturbações mecânicas tais como a agitação. A viscosidade de uma suspensão de argila depende também da sua densidade: quanto maior for a densidade, maior será a viscosidade e vice-versa. O uso da fluidez, inversa da viscosidade, usado tradicionalmente na indústria cerâmica, é uma designação lógica pois fluidez é a facilidade com que se processa o movimento, isto é, a facilidade com que flui uma determinada suspensão. Associado à fluidez surge a propriedade reológica da tixotropia, com a qual está intimamente correlacionada. A capacidade de endurecimento das barbotinas quando em repouso denomina-se tixotropia. Esta propriedade é muito importante pois é um factor determinante da velocidade de formação de parede no fabrico de peças por enchimento. Se todos os factores se mantiverem constantes, quanto maior for a densidade de uma barbotina maior será a sua tixotropia. À medida que a fluidez aumenta a tixotropia diminui pelo que se deve utilizar um valor de fluidez que seja um compromisso entre o bom

MSR: Uma Investigação no âmbito da Indústria Cerâmica

escorrimento e uma tixotropia que garanta uma boa velocidade de formação de parede. Qualquer barbotina deve possuir uma viscosidade e tixotropia constantes, para que a peça moldada tenha propriedades uniformes.

Técnica de medição: Viscosímetro de torção. Antes de introduzir uma amostra deve-se agitá-la vigorosamente. A fluidez é lida ao fim da rotação máxima do disco; é depois feita uma segunda leitura após 5 minutos. A tixotropia é a diferença entre os valores da 1ª e da 2ª leitura.

Caraterísticas Reológicas	Defeito de Enchimento
Fluidez demasiado alta	<ul style="list-style-type: none">• Baixa velocidade de formação de parede• Rachas• Cordões
Fluidez demasiado baixa	<ul style="list-style-type: none">• Borbulhas• Mau escorrimento
Tixotropia demasiado alta	<ul style="list-style-type: none">• Paredes moles (gelatinosas)• Mau escoamento• Secagem muito lenta
Tixotropia demasiado baixa	<ul style="list-style-type: none">• Paredes frágeis• Baixa velocidade de formação de parede• Cordões• Rachas• Dificuldade no corte e apara

Quadro 14 - Defeito de Enchimento/ características reológicas

Adição	Fluidez	Tixotropia
Água	Aumenta	Diminui ligeiramente
Desfloculante	Aumenta	Diminui
Barro	Diminui	Aumenta

Quadro 15 - soluções para os defeitos de enchimento

6.3.2 Variáveis independentes

A seleção das variáveis independentes ficou subordinada ao espaço, ao tempo e à vontade de sermos pouco intrusivos na atividade laboral da pequena empresa que

MSR: Uma Investigação no âmbito da Indústria Cerâmica

acolheu a nossa experiência. Já observámos, na introdução, que uma pasta cerâmica é função de muitas variáveis nomeadamente da quantidade de materiais plásticos e não plásticos e sua moagem, da quantidade de água, de materiais desfloculantes e secantes, entre outros. Fixamos a nossa observação em dois destes elementos: a quantidade de água, elemento susceptível de medida exata quando calculada a % de água nos materiais húmidos e a água adicionada; a quantidade de desfloculante adicionado. Ambas as variáveis produzem efeitos reológicos suscetíveis de ser medidos com precisão: **Água e Desfloculante**

Água. Toda a barbotina contém uma determinada quantidade de água, uma parte que é adicionada, outra que vem incorporada nos materiais húmidos. O procedimento para calcular a % de humidade das matérias-primas: - pese 0,5kg de cada uma das matérias-primas e ponha-as a secar na estufa. Depois de secas pese-as de novo e registe os valores obtidos para cada uma delas num quadro tipo:

Matéria Prima	Peso Húmido	Peso Seco	% de Humidade	Data
Barro A	500g	468g	6,4	
Caulino K	500g	420g	16	

Quadro 16 - matérias-primas

Suponha que tem uma dada formulação de pasta, na sua fábrica, expressa em % de material seco. Como os materiais não estão todos secos é necessário fazer os acertos – suponhamos que para 1500kg.

Matéria Prima	Formulação	Carga de 1500kg	% de Humidade	Carga Húmida
Barro A	30%	450	6,4	481
Caulino K	28%	420	16	500
Areia moída	30%	450	0	450
Calcite moída	12%	180	0	180

Quadro 17 - cálculo da percentagem de humidade das matérias-primas

Sabe-se por experiência que para produzir uma barbotina com densidade de 1,7 se deve utilizar uma proporção aproximada de 4,8 litros de água para cada 10 kg de pasta seca. Assim para 1500 kg de pasta seca serão necessários 720 litros de água; mas

como se utilizam matérias-primas húmidas, uma parte desta água já vai incorporada nelas. A água a adicionar será somente a diferença entre a quantidade calculada e a quantidade já incorporada. Neste caso 111 litros já estavam incorporados, seriam necessários mais 609 litros.

Associada à quantidade de água está a quantidade de outros materiais que compõem a pasta cerâmica e que poderiam ser também objeto de tratamento como variáveis independentes. As matérias-primas não plásticas – areias, calcite, feldspato – são submetidas a um prévio tratamento de moagem. Para uma dada matéria-prima, à partida com uma dada granulometria, o tempo de moagem para obtenção da granulometria desejada na pasta, determina-se por tentativa de erro, isto é, arbitra-se um primeiro número de moagem, recolhe-se uma amostra e leva-se ao laboratório para analisar. Se o tamanho médio dos grãos for ainda muito grande, o laboratório arbitrará mais um dado número de horas de moagem, ao fim das quais se repetirá a análise. Determina-se desta forma um tempo de moagem que, no entanto está sujeito a pequenos ajustes em cada moagem. A análise referida será uma determinação do resíduo num peneiro com uma dada malha. A distribuição granulométrica de uma pasta tem grande importância pois condiciona as reações no estado sólido entre os diferentes componentes da pasta: quanto menores forem as partículas sólidas e portanto quanto maior for a superfície de contacto total entre estas, mais facilmente se verificará a reação entre elas e mais completa esta será. A dimensão das partículas, ao influenciar a superfície de contacto, **influencia a relação com a água**, condicionando a plasticidade, a viscosidade e a contração; mas influencia ainda a resistência em seco e a porosidade depois da cozedura; influencia também o tempo de formação de parede, pois um tamanho de partícula demasiado pequeno conduz a um tempo excessivo de formação de parede. Os resíduos ao fim de determinado tempo de moagem, devem situar-se na ordem dos 1% a 3% do peso seco da amostra.

Desfloculantes. A fluidez de uma barbotina é grandemente afetada pelos desfloculantes: uma barbotina sem desfloculantes tem uma fluidez muito baixa, fluidez esta que, por adição de um desfloculante normal, aumenta até um valor máximo (ótimo), mas diminuindo se a adição continuar a partir desse ponto. A fluidez poderia ser aumentada, diminuindo a sua densidade, adicionando água; mas este processo

prolongaria o tempo de formação de parede e daria origem a paredes muito finas, além de encharcar os moldes. É por isso que todas as barbotinas são desfloculadas.

Para uma dada composição e densidade, a fluidez ótima deve ser determinada previamente através da execução da curva de desfloculação da barbotina. Esta obtém-se por adições sucessivas de desfloculante e determinação da fluidez a cada adição. Estas curvas podem ter formas variadas: Exemplo abaixo,

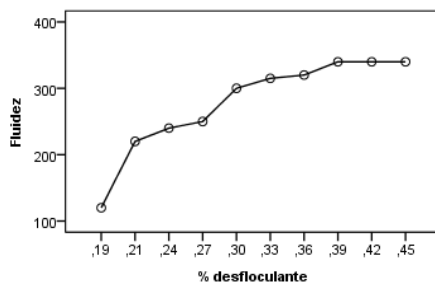


Gráfico 6 – Curva 1 desfloculação da barbotina 1

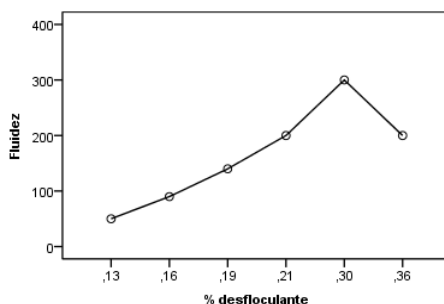


Gráfico 7 – Curva 2 desfloculação da barbotina 2

Se a curva de desfloculação da barbotina for do primeiro tipo, torna-se mais prudente trabalhar com uma % de desfloculante à volta dos 0,25% pois qualquer pequeno engano pode dar início à floculação; se for do segundo tipo, pode-se trabalhar com 0,3% de desfloculante que dará barbotinas com uma fluidez à volta de 300°.

A execução da curva de desfloculação da barbotina obtém-se experimentalmente por adições sucessivas de desfloculante e determinação da viscosidade (fluidez) a cada adição. Normalmente trabalha-se com [0,25%, 0,3%] de desfloculante que dão barbotinas com viscosidade próxima dos 300°.

7. Superfícies de Resposta: Método e Planeamento

O delineamento experimental envolve a realização de um conjunto de experiências que, dependendo do problema, da forma e da complexidade, visam a representatividade de uma dada questão a que se tenta dar resposta. O método das superfícies de resposta assenta em delineamentos e modelos para trabalhar variáveis contínuas na busca do ótimo; quando há mais de uma resposta, então é importante encontrar o equilíbrio, não otimizando uma só resposta; havendo restrições, cumpre que o projeto experimental as satisfaça. Um outro objectivo é compreender as mudanças de resposta numa dada direção ajustando as variáveis do projeto. Normalmente a superfície de resposta pode ser visualizada graficamente quer através de gráficos 3D quer pelas suas projeções no plano, as curvas de nível.

O primeiro gráfico, a 3 dimensões, que a seguir apresentamos, mostra a superfície de resposta onde cada valor de x_1 e x_2 gera um valor para y . O segundo é um gráfico de contornos onde cada contorno corresponde a uma altura particular da superfície de resposta: trata-se de uma projeção no plano onde em cada linha a resposta é constante. Estes gráficos foram obtidos no software estatístico MiniTab e têm por base os dados recolhidos ao longo da execução do projeto, durante a preparação da pasta cerâmica produzida. A produção normalizada de 1000 kg desta pasta cerâmica específica exige o consumo aproximado de 325 litros de água e 3 kg de desfloculante, que constituem o ponto central da nossa experiência. Os níveis, baixo e superior, com base no ponto central, fizemo-los variar: a quantidade de água entre 300 e 350 l e o desfloculante entre 2 e 4 kg.

