

PLANOS EXPERIMENTAIS COM RESTRIÇÕES NA ALEATORIZAÇÃO ENVOLVENDO MISTURAS

ALLAN ROBERT DA SILVA

Orientadora: Profa. Dra. Carla Almeida Vivacqua

Linha de Pesquisa: Métodos Estatísticos

Dissertação submetida como requerimento parcial para obtenção
de grau de Mestre em Matemática Aplicada e Estatística
pela Universidade Federal do Rio Grande do Norte

Natal, Abril de 2009

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

Agradecimentos

"O bater de asas de uma simples borboleta poderia influenciar o curso natural das coisas e, assim, talvez provocar um tufão do outro lado do mundo". Edward Norton Lorenz, meteorologista e **matemático** norte-americano.

Encaro essa teoria de um ponto de vista mais humano acreditando que tudo está ligado e que cada pessoa que passa na nossa vida tem um parcela de responsabilidade, mesmo que pequena, sobre a sua personalidade. Por isso aqui vão meus sinceros agradecimentos para essas pessoas. Devo citar poucos nomes a fim de não falar muitos e esquecer alguém.

Primeiramente a Deus por ter me dado a oportunidade de conhecer pessoas maravilhosas na minha vida;

Agradeço especialmente aos meus pais e minha família que me apoiaram em toda essa caminhada, apesar de todas as dificuldades encontradas de uma vida cheia de desencontros.

Aos professores do curso de Estatística da UFRN, que me deram a base necessária para o conhecimento que tenho hoje. Em especial minha orientadora e seu esposo: Carla Almeida e André Pinho, na qual não vou falar do lado profissional, pois não é necessário, mas do lado humano que, pra mim, são exemplos de humildade, respeito e sempre me servirão de admiração.

Aos colegas da graduação, pelos estudos em grupo, discussões e momentos de descontração, em especial destaque Carlos Raphael, Nelson Alessandro e Clemer Caline.

Aos professores do PPGMAE, pela atenção, dedicação e carinho que não se vê em nenhum outro mestrado. Sou bastante feliz por ter sido aluno desse mestrado.

Ao professor Cristiano Ferraz que participa dessa banca e que também tive a oportunidade de ser seu aluno do mestrado em Estatística da UFPE, no qual pude confirmar porque esse mestrado é tão bem reconhecido.

Aos colegas da mestrado, pelas várias noites perdidas em estudos e também pelos momentos de alegria, em especial destaco a Renata Fonseca, Renata Santana e o Cecílio, que foram pessoas com quem tive um convívio mais forte e se tornaram muito importantes na minha vida particular e acadêmica.

A CAPES e ao REUNI pelo apoio financeiro.

A todas as outras pessoas que passaram pela minha vida e que me fizeram crescer como pessoa e como profissional.

Resumo

Durante a fase de planejamento de um experimento, um dos grandes desafios é lidar com a limitação de recursos. Assim, planos completamente aleatorizados nem sempre representam uma alternativa viável de realizar um experimento. Naturalmente, planos experimentais com restrições na aleatorização, como os da classe *split-plot* e *strip-plot* tornam-se práticos e úteis. Neste trabalho, aborda-se restrições na aleatorização no contexto de experimentos com misturas, cuja motivação inicial foi um plano experimental inovador proposto para um estudo comparativo de telhas cerâmicas no estado do Rio Grande do Norte. O plano é considerado inovador, pois envolve misturas com restrição na aleatorização do tipo *strip-plot* possuindo réplicas somente do ponto central das variáveis de processo. A partir de uma revisão da literatura identificou-se um artigo com um plano similar ao plano proposto, mas com restrição na aleatorização do tipo *split-plot* com réplicas do ponto central das misturas e das variáveis de processo. Neste artigo, os autores investigam a eficiência de três métodos de estimação de parâmetros: Mínimos Quadrados Ordinários, Máxima Verossimilhança Restrita e o Erro Puro, em que os dois últimos apresentaram bons resultados na estimativa dos parâmetros. Nessa dissertação verifica-se a eficiência desses dois último métodos considerando agora um plano experimental com mistura e com restrição na aleatorização do tipo *strip-plot* com réplicas do ponto central das misturas e das variáveis de processo, o que difere um pouco do plano proposto, pois este não possui réplica do ponto central da mistura. Observa-se que neste plano experimental os métodos de estimação não apresentaram bom desempenho. Esse resultado foi obtido a partir de simulações utilizando o pacote estatístico R. Ao final,

discussões sobre os resultados obtidos e sugestões para novas investigações são dadas.

Palavras-Chave: Superfície de resposta, Experimentos não Replicados, Modelos Lineares Mistos, Restrições na aleatorização.

Abstract

During the planning of an experiment, one of the major challenges is dealing with limited resources. Thus, completely randomized plans do not always represent a viable alternative to conduct an experiment. Naturally, experimental designs with restrictions on randomization, such those in the class of *split-plot* and *strip-plot* become practical and useful. In this work, deals about restrictions on randomization in the context of experiments with mixtures, whose initial motivation was an experimental innovator proposed for a comparative study of ceramic tiles in the state of Rio Grande do Norte. The design is considered innovative because it involves mixing with restriction in randomization of the type) (strip-plot has only replicas of the central point of the process variables. From a review of the literature identified was an article on a design similar to the design proposed, but with restriction in randomization of the type *split-plot* with replicas of the central point of the mixtures and of process variables. In this article, the authors investigate the effectiveness of three methods of estimation of parameters: Ordinary Least Squares, Restricted Maximum Likelihood and the Pure Error in the last two had good results in the estimation of parameters. In this dissertation there is the efficiency of these two latter methods now considering a experimental design with mixture and with restriction in randomization of the type *strip-plot* with replicas of the central point of the mixture and of the process variables, which differs slightly the proposed plan because it does not have the central point of the mixture. Observe that in this design experimental the methods of estimation did not show good performance. This result was obtained from simulations using the statistical package R. Finally, discussions on the results and suggestions for

further research are given.

Keywords: Response Surface, Experiments non Replicate, Linear Mixed Models, Restrictions on Randomization.

Sumário

Lista de Figuras	ix
Lista de Tabelas	x
1 Introdução	1
1.1 Objetivos da Dissertação	3
1.2 Organização da Dissertação	4
1.3 Plataforma Computacional	4
2 Experimentos com misturas e variáveis de processo	5
3 Experimentos em Telhas Cerâmicas	13
4 Análise de Experimentos com Mistura e Variáveis de Processo com Restrição na Aleatorização	21
4.1 Modelos de misturas e variáveis de processo	22
4.2 Estimação do Modelo	26
4.3 Resultados	35
5 Considerações finais	38
5.1 Conclusões	38

5.2	Perspectiva de trabalhos futuros	39
A	Código do R estimação dos parâmetros de um experimento <i>strip-plot</i>	40
	Apêndice	40
	Referências Bibliográficas	44

Lista de Figuras

2.1	Planos <i>split-plot</i>	7
3.1	Superfície de resposta: Tetraedo com e sem restrições	15
3.2	Alocação de corpos de prova - Plano 1: <i>split-plot</i> Tipo 1	17
3.3	Alocação de corpos de prova - Plano 2 : <i>split-plot</i> Tipo 2	18
3.4	Alocação de corpos de prova - Plano 3: <i>strip-plot</i>	19
4.1	Plano <i>split-plot</i>	23
4.2	Plano <i>strip-plot</i> proposto para o ajuste da equação (4.3)	24

Lista de Tabelas

2.1	ANOVA para o plano <i>Split-Plot</i> Figura 2.1(a)	9
2.2	ANOVA para um plano <i>strip-plot</i>	11
2.3	Alterações nos efeitos aleatórios para planos não replicados	12
3.1	Formulações, em percentual, sugeridas para cada tipo de argila	14
3.2	Variáveis de Processo	16
4.1	Tratamentos utilizados no plano sugerido em 4.2	25
4.2	Unidades experimentais	27
4.3	Plano <i>strip-plot</i> - Exemplo 1	29
4.4	Plano <i>strip-plot</i> - Exemplo 2	31
4.5	Tratamentos destacados usados para estimar o Erro Puro	34
4.6	Estimativas de $ \widehat{Var}(\hat{\beta}) $	35
4.7	Variâncias Estimadas em 5.000 simulações	36

Capítulo 1

Introdução

A fase mais importante de um experimento é o seu planejamento. É nesse momento que o experimentador se depara com alguns desafios, como obter informações úteis sobre o processo em estudo com recursos limitados, incluindo tempo e dinheiro. Nesse sentido, é de interesse para o pesquisador a diminuição do número de provas. Em muitas aplicações, especialmente na área de Engenharia, nem sempre é possível utilizar planos completamente aleatorizados devido ao alto custo envolvido, seja pela quantidade excessiva de unidades experimentais que devam ser produzidas ou devido ao uso de testes em que é necessário a destruição das unidades experimentais. Para reduzir o esforço experimental, em geral, são utilizados planos experimentais com restrições na aleatorização, como os planos *split-plot* e *strip-plot*. Além disso, no meio industrial, a urgência de respostas e a necessidade de redução de custos tornam necessária a realização de experimentos sem a presença de réplicas.

Outro desafio é quando os fatores são proporções de diferentes componentes de uma mistura. Por exemplo, se é desejável otimizar a resistência à tração de um aço inoxidável, os fatores de interesse podem ser as proporções de ferro, cobre, níquel e cromo na liga. O fato da proporção dos diferentes fatores somarem 100% caracteriza este experimento

como pertencente a uma classe particular denominada experimentos com misturas.

Diante deste cenário, o presente trabalho foi motivado inicialmente por um plano experimental desenvolvido na UFRN, como parte integrante de uma tese de doutorado em Ciências de Materiais. O intuito do experimento era identificar condições ideais para a produção de telhas cerâmicas de alta qualidade. Durante o planejamento foi discutido com o pesquisador como realizar o experimento de forma a reduzir os custos experimentais, otimizando, assim, os recursos disponíveis. O plano proposto envolve experimentos com misturas e com restrição na aleatorização do tipo *strip-plot* com réplica somente das variáveis de processo. Não há na literatura um método de análise ou um método que garanta a qualidade da estimação dos parâmetros de um modelo para o plano proposto. Neste trabalho será dado o primeiro passo para a análise deste plano. Para isso, será realizada uma simulação com objetivo de verificar a eficiência de métodos de estimação para um plano experimental que envolve experimentos com misturas e com restrição na aleatorização do tipo *strip-plot*, mas com réplicas das variáveis de processo e das formulações das misturas. A análise dos métodos aplicados neste plano experimental servirá de base para análise do plano inovador proposto.

1.1 Objetivos da Dissertação

Objetivo geral:

O objetivo geral dessa dissertação é verificar a eficiência de métodos de estimação de parâmetros para um plano experimental que envolve a combinação de um experimento fatorial e um experimento com misturas com restrição na aleatorização do tipo *strip-plot*.

Objetivos específicos:

- Verificar a eficiência dos métodos do Erro Puro e Máxima verossimilhança Restrita na estimativa de parâmetros em planos experimentais com restrições na aleatorização.
- Diferenciar formas de aleatorização em planos experimentais envolvendo misturas com restrições na aleatorização do tipo *split-plot* e *strip-plot*;
- Mostrar como obter estimativas de parâmetros para planos experimentais replicados ou não replicados com restrições na aleatorização do tipo *split-plot* e *strip-plot*, usando os softwares R e SAS®.

Espera-se, desta forma, contribuir para a divulgação dos planos *split-plot*, *strip-plot* e planos experimentais envolvendo misturas.

Para a familiarização com a teoria de planejamento de experimentos utilizada constantemente durante esta dissertação, o leitor interessado pode consultar, por exemplo, Montgomery (2002) e Cornell (1990), que apresentam conceitos e definições fundamentais que não foram necessariamente descritas nesta dissertação.