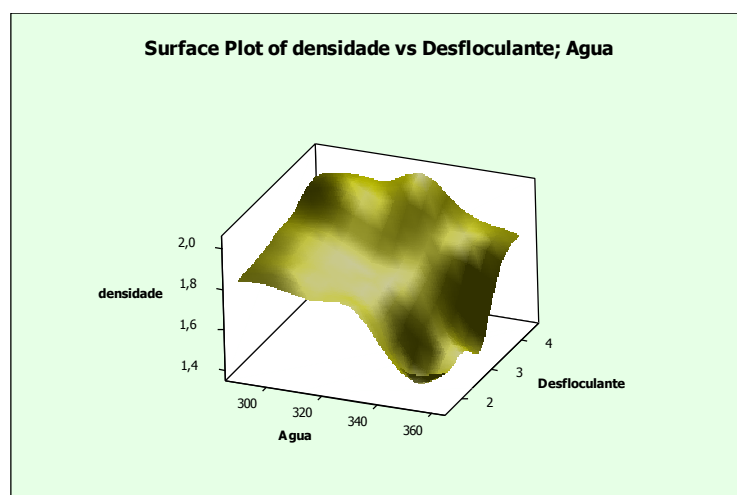


Gráfico 8 – Superfície de Resposta Densidade

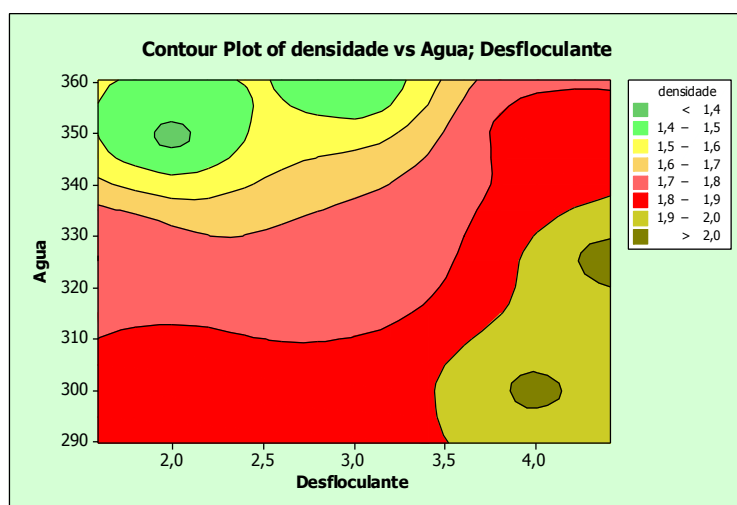


Gráfico 9 - Curvas de nível densidade

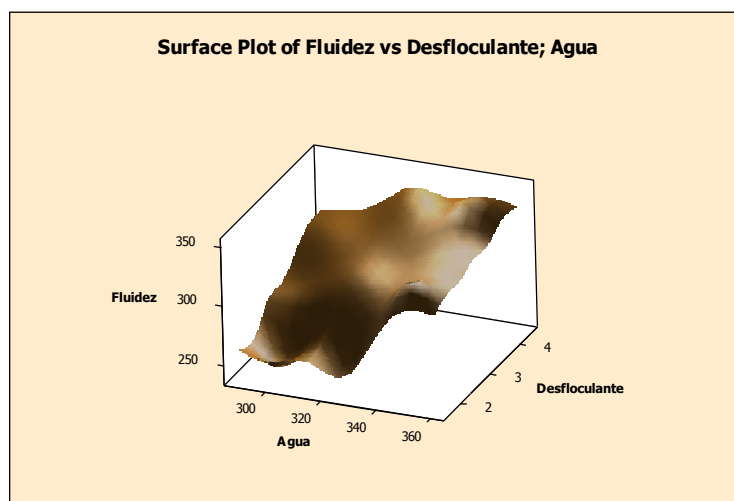


Gráfico 10 - Superfície de Resposta Fluidiez

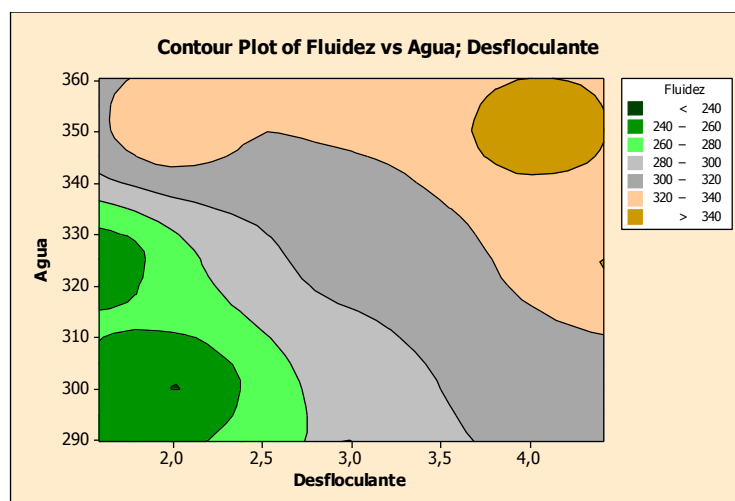


Gráfico 11 - Curvas de nível fluidiez

MSR: Uma Investigação no âmbito da Indústria Cerâmica

Uma abordagem comum em qualquer MSR consiste em definir uma referência experimental estandardizada e depois realizar, em seu torno, novas e representativas experiências. Vamos começar com uma experiência simples, envolvendo dois factores para percebermos o algoritmo do processo.

Sabemos que há muitos parâmetros que podem afetar a produção de uma pasta cerâmica, nomeadamente a quantidade de água, a composição química dos materiais plásticos e não plásticos utilizados, o tempo de moagem, a quantidade de desfloculantes e secantes entre outros, mas definimos como variáveis independentes apenas duas: a quantidade de água (x_1) e a quantidade de desfloculante (x_2); e como variável dependente as características reológicas da pasta cerâmica (y) (que decompos em (y_1) densidade e (y_2) fluidez).

Os 1000 kg de pasta cerâmica, que pretendemos produzir, destinam-se ao fabrico de “chávenas de porcelana”. Foi definido laboratorialmente que as características reológicas ideais para este processo se cifram em densidade = 1.7, fluidez = 300° e tixotropia = 80°. As amostras foram recolhidas em dois dias sucessivos e a pasta seca utilizada foi a mesma.

A nossa função-objectivo é definida pela distância dos valores experimentalmente encontrados a estes valores pré-fixados. Considerando constantes todos os restantes factores, vamos observar o efeito provocado pelas variações das quantidades de água e desfloculante em cada uma das propriedades reológicas e tentar obter o valor mínimo da função objetivo.

O ponto central, correspondente à produção normalizada, corresponde a 325 litros de água e 3 kg de desfloculante. Definimos os níveis baixo e superior da nossa experiência com base no ponto central, decidindo fazer variar a quantidade de água entre 300 e 350 l e a de desfloculante entre 2 e 4 kg.

Adotamos o modelo fatorial 2^2 aumentado com 3 pontos centrais. Temos assim dois factores: água x_1 e desfloculante x_2 ; como variável dependente as características reológicas da pasta cerâmica (y).

MSR: Uma Investigação no âmbito da Indústria Cerâmica

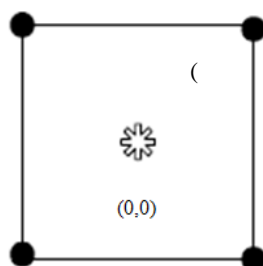


Gráfico 12 - Planejamento fatorial 2^2 com pontos centrais

Os dados recolhidos na experiência programada encontram-se descritos no seguinte quadro:

ID	Variáveis naturais		densidade $y_1 = d-1,7 $	Fluidez $y_2 = f-300 $
	Água	Desfloculante		
1	300	2	0,15	60
2	350	2	0,31	30
3	300	4	0,31	10
4	350	4	0,13	50
5	325	3	0,11	10
6	325	3	0,02	10
7	325	3	0,08	0
8	325	3	0,03	20
9	325	3	0,05	10
10	325	3	0,1	10
11	325	4,414	0,32	40
12	325	1,586	0,1	50
13	360,355	3	0,28	30
14	289,645	3	0,18	20

Quadro 18 - Registo de observações reais

O quadro seguinte mostra as estatísticas descritivas das variáveis resposta, que denominamos Densidade e Fluidez mas que representam as distâncias aos valores pré-fixados de Densidade=1,7 e Fluidez=300.

MSR: Uma Investigação no âmbito da Indústria Cerâmica

		Statistic	Std. Error	
Densidade	Mean	,1550	,02878	
	95% Confidence Interval for Mean	Lower Bound	,0928	
		Upper Bound	,2172	
	5% Trimmed Mean	,1533		
	Median	,1200		
	Variance	,012		
	Std. Deviation	,10769		
	Minimum	,02		
	Maximum	,32		
	Range	,30		
	Interquartile Range	,22		
	Skewness	,537	,597	
	Kurtosis	-1,209	1,154	
Fluidez	Mean	25,0000	5,00000	
	95% Confidence Interval for Mean	Lower Bound	14,1982	
		Upper Bound	35,8018	
	5% Trimmed Mean	24,4444		
	Median	20,0000		
	Variance	350,000		
	Std. Deviation	18,70829		
	Minimum	,00		
	Maximum	60,00		
	Range	60,00		
	Interquartile Range	32,50		
	Skewness	,617	,597	
	Kurtosis	-,865	1,154	

Quadro 19 - Estatísticas descritivas

Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Densidade	,163	14	,200	,881	14	,060
Fluidez	,217	14	,072	,899	14	,109

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Quadro 20 - Testes de normalidade

Não rejeitamos a hipótese de normalidade das variáveis Densidade (distância ao valor de Densidade=1,7) e Fluidez (distância ao valor de Fluidez=300); a primeira

apresenta uma distância média ao valor pré-fixado que se cifra em 0,155 com um desvio padrão de 0,11 e Erro standardizado de 0,03 com uma distribuição simétrica ($g_1/SEg_1 = 0,9$) e mesocúrtica ($k/SEk = 1,05$); a segunda apresenta uma distância média ao valor pré-fixado de Fluidez que se cifra em 25° com desvio padrão de 18,7° e Erro standardizado de 5° com uma distribuição simétrica e mesocúrtica.

Procedemos de seguida à codificação das variáveis. O desenho de superfície de resposta mais popular e que usamos no caso de estudo é o ccd. Os blocos em ccd são de dois tipos: os blocos “*cube*” que contêm os pontos de um factorial 2^K mais os pontos centrais; os blocos “*star*” que contêm os pontos axiais, mais pontos centrais. Os pontos centrais situam-se em (0,0), os pontos fatoriais têm coordenadas ± 1 , os pontos axiais $\pm \alpha$. O valor de α , a escolha dos pontos a replicar e dos pontos centrais são normalmente selecionados em função de necessidades de rotatividade e ortogonalidade do modelo. A rotatividade está associada à variância do valor predito que depende somente da sua distância ao centro; ortogonalidade dos blocos está associada aos coeficientes da equação da superfície de resposta evitando que estes estejam correlacionados com os efeitos dos blocos.

7.1 Codificação dos dados e análise

A codificação dos dados é o primeiro passo e constitui um elemento importante na análise das superfícies de resposta. A forma como os dados são codificados afecta o resultado da análise canónica e da *steepest-descent path*. Usando um método de codificação que faça com que todas as variáveis codificadas da experiência variem de acordo com o mesmo padrão, impomos a cada preditor a mesma capacidade de determinar o percurso até à superfície de resposta.

O ponto central, correspondente à produção normalizada, equivalente ao consumo de 325 l de água e 3 kg de desfloculante para a produção de 1000 kg de pasta cerâmica. Definimos os níveis baixo e superior da nossa experiência com base no ponto central, decidindo fazer variar a quantidade de água entre 300 e 350 litros e o desfloculante entre 2 e 4 kg. A codificação $x_1 = \frac{F-325}{25}$; $x_2 = \frac{M-3}{1}$

MSR: Uma Investigação no âmbito da Indústria Cerâmica

ID	Bloco	Variáveis naturais		Variáveis codificadas		densidade y1= d-1,7	Fluidez y2= f-300°
		Água	Desfloculante	A	D		
1	B1	300	2	-1	-1	0,15	60
2	B1	350	2	1	-1	0,31	30
3	B1	300	4	-1	1	0,31	10
4	B1	350	4	1	1	0,13	50
5	B1	325	3	0	0	0,11	10
6	B1	325	3	0	0	0,02	10
7	B1	325	3	0	0	0,08	0
8	B2	325	3	0	0	0,03	20
9	B2	325	3	0	0	0,05	10
10	B2	325	3	0	0	0,1	10
11	B2	325	4,414	0	1,414	0,32	40
12	B2	325	1,586	0	-1,414	0,1	50
13	B2	360,355	3	1,414	0	0,28	30
14	B2	289,645	3	-1,414	0	0,18	20

Quadro 21 – Observações Codificadas (1)

Nesta experiência, os dados do bloco B1 foram os primeiros a ser colhidos e analisados, após o que foi acrescentado o Bloco B2 para ser feita uma nova análise. As seis experiências, correspondentes aos bolos 5, 6, 7, 8, 9 e 10 servem para determinar o erro experimental.

Vamos agora analisar os dados para descobrir que factores mais influenciam a resposta. A modelação é feita por um modelo polinomial do primeiro grau ajustado aos dados:

$$\hat{y} = \hat{\beta}_0 + \sum_{i=1}^k \hat{\beta}_i x_i$$

Recorremos aos programas informáticos, SPSS e Minitab, donde obtivemos os valores que a seguir reproduzimos:

```

Response Surface Regression: densidade versus A; D

The analysis was done using coded units.
Estimated Regression Coefficients for densidade

Term          Coef  SE Coef    T      P
Constant    0,15500  0,02997   5,172  0,000
A            0,02146  0,05606   0,383  0,709
D            0,05146  0,05606   0,918  0,378

S = 0,112134  PRESS = 0,244177
    
```

MSR: Uma Investigação no âmbito da Indústria Cerâmica

R-Sq = 8,25% R-Sq(pred) = 0,00% R-Sq(adj) = 0,00%

Analysis of Variance for densidade

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	2	0,012435	0,012435	0,006218	0,49	0,623
Linear	2	0,012435	0,012435	0,006218	0,49	0,623
A	1	0,001843	0,001843	0,001843	0,15	0,709
D	1	0,010593	0,010593	0,010593	0,84	0,378
Residual Error	11	0,138315	0,138315	0,012574		
Lack-of-Fit	6	0,131365	0,131365	0,021894	15,75	0,004
Pure Error	5	0,006950	0,006950	0,001390		
Total	13	0,150750				

Quadro 22 - Cálculo dos coeficientes de regressão linear para “densidade”

Response Surface Regression: Fluidez versus A; D

The analysis was done using coded units.
Estimated Regression Coefficients for Fluidez

Term	Coef	SE Coef	T	P
Constant	25,000	5,243	4,768	0,001
A	4,267	9,808	0,435	0,672
D	-7,803	9,808	-0,796	0,443

S = 19,6181 PRESS = 7625,91
R-Sq = 6,95% R-Sq(pred) = 0,00% R-Sq(adj) = 0,00%

Analysis of Variance for Fluidez

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	2	316,43	316,43	158,22	0,41	0,673
Linear	2	316,43	316,43	158,22	0,41	0,673
A	1	72,85	72,85	72,85	0,19	0,672
D	1	243,58	243,58	243,58	0,63	0,443
Residual Error	11	4233,57	4233,57	384,87		
Lack-of-Fit	6	4033,57	4033,57	672,26	16,81	0,004
Pure Error	5	200,00	200,00	40,00		
Total	13	4550,00				

Quadro 23 - Cálculo dos coeficientes de regressão linear para “fluidez”

Verificamos que o modelo de primeira ordem não se ajusta aos dados, não sendo significativa a regressão linear em nenhuma das duas situações. O modelo de regressão linear aplica-se em condições experimentais onde a natureza dos erros, ε_i tem distribuição normal de esperança 0 e variância σ^2 ; a linearidade do modelo não se presta a confusão, aqui o modelo não é linear e a relação entre as variáveis não é igualmente linear. A regressão linear significa encontrar a linha reta que melhor se ajuste aos dados ou seja encontrar a linha reta para a qual as diferenças entre os valores reais de Y_i e os valores previstos da linha de regressão ajustada \hat{Y}_i sejam os menores possíveis. O método dos mínimos quadrados é o mais vulgarmente utilizado na estimação dos

parâmetros. Os quatro principais pressupostos da regressão são: Normalidade, Homocedasticidade, Independência dos erros e Linearidade a que acresce, na regressão linear múltipla, a não multicolinearidade.

No caso da “densidade”, o Teste F = 4,94 com p = 0,623, traduz que o modelo linear não é significativo; para além da constante, também os parâmetros estimados não são significativamente diferentes de zero. A verificação dos pressupostos é normalmente feita (completada) com a análise de resíduos. A análise dos resíduos é uma técnica de leitura da qualidade do ajustamento que também nos permite estudar potenciais violações dos pressupostos do modelo de regressão. Lembremos que ε é uma variável aleatória de esperança nula e variância σ^2 e, assim sendo, os resíduos devem ter um comportamento caótico no sentido de não serem detetáveis quaisquer regularidades, tendências ou padrões.

Uma estimativa do erro amostral pode ser obtida através das 6 repetições que incluímos no modelo. $\hat{\sigma}^2 = \frac{(0,11)^2 + (0,02)^2 + (0,08)^2 + (0,03)^2 + (0,05)^2 + (0,1)^2 - \frac{(0,39)^2}{6}}{6-1} = 0,00139$ (Valor igualmente inscrito no “pure error” do teste lack of fit).

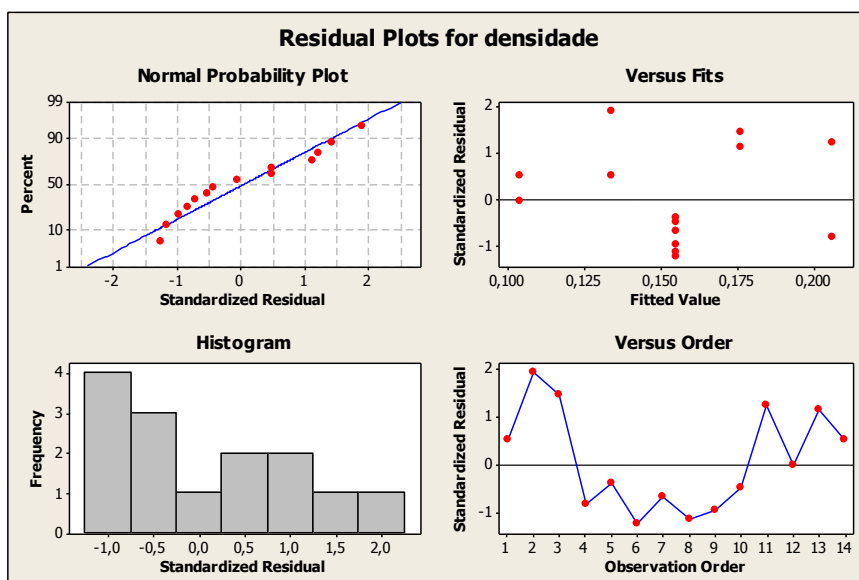


Gráfico 13 - Resíduos da regressão linear “densidade”

A análise dos resíduos através dos gráficos leva-nos também a concluir pelo não ajustamento do modelo: no primeiro gráfico verificamos que os desvios à normal

apresentam uma dada estrutura, estando alguns dos valores afastados da reta; o histograma revela uma maior concentração dos dados à esquerda, assimetria positiva e a distribuição dos resíduos em torno de zero revela um padrão.

Uma forma de nos pronunciarmos, com acuidade, sobre qualquer falta de ajustamento dos dados, consiste na realização do teste “*lack of fit*”, que também não revela significância. A inclusão dos pontos centrais no modelo tem como um dos objetivos a capacidade de proporcionar elementos à realização deste teste.