1.2 Organização da Dissertação

A presente dissertação encontra-se dividida em cinco capítulos. O primeiro faz uma breve introdução dos temas discutidos e objetivos da dissertação. No segundo capítulo é mostrado como obter estimativas de fontes de erros de planos experimentais com restrições na aleatorização do tipo *split-plot* e *strip-plot* usando os softwares R e SAS[®]. O terceiro capítulo descreve o plano experimental que serviu de motivação inicial desta dissertação. O quarto capítulo apresenta a metodologia usada neste trabalho. O último capítulo apresenta as conclusões obtidas no estudo e indica sugestões para trabalhos futuros.

1.3 Plataforma Computacional

Esta dissertação foi digitada utilizando o sistema de tipografia LATEX que consiste em uma série de macros ou rotinas do sistema TEX que facilitam a edição do texto. Mais informações sobre o sistema de tipografia LATEX podem ser encontrados no site <http://www.tex.ac.uk/CTAN/latex>. Para a realização das simulações e obtenção dos resultados numéricos foi utilizado o ambiente de programação, análise de dados e gráficos R (Team 2008) para sistema operacional Microsoft Windows, que se encontra disponível gratuitamente através do site <http://www.R-project.org>. Alguns resultados foram obtidos do ambiente gráfico e de programação SAS[®] para o sistema operacional Microsoft Windows, cujo intuito foi comparar os comandos e saídas computacionais com o software R.

Capítulo 2

Experimentos com misturas e variáveis de processo

Em muitos casos, os produtos comercializados são criados a partir da mistura de mais de um ingrediente em conjunto. Neste caso, o fabricante está interessado em uma ou mais propriedades do produto final, que depende das proporções dos ingredientes utilizados. Por exemplo, a criação de uma nova tonalidade de tinta, formulada pela mistura de outras duas cores, em que o interesse é criar uma cor mais chamativa para atrair mais clientes.

Neste sentido surgiu uma necessidade de identificar qual percentual de cada substância usada é necessário para criar um produto de melhor qualidade. Para responder essa questão, surge uma nova classe de experimentos denominada experimentos com misturas. Nesta classe, existem q fatores dados por x_1, x_2, \dots, x_q , em que cada nível desses fatores representa uma proporção cuja a construção dos tratamentos nada mais é do que a combinação de níveis das q proporções. A diferença deste tipo de experimento é que os tratamentos são restritos ao fato de que a soma dos níveis das proporções deve ser igual a uma constante (1 ou 100%), ou seja, $x_1 + x_2 + \dots + x_q = 1$. Muitos autores descrevem sobre experimentos com misturas, por exemplo, Myers & Montgomery (1995), Smith (2005)

e Cornell (1990). Este último autor é o mais citado entre aqueles que escrevem sobre experimentos com misturas devido às inúmeras contribuições na área.

Observe que, diferente de um experimento fatorial, a modificação de um nível de um fator implica na modificação dos demais, já que a soma da proporção total é fixa. A superfície de resposta é normalmente representada a partir de desenhos geométricos $q - 1$ dimensionais. Os pontos experimentais estão sobre extremidades ou interior dos desenhos. Com dois componentes a representação é uma reta, para três é um triângulo e para quatro um tetraedro.

Em geral, a qualidade do produto final não depende somente da formulação do produto, mas também das condições de transformação deste produto. As variáveis que fazem parte do processo de transformação do produto são denominadas variáveis de processo, que quando modificadas alteram as propriedades dos ingredientes das formulações. Por exemplo, a preparação de um tipo de café, formulado pela mistura do pó do café, açúcar e água. Esta composição deverá ser exposta a uma determinada temperatura que pode afetar o sabor do café caso seja modificada. Muitos autores escrevem sobre modelos de experimentos com mistura e variáveis independentes do processo de formulação. O artigo pioneiro a considerar este tipo de experimento foi escrito por Scheffé em 1963, intitulado como *The simplex-centroid designs for experiments with mixtures*.

Na prática, o número de variáveis de processo é pequeno, assim como o número de combinações de misturas. Em geral, considera-se apenas dois níveis para cada variável de processo. Para a l -ésima variável de processo a notação usada será $z_l = -1$ e $z_l = +1$, que representam, respectivamente, os níveis baixo e alto desta variável. Essa representação, nada mais é do que a representação de um modelo fatorial 2^k . A Figura 2.1 ilustra a estrutura de fatores relativos aos planos *split-plot*.

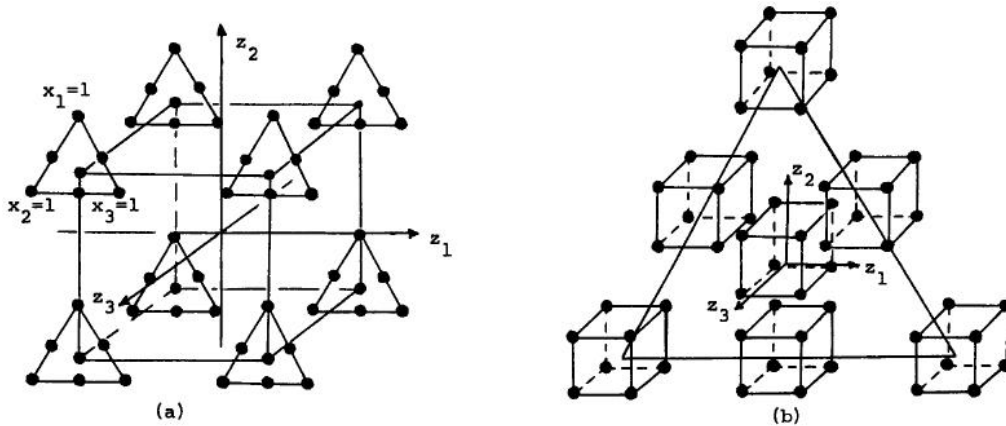


Figura 2.1: Planos *split-plot*

A Figura 2.1(a) sugere que, de forma aleatória, sete unidades experimentais foram produzidas com diferentes tipos de formulações da mistura e expostas, ao mesmo tempo e de forma aleatória, a uma mesma combinação das variáveis de processo. Em seguida, mais sete unidades experimentais foram aleatoriamente produzidas com diferentes tipos de formulações e mais uma vez expostas, aleatoriamente e de uma só vez, a outra combinação das variáveis de processo. O processo se repete até o final. Assim, serão produzidos 56 combinações de misturas e a exposição às variáveis de processo ocorre somente oito vezes. Neste caso, este experimento envolve misturas e variáveis de processo com restrição do tipo *split-plot*, em que as combinações das variáveis de processo são classificadas como plots, pois as unidades experimentais expostas ao mesmo tempo a uma combinação das variáveis de processo são naturalmente correlacionadas, e as combinações de misturas são classificadas como subplots.

A Figura 2.1(b) indica que, aleatoriamente, sete unidades experimentais foram produzidas com diferentes tipos de formulações da mistura e que cada uma possui tamanho suficiente para ser dividida em oito partes, na qual cada parte das sete formulações será aleatoriamente exposta a uma combinação das variáveis de processo. Assim, serão pro-

duzidos apenas sete formulações das misturas e a exposição às variáveis de processo ocorrerá 56 vezes. Neste caso, este experimento envolve misturas e variáveis de processo com restrição do tipo *split-plot*, em que as formulações das misturas são classificadas como plots, pois as unidades experimentais criadas ao mesmo tempo são naturalmente correlacionadas, e as combinações das variáveis de processo são classificadas como subplots.

Se existe uma réplica do experimento é possível fazer uma análise a partir de uma ANOVA. O modelo linear geral associado ao plano *split-plot* para Figura 2.1(a) pode ser escrito segundo a forma:

$$y_{ijk} = \mu + R_i + Z_j + e_{ij}^{RZ} + X_k + ZX_{jk} + e_{ijk}$$

na qual:

- y_{ijk} é a resposta observada na i -ésima réplica sob o j -ésimo nível do *plot* e do k -ésimo nível do *subplot*;
- μ é a média geral;
- R_i é o efeito da i -ésima réplica, em que $R_i \sim N(0, \sigma_r^2)$;
- Z_j é o efeito da j -ésimo nível do *plot*;
- e_{ij}^{RZ} é o termo residual relativo ao valor observado na i -ésima réplica do j -ésimo nível do *plot*, em que $e_{ij}^{RZ} \sim N(0, \sigma_{RZ}^2)$;
- X_k é o efeito do k -ésimo tratamento de *subplot*;
- ZX_{jk} é o efeito da interação entre o j -ésimo nível do *plot* e o k -ésimo nível do *subplot*;
- e_{ijk} é o termo residual relativo ao valor observado na i -ésima réplica sob o j -ésimo nível do *plot* e do k -ésimo nível do *subplot*, em que $e_{ijk} \sim N(0, \sigma_e^2)$.

Seja n e m , respectivamente, o número de combinações variáveis de processo e o número de formulações da mistura e assumindo não haver entre interação réplicas \times subplot e os demais (pressupondo aditividade entre unidade e tratamentos) a ANOVA deste plano é dada na Tabela 2.1 abaixo.

Tabela 2.1: ANOVA para o plano *Split-Plot* Figura 2.1(a)

Fonte	G.L.	(SQ)	E(QM)
Réplicas(R)	$r - 1$	SQR	$mn\sigma_r^2 + \sigma_e^2 + m\sigma_{RZ}^2$
<i>plot</i> (Z)	$n - 1$	SQZ	$m\sigma_{RZ}^2 + \sigma_e^2 + rm \sum_j Z^2 / (n - 1)$
Erro (RZ)	$(r - 1)(n - 1)$	$SQRZ$	$m\sigma_{RZ}^2 + \sigma_e^2$
<i>sub-plot</i> (X)	$(m - 1)$	SQX	$\sigma_e^2 + rn \sum_k X^2 / (m - 1)$
Interação (ZX)	$(n - 1)(m - 1)$	$SQZX$	$\sigma_e^2 + r \sum_k (ZX)^2 / (m - 1)(n - 1)$
Erro	$n(r - 1)(m - 1)$	SQE	σ_e^2
Total	$rn - 1$	SQT	

A ANOVA apresentada na Tabela 2.1 pode ser obtida nos softwares SAS® e R. No SAS é usado um procedimento é denominado *PROC MIXED*. Os comandos necessários são dados abaixo, no qual o símbolo * indica um comentário:

```
\* Início
proc mixed data=dados method=reml; \* Indica o pacote, onde estão os dados e o método de estimação.
class r z x; \* Define quais variáveis são fatores
model y=z*x/noit; \* define o modelo
random r r*x; \* define os efeitos aleatórios
run;
\* FIM
\* r indica a i-ésima réplicas, z indica o j-ésima combinação das variável de processo
\* x indica a k-ésima formulação das misturas e y indica o vetor de respostas
```

De maneira similar no software R essa ANOVA é obtida usando um pacote denominado *nlme* os comandos são dados abaixo, no qual # indica um comentário:

```
require(nlme) # Carregando o pacote necessário
dados=as.data.frame(dados) # 0 local onde estão os dados deve ter formato de data.frame
dados$r=as.factor(dados$r) # r fator indicando a i-ésima réplica
dados$z=as.factor(dados$z) # z fator indicando a j-ésima combinação das variável de processo
dados$x=as.factor(dados$x) # x fator indicando a k-ésima formulação das misturas
fm1 <- lme(y~z*x, random=~1|r/z) # y é o vetor de respostas
anova(fm1) # ANOVA
```

Observe que os comandos usados no programa SAS[®] são escritos de forma mais simples para o usuário, sobretudo na maneira de definir quais são os efeitos aleatórios. Recentemente alguns autores têm trabalhado esse tipo de experimento usando uma restrição na aleatorização do tipo *strip-plot*, ver, por exemplo, Milliken & Johnson (2000). Neste caso existem dois plots, um associado às variáveis de processo e outro associado às formulações da mistura.