Para completar esta análise de ajustamento da reta ao modelo falta calcular uma estimativa de mínimos quadrados do coeficiente da interação.

Response Surface Regression: densidade versus Agua_r; Desfloculante_r

The analysis was done using coded units.

Estimated Regression Coefficients for densidade

Term	Coef	SE Coef	T	P
Constant	0,15500	0,02796	5,544	0,000
Agua_r	0,01518	0,03699	0,410	0,690
Desfloculante_r	0,03639	0,03699	0,984	0,348
Agua_r*Desfloculante_r	-0,08500	0,05230	-1,625	0,135

S = 0,104602 PRESS = 0,273328

R-Sq = 27,42% R-Sq(pred) = 0,00% R-Sq(adj) = 5,65%

Analysis of Variance for densidade

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	3	0,041335	0,041335	0,013778	1,26	0,340
Linear	2	0,012435	0,012435	0,006218	0,57	0,584
Agua_r	1	0,001843	0,001843	0,001843	0,17	0,690
Desfloculante_r	1	0,010593	0,010593	0,010593	0,97	0,348
Interaction	1	0,028900	0,028900	0,028900	2,64	0,135

Quadro 24 - Resíduos da regressão linear múltipla “densidade”

Conclusão: o efeito da interação, no que concerne à densidade, não foi significativo, tal como o efeito linear não é significativo. Também na fluidez o efeito de interação não é significativo. O modelo de primeira ordem não se revela ajustado em nenhuma das variáveis resposta como indica a significância do *Lack-of-Fit*.

7.2 Direção da máxima inclinação descendente

Quando estamos distantes do ótimo, assumimos um modelo de primeira ordem como aproximação da verdadeira superfície de resposta em uma pequena região das variáveis independentes (x_i): se procurássemos o máximo incremento na resposta,

MSR: Uma Investigação no âmbito da Indústria Cerâmica

teríamos o método da máxima inclinação ascendente (steepest ascent); buscamos um ponto de mínimo, o método chama-se máxima inclinação descendente (steepest descent). Os gráficos seguintes, superfícies de resposta e respetivas curvas de nível, foram construídos com recurso ao software estatístico MiniTab e têm por base os valores codificados das variáveis Água e Desfloculante e os valores das distâncias aos valores pré-fixados de densidade e fluidez. Objetivamente procuramos um mínimo, já que pretendemos que o ótimo se aproxime o mais possível da distância zero.

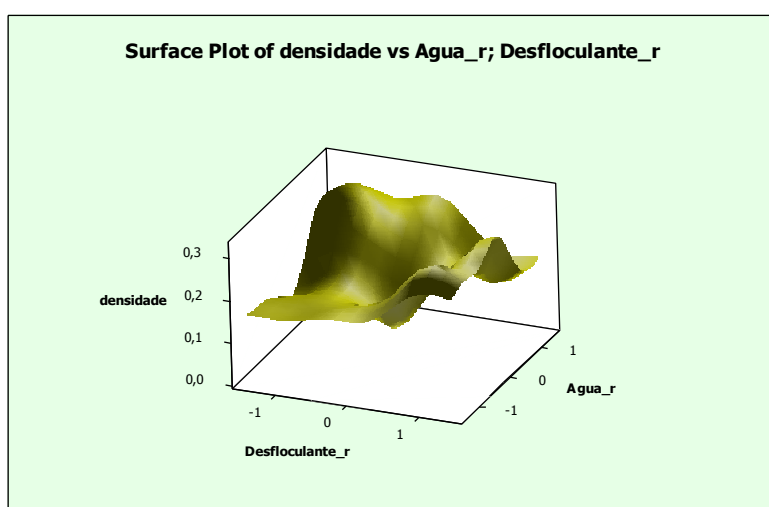


Gráfico 14 – Superfície de Resposta variável densidade

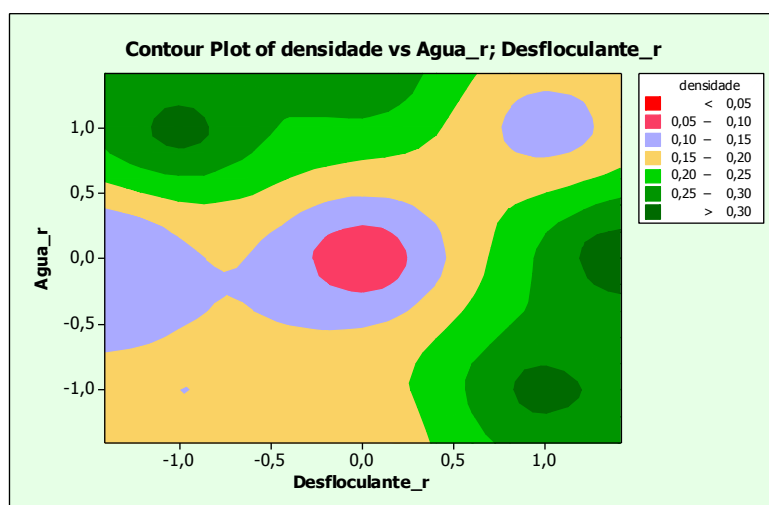


Gráfico 15 - Curvas de nível variável densidade

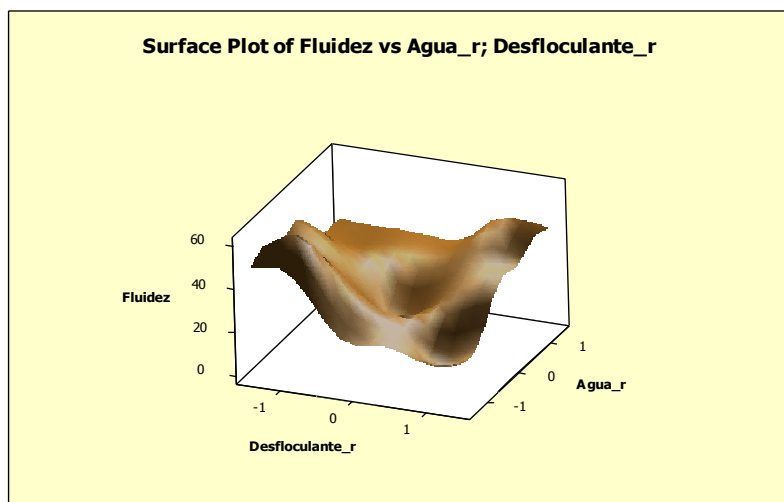


Gráfico 16- Superfície de Resposta variável Fluidiez

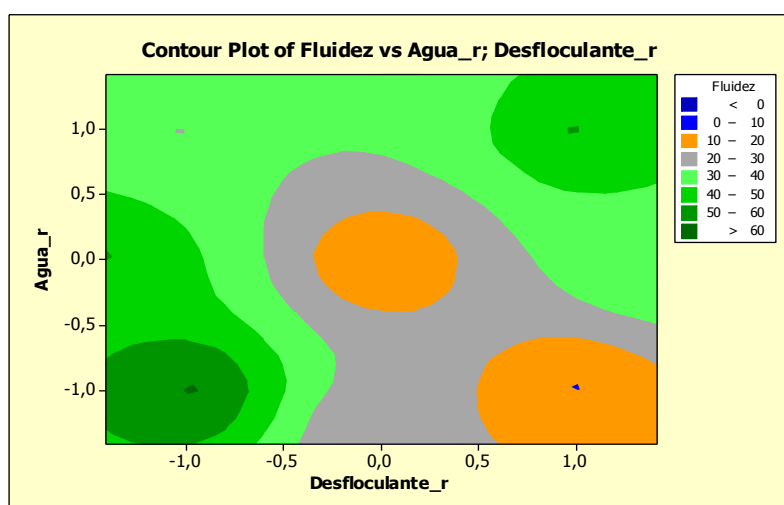


Gráfico 17 - Curvas de nível variável fluidiez

Os gráficos dão-nos uma ideia do caminho que vamos agora começar a percorrer em direção aos valores mínimos da resposta.

O gráfico de contornos dos valores preditos da variável resposta é, quando assumimos o modelo de primeira ordem, uma série de linhas paralelas. O caminho de inclinação descendente – linha perpendicular às paralelas – é a direção em que os valores ajustados diminuem mais rapidamente. Os passos ao longo do caminho são proporcionais aos sinais e grandeza dos coeficientes de regressão β_i , sendo o tamanho

real do passo definido pelo pesquisador. Os tratamentos são conduzidos ao longo do trajeto descendente até que não ocorram mais diminuições na variável resposta; chegados a esta área, ou ajustamos um novo modelo de primeira ordem ou realizamos experiências adicionais, particularmente ajustando um polinómio de ordem dois, para obter uma estimativa mais precisa do ótimo.

Observando agora o caso da fluidez. Para nos movermos do centro do delineamento ($x_1=0$ e $x_2=0$) no caminho da inclinação descende, deveremos mover (+4,267) unidades na direção x_1 para cada (-7,803) unidades na direção de x_2 . Assim, a reta da inclinação descendente contém o ponto central ($x_1=0$ e $x_2=0$) e tem inclinação $\frac{-7,803}{4,267} = -1,83$. Decidimos usar a unidade 10 (litros de água) como tamanho do passo inicial que correspondem na variável codificada a $\Delta x_1 = \frac{10}{25} = 0,4$. Assim os passos no caminho da inclinação descendente são: $\Delta x_1 = 0,4$ $\Delta x_2 = \frac{-7,803}{4,267} * \Delta x_1 = -0,731$.

Aqui é feito um ponto de ordem para explicar ao operacional ceramista que, nesta fase, seria importante procedermos a mais experiências que nos levassem ao caminho mais curto na busca do ótimo: na produção da pasta deveríamos experimentar a mistura progressiva de mais 10 litros de água de cada vez e simultaneamente retirar 0,731kg de desfloculante, medindo o resultado, mas tal seria inviável sem começarmos a experiência novamente com novos valores. Teoricamente deveríamos calcular os pontos experimentais e medir o grau de “fluidez” para cada um dos decréscimos a observar até que houvesse um acréscimo. Como o modelo de primeira ordem se revelou desajustado, prescindimos desta fase experimental, dado o incómodo e acréscimo de trabalho que provocaria no laboratório.