O modelo linear associado ao plano *strip-plot* pode ser escrito segundo a forma:

$$y_{ijk} = \mu + R_i + Z_j + e_{ij}^{RZ} + X_k + e_{ik}^{RX} + ZX_{jk} + e_{ijk}$$

na qual:

- y_{ijk} é a resposta observada na i -ésima réplica sob o j -ésimo nível do plot associado às variáveis de processo do k -ésimo nível do plot associado às misturas.
- μ é a média geral;
- R_i é o efeito da i -ésima réplica, em que $R_i \sim N(0, \sigma_r^2)$
- Z_j é o efeito da j -ésimo nível do plot associado às misturas;
- e_{ij}^{RZ} é o termo residual relativo ao valor observado na i -ésima réplica do j -ésimo nível do plot associado às variáveis de processo, em que $e_{ij}^{RZ} \sim N(0, \sigma_{RZ}^2)$;
- X_k é o efeito do k -ésimo tratamento do plot associado às misturas;
- e_{ik}^{RX} é o termo residual relativo ao valor observado na i -ésima réplica do k -ésimo nível do plot associado às misturas, em que $e_{ik}^{RX} \sim N(0, \sigma_{RX}^2)$;
- ZX_{jk} é o efeito da interação entre o j -ésimo nível do plot associado ao k -ésimo nível do plot associado às misturas;
- e_{ijk} é o termo residual relativo ao valor observado na i -ésima réplica sob o j -ésimo nível do plot associado às variáveis de processo, e do k -ésimo nível do plot associado às misturas, em que $e_{ijk} \sim N(0, \sigma_e^2)$.

Supondo, mais uma vez, que n e m , representam, respectivamente, o números de combinações das variáveis de processo e o número de formulações da mistura e assumindo não haver a interação réplicas \times subplot (pressupondo aditividade entre unidade e tratamentos) a ANOVA deste plano é dada na Tabela 2.2 abaixo.

Tabela 2.2: ANOVA para um plano *strip-plot*

Fonte	G.L.	(SQ)	E(QM)
Réplicas(R)	$r - 1$	SQR	$mn\sigma_r^2 + \sigma_e^2 + n\sigma_{RZ}^2 + m\sigma_{RX}^2$
Plot(Z)	$n - 1$	SQZ	$m\sigma_{RZ}^2 + \sigma_e^2 + rm \sum_j Z^2 / (n - 1)$
Erro (RZ)	$(r - 1)(n - 1)$	$SQRZ$	$m\sigma_{RZ}^2 + \sigma_e^2$
Plot(X)	$(m - 1)$	SQX	$\sigma_e^2 + n\sigma_{RX}^2 + rn \sum_k X^2 / (m - 1)$
Erro (RX)	$(r - 1)(m - 1)$	$SQRX$	$\sigma_e^2 + n\sigma_{RX}^2$
Interação (ZX)	$(n - 1)(m - 1)$	$SQZX$	$\sigma_e^2 + r \sum_k (ZX)^2 / (m - 1)(n - 1)$
Erro	$(r - 1)(m - 1)(n - 1)$	SQE	σ_e^2
Total	$rn m - 1$	SQT	

No software SAS[®] para se obter a ANOVA da Tabela 2.2 é usado, mais uma vez, o pacote *PROC MIXED*, os comandos são dados abaixo:

```
proc mixed data=dados method=reml;
class r z x;          \* r indica a i-ésima réplica e z a j-ésima combinação das variável de processo
model y=z*x/noint;    \* x é um vetor indicando a k-ésima formulação das misturas
random r r*x r*z;     \* define os efeitos aleatórios
run;
```

De maneira similar no software R a ANOVA é obtida usando o pacote *nlme* com os comandos dados abaixo:

```
require(nlme) # Instalando o pacote necessário
dados=as.data.frame(dados) # Local onde estão os dados deve ter formato de dataframe
dados$r=as.factor(dados$r) # r fator indicando a i-ésima réplicas
dados$z=as.factor(dados$z) # z fator indicando a j-ésima tratamento das variável de processo
dados$x=as.factor(dados$x) # x fator indicando a k-ésima tratamento das misturas
fm2=lme(y~z*x,random=list(r=pdBlocked(list(pdIdent(~1), pdIdent(~z-1),pdIdent(~ordered(x)-1)))),data=dados)
anova(fm2)
```

Mesmo não havendo um réplica do experimento é possível estimar as demais fontes de erro do modelo usando praticamente as mesmas linhas de comando apresentadas, mas fazendo alterações somente nas linhas referentes aos efeitos aleatórios. A Tabela 2.3 mostra as alterações necessárias para obter estas estimativas.

Tabela 2.3: Alterações nos efeitos aleatórios para planos não replicados

<i>split-plot</i>		<i>strip-plot</i>	
R	SAS [®]	R	SAS [®]
random = $\sim 1 r$	random = r	Não se altera	random = r x

Os pacotes do R e SAS[®] usam um mesmo método iterativo para estimar os parâmetros denominado método da Máxima Verossimilhança Restrita (MVR), ver Patterson & Thompson (1971) e Harville (1977). Alguns autores descrevem e exemplificam esses comandos ilustrando inclusive as saídas computacionais. Algumas boas referências são Pinheiro & Bates (2000) que são os autores do pacote *nlme* do software R. Littell, Milliken, Stroup & Wolfinger (1996) comentam sobre modelos com mais de um efeito aleatório usando o programa SAS[®]. Milliken & Johnson (2000) também ilustram modelos usando o SAS[®] e, por fim, West, Welch & Galecki (2006) que comentam sobre modelos com mais de um efeito aleatório usando os programas R, SAS[®], Stata, HLM e SPSS.

Capítulo 3

Experimentos em Telhas Cerâmicas

Essa dissertação teve por motivação inicial um plano experimental proposto a um pesquisador doutorando do curso de Ciências de Materiais da UFRN. A seguir serão apresentados os detalhes da construção do plano experimental.

Toda informação de profissionais ligados à área do experimento é de importância fundamental para o seu planejamento. Para o processo de fabricação de telhas cerâmicas considerado, existem dois tipos de argila que são misturados para formulação de telhas: a argila gorda e a magra. Essa classificação leva em conta a granulometria da argila. Além disso, existem duas jazidas no qual será retirada a argila. Assim, existe uma argila magra da jazida um ($M1$), uma gorda da jazida um ($G1$), analogamente existe uma argila magra da jazida dois ($M2$) e uma argila gorda da jazida dois ($G2$). Para a produção de telhas cerâmicas apropriadas para uso, é necessário que a mistura tenha, pelo menos, um tipo de argila gorda e um tipo de argila magra. O objetivo do experimento é identificar as proporções de cada tipo de argila que otimizam diversas características de qualidade do produto final, como absorção de água e retração linear. A Tabela 3.1 mostra as formulações sugeridas para o experimento em questão.

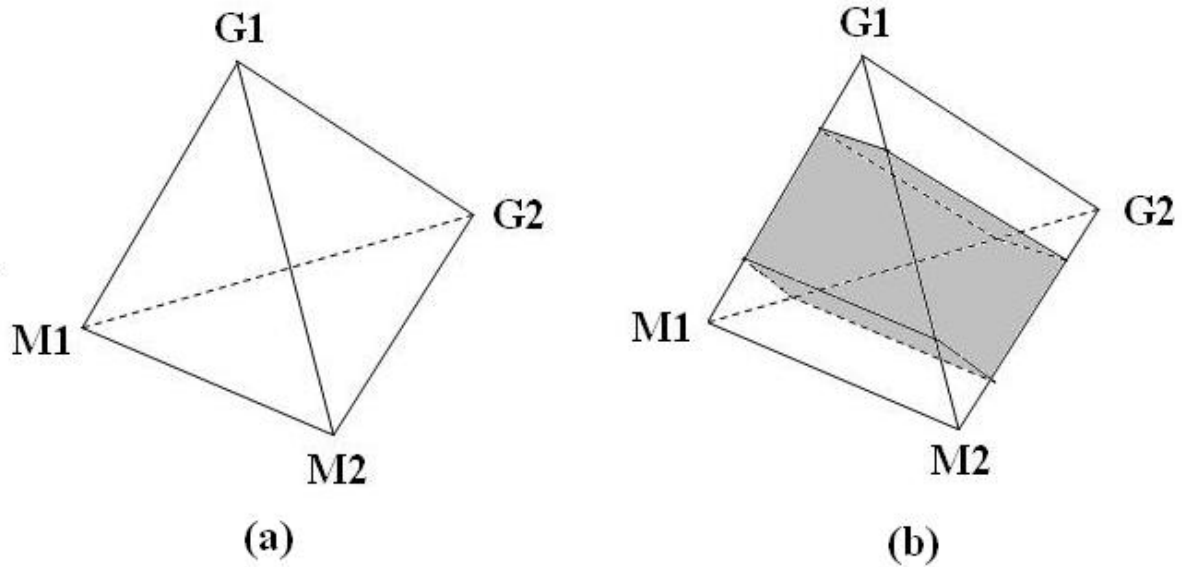
Tabela 3.1: Formulações, em percentual, sugeridas para cada tipo de argila

Formulação	(G1)	(M1)	(G2)	(M2)
1	0	0	50	50
2	0	12,5	75	12,5
3	25	37,5	0	37,5
4	0	0	75	25
5	50	0	0	50
6	25	0	0	75
7	0	37,5	25	37,5
8	12,5	0	12,5	75
9	0	0	25	75
10	75	25	0	0
11	37,5	0	37,5	25
12	12,5	75	12,5	0
13	25	25	25	25
14	0	75	25	0
15	75	12,5	0	12,5
16	0	50	50	0
17	25	75	0	0
18	0	25	75	0
19	75	0	0	25
20	37,5	25	37,5	0
21	50	50	0	0

Vale ressaltar que os fatores $G1$, $M1$, $G2$ e $M2$, ou seja, as quantidades dos tipos de argila são dependentes e que cada formulação tem a soma de suas respectivas proporções igual a 100%. Além disso, a condição de que cada formulação deve conter pelo menos uma argila gorda e uma magra é incorporada nesta sugestão através da restrição: $25\% \leq M_1 + M_2 \leq 75\%$. Essa descrição caracteriza um experimento com misturas com restrições na região experimental. Pelo fato do esquema ter quatro componentes

(ou fatores), sua representação dever ser realizada através de uma figura tridimensional denominada tetraedro, sendo este um objeto de quatro faces triangulares como mostrado na Figura 3.1.

Figura 3.1: Superfície de resposta: Tetraedo com e sem restrições



A Figura 3.1(a) ilustra a região experimental sem restrições, enquanto a Figura 3.1(b) representa a região experimental com a restrição sugerida.

Além das proporções de cada tipo de argila, outros fatores podem afetar as características de qualidade das telhas cerâmicas. Na produção de telhas, também é de interesse observar quais níveis de temperatura e taxa de aquecimento do forno utilizados são ideais para a queima. Daqui em diante essas variáveis que não fazem parte da formulação serão denominadas de variáveis de processo. A Tabela 3.2 ilustra os valores escolhidos para esses dois fatores.

Tabela 3.2: Variáveis de Processo			
Temperatura	Taxa de aquecimento		
	$6^{\circ}C/min$	$8^{\circ}C/min$	$10^{\circ}C/min$
$950^{\circ}C$	Tratamento I	Tratamento III	Tratamento II
$1000^{\circ}C$			
$1050^{\circ}C$	Tratamento IV		Tratamento V

Após a definição dos fatores e níveis a serem utilizados, as próximas etapas do planejamento são definir como os corpos de prova serão produzidos e como os tratamentos serão alocados aos corpos de prova, ou seja, a forma de aleatorização do experimento. Esse experimento possui dois estágios. O primeiro estágio é a produção dos corpos de prova, ao qual está associado o experimento com misturas. O segundo estágio é a queima dos corpos de prova, que está associado ao experimento fatorial.