7.3 Conclusões do modelo de primeira ordem

Um modelo de primeira ordem usa os recursos da regressão múltipla com polinómios de grau 1 para descrever alguma parte da superfície de resposta; é um modelo apropriado para descrever superfícies planas, com ou sem inclinação.

A regressão múltipla permite modelar as observações e, pelo método dos mínimos quadrados, permite o cálculo de estimadores para os parâmetros do modelo e a formulação da equação correspondente.

MSR: Uma Investigação no âmbito da Indústria Cerâmica

Conhecida a equação de regressão, o investigador fica capacitado para observar os efeitos principais, para o cálculo das interações, para se pronunciar sobre a adequação do modelo através de ANOVA (teste F, t-test, R^2 e R^2 ajustado), da significância dos estimadores, da análise de resíduos e do teste lack of fit que o conjunto de observações de pontos centrais permite. Quando o modelo de primeira ordem se revela desajustado, o investigador deve prosseguir a investigação através de modelos mais estruturados como por exemplo o modelo de segunda ordem.

8. Análise de uma superfície de resposta de segunda ordem

Quando existe uma curvatura na superfície de resposta, o modelo de primeira ordem revela-se desajustado. O modelo de segunda ordem inclui todos os termos do modelo de primeira ordem mais os termos quadráticos e todas as interações (*cross product terms*):

$$y = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_{ii} x_i^2 + \sum \sum \beta_{ij} x_i x_j + \varepsilon$$

ou

$$y = \beta_0 + x_i^2 \beta + x_i^2 \beta x_i + \varepsilon_{ij}$$

Existem muitos delineamentos para a implementação de um modelo de segunda ordem. O mais generalizado é o delineamento CCD – *central composite design*, introduzido por Box e Wilson que consiste em pontos fatoriais (a partir de um planeamento fatorial 2^n ou fatorial fracionado 2^{n-k}), pontos centrais e pontos axiais.

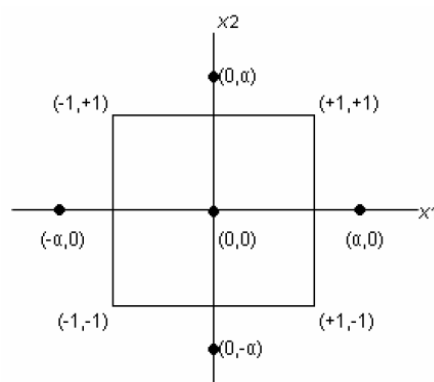


Gráfico 18 - Planeamento fatorial 2^2 com pontos centrais e axiais

$\{(-1,-1);(-1,+1);(+1,+1);(+1,-1)\}$ – 2^2 pontos fatoriais. $\{(0,0)\}$ – pontos centrais.
 $\{(-a,0); (0,+a); (+a,0); (0,-a)\}$ – 2×2 pontos axiais.

Quando o modelo de primeira ordem evidencia *lack of fit* (falta de ajustamento), podemos acrescentar aos pontos fatoriais, os pontos axiais e mais pontos centrais para desenvolver CCD. Quando a curvatura da superfície é significativa, os pontos axiais adicionados permitem, ao investigador, uma estimativa eficiente dos termos quadráticos.

O objectivo agora, neste modelo de segunda ordem, é encontrar os níveis de x_1, x_2, \dots, x_k que maximizem a resposta. Este ponto, se existir, será (x_1, x_2, \dots, x_k) , para o qual as derivadas parciais são iguais a zero, $\frac{\delta \hat{y}}{\delta x_1} = \frac{\delta \hat{y}}{\delta x_2} = \dots = \frac{\delta \hat{y}}{\delta x_k} = 0$ e é chamado “ponto de estacionaridade”. O ponto de estacionaridade pode representar um máximo, um mínimo ou um ponto de sela.

O modelo de segunda ordem escrita na forma matricial fica: $\hat{y} = \hat{\beta}_0 + x'b + x'Bx$

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_k \end{bmatrix} \quad b = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \cdot \\ \cdot \\ \cdot \\ \hat{\beta}_k \end{bmatrix} \quad B = \begin{bmatrix} \hat{\beta}_{11} & \hat{\beta}_{12}/2 & \cdot & \cdot & \cdot & \hat{\beta}_{1k}/2 \\ & \hat{\beta}_{22} & \cdot & \cdot & \cdot & \hat{\beta}_{2k}/2 \\ & & \cdot & & & \\ & & & \cdot & & \\ & & & & \cdot & \\ \text{sim.} & & & & & \hat{\beta}_{kk} \end{bmatrix}$$

b é um vetor ($k \times 1$) dos coeficientes de regressão de primeira ordem e B é uma matriz simétrica ($k \times k$) que, na diagonal, tem os coeficientes quadráticos de regressão de segunda ordem e fora da diagonal a metade de cada coeficiente quadrático da interação $\hat{\beta}_{ij}/2$ com $i \neq j$.

As derivadas parciais dos valores estimados da resposta (\hat{y}) em relação a x são igualadas a zero: $\frac{\delta \hat{y}}{\delta x} = b + 2Bx = 0$

O ponto estacionário é a solução das equações: $x_s = -\frac{1}{2} B^{-1} b$

O valor predito da variável resposta no ponto estacionário é: $\hat{y}_s = \hat{\beta}_0 + \frac{1}{2} x'_s b$

Matematicamente, a forma de encontrar um valor mínimo de uma função passa por obter a derivada de primeira ordem dessa função, procurando os seus zeros e verificando que os valores obtidos na sua proximidade passam de negativos a positivos.

8.1 Caso de estudo: Modelo de 2ª ordem

Vamos continuar com a análise do processo de produção cerâmica, procurando os níveis de densidade e fluidez ótimos (segunda fase do estudo), pois já verificamos que o modelo de primeira ordem se revela desajustado e tal se deve à possível curvatura

MSR: Uma Investigação no âmbito da Indústria Cerâmica

da superfície de resposta: para atingirmos ótimo teremos de encontrar a expressão da curva que nos permita lá chegar.

Para ajustar o modelo, decidimos usar as observações do Bloco B2: 7 observações adicionais - 4 pontos axiais e 3 pontos centrais. Os quatro pontos axiais foram: $(x_1=0; x_2=+1,414)$; $(0; -1,414)$; $(1,414; 0)$; $(-1,414; 0)$. O delineamento completo é mostrado na tabela e figura a seguir e denomina-se, como vimos, delineamento central composto (CCD).

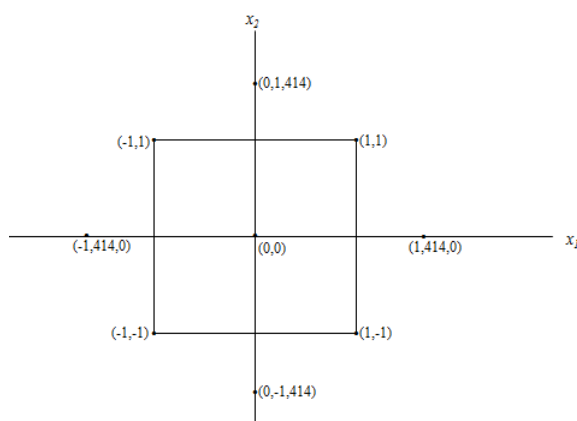


Gráfico 19 - Planejamento fatorial Modelo

Vamos usar análise de variância, análise de regressão e o teste lack of fit para examinar o modelo de regressão polinomial de segunda ordem que pretendemos ajustar. Os resultados, obtidos em Minitab, são mostrados nas tabelas seguintes:

Response Surface Regression: Densidade versus Agua_r; Desfloculante_r

The analysis was done using coded units.

Estimated Regression Coefficients for densidade

Term	Coef	SE Coef	T	P
Constant	0,06500	0,02234	2,909	0,020*
Agua_r	0,02146	0,02736	0,784	0,455
Desfloculante_r	0,05146	0,02736	1,880	0,097
Agua_r*Agua_r	0,16748	0,04028	4,158	0,003**
Desfloculante_r*Desfloculante_r	0,14748	0,04028	3,662	0,006**
Agua_r*Desfloculante_r	-0,16995	0,05471	-3,106	0,015*

S = 0,0547311 PRESS = 0,130996

R-Sq = 84,10% R-Sq(pred) = 13,10% R-Sq(adj) = 74,17%

Analysis of Variance for densidade

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	5	0,126786	0,126786	0,025357	8,47	0,005
Linear	2	0,012435	0,012435	0,006218	2,08	0,188
Agua_r	1	0,001843	0,001843	0,001843	0,62	0,455
Desfloculante_r	1	0,010593	0,010593	0,010593	3,54	0,097
Square	2	0,085451	0,085451	0,042725	14,26	0,002**
Agua_r*Agua_r	1	0,045292	0,051790	0,051790	17,29	0,003**
Desfloculante_r*Desfloculante_r	1	0,040159	0,040159	0,040159	13,41	0,006**
Interaction	1	0,028900	0,028900	0,028900	9,65	0,015*

MSR: Uma Investigação no âmbito da Indústria Cerâmica

Agua_r*Desfloculante_r	1	0,028900	0,028900	0,028900	9,65	0,015*
Residual Error	8	0,023964	0,023964	0,002995		
Lack-of-Fit	3	0,017014	0,017014	0,005671	4,08	0,082
Pure Error	5	0,006950	0,006950	0,001390		
Total	13	0,150750				

Estimated Regression Coefficients for densidade using data in uncoded units

Term	Coef
Constant	0,0649995
Agua_r	0,0151773
Desfloculante_r	0,0363905
Agua_r*Agua_r	0,0837638
Desfloculante_r*Desfloculante_r	0,0737608
Agua_r*Desfloculante_r	-0,0850000

Quadro 25- Cálculo dos coeficientes de regressão quadrática para “densidade”

Response Surface Regression: Fluidez versus Agua_r; Desfloculante_r

The analysis was done using coded units.
Estimated Regression Coefficients for Fluidez

Term	Coef	SE Coef	T	P
Constant	10,000	2,264	4,417	0,002
Agua_r	4,267	2,772	1,539	0,162
Desfloculante_r	-7,803	2,772	-2,815	0,023
Agua_r*Agua_r	16,246	4,081	3,981	0,004
Desfloculante_r*Desfloculante_r	36,246	4,081	8,882	0,000
Agua_r*Desfloculante_r	34,989	5,543	6,312	0,000

S = 5,54513 PRESS = 615,027
R-Sq = 94,59% R-Sq(pred) = 86,48% R-Sq(adj) = 91,21%

Analysis of Variance for Fluidez

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	5	4304,01	4304,01	860,80	27,99	0,000
Linear	2	316,43	316,43	158,22	5,15	0,037*
Agua_r	1	72,85	72,85	72,85	2,37	0,162
Desfloculante_r	1	243,58	243,58	243,58	7,92	0,023*
Square	2	2762,58	2762,58	1381,29	44,92	0,000**
Agua_r*Agua_r	1	336,69	487,37	487,37	15,85	0,004**
Desfloculante_r*Desfloculante_r	1	2425,88	2425,88	2425,88	78,89	0,000**
Interaction	1	1225,00	1225,00	1225,00	39,84	0,000**
Agua_r*Desfloculante_r	1	1225,00	1225,00	1225,00	39,84	0,000**
Residual Error	8	245,99	245,99	30,75		
Lack-of-Fit	3	45,99	45,99	15,33	0,38	0,770
Pure Error	5	200,00	200,00	40,00		
Total	13	4550,00				

Estimated Regression Coefficients for Fluidez using data in uncoded units

Term	Coef
Constant	9,99975
Agua_r	3,01796
Desfloculante_r	-5,51833
Agua_r*Agua_r	8,12569
Desfloculante_r*Desfloculante_r	18,1287
Agua_r*Desfloculante_r	17,5000

Quadro 26 - Cálculo dos coeficientes de regressão quadrática para “fluidez”

MSR: Uma Investigação no âmbito da Indústria Cerâmica

Estes resultados, R^2 ajustado $> 74,17$, significa mais que 74,17% da variabilidade total em y (ao nível de densidade e fluidez) é explicada pelas variáveis independentes presentes no modelo de regressão ajustado.

Verificamos que o modelo de regressão polinomial do segundo grau é significativo e, pelo teste lack of fit, concluímos que não há evidências que levem à sua rejeição.

Verificamos que todos os parâmetros dos modelos quadráticos são significativamente diferentes de zero.

Uma observação da distribuição dos resíduos vem confirmar a opinião positiva sobre o ajustamento do modelo.

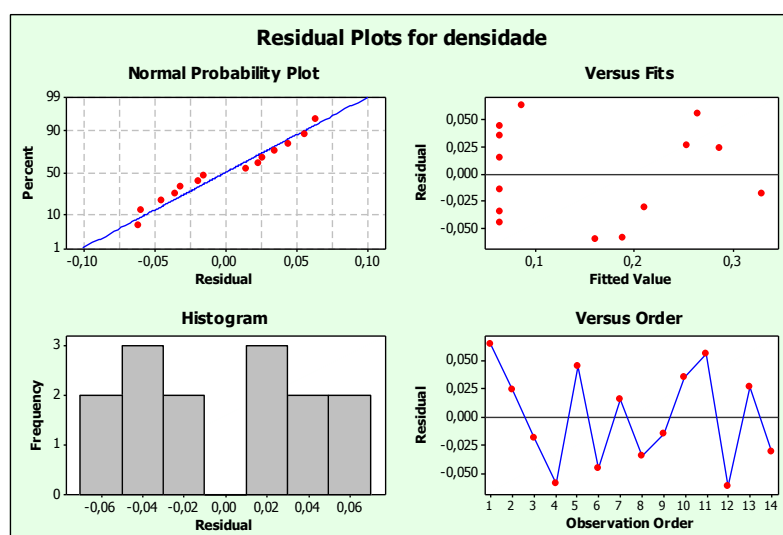


Gráfico 20- Resíduos da regressão quadrática “densidade”

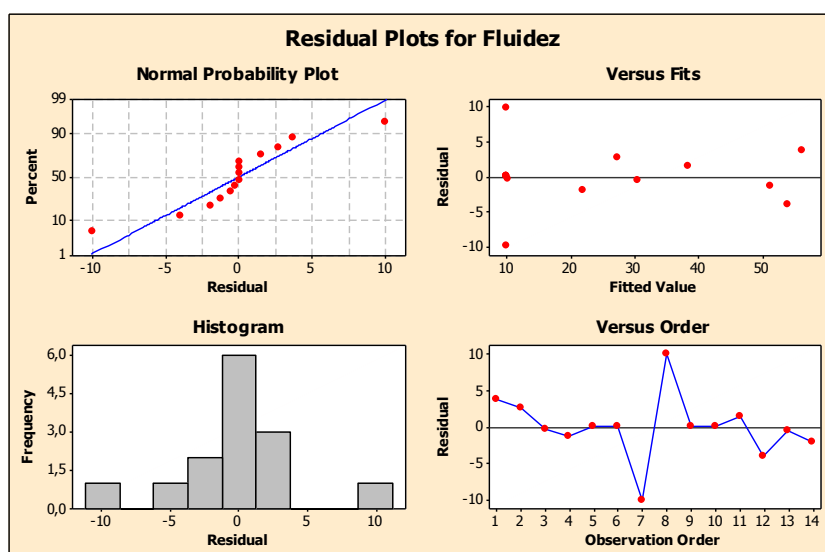


Gráfico 21 - Resíduos da regressão quadrática “fluidez”

As equações de regressão são:

Densidade

$$Y_1 = 0,065 + 0,015 x_1 + 0,036 x_2 - 0,084 x_1^2 - 0,085 x_1 x_2 + 0,074 x_2^2$$

Fluidez

$$Y_2 = 10 + 3,018 x_1 - 5,518 x_2 + 8,126 x_1^2 - 17,5 x_1 x_2 + 18,129 x_2^2$$

A densidade Y_1 , a nossa variável resposta, assume valores diferentes à medida que adicionamos (ou reduzimos) a quantidade de água e/ou defloculante que acrescentamos – quando adicionamos uma unidade de água (mais 25 litros) não acrescentando defloculante, a densidade da pasta ficará reduzida de $(0,015-0,084 =) 0,069$; Se adicionarmos 1 unidade de defloculante (1kg) e zero unidades de água, a pasta ficará com densidade acrescida de $(0,036+0,074=) 0,11$; se adicionarmos simultaneamente 1 unidade de água e 1 unidade de defloculante, então a densidade da pasta ficará reduzida em $(0,015+0,036-0,084-0,085+0,074 =) - 0,044$. Obviamente que esta inferência só é válida quando respeitamos os limites experimentalmente impostos e verificados $[-1,44,+1,44]$ (– unidades codificadas: $x_1=1 \Leftrightarrow x_1 = 25$ litros de água; $x_2=1 \Leftrightarrow x_2 = 1$ kg de defloculante).

Pelo mesmo processo de substituição de valores obteríamos as alterações da fluidez.

8.2 Determinação do ponto de estacionaridade

Apresentamos as tabelas dos valores previstos, computacionalmente, para cada observação registada ao nível de desvios ao padrão pré-fixado de densidade e fluidez e calculamos os valores óptimos para cada uma das situações, especificando os valores da quantidade de água e de defloculante ideais para o cumprimento das metas definidas (densidade=1,7 e fluidez=300°).

Predicted Response for New Design Points Using Model for densidade				
Point	Fit	SE Fit	95% CI	95% PI
1	0,085956	0,0432714	(-0,013828; 0,185740)	(-0,074934; 0,246847)

MSR: Uma Investigação no âmbito da Indústria Cerâmica

2	0,286311	0,0432714	(0,186527; 0,386095)	(0,125420; 0,447202)
3	0,328737	0,0432714	(0,228953; 0,428521)	(0,167847; 0,489628)
4	0,189092	0,0432714	(0,089308; 0,288876)	(0,028201; 0,349983)
5	0,065000	0,0223439	(0,013474; 0,116525)	(-0,071323; 0,201322)
6	0,065000	0,0223439	(0,013474; 0,116525)	(-0,071323; 0,201322)
7	0,065000	0,0223439	(0,013474; 0,116525)	(-0,071323; 0,201322)
8	0,065000	0,0223439	(0,013474; 0,116525)	(-0,071323; 0,201322)
9	0,065000	0,0223439	(0,013474; 0,116525)	(-0,071323; 0,201322)
10	0,065000	0,0223439	(0,013474; 0,116525)	(-0,071323; 0,201322)
11	0,263933	0,0432662	(0,164161; 0,363705)	(0,103049; 0,424816)
12	0,161020	0,0432662	(0,061248; 0,260792)	(0,000137; 0,321904)
13	0,253937	0,0432662	(0,154165; 0,353709)	(0,093054; 0,414821)
14	0,211016	0,0432662	(0,111244; 0,310788)	(0,050133; 0,371899)

Quadro 27 - Valores previstos para a variável “densidade”

Predicted Response for New Design Points Using Model for Fluidez				
Point	Fit	SE Fit	95% CI	95% PI
1	56,2545	4,38408	(46,1448; 66,3642)	(39,9537; 72,5553)
2	27,2904	4,38408	(17,1807; 37,4001)	(10,9896; 43,5912)
3	10,2179	4,38408	(0,1082; 20,3276)	(-6,0829; 26,5187)
4	51,2538	4,38408	(41,1441; 61,3635)	(34,9530; 67,5546)
5	9,9997	2,26379	(4,7794; 15,2201)	(-3,8119; 23,8114)
6	9,9997	2,26379	(4,7794; 15,2201)	(-3,8119; 23,8114)
7	9,9997	2,26379	(4,7794; 15,2201)	(-3,8119; 23,8114)
8	9,9997	2,26379	(4,7794; 15,2201)	(-3,8119; 23,8114)
9	9,9997	2,26379	(4,7794; 15,2201)	(-3,8119; 23,8114)
10	9,9997	2,26379	(4,7794; 15,2201)	(-3,8119; 23,8114)
11	38,4433	4,38355	(28,3348; 48,5518)	(22,1433; 54,7433)
12	54,0491	4,38355	(43,9407; 64,1576)	(37,7491; 70,3492)
13	30,5136	4,38355	(20,4051; 40,6221)	(14,2136; 46,8137)
14	21,9788	4,38355	(11,8704; 32,0873)	(5,6788; 38,2789)

Quadro 28 - Valores previstos para a variável “Fluidez”