Para a produção dos corpos de prova, as argilas ainda separadas são peneiradas, pois o corpo de prova pesa somente 13 gramas e qualquer impureza pode danificá-lo. Em seguida, fazem-se as misturas, que são acrescidos 10% de água. Logo após, as misturas são prensadas a 25 mpa (aproximadamente 3 toneladas). Depois de prensadas são retiradas algumas medidas (como peso e tamanho) do corpo e, por último, são secadas a uma temperatura de $110^{\circ}C$ durante 24 h. Após a produção dos corpos de prova, a próxima etapa é a queima, submetendo os corpos de prova às temperaturas e taxas de aquecimento especificadas. Além das quatro combinações de temperatura e taxa de aquecimento, acrescentou-se um ponto central com uma taxa de aquecimento e temperatura intermediária.

No plano proposto existem $21 \times 5 = 105$ tratamentos a serem observados, e como existe réplica do ponto central das variáveis de processo seria necessário criar $21 \times 6 = 126$ formulações e realizar 126 operações com o forno se o plano experimental fosse completamente

aleatorizado. Devido a inviabilidade prática e econômica deste plano surge a necessidade de usar diferentes formas de aleatorização de forma a minimizar o tempo de execução e o custo envolvido. A seguir serão discutidas três diferentes formas de aleatorização.

Plano 1: *split-plot* Tipo 1

Um maneira de tornar mais prática a execução desse experimento é alocar todos os corpos de provas que vão ser receber a mesma combinação das variáveis de processo e queimá-los ao mesmo tempo. Em cada combinação das variáveis de processo seriam alocados 21 corpos de provas e como existem seis combinações de variáveis de processo seria necessário criar $6 \times 21 = 126$ corpos de prova. Esse tipo de experimento é classificado como *split-plot*. Observe que os corpos de prova contidos na mesma combinação das variáveis de processo são correlacionados. Assim, as variáveis de processo são classificadas como plots e as formulações são classificadas como subplots. A Figura 3.2 abaixo ilustra esse plano.

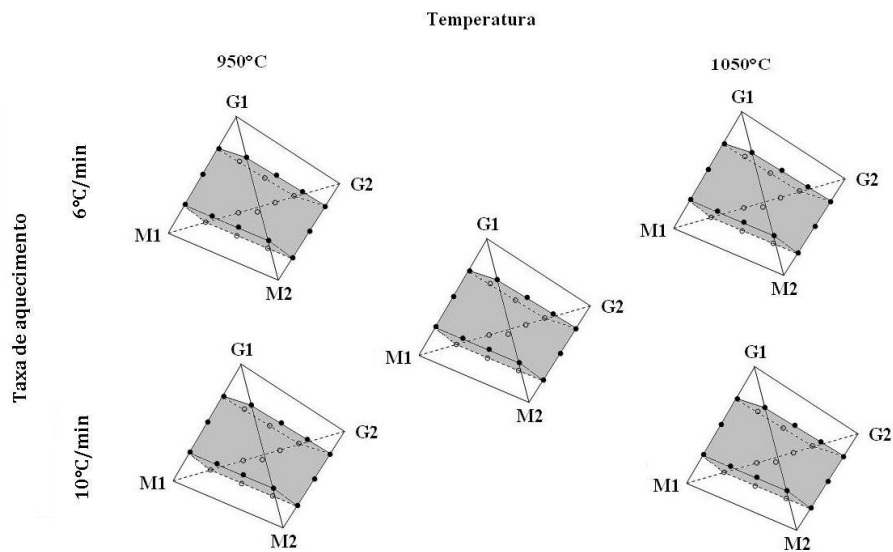


Figura 3.2: Alocação de corpos de prova - Plano 1: *split-plot* Tipo 1

Plano 2: *split-plot* Tipo 2

Um problema que pode ocorrer no plano *split-plot* citado anteriormente seria a impossibilidade de alocar os 21 corpos de prova no forno por questão de espaço. Uma maneira de tornar mais prático o experimento seria agora criar cada formulação de uma só vez. Esse tipo de experimento também é denominado *split-plot*, mas, agora, as formulações são os plots, pois são os corpos de prova de cada formulação que são correlacionados. Será necessário produzir somente 21 formulações e realizar $21 \times 6 = 126$ operações com o forno. Ao contrário do plano anterior, a desvantagem prática é o tempo levado para as operações com o forno, contudo poucas formulações deverão ser criadas. Observe que cada formulação deve ter um tamanho suficiente para que possa ser dividida em 6 partes. A Figura ilustra esse tipo de aleatorização.

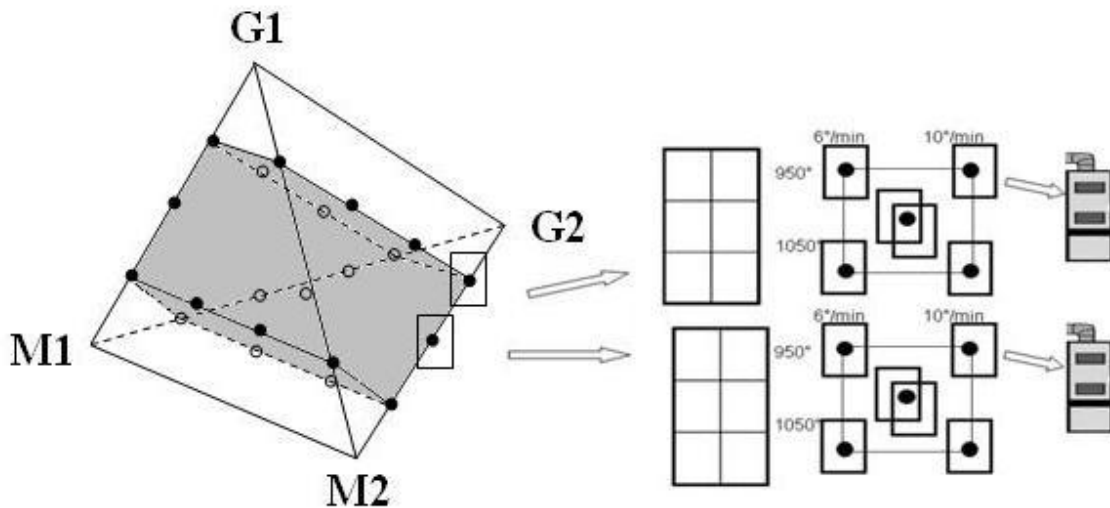


Figura 3.3: Alocação de corpos de prova - Plano 2 :*split-plot* Tipo 2

Plano 3: *strip-plot*

Se as formulações são suficientemente grandes para serem divididas em 6 partes e no forno couber as 21 diferentes formulações é bem mais prático criar as formulações e realizar as operações com o forno simultaneamente. Com isso são necessários criar apenas 21 formulações e realizar apenas 6 operações com o forno. Esse tipo de plano é denominado *strip-plot*. Observe que existe uma correlação entre os corpos de provas que são da mesma formulação e também existe uma correlação entre os corpos de provas que são submetidos a mesma combinação de variáveis de processo. Assim, neste plano, existem dois plots um referente às formulações e outro às variáveis de processo. Comparado este plano aos planos anteriores, o tempo levado para as operações com o forno e para criação de formulações é o bem menor, tornando-o mais prático e econômico. A Figura 3.4 ilustra a execução para esse tipo de plano.

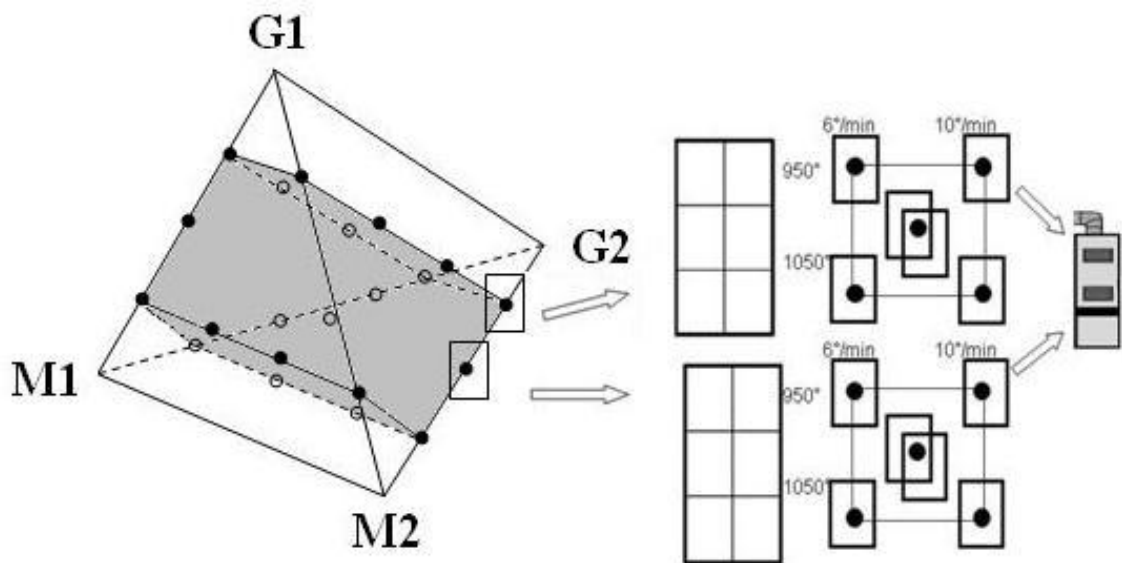


Figura 3.4: Alocação de corpos de prova - Plano 3: *strip-plot*

Mas na prática como foram alocados os corpos de prova?

Na prática, foram produzidas, de uma só vez, uma quantidade de cada formulação suficiente para moldar 24 corpos de prova com as mesmas características. No total, foram utilizados $24 \times 21 = 504$ corpos de prova. O forno foi utilizado seis vezes: quatro utilizando cada uma das combinações de temperatura e taxa de aquecimento e duas vezes usando temperatura e a taxa de aquecimento intermediárias. Cada vez que o forno for utilizado, 84 corpos de prova foram submetidos àquela determinada condição de temperatura e taxa de aquecimento, sendo quatro corpos de prova de cada uma das 21 formulações.

Para cada combinação dos níveis de temperatura e taxa de aquecimento, todos os tipos de formulações sugeridas foram executados simultaneamente. O fato de vários corpos de prova serem submetidos a uma determinada temperatura e taxa de aquecimento conjuntamente, e as argilas de uma determinada concentração serem produzidas também simultaneamente caracterizam a forma de aleatorização do experimento como *strip-plot*. Maiores detalhes sobre estes planos podem ser vistos em Daniel (2008), Vivacqua & Bisgaard (2004) e Box & Jones (1992).

Capítulo 4

Análise de Experimentos com Mistura e Variáveis de Processo com Restrição na Aleatorização

Com base no experimento citado no capítulo anterior, foi realizada uma revisão literária a fim de encontrar alguma bibliografia que descreva um plano com características semelhantes ao plano proposto. Observou-se na pesquisa que Kowalski et. al. (2000) propuseram executar somente uma fração da mistura em cada combinação das variáveis de processo, usando um plano completamente aleatorizado. Em 2002, os mesmos autores estendem esse trabalho para o caso de um experimento com restrição na aleatorização do tipo *split-plot*, comparando três métodos de estimação do modelo. Entre os métodos testados dois se destacaram o método do Erro Puro e o da Máxima Verossimilhança Restrita (MVR). Com base neste resultado, surgiu o interesse em verificar a eficiência desses métodos de estimação em um experimento com restrição na aleatorização do tipo *strip-plot*. Neste capítulo, será discutido um novo plano que terá uma réplica dos pontos centrais referentes às misturas e às variáveis de processo, o que difere um pouco do plano proposto no capítulo anterior que possui réplica somente das variáveis de processo. O

motivo de incluir uma réplica do ponto central das formulações de misturas é possibilitar a verificação da eficiência do método do Erro Puro, que depende da réplica de pelo menos um dos tratamentos em cada plot para estimar o erro proveniente da aleatorização destes.