8.2.1 Densidade

$$Y_1 = 0,065 + 0,015 x_1 + 0,036 x_2 - 0,084 x_1^2 - 0,085 x_1 x_2 + 0,074 x_2^2$$

Uma vez que a superfície de resposta é explicada pelo modelo de segunda ordem, é necessário analisar o ponto de ótimo: se na superfície existe um ótimo, então é um ponto de estacionaridade.

$$\hat{y} = \hat{\beta}_0 + x'b + x'Bx$$

$$X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$b = \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = \begin{pmatrix} 0,015 \\ 0,036 \end{pmatrix}$$

$$B = \begin{pmatrix} \hat{\beta}_{11} & \hat{\beta}_{12}/2 \\ \hat{\beta}_{12}/2 & \hat{\beta}_{22} \end{pmatrix} = \begin{pmatrix} 0,084 & -\frac{0,085}{2} \\ -0,0425 & 0,074 \end{pmatrix}$$

O ponto estacionário é dado por: $x_s = -\frac{1}{2} B^{-1} b$

$$B^{-1} = \begin{pmatrix} 16,78 & 9,64 \\ 9,64 & 19,05 \end{pmatrix}$$

Então,

$$x_s = -\frac{1}{2} \begin{pmatrix} 16,78 & 9,64 \\ 9,64 & 19,05 \end{pmatrix} * \begin{pmatrix} 0,015 \\ 0,036 \end{pmatrix} = \begin{pmatrix} -0,299 \\ -0,415 \end{pmatrix}$$

Em termos das **variáveis naturais**, o ponto estacionário é dado por:

$$-0,299 = \frac{A-325}{25} \Leftrightarrow \mathbf{A = 317,5 \text{ litros}}$$

$$-0,415 = \frac{D-3}{1} \Leftrightarrow \mathbf{D = 2,585 \text{ kg}}$$

O **máximo da resposta** pode ser agora calculado:

$$\hat{y}_s = \hat{\beta}_0 + \frac{1}{2} x'_s b$$

$$\hat{y} = 0,065 + \frac{1}{2} [-0,299 \quad -0,415] \times \begin{bmatrix} 0,015 \\ 0,036 \end{bmatrix} = \mathbf{0,055}$$

MSR: Uma Investigação no âmbito da Indústria Cerâmica

Encontrámos, neste patamar da superfície de resposta, os valores de Água e de Desfloculante que minimizam o desvio para o valor de Densidade pré-fixado.

$$|y_1 - 1,7| = 0,055 \rightarrow y_1 = 1,755 \vee y_1 = 1,645$$

8.2.2 Fluidez

$$Y_2 = 10 + 3,018 x_1 - 5,518 x_2 + 8,126 x_1^2 - 17,5 x_1 x_2 + 18,129 x_2^2$$

$$b = \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = \begin{pmatrix} 3,018 \\ -5,518 \end{pmatrix}$$

$$B = \begin{pmatrix} \hat{\beta}_{11} & \hat{\beta}_{12}/2 \\ \hat{\beta}_{12}/2 & \hat{\beta}_{22} \end{pmatrix} = \begin{pmatrix} 8,126 & -\frac{17,5}{2} \\ -8,75 & 18,129 \end{pmatrix}$$

O ponto estacionário é dado por: $x_s = -\frac{1}{2} B^{-1} b$

$$B^{-1} = \begin{pmatrix} 0,256 & 0,124 \\ 0,124 & 0,115 \end{pmatrix}$$

Então,

$$\underline{x}_s = -\frac{1}{2} \begin{pmatrix} 0,256 & 0,124 \\ 0,124 & 0,115 \end{pmatrix} * \begin{pmatrix} 3,018 \\ -5,518 \end{pmatrix} = \begin{pmatrix} -0,0454 \\ -0,1303 \end{pmatrix}$$

Em termos das **variáveis naturais**, o ponto estacionário é dado por:

$$-0,0454 = \frac{A-325}{25} \Leftrightarrow \mathbf{A} = \mathbf{323,9 \text{ litros}}$$

$$0,1303 = \frac{D-3}{1} \Leftrightarrow \mathbf{D} = \mathbf{3,13 \text{ kg}}$$

O **máximo da resposta** pode ser agora calculado:

$$\hat{y}_s = \hat{\beta}_0 + \frac{1}{2} x'_s b$$

$$\hat{y}_2 = 10 + \frac{1}{2} [-0,0454 \quad 0,1303] \times \begin{bmatrix} 3,018 \\ -5,518 \end{bmatrix} = \mathbf{9,572}$$

Encontrámos, neste patamar da superfície de resposta, os valores de Água e de Desfloculante que minimizam o desvio para o valor de Fluidiez pré-fixado.

$$|y_2 - 300^\circ| = 9,572 \rightarrow y_2 = 309,6^\circ \vee y_2 = 290,4^\circ$$

Tínhamos partido para esta experiência buscando valores de densidade e fluidez que foram declarados os ideais para a pasta cerâmica pretendida; para tal a experiência do mestre ceramista ajudou-nos a definir os valores altos e baixos de água e desfloculante a misturar aos materiais sólidos. O ótimo seria encontrado quando a distância entre os valores de densidade e fluidez definidos *a priori* e os valores experimentais fosse zero. Para isto era necessário determinar exatamente a quantidade de água e desfloculante a utilizar no processo para se obter uma pasta o mais próximo possível dos valores de densidade 1,7 e fluidez 300°.

Assim, para a densidade, verificamos que quando misturamos 317,5 litros de água e 2,585kg de desfloculante obtemos os 1000 kg de pasta cerâmica pretendida com uma densidade de 1,755 (1,645), isto é, distante 0,055 do valor de densidade 1,7 que pretendíamos. Para a fluidez verificamos que a mistura de 323,9 litros de água e 3,13kg de desfloculante daria lugar a uma pasta com fluidez ótima, próxima 9.6° do valor de 300° pretendido.

8.3 Conclusão do Modelo de Segunda Ordem

Quando o modelo de primeira ordem mostra uma significativa falta de ajuste, o analista pode utilizar um modelo de segunda ordem para descrever a superfície de resposta. Há muitos desenhos disponíveis para realizar um projeto de segunda ordem, como já referido, sendo o CCD o mais vulgar e mais utilizado nomeadamente quando os fatores têm dois níveis; é de fácil aplicação e, quando são convenientemente escolhidos os pontos axiais, tem propriedades de ortogonalidade e rotacionalidade que facilitam a verificação dos pressupostos de aplicação da MSR.

Neste tipo de desenho o analista pode começar com os pontos fatoriais 2^k , adicionando em seguida pontos centrais e pontos axiais para obter o Delineamento Central Composto. Ao adicionar os pontos axiais obterá termos quadráticos para incluir no modelo de segunda ordem.

MSR: Uma Investigação no âmbito da Indústria Cerâmica

Os gráficos de contorno são úteis para a visualização da superfície de resposta quando os fatores não são mais que três, porque a partir daí é quase impossível visualizar a superfície.

Quando o ponto estacionário é localizado, o analista fica em condições de tirar conclusões sobre o resultado que satisfaz a investigação, ou então em condições de promover nova pesquisa e explorar novos resultados.

9. Conclusões

A Metodologia de Superfícies de Resposta, pela sua natureza e simplicidade, ocupa um espaço relevante no campo do delineamento experimental. Baseando-se na Análise de Variância e nos modelos de Regressão linear, múltipla e polinomial, exige do analista um conhecimento aprofundado dos temas e um alerta permanente para os pressupostos de aplicação dos modelos.

O método consiste em encontrar o caminho certo que leve o investigador a atingir um patamar, uma superfície onde provavelmente se encontra a solução ótima do problema em estudo. Sempre aprendemos que o caminho mais próximo entre dois pontos é o segmento de reta, também aqui a estratégia começa por definir uma orientação que nos leve a direito para a superfície de resposta, usando a regressão múltipla linear. Chegados à superfície de resposta, cumpre encontrar o ponto de estacionaridade.

Nem sempre é possível encontrar um ponto ótimo quando há restrições. Nestes casos importa estudar as superfícies e encontrar a melhor solução possível. Este é um tópico que extravasa o âmbito deste trabalho que se subordinou à metodologia.

O exemplo que apresentamos neste trabalho, a produção de uma determinada pasta cerâmica com características definidas como as adequadas ao fabrico de “chávenas de porcelana”- densidade 1.7 e Fluidez 300° - permitiu-nos aprofundar a metodologia MSR e mostrar que a sua aplicação é fácil e de grande utilidade prática. A definição das variáveis dependentes como distâncias ao valor pré-fixado, permitiu-nos definir o problema como a busca do valor ótimo, mínimo que nos levasse ao desiderato, distância nula.

As variáveis dependentes, quantidade de água e de desfloculante, necessários à produção de 1000kg de pasta cerâmica, foram definidas num intervalo que os artesãos indicaram como os mais apropriados e, de facto, verificamos que cumpriram, como o esperado, tal requisito.

Centrámos a nossa experiência na produção da barbotina, mas, obviamente, poderíamos ter optado por outra ou outras quaisquer fases da produção, já que toda a laboração é um *continuum* e importante mesmo é a qualidade do produto final.

MSR: Uma Investigação no âmbito da Indústria Cerâmica

O objetivo deste trabalho cingiu-se à apresentação e justificação do modelo MSR: delimitamos o estudo às respostas individuais de cada uma das variáveis dependentes definidas. Uma proposta de trabalho no futuro passaria pela otimização conjunta de todas as variáveis de resposta; a qualidade é um somatório de requisitos, muitos dos quais estão correlacionados, quando buscamos o ótimo global, novas técnicas estatísticas, novos métodos de otimização múltipla, serão trazidas à colação, nomeadamente o método *desirability*, a regressão múltipla e a análise em componentes principais.

Os gráficos 3D e as curvas de nível revelam-se como uma preciosa ajuda na interpretação do problema, permitindo uma visualização imediata das áreas dos valores ótimos da função objetivo. Podem ainda ser úteis na indicação de áreas onde outros valores, fixados, de propriedades reológicas poderão ter valores ótimos.