4.1 Modelos de misturas e variáveis de processo

Cornell (1990) apresenta um modelo combinado de misturas e variáveis de processo (MVP) cujo objetivo é expressar a variável resposta como uma função das propriedades da mistura, dos efeitos das variáveis de processo e das interações entre ambos. A forma deste modelo depende da complexidade e limitações do experimento. Para q misturas representadas por x_1, x_2, \dots, x_q um modelo bastante citado na literatura é o modelo quadrático de Scheffé dado por:

$$\eta = \sum_{i=1}^q \beta_i x_i + \sum_{i < j} \sum \beta_{ij} x_i x_j \quad (4.1)$$

que contém $q(q+1)/2$ termos.

Supondo que existam n variáveis de processo dadas por z_1, z_2, \dots, z_n . Um típico modelo de segunda ordem de um plano fatorial com dois níveis é dado por

$$\eta_{vp} = \alpha_0 + \sum_{i=1}^n \alpha_i z_i + \sum_{i < j} \sum \alpha_{ij} z_i z_j \quad (4.2)$$

que contém $1 + [n(n+1)]/2$ termos.

Quando um modelo quadrático da mistura é combinado com um modelo de variáveis de processo, em que somente os efeitos principais das variáveis de processo são cruzados com os efeitos lineares da mistura, o modelo combinado MVP é dado por:

$$\eta(x, z) = \sum_{i=1}^q \beta_i x_i + \sum_{i < j}^q \beta_{ij} x_i x_j + \sum_{k < l}^n \alpha_{kl} z_k z_l + \sum_{i=1}^q \sum_{k=1}^n \gamma_{ik} x_i z_k \quad (4.3)$$

Esse modelo contém $[(q + n)^2 + q - n] / 2$ termos.

Para um plano *strip-plot* ser considerado balanceado deve existir o mesmo número de formulações da mistura em cada combinação das variáveis de processo e, ao mesmo tempo, cada combinação da mistura deve ser exposta a um mesmo número de combinações de variáveis de processo. Apenas para ilustração, observe a Figura 4.1 que mostra um plano experimental descrito por descrito por Kowalski, Cornell & Vining (2002).

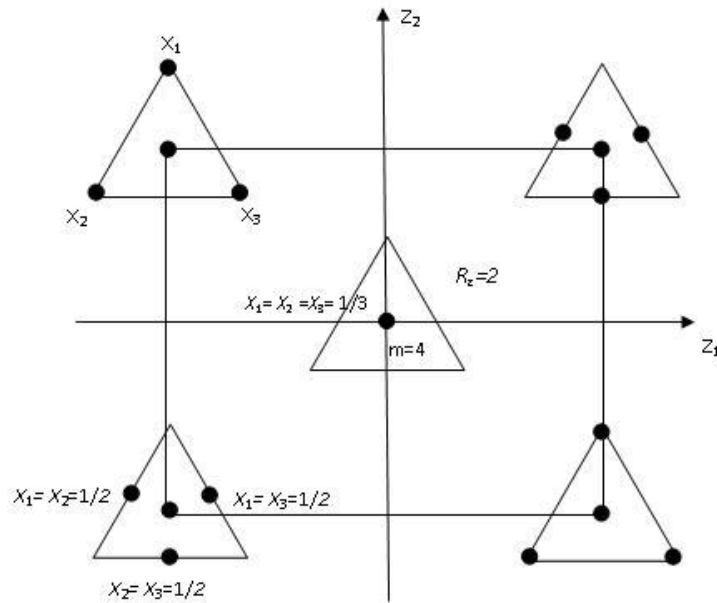


Figura 4.1: Plano *split-plot*

Este exemplo, ilustra um plano *split-plot*, em que os plots são combinações das variáveis de processo e os subplots são formulações da mistura. O Valor $R_z = 2$ indica que existem duas réplicas do ponto central das variáveis de processo, enquanto o valor $m = 4$ indica que o ponto $X_1 = X_2 = X_3 = 1/3$ é replicado 4 vezes em cada uma das R_z réplicas do

ponto central das variáveis de processo. Este experimento é balanceado para este plano, pois, cada plots tem o mesmo número de subplots. Contudo, não seria balanceado se a restrição na aleatorização deste experimento fosse do tipo *strip-plot*, pois, além de cada combinação de variáveis de processo terem o mesmo número de formulações de mistura é necessário que cada formulação das misturas seja exposta a um mesmo número de combinações de variáveis de processo. Por exemplo, a formulação que representa 100% de X_1 está sendo exposta somente em duas combinações das variáveis de processo, enquanto a formulação composta de 1/3 de cada componente da mistura está presente em todas as seis combinações das variáveis de processo. A Figura 4.2, abaixo, é um exemplo de experimento balanceado para o plano *strip-plot*.

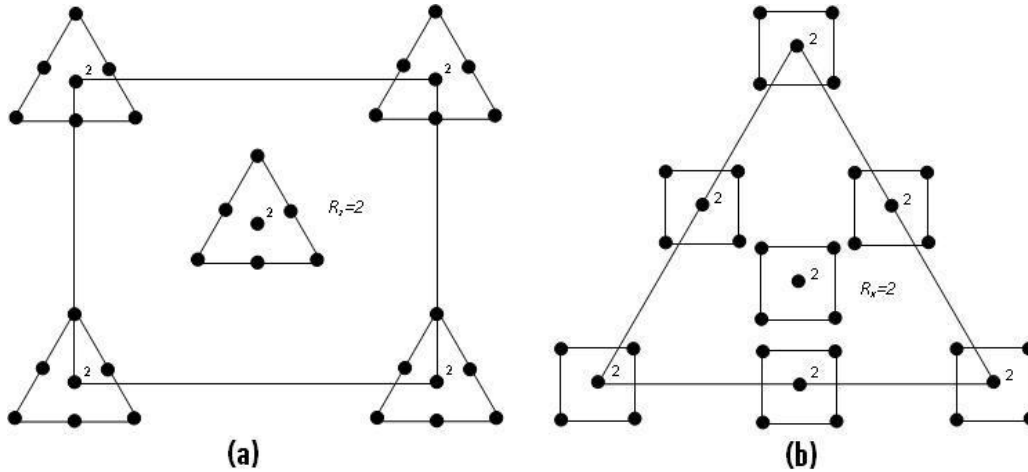


Figura 4.2: Plano *strip-plot* proposto para o ajuste da equação (4.3)

Observe que na Figura 4.2(a) cada uma das seis combinações de variáveis de processo possui oito formulações da mistura. Consideram-se oito formulações da mistura, pois existem sete misturas e mais uma réplica da mistura central, além disso existem seis variáveis de processo, pois existe uma réplica do ponto central das variáveis de processo, denotada por $R_z = 2$. Olhando por outro ângulo, em 4.2(b) é possível, mais uma vez,

verificar que cada uma das oito formulações da mistura é exposta em seis combinações das variáveis de processo. Consideram-se oito formulações de misturas, pois existe uma réplica do ponto central das formulações da mistura, denotada por $R_x = 2$. A Tabela 4.1 indica os tratamentos utilizados neste plano.

Tabela 4.1: Tratamentos utilizados no plano sugerido em 4.2						
Misturas (%)	Variáveis de Processo (Z_1, Z_2)					
(X_1, X_2, X_3)	(1,-1)	(1, 1)	(-1, 1)	(-1,-1)	(0, 0)	(0, 0)
(1, 0, 0)	y_{11}	y_{12}	y_{13}	y_{14}	y_{15}	y_{16}
(0, 1, 0)	y_{21}	y_{22}	y_{23}	y_{24}	y_{25}	y_{26}
(0, 0, 1)	y_{31}	y_{32}	y_{33}	y_{34}	y_{35}	y_{36}
(0, 1/2, 1/2)	y_{41}	y_{42}	y_{43}	y_{44}	y_{45}	y_{46}
(1/2, 0, 1/2)	y_{51}	y_{52}	y_{53}	y_{54}	y_{55}	y_{56}
(1/2, 1/2, 0)	y_{61}	y_{62}	y_{63}	y_{64}	y_{65}	y_{66}
(1/3, 1/3, 1/3)	y_{71}	y_{72}	y_{73}	y_{74}	y_{75}	y_{76}
(1/3, 1/3, 1/3)	y_{81}	y_{82}	y_{83}	y_{84}	y_{85}	y_{86}

Observando a Tabela 4.1 fica mais claro observar que somente os pontos centrais das misturas e das variáveis de processo são replicados. Na prática, existem $7 \times 5 = 35$ tratamentos a serem testados. Se todo o experimento fosse replicado somente uma vez seriam necessários 70 unidades experimentais. No plano descrito acima são necessárias somente 42 unidades experimentais, pois existe réplica somente do ponto central de cada plot.

4.2 Estimação do Modelo

Em um experimento do tipo *split-plot* a estrutura de erro conta com duas fonte de erro, uma associada ao plot e outra associada ao subplot, já em um plano *strip-plot* existem três fontes de erro, pois existem um erro associado a cada plot, e um terceiro erro associado às unidades.

O modelo do experimento pode ser descrito por:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\delta} + \boldsymbol{\gamma} + \boldsymbol{\epsilon}$$

Em que

$$\boldsymbol{\delta} + \boldsymbol{\gamma} + \boldsymbol{\epsilon} \sim N(0, \mathbf{V})$$

Se existem N variáveis respostas, a matriz \mathbf{V} representa a matriz de variância-covariância de um vetor de resposta \mathbf{y} de dimensão $N \times 1$.

A matriz $\boldsymbol{\delta}$, de dimensão $N \times N$, representa a matriz de variância-covariância dos plots das variáveis de processo. Esta matriz é composta de zeros e uns. Será um para aquelas unidades experimentais que são provenientes da mesma combinação de variáveis de processo e zero caso contrário, sendo escrita como:

$$\boldsymbol{\delta} * \sigma_z^2$$

Em que σ_z^2 é um parâmetro referente ao erro proveniente da aleatorização das combinações das variáveis de processo. Analogamente, A matriz $\boldsymbol{\gamma}$, representa a variância-covariância do *plots* referente às formulações da mistura. Esta matriz também é composta de zeros e uns. Será um para aquelas unidades experimentais que pertencem a mesma formulação das misturas e zero caso contrário, sendo escrita como:

$$\gamma * \sigma_x^2$$

Em que σ_x^2 é um parâmetro referente ao erro proveniente da aleatorização das formulações da mistura. Por fim, a matriz ϵ representa a matriz de variância-covariância das unidades, na qual é escrita como.

$$\epsilon * \sigma_e^2$$

No qual σ_e^2 é um parâmetro referente ao erro proveniente de uma terceira aleatorização referente a alocação de uma unidade experimental a uma combinação dos componentes de mistura e das variáveis de processo.

Sejam n e m , respectivamente, o número de combinações de variáveis de processo e o número de combinações das misturas, então o número total de unidades experimentais para um experimento não replicado é dado por $N = n * m$. A Tabela 4.2 indica as N unidades experimentais deste tipo de experimento.

Tabela 4.2: Unidades experimentais

Misturas	Variáveis de Processo				
	1	2	3	...	n
1	*	*	*	...	*
2	*	*	*	...	*
3	*	*	*	...	*
4	*	*	*	...	*
⋮	⋮	⋮	⋮	⋮	⋮
m	*	*	*	...	*

Genericamente, o formato da matrix \mathbf{V} das N unidades experimentais é dado por:

[illegible]

Observando a diagonal principal da matriz apresentada acima, é possível verificar que existem n submatrizes de ordem $m \times m$ ao seu redor, estas matrizes são denominadas de blocos diagonais, na qual, neste caso, indicam as m unidades experimentais provenientes de cada n combinações das variáveis de processo e, além disso, existem $(n - 1) * 2$ diagonais paralelas à diagonal principal que indicam às n unidades experimentais provenientes de cada m combinações das misturas. Observa-se, também, que na diagonal principal existe um erro, σ_t , que representa os três erros relativos a cada uma das três aleatorizações.

Para exemplificar esta estrutura de erro, suponha um experimento com restrição do tipo *strip-plot* com apenas uma variável de processo com dois níveis baixo(-) e alto(+), e outras duas variáveis de mistura, no qual são observados somente três formulações $(X_1 = 1, X_2 = 0)$, $(X_1 = 0, X_2 = 1)$ e $(X_1 = 1/2, X_2 = 1/2)$. Suponha, também, que foram usadas somente seis unidades experimentais. Uma possibilidade de construção desse experimento é apresentada na tabela 4.3 abaixo:

Tabela 4.3: Plano *strip-plot* - Exemplo 1

Misturas (%)	Variáveis de Processo	
	Baixo(-)	Alto(+)
$(X_1 = 1, X_2 = 0)$	y_1	y_4
$(X_1 = 0, X_2 = 1)$	y_2	y_5
$(X_1 = 1/2, X_2 = 1/2)$	y_3	y_6

A matriz de variância-covariância desse exemplo é dada por:

$$\mathbf{V} = \begin{matrix} & \begin{matrix} y_1 & y_2 & y_3 & y_4 & y_5 & y_6 \end{matrix} \\ \begin{matrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \end{matrix} & \begin{pmatrix} \sigma_t & \sigma_z & \sigma_z & \sigma_x & 0 & 0 \\ \sigma_z & \sigma_t & \sigma_z & 0 & \sigma_x & 0 \\ \sigma_z & \sigma_z & \sigma_t & 0 & 0 & \sigma_x \\ \sigma_x & 0 & 0 & \sigma_t & \sigma_z & \sigma_z \\ 0 & \sigma_x & 0 & \sigma_z & \sigma_t & \sigma_z \\ 0 & 0 & \sigma_x & \sigma_z & \sigma_z & \sigma_t \end{pmatrix} \end{matrix}$$

Em que $\sigma_t = \sigma_z + \sigma_x + \sigma_e$

A matriz \mathbf{V} pode ser expressa como soma das matrizes $\boldsymbol{\delta}$, $\boldsymbol{\gamma}$ e $\boldsymbol{\epsilon}$, da seguinte maneira:

$$\mathbf{V} = \sigma_z * \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix} + \sigma_x * \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix} + \sigma_e * \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Considere agora um experimento com restrição do tipo *strip-plot* com duas variável de processo com dois níveis baixo(-) e alto(-), e outras duas variáveis de mistura, no qual são observados somente três formulações ($X_1 = 1, X_2 = 0$), ($X_1 = 0, X_2 = 1$) e ($X_1 = 1/2, X_2 = 1/2$). Suponha, também, que foram usadas somente seis unidades experimentais. Uma possibilidade de construção desse experimento é apresentada na tabela 4.4 abaixo:

Tabela 4.4: Plano *strip-plot* - Exemplo 2

Misturas (%)	Variáveis de Processo			
	1(-)	1(+)	2(-)	2(+)
$(X_1 = 1, X_2 = 0)$	y_1	y_4	y_7	y_{10}
$(X_1 = 0, X_2 = 1)$	y_2	y_5	y_8	y_{11}
$(X_1 = 1/2, X_2 = 1/2)$	y_3	y_6	y_9	y_{12}

A matriz de variância-covariância desse exemplo é dada por:

$$\mathbf{V} = \begin{matrix} & \begin{matrix} y_1 & y_2 & y_3 & y_4 & y_5 & y_6 & y_7 & y_8 & y_9 & y_{10} & y_{11} & y_{12} \end{matrix} \\ \begin{matrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \\ y_8 \\ y_9 \\ y_{10} \\ y_{11} \\ y_{12} \end{matrix} & \begin{pmatrix} \sigma_t & \sigma_z & \sigma_z & \sigma_x & 0 & 0 & \sigma_x & 0 & 0 & \sigma_x & 0 & 0 \\ \sigma_z & \sigma_t & \sigma_z & 0 & \sigma_x & 0 & 0 & \sigma_x & 0 & 0 & \sigma_x & 0 \\ \sigma_z & \sigma_z & \sigma_t & 0 & 0 & \sigma_x & 0 & 0 & \sigma_x & 0 & 0 & \sigma_x \\ \sigma_x & 0 & 0 & \sigma_t & \sigma_z & \sigma_z & \sigma_x & 0 & 0 & \sigma_x & 0 & 0 \\ 0 & \sigma_x & 0 & \sigma_z & \sigma_t & \sigma_z & 0 & \sigma_x & 0 & 0 & \sigma_x & 0 \\ 0 & 0 & \sigma_x & \sigma_z & \sigma_z & \sigma_t & 0 & 0 & \sigma_x & 0 & 0 & \sigma_x \\ \sigma_x & 0 & 0 & \sigma_x & 0 & 0 & \sigma_t & \sigma_z & \sigma_z & \sigma_x & 0 & 0 \\ 0 & \sigma_x & 0 & 0 & \sigma_x & 0 & \sigma_z & \sigma_t & \sigma_z & 0 & \sigma_x & 0 \\ 0 & 0 & \sigma_x & 0 & 0 & \sigma_x & \sigma_z & \sigma_z & \sigma_t & 0 & 0 & \sigma_x \\ \sigma_x & 0 & 0 & \sigma_x & 0 & 0 & \sigma_x & 0 & 0 & \sigma_t & \sigma_z & \sigma_z \\ 0 & \sigma_x & 0 & 0 & \sigma_x & 0 & 0 & \sigma_x & 0 & \sigma_z & \sigma_t & \sigma_z \\ 0 & 0 & \sigma_x & 0 & 0 & \sigma_x & 0 & 0 & \sigma_x & \sigma_z & \sigma_z & \sigma_t \end{pmatrix} \end{matrix}$$

Se a matriz \mathbf{V} fosse conhecida a variância de $\hat{\beta}$ seria dada por:

$$\begin{aligned} Var(\hat{\beta}) &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(var(\mathbf{y}))\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \\ &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \end{aligned} \quad (4.4)$$

A presença da matriz \mathbf{V} na equação acima indica que além dos coeficientes do modelo é necessário estimar três fontes de erro: σ_z^2, σ_x^2 e σ_e^2 , que são parâmetros presentes em \mathbf{V} . O método de Mínimos Quadrados Ordinário(MQO) não será usado para estimar \mathbf{V} , pois este não leva em conta que existem diferentes fontes de erro. Outros dois métodos serão utilizados para estimar as três fontes de erro, são eles o método da Máxima Verossimilhança Restrita (MVR) e o método do Erro Puro.

O método da Máxima Verossimilhança Restrita é um método popular de estimação que é usado em modelos mistos, ou seja, modelos com fatores fixos e fatores aleatórios. Uma suposição deste método é que os efeitos fixos do modelo ajustado são corretos e, com isso, os componentes de variância são estimados pelos resíduos do modelo obtido no método de MQO. O Método do Erro Puro é um método alternativo independente do modelo que usa as réplicas do experimento para estimar as fontes de variação.

Será realizada uma simulação para obter estimativas das matrizes \mathbf{V} e $Var(\hat{\beta})$. A maneira mais popular de representar uma matriz de variância-covariância é a partir do seu determinante, essa representação é denominada de Variância Generalizada. A informação obtida em $\widehat{Var}(\hat{\beta})$ será resumida a partir de seu determinante. O verdadeiro valor do determinante de $Var(\hat{\beta})$ será comparado com o valor obtido da matriz estimada pelos dois métodos. Na prática, \mathbf{V} é desconhecida, seu valor é conhecido somente nesta simulação.

No método da MVR a estimativa das fontes de variação são imputadas em \mathbf{V}_c que possui a mesmas características da matriz \mathbf{V} , a única diferença é que os valores contidos

são as variâncias estimadas pelo método. O vetor β é estimado por

$$\hat{\beta} = (\mathbf{X}'\mathbf{V}_c\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}_c^{-1}\mathbf{y}$$

Enquanto a variância desses parâmetros é dada por:

$$\begin{aligned}\widehat{Var}(\hat{\beta}) &= (\mathbf{X}'\mathbf{V}_c^{-1}\mathbf{X})^{-1}\mathbf{X}'(var(\mathbf{y}))\mathbf{X}(\mathbf{X}'\mathbf{V}_c^{-1}\mathbf{X})^{-1} \\ &= (\mathbf{X}'\mathbf{V}_c^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}_c^{-1}\mathbf{V}_c\mathbf{V}_c^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}_c^{-1}\mathbf{X})^{-1}\end{aligned}$$

O vetor de resposta de todo o experimento será simulado. Estes serão simulados 5.000 vezes em cinco diferentes combinações dos parâmetros σ_z , σ_x e σ_e . A cada simulação os dois métodos serão usados obter os estimadores $\hat{\sigma}_z$, $\hat{\sigma}_x$, $\hat{\sigma}_e$ e $\widehat{Var}(\hat{\beta})$. No Método do Erro Puro, a estimativa da variância de β é idêntica ao método da MVR.

Vale salientar que no plano *strip-plot* em discussão existem duas réplicas de cada plot e, com isso, existem quatro unidades experimentais que serão usadas para estimar o Erro Puro a partir da variância amostral destas unidades. A seguir a Tabela 4.5 destaca os unidades usadas para estimar o Erro Puro.

Tabela 4.5: Tratamentos destacados usados para estimar o Erro Puro

Misturas (X_1, X_2, X_3)	Variáveis de Processo (Z_1, Z_2)					
	(1,-1)	(1, 1)	(-1, 1)	(-1,-1)	(0, 0)	(0, 0)
(1, 0, 0)	y_{11}	y_{12}	y_{13}	y_{14}	y_{15}	y_{16}
(0, 1, 0)	y_{21}	y_{22}	y_{23}	y_{24}	y_{25}	y_{26}
(0, 0, 1)	y_{31}	y_{32}	y_{33}	y_{34}	y_{35}	y_{36}
(1/2, 1/2, 0)	y_{41}	y_{42}	y_{43}	y_{44}	y_{45}	y_{46}
(1/2, 0, 1/2)	y_{51}	y_{52}	y_{53}	y_{54}	y_{55}	y_{56}
(0, 1/2, 1/2)	y_{61}	y_{62}	y_{63}	y_{64}	y_{65}	y_{66}
(1/3, 1/3, 1/3)	y_{71}	y_{72}	y_{73}	y_{74}	y_{75}	y_{76}
(1/3, 1/3, 1/3)	y_{81}	y_{82}	y_{83}	y_{84}	y_{85}	y_{86}

O motivo dessas unidades serem usadas para o cálculo do Erro Puro é que além de serem réplicas do ponto central do plot associado às formulações da mistura, também são réplicas do ponto central associado às variáveis de processo. A estimativa da variância do plot associado às variáveis de processo é dado por:

$$\hat{\sigma}_z = \frac{\sum_{i=1}^r (\bar{y}_{i.} - \bar{y}_{z..})^2}{R_z - 1}$$

Em que $\bar{y}_{.j}$ é a média do i-ésimo plot referente ao ponto central das variáveis de processo e $\bar{y}_{z..}$ é a média de todos os plots referente ao ponto central das variáveis de processo. Analogamente, a estimativa da variância do plot associado às variáveis de mistura é dada por:

$$\hat{\sigma}_x = \frac{\sum_{j=1}^r (\bar{y}_{.j} - \bar{y}_{x..})^2}{R_x - 1}$$

Em que $\bar{y}_{.j}$ é a média do i-ésimo plot referente ao ponto central das variáveis de processo e $\bar{y}_{x..}$ é a média de todos os plots referente ao ponto central das variáveis de processo.