Por último uma referência, mesmo que curta, à enorme quantidade de estudos, artigos e trabalhos no âmbito da metodologia das superfícies de resposta. A revisão bibliográfica levou-me à leitura de um sem número de *sites*: a busca começou pelos sites avalizados nomeadamente em *B-on* e *Google Student*. São tantos e tão diferenciados os campos experimentais que a compreensão de muitos dos conceitos específicos e da nomenclatura me ultrapassaram, contudo em cada um deles pude observar como a metodologia foi implementada e de facto é fascinante: quando estamos no “bom” podemos sempre tentar atingir o “ótimo”.

Bibliografia

- Anderson-Cook, C.M., Borror, C. M., Montgomery, D. C. (2009). *Response surface design evaluation and comparison (with discussion)*. Journal of Statistical Planning and Inference 139; 629 - 674.
- Anderson, M., Whitcomb, P..*RSM Simplified, Optimizing Processes Using Response surface Methods for Design of Experiments* (2005). Taylor Francis Group, LLC
- Allen, T. T. (2006). *Introduction to Engineering Statistics and Six Sigma – Statistical Quality Control and Design of Experiments and Systems*. Springer.
- Atkinson, A. C. and Bailey, R. A. (2001).*One Hundred Years of the Design of Experiments on and off the Pages of Biometrika: Biometrika*, Vol. 88, No. 1 (Mar.), pp. 53-97;Biometrika Trust; DOI: <http://www.jstor.org/stable/2673675>
- Bailey, R. A. (2008). *Design of Comparative Experiments*. Cambridge University Press.
- Box, G. E. P., Hunter, W. G., Hunter, J. S. (1978). *Statistics for Experimenters*. John Wiley & Sons, 1 edição.
- Box, G. E. P., Draper, N. R. (1987). *Empirical Model-Building and Response Surfaces*. John Wiley & Sons, 1 edição.
- Cheng, S.,Wu, C. F. J.(2001). Factor Screening and Response Surface Exploration. *Statistica Sinica* 11(2001), 553-604.
<http://www3.stat.sinica.edu.tw/statistica/oldpdf/A11n31.pdf>
- Costa, N. (2010). *Multiple response optimization: methods and results*. International Journal of Industrial and Systems Engineering, Vol. 5, No.4 pp. 442 - 459.
- Dean, A., Voss, D. (1999). *Design and Analysis of Experiments*. Springer.
- Douglas. C. MONTGOMERY (2004): *Design and Analysis of Experiments*, John Wiley & Sons, New York.
- Fiocco, Carola (1986): *Storia dell'arte Ceramica*, Zanichelli, Itália.
- Fonseca, J. (2001). *Estatística Matemática, vol.2*. Lisboa: Edições Sílabo. ISBN: 972-618-253-0
- Guimarães, R.Campos e Cabral, J.A.Sarsfield (1999). *Estatística – Edição Revista*. Amadora: Macgraw-Hill. ISBN: 972-8298-45-5
- Hinkelmann, Klaus e Kempthorne, Oscar (2008). *Design and Analysis of Experiments*. John Wiley & Sons,2ª edição.
- Marôco, J. (2010). *Análise Estatística com o PASW Statistics (ex-SPSS)*. Pero Pinheiro: Report Number. ISBN: 978-989-96763-0-5
- Montgomery, D. C.(2005). *Design and Analysis of Experiments*. 5ª edição; John Wiley & Sons
- Montgomery, D. C.(2005). *Introduction to Statistical Control*. 5ª edição. John Wiley&Sons.

- Montgomery, D. C. (2001). *Design and Analysis of Experiments*. John Wiley & Sons, 5ª Edição.
- Montgomery, D. C., Runger, G. C. (2003). *Estatística Aplicada e Probabilidade para Engenheiros*. LTC – Livros Técnicos e Científicos Editora S.A., 2ª Edição.
- Myers, R. H., Khuri, A., Vining, G. (1992). *Response Surface Alternatives to the Taguchi Robust Parameter Design Approach*. The American Statistician, Vol. 46, No. 2 (May), pp. 131-139. American Statistical Association. DOI: <http://www.jstor.org/stable/2684183>.
- Myers, R. H., Montgomery, D., Anderson-Cook, C. (2009). *Response Surface Methodology, Process and Designed Experiments*. John Wiley Sons., Inc 3ª Edição.
- Myers, R. H., Montgomery, D. C. (1995). *Response Surface Methodology: Process and Product Optimization Using Design of Experiments*. Wiley – Interscience. New York. 2ª Edição.
- Narimani, E. and Alaei Kadijani, J. (2014), Optimization a Gasoil Vacuum Dryer Using Response Surface Methodology. Chem. Eng. Technol., 37: 229–239. doi: 10.1002/ceat.201300123
- Oliveira, T.P.C.Azinheira (2004). *Estatística Aplicada*. Lisboa: Universidade Aberta. ISBN: 972-674-444-X.
- Paiva, A. P. (2004). *Estudo da Minimização de Erro nas Medições de Concentração de Emulsões por Titração Karl-Fischer utilizando-se Projeto de Experiências*. Dissertação de Mestrado em Engenharia de Produção. Universidade Federal de Itajubá.
- Pedrosa, A.C. e Gama, S.M.A. (2004). *Introdução Computacional à Probabilidade e Estatística*. Porto: Porto Editora. ISBN: 978-972-0-06056-3
- Pereira, Z. L; Requeijo, J. G. (2008). *QUALIDADE: Planeamento e Controlo Estatístico de Processos*. FCT-UNL e Prefácio, Lisboa.
- Pestana, D.D e Velosa, S.F. (2010). *Introdução à probabilidade e à Estatística*. 4th Ed., Lisboa: Fundação Calouste Gulbenkian. ISBN: 978-972-31-1150-7
- R Development Core Team (2012). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0
- Raissi, S. (2009). *Developing New Processes and Optimizing Performance Using Response Surface Methodology*. World Academy of Science, Engineering and Technology 49.
- Ross, Phillip. J. (1991). *Aplicações das técnicas de Taguchi na engenharia de qualidade*, São Paulo, Makkm, 1991.
- Wu, X., Liu, P., Ren, L., Tong, J. and Zhou, J. (2013), Optimization of corn starch succinylation using response surface methodology. Starch/Stärke. doi: 10.1002/star.201300194

MSR: Uma Investigação no âmbito da Indústria Cerâmica

Youn, B. D., Choi, K. K. (2003). A new response surface methodology for reliability-based design optimization. *Computers and Structures* 82(2004)241-256, Elsevier.

WEBSITES

http://www.cs.iusb.edu/thesis/NBradley_thesis.pdf # trabalho de dissertação em MSR - visitado em 19.10.2013

www.inf.ufsc.br/~ogliari/metodologiasuperficiesresposta.ppt # Imagens simples que, de forma muito acessível nos permitem a compreensão do tema MSR - visitado em 23.10.2013

http://www JMP.com/support/downloads/pdf/jmp_design_of_experiments.pdf # Experiências - visitado em 26.10.2012

http://web.iyte.edu.tr/food/Course_Web_Pages/DOE_introduction.pdf # Delineamento de experiências - visitado em 30.10.2012

<http://www.carinafagiani.com.br/resumos/index.php/2009/08/23/tipos-de-pesquisa/> # tipos de pesquisa - visitado em 02.11.2013

<http://adm-net-a.unifei.edu.br/phl/pdf/0030715.pdf> # MSR e ACP – tese de doutoramento # visitado em 13.11.2013

<http://adm-net-a.unifei.edu.br/phl/pdf/0032819.pdf> # MSR e ACP – texto didático - visitado em 09.11.2013

http://www.cs.iusb.edu/thesis/NBradley_thesis.pdf # visitado em 13.11.2013

[*Metodologia de superfície de resposta e análise de componentes principais em otimização de processos de manufatura com múltiplas respostas*](#) # visitado em 16.11.2013

[*Logurte” de soja suplementado com oligofrutose e inulina*](#) # visitado em 20.11.2013

[*Fabricação de pães com reduzido teor calórico e modificações reológicas ocorridas durante o armazenamento*](#) # visitado em 20.11.2013

<http://www.doe.soton.ac.uk/elearning/section1.jsp> # visitado em 23.11.2013

http://www.cs.iusb.edu/thesis/NBradley_thesis.pdf # trabalho de dissertação em MSR , visitado em 27.11.2013

<http://www.jstatsoft.org/> #Journal of Statistical Software. October 2009, Volume 32, Issue 7. Visitado em 30.11.2013

http://www.scielo.br/scielo.php?pid=S141543662012000700013&script=sci_arttext&lng=es # agro-alimentar. Visitado em 30.11.2013

http://www.latamjpharm.org/trabajos/22/2/LAJOP_22_2_1_8_S4236T5713.pdf # farmácia. Visitado em 30.11.2013

<http://acervo.ufvjm.edu.br:8080/jspui/handle/1/246> # química. Visitado em 30.11.2013

MSR: Uma Investigação no âmbito da Indústria Cerâmica

<http://www.scielo.br/pdf/qn/v31n3/a23v31n3.pdf> # Indústria Alimentar. Visitado em 30.11.2013

http://www.bv.fapesp.br/en/pesquisa/?q=%22Superf%C3%ADcies%20de%20resposta%22&count=10&q2=&ano_inicio_de=&ano_inicio_ate=&sort=-data_inicio&page=4 # Indústria química. Visitado em 30.11.2013

<http://www.bv.fapesp.br/en/bolsas/54227/the-weddell-sea-circulation-and-its-interaction-with-the-antarctic-sea-ice-implications-to-large-sca/> # Ambiente. Visitado em 30.11.2013

<http://www.iwaponline.com/wst/up/wst2013733.htm> # visitado em 29.01.2014

<http://www.sciencedirect.com/science/article/pii/S0304389408008972> # visitado em 30.01.2014

<http://pubs.acs.org/doi/abs/10.1021/es0498080> # visitado em 29.01.2014

<http://www.sciencedirect.com/science/article/pii/S0377042713006262> # visitado em 29.01.2014

<http://onlinelibrary.wiley.com/doi/10.1002/stc.1642> # visitado em 31.01.2014

<http://www.sciencedirect.com/science/article/pii/S0961953410005076> # visitado em 03.02.2014

<http://www.sciencedirect.com/science/article/pii/S0016236112001895> # visitado em 10.02.2014

<https://pangea.stanford.edu/ERE/db/GeoConf/papers/SGW/2014/Quinao.pdf> # visitado em 12.02.2014