4.3 Resultados

Os resultados a seguir foram obtidos a partir de cinco combinações das três fontes de erro. Na prática é comum que as fontes de erro associados aos plots sejam maiores que o erro associado às unidades, logo nos cenários escolhidos este erro será sempre menor que os demais. Para comparação dos métodos de estimação, uma medida da eficiência relativa (E.R) é definida como:

$$ER = \frac{\text{Média dos } |(\mathbf{X}'\mathbf{V}_c^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}_c^{-1}\mathbf{V}_c\mathbf{V}_c^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}_c^{-1}\mathbf{X})^{-1}|}{|(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}|}$$

A Tabela 4.6 mostra um resumo das estimativas obtidas para a eficiência realtiva, junto com o coeficiente de variação do determinante da $\widehat{Var}(\hat{\beta})$ em 5.000 simulações.

Tabela 4.6: Estimativas de $ \widehat{Var}(\hat{\beta}) $		
Parâmetros	Método do Erro Puro	
$(\sigma_z, \sigma_x, \sigma_e)$	E.R	C.V
(6, 4, 2)	220066.1	4.535831
(4, 6, 2)	23053614	8.950964
(10, 5, 2)	894239590	36.40055
(5, 10, 2)	540501788	25.70046
(10, 10, 2)	3056715445	43.95888
Parâmetros	Método da MVR	
(6, 4, 2)	794.5374	8.69834
(4, 6, 2)	1127.927	8.939697
(10, 5, 2)	145158.3	20.03478
(5, 10, 2)	26088.98	19.59445
(10, 10, 2)	704154.1	39.81915

A Tabela 4.7, abaixo, mostra um resumo das estimativas das fontes de erro nas 5000 simulações obtidas em cada um dos cinco cenários observados.

Tabela 4.7: Variâncias Estimadas em 5.000 simulações

Método do Erro Puro										
Parâmetros ($\sigma_z, \sigma_x, \sigma_e$)	$\hat{\sigma}_z$			$\hat{\sigma}_x$			$\hat{\sigma}_e$			
	Média	D. P.	Mediana	Média	D. P.	Mediana	Média	D. P.	Mediana	
(6, 4, 2)	6.0828	7.9754	2.5989	3.6222	4.4047	1.1348	1.9982	1.9057	1.3676	
(4, 6, 2)	5.62018	6.8007	3.5266	5.9943	7.5094	3.0316	2.3762	2.0234	1.7556	
(10, 5, 2)	9.7230	14.0855	4.4198	5.1964	7.0944	2.2491	2.8649	2.8904	2.0244	
(5, 10, 2)	5.2069	7.4270	2.3840	10.5304	15.0878	4.7369	2.9948	2.9733	2.0912	
(10, 10, 2)	10.9723	15.3321	5.0129	10.1267	14.0497	4.7305	3.9008	3.6842	2.8327	
Método da MVR										
(6, 4, 2)	5.8232	12.3050	1.9508	0.8648	1.9212	0.0308	3.1181	1.4043	2.7759	
(4, 6, 2)	4.3461	9.4669	1.1541	1.5681	2.8826	0.0497	3.9361	2.6525	3.1118	
(10, 5, 2)	10.7793	20.9142	3.7225	2.0003	3.5691	0.1614	3.3825	1.9001	2.8754	
(5, 10, 2)	4.3114	10.2796	1.4609	3.7943	7.3956	0.0801	4.5373	3.3782	3.4644	
(10, 10, 2)	11.7852	24.2768	3.7974	3.8593	7.0054	0.3439	4.5421	3.4227	3.3711	

A partir da Tabela 4.7, observa-se que, em média, o método do Erro Puro estima com boa qualidade as fontes de erro, mas o desvio-padrão e a mediana indicam que a variabilidade destas estimativas é alta. Enquanto isso, o método da MVR estima bem o erro associado às variáveis de processo e o erro associado às unidades, mas subestima o erro associado às misturas. Neste método também existe uma grande variabilidade nas estimativas das fontes de erro.

Observando a Tabela 4.6 é interessante destacar que apesar do método do Erro Puro apresentar uma estimativa centrada para os parâmetros referentes às fontes de erro o mesmo não é confirmado na estimativa do determinante de $Var(\hat{\beta})$. A explicação para isso é que o determinante de $Var(\hat{\beta})$ indica uma estimativa conjunta para as três fontes de erro em cada uma das 5.000 simulações. Logo, em uma determinada simulação, o método pode estimar bem uma das fontes de erro, mas estimar com má qualidade as demais. Apesar de subestimar a fonte de erro associado às misturas, o método da MVR produz uma estimativa do determinante de $Var(\hat{\beta})$ com menor variabilidade se comparado com o método do Erro Puro, contudo os dois métodos produzem estimativas viesadas e com pouca precisão. Este resultado indica que para este tipo de experimento, com restrição na aleatorização do tipo *strip-plot* e com somente uma réplica dos pontos centrais das formulações da mistura e das variáveis de processo componentes, os métodos não conseguem estimar com boa qualidade as fontes de variação. É bem provável que isto ocorra devido a pouca quantidade de réplicas disponíveis.

Vale salientar que em algumas amostras geradas o método da MVR, por ser um método iterativo, não consegue uma convergência para as estimativas dos parâmetros das variâncias. Essa não convergência acontece cerca de 1 vez a cada 500 amostras simuladas. Nos casos em que não houve convergência a amostra foi desconsiderada, por não ser possível uma comparação adequada dos dois métodos.

Capítulo 5

Considerações finais

Este trabalho focou a estimativa de parâmetros em experimentos *strip-plot* com réplicas somente do ponto central das misturas e das variáveis de processo, considerando dois métodos de estimação. O interesse em analisar especificamente essa situação se deve à economia e praticidade proporcionadas pela utilização do plano *strip-plot* e pela ausência da réplica do experimento.

5.1 Conclusões

A partir das simulações a conclusão obtida é que os métodos não conseguem uma boa estimativa conjunta para as fontes de erro. O resultado obtido mostra falhas nos dois métodos de estimação dos parâmetros para um experimento *strip-plot* com réplicas somente dos pontos centrais de cada plot, o que pode não acontecer em um outro tipo de aleatorização, como, por exemplo, um plano com restrição do tipo *split-plot* com réplica somente dos pontos centrais do plot e dos subplots.

Uma grande vantagem do método da MVR comparado ao método do Erro Puro é que este depende de réplicas do experimento, enquanto aquele não depende. Lembrando que nem sempre há convergência no processo iterativo para o método da MVR.

5.2 Perspectiva de trabalhos futuros

Uma proposta para trabalhos futuros, que poderia até ser um estudo anterior a este, é observar estes métodos no caso em que há réplicas de todo o experimento, já que nas pesquisas realizadas não foi encontrado nada semelhante. Outra proposta seria aumentar mais uma réplica do ponto central de cada plot. Neste caso seriam necessárias 49 unidades experimentais, o que ainda tornaria o experimento viável, pois para se ter uma réplica de todo experimento seriam necessários 70 unidades experimentais. Outro possível trabalho seria tentar melhorar as estimativas dos parâmetros no método da MVR. Para isso, West et al. (2006) sugerem o uso do algoritmo E.M. como forma de encontrar um vetor inicial mais adequado para o processo iterativo, outra proposta, sugerida pela banca examinadora desta dissertação, seria tentar modificar o procedimento da MRV para que o vetor inicial use as estimativas obtidas no método do Erro Puro.

Existe, também, a possibilidade de observar outros métodos de estimação que proporcionem estimativas mais precisas para os parâmetros. Em fim, a maior pergunta a responder seria como obter um estimador para os parâmetros do modelo que seja não viesado e com boa precisão. Sendo respondida essa questão, provavelmente será mais fácil chegar uma maneira correta de analisar o experimento citado no capítulo três, no qual foi a motivação inicial de todo este trabalho.

Apêndice A

Código do R estimação dos parâmetros de um experimento *strip-plot*

```
require(nlme)
require(Matrix)
require(MASS)

##### Simulação da dissertação #####

# z1 negativo z2 positivo
l1=c(1, 0, 0, 0, 0, 0, -1, -1, 0, 0, 1, 0, 0)
l2=c(0, 1, 0, 0, 0, 0, -1, 0, -1, 0, 0, 1, 0)
l3=c(0, 0, 1, 0, 0, 0, -1, 0, 0, -1, 0, 0, 1)
l4=c(1/2, 1/2, 0, 1/4, 0, 0, -1, -1*1/2, -1*1/2, 0, 1/2, 1/2, 0)
l5=c(1/2, 0, 1/2, 0, 1/4, 0, -1, -1/2, 0, -1/2, 1/2, 0, 1/2)
l6=c(0, 1/2, 1/2, 0, 0, 1/4, -1, 0, -1/2, -1/2, 0, 1/2, 1/2)
l7=c(1/3, 1/3, 1/3, 1/9, 1/9, 1/9, -1, -1/3, -1/3, -1/3, 1/3, 1/3, 1/3) # Rx1
l8=c(1/3, 1/3, 1/3, 1/9, 1/9, 1/9, -1, -1/3, -1/3, -1/3, 1/3, 1/3, 1/3) # Rx2

# z1 positivo z2 positivo

l9=c(1, 0, 0, 0, 0, 0, 1, 1, 0, 0, 1, 0, 0)
l10=c(0, 1, 0, 0, 0, 0, 1, 0, 1, 0, 0, 1, 0)
l11=c(0, 0, 1, 0, 0, 0, 1, 0, 0, 1, 0, 0, 1)
l12=c(1/2, 1/2, 0, 1/4, 0, 0, 1, 1*1/2, 1*1/2, 0, 1/2, 1/2, 0)
l13=c(1/2, 0, 1/2, 0, 1/4, 0, 1, 1/2, 0, 1/2, 1/2, 0, 1/2)
l14=c(0, 1/2, 1/2, 0, 0, 1/4, 1, 0, 1/2, 1/2, 0, 1/2, 1/2)
l15=c(1/3, 1/3, 1/3, 1/9, 1/9, 1/9, 1, 1/3, 1/3, 1/3, 1/3, 1/3, 1/3) # Rx1
l16=c(1/3, 1/3, 1/3, 1/9, 1/9, 1/9, 1, 1/3, 1/3, 1/3, 1/3, 1/3, 1/3) # Rx2

# z1 negativo z2 negativo
```

```

117=c(1, 0, 0, 0, 0, 0, 1, -1, 0, 0, -1, 0, 0)
118=c(0, 1, 0, 0, 0, 0, 1, 0, -1, 0, 0, -1, 0)
119=c(0, 0, 1, 0, 0, 0, 1, 0, 0, -1, 0, 0, -1)
120=c(1/2, 1/2, 0, 1/4, 0, 0, 1, -1*1/2, -1*1/2, 0, -1/2, -1/2, 0)
121=c(1/2, 0, 1/2, 0, 1/4, 0, 1, -1/2, 0, -1/2, -1/2, 0, -1/2)
122=c(0, 1/2, 1/2, 0, 0, 1/4, 1, 0, -1/2, -1/2, 0, -1/2, -1/2)
123=c(1/3, 1/3, 1/3, 1/9, 1/9, 1/9, 1, -1/3, -1/3, -1/3, -1/3, -1/3, -1/3) # Rx1
124=c(1/3, 1/3, 1/3, 1/9, 1/9, 1/9, 1, -1/3, -1/3, -1/3, -1/3, -1/3, -1/3) # Rx2

```

z1 positivo z2 negativo

```

125=c(1, 0, 0, 0, 0, 0, -1, 1, 0, 0, -1, 0, 0)
126=c(0, 1, 0, 0, 0, 0, -1, 0, 1, 0, 0, -1, 0)
127=c(0, 0, 1, 0, 0, 0, -1, 0, 0, 1, 0, 0, -1)
128=c(1/2, 1/2, 0, 1/4, 0, 0, -1, 1*1/2, 1*1/2, 0, -1/2, -1/2, 0)
129=c(1/2, 0, 1/2, 0, 1/4, 0, -1, 1/2, 0, 1/2, -1/2, 0, -1/2)
130=c(0, 1/2, 1/2, 0, 0, 1/4, -1, 0, 1/2, 1/2, 0, -1/2, -1/2)
131=c(1/3, 1/3, 1/3, 1/9, 1/9, 1/9, -1, 1/3, 1/3, 1/3, -1/3, -1/3, -1/3) # Rx1
132=c(1/3, 1/3, 1/3, 1/9, 1/9, 1/9, -1, 1/3, 1/3, 1/3, -1/3, -1/3, -1/3) # Rx2

```

Ponto Central Rz1

```

133=c(1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)
134=c(0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)
135=c(0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)
136=c(1/2, 1/2, 0, 1/4, 0, 0, 0, 0, 0, 0, 0, 0, 0)
137=c(1/2, 0, 1/2, 0, 1/4, 0, 0, 0, 0, 0, 0, 0, 0)
138=c(0, 1/2, 1/2, 0, 0, 1/4, 0, 0, 0, 0, 0, 0, 0)
139=c(1/3, 1/3, 1/3, 1/9, 1/9, 1/9, 0, 0, 0, 0, 0, 0, 0) # Rx1
140=c(1/3, 1/3, 1/3, 1/9, 1/9, 1/9, 0, 0, 0, 0, 0, 0, 0) # Rx2

```

Ponto Central Rz2

```

141=c(1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)
142=c(0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)
143=c(0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)
144=c(1/2, 1/2, 0, 1/4, 0, 0, 0, 0, 0, 0, 0, 0, 0)
145=c(1/2, 0, 1/2, 0, 1/4, 0, 0, 0, 0, 0, 0, 0, 0)
146=c(0, 1/2, 1/2, 0, 0, 1/4, 0, 0, 0, 0, 0, 0, 0)
147=c(1/3, 1/3, 1/3, 1/9, 1/9, 1/9, 0, 0, 0, 0, 0, 0, 0) # Rx1
148=c(1/3, 1/3, 1/3, 1/9, 1/9, 1/9, 0, 0, 0, 0, 0, 0, 0) # Rx2

```

```

matx=matrix(c(11,12,13,14,15,16,17,18,19,110,111,112,113,114,
115,116,117,118,119,120,121,122,123,124,125,126,127,128,
129,130,131,132,133,134,135,136,137,138,139,140,141,142,
143,144,145,146,147,148),ncol=13,byrow=T,

```

```

dimnames = list(seq(1,48,1), c("X1", "X2", "X3","X1X2","X1X3",
                                "X2X3","Z1Z2","Z1X1","Z1X2","Z1X3","Z2X1","Z2X2","Z2X3"))

m=8; w=6
# Cenários escolhidos

#sigw1=6 ; sigw2=4 ; sige=2
#sigw1=4 ; sigw2=6 ; sige=2
#sigw1=10 ; sigw2=5 ; sige=2
#sigw1=5 ; sigw2=10 ; sige=2
#sigw1=10 ; sigw2=10 ; sige=2

# wholeplot como variável de processo + erro comum

jm=matrix(rep(1,m*m),ncol=m)
im=diag(sige,m)
qq1=sigw1*jm+im
n1vv=bdiag(qq1,qq1,qq1,qq1,qq1,qq1) # matriz de blocos diagonais
n1vv=as.matrix(n1vv)

# whole-plot como misturas

nvv=matrix(rep(0,(w*m)^2), ncol=m*w)
for(i in 1:(m*w-m)){
  k=rep((w-1):1,each=m)
  for(j in 1:k[i]){
    nvv[i,i+m*j]=nvv[i+m*j,i]=sigw2}}
nvv=nvv+diag(sigw2,48)

vv=n1vv+nvv # Matriz de covariância

## verdadeiro valor de variância de beta estimado

verd=solve(t(matx)%*%matx)%*%t(matx)%*%vv%*%matx)%*%solve(t(matx)%*%matx)
parametros=c(8.5,15,10,1,5,2,3,2,4,1.5,5,2,1)

wp1=c(rep(1:8, each=6))
wp2=rep(1:6,8)
wp1=as.factor(wp1)
wp2=as.factor(wp2)

z1=c(rep(-1,8),rep(1,8),rep(-1,8),rep(1,8),rep(0,16))
z2=c(rep(1,8),rep(1,8),rep(-1,8),rep(-1,8),rep(0,16))
dad2=cbind(matx,z1,z2)
dad2=data.frame(dad2) # organizando para usar em nlme

```

```

nrep=rep(1,48)
detvbep=NULL ; detvbr=NULL
vwple=NULL ; vwp2e=NULL ; vep=NULL
vwp1r=NULL ; vwp2r=NULL ; ver=NULL

for(i in 1:5000){
  erro=mvrnorm(n=1, rep(0,48), vv)
  dados=matx%%parametros+erro

  ##### Erro Puro

  ## estimando as variâncias

  # Whole-plot associado aos fatores de processo

  mediag1=mean(dados[33:48])
  vwp1e[i]=(mean(dados[33:40])-mediag1)^2+(mean(dados[41:48])-mediag1)^2

  # Whole-plot associado aos fatores de mistura

  mediag2=mean(c(dados[seq(7,47,8)],dados[seq(8,48,8)]))
  vwp2e[i]=(mean(dados[seq(7,47,8)])-mediag2)^2+(mean(dados[seq(8,48,8)])-mediag2)^2

  # Erro associado às unidades

  vard=c(dados[39],dados[40],dados[47],dados[48])
  vep[i]=var(vard)

  ### Construindo a matriz de variância-covariância estimada usando EP

  jmep=matrix(rep(1,m*m),ncol=m)*vwp1e[i]
  imep=diag(vep[i],m)
  qep=jmep+imep
  n1vvep=bdiag(qep,qep,qep,qep,qep,qep)
  n1vvep=as.matrix(n1vvep)

  nvve=matrix(rep(0,(w*m)^2), ncol=m*w)
  for(l in 1:(m*w-m)){
    k=rep((w-1):1,each=m)
    for(j in 1:k[1]){
      nvve[l,1+m*j]=nvve[l+m*j,1]=vwp2e[i]}
    nvve=nvve+diag(vwp2e[i],48)

  vvep=nvve+n1vvep

  varbep=solve(t(matx)%%solve(vvep)%%matx)%%t(matx)%%solve(vvep)%%
  vvep%%solve(vvep)%%matx%%solve(t(matx)%%solve(vvep)%%matx)

```

```

detvbep[i]=det(varbep) # ok

##### MVR

spb4=lme( dados ~ 0 + X1 + X2 + X3 + X1X2 + X1X3 + X2X3 + Z1Z2
+ Z1X1 + Z1X2 + Z1X3 + Z2X1 + Z2X2 + Z2X3 , random=list(nrep=pdBlocked(list(pdIdent(~1),
pdIdent(~wp1-1),pdIdent(~ordered(wp2)-1))))),data=dad2)

### estraindo variâncias
vwp1r[i]=as.numeric(VarCorr(spb4)[2]) # retorna a estimativa para wp1
vwp2r[i]=as.numeric(VarCorr(spb4)[10]) # retorna a estimativa para wp2
ver[i]=as.numeric(VarCorr(spb4)[16]) # retorna a estimativa para o erro geral

### Construindo a matriz de variância-covariância estimada usando MVR

jmr=matrix(rep(1,m*m),ncol=m)*vwp1r[i]
imr=diag(ver[i],m)
qqr=jmr+imr
vvr1=bdiag(qqr,qqr,qqr,qqr,qqr,qqr)
vvr1=as.matrix(vvr1)

nvvr2=matrix(rep(0,(w*m)^2), ncol=m*w)
for(l in 1:(m*w-m)){
k=rep((w-1):1,each=m)
for(j in 1:k[l]){
nvvr2[l,l+m*j]=nvvr2[l+m*j,l]=vwp2r[i]}
nvvr2=nvvr2+diag(vwp2r[i],48)

vvr=vvr1+nvvr2

varbr=solve(t(matx)%*%solve(vvr)%*%matx)%*%t(matx)%*%solve(vvr)%*%
vvr)%*%solve(vvr)%*%matx)%*%solve(t(matx)%*%solve(vvr)%*%matx)
detvbr[i]=det(varbr)
}

# Resumos
mean(vwp1e); sd(vwp1e) ; median(vwp1e)
mean(vwp2e); sd(vwp2e) ; median(vwp2e)
mean(vep); sd(vep) ; median(vep)
mean(vwp1r); sd(vwp1r) ; median(vwp1r)
mean(vwp2r); sd(vwp2r) ; median(vwp2r)
mean(ver); sd(ver) ; median(ver)
mean(detvbep); sd(detvbep);median(detvbep)
mean(detvbr); sd(detvbr);median(detvbr)

```


Referências Bibliográficas

- Box, G. & Jones, S. (1992), ‘Split-plot designs for robust product experimentation’, *Journal of Applied Statistics* **19**(1), 3–26.
- Cornell, J. (1990), *Experiments with mixtures, designs, models, and the analysis of mixture data*, John Wiley.
- Daniel, C. (2008), *Estudo do desempenho de métodos Lenth na análise do tipo Split-Plot na ausência de réplicas*, Dissertação (Mestrado em Estatística), UFPE, Recife.
- Harville, D. (1977), ‘Maximum likelihood approaches to variance component estimation and to related problems’, *Journal of the American Statistical Association* pp. 320–338.
- Kowalski, S., Cornell, J. & Vining, G. (2002), ‘Split-plot designs and estimation methods for mixture experiments with process variables’, *Technometrics* **44**(1), 72–79.
- Littell, R., Milliken, G., Stroup, W. & Wolfinger, R. (1996), *SAS system for mixed models*, SAS Institute, Cary, NC.
- Milliken, A. & Johnson, D. (2000), *Analysis of Messy Data: Analysis of Covariance*, Chapman Hall/CRC.

- Montgomery, D. (2002), *Design and Analysis of Experiments, Student Solutions Manual*, John Wiley Sons.
- Myers, R. & Montgomery, D. (1995), *Response Surface Methodology: Process and Product in Optimization Using Designed Experiments*, John Wiley and Sons, Inc. New York, NY, USA.
- Patterson, H. & Thompson, R. (1971), ‘Recovery of inter-block information when block sizes are unequal’, *Biometrika* **58**(3), 545–554.
- Pinheiro, J. & Bates, D. (2000), *Mixed-effects models in S and S-PLUS*, Springer.
- Smith, W. F. (2005), *Experimental design for formulation*, Society for Industrial Mathematics.
- Team, R. (2008), ‘R: A language and environment for statistical computing’, *Vienna, Austria: R Foundation for Statistical Computing* .
- Vivacqua, C. A. & Bisgaard, S. (2004), ‘Strip-block experiments for process improvement and robustness’, *Quality Engineering* **16**(3), 495–500.
- West, B., Welch, K. & Galecki, A. (2006), *Linear mixed models: a practical guide using statistical software*, Chapman Hall/CRC.

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)