

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA
DEPARTAMENTO DE QUÍMICA

NERI GODOY TOBIAS JÚNIOR

LINGUAGEM R APLICADA AO PLANEJAMENTO FATORIAL EM QUÍMICA

SÃO CARLOS -SP

2022

NERI GODOY TOBIAS JÚNIOR

LINGUAGEM R APLICADA AO PLANEJAMENTO FATORIAL EM QUÍMICA

Trabalho de conclusão de curso apresentado ao Departamento de Química da Universidade Federal de São Carlos, para obtenção do título de bacharel em Química Tecnológica.

Orientador: Edenír Rodrigues Pereira-Filho

SÃO CARLOS -SP

2022



FUNDAÇÃO UNIVERSIDADE FEDERAL DE SÃO CARLOS

DEPARTAMENTO DE QUÍMICA - DQ/CCET

Rod. Washington Luís km 235 - SP-310, s/n - Bairro Monjolinho, São Carlos/SP, CEP
13565-905

Telefone: (16) 33518206 - <http://www.ufscar.br>

DP-TCC-FA nº 20/2022/DQ/CCET

Graduação: Defesa Pública de Trabalho de Conclusão de Curso

Folha Aprovação (GDP-TCC-FA)

FOLHA DE APROVAÇÃO

NERI GODOY TOBIAS JÚNIOR

LINGUAGEM R APLICADA AO PLANEJAMENTO FATORIAL EM QUÍMICA

Trabalho de Conclusão de Curso

Universidade Federal de São Carlos - Campus São Carlos

São Carlos, 09 de setembro de 2022

ASSINATURAS E CIÊNCIAS

Cargo/Função	Nome Completo
Orientador	Prof. Dr. Edenir Rodrigues Pereira Filho
Membro da Banca 1	Dr. Marco Aurelio Sperança
Membro da Banca 2	Prof. Dr. Alex Virgílio



Documento assinado eletronicamente por **Caio Marcio Paranhos da Silva, Professor(a)**, em 27/09/2022, às 09:23, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site <https://sei.ufscar.br/autenticacao>, informando o código verificador **0828902** e o código CRC **6C367A90**.

Referência: Caso responda a este documento, indicar expressamente o Processo nº 23112.035430/2022-02

SEI nº 0828902

Modelo de Documento: Grad: Defesa TCC: Folha Aprovação, versão de 02/Agosto/2019

Lista de abreviações:

GL - Grau de Liberdade

SQReg - Soma quadrática da regressão

SQRes - Soma quadrática dos resíduos

SQT - Soma quadrática total

SQEP - Soma quadrática do erro puro

SQFaj - Soma quadrática da falta de ajuste

MQ - Média quadrática

MQReg - Média quadrática da regressão

MQRes - Média quadrática do resíduo

MQTot - Média quadrática total

MQEP - Média quadrática do erro puro

MQFaj - Média quadrática da falta de ajuste

ANOVA – Analysis of Variance

IF - Intensidade de Fluorescência

PEG - Polietileno Glicol

PVA - Álcool Polivinílico

CLF - Clorofórmio

DCM - Diclorometano

ACN - Acetonitrila

DIA -Desisopropilatrazina

DEA - Desetilatrazina

ATZ - Atrazina

DIU - Diuron

TFB - Teflubenzuron

Lista de Tabelas

Tabela 1 - Tabela ANOVA	14
Tabela 2 - Polinômios utilizados no planejamento de misturas.....	16
Tabela 3 - Valores codificados e reais para as variáveis e a resposta.....	18
Tabela 4 - Valores codificados e reais para as variáveis de planejamento fatorial fracionário e respostas.....	19
Tabela 5 - Valores codificados e reais para as variáveis de planejamento fatorial do tipo Doehlert e respostas.....	20
Tabela 6 - Valores codificados das variáveis e respostas.....	23
Tabela 7 - Médias quadráticas, testes F e coeficientes de determinação para o modelo obtido.....	27
Tabela 8 - Médias quadráticas, testes F e coeficientes de determinação para o modelo recalculado.....	30
Tabela 9 - Pesos atribuídos à desejabilidade de resposta.....	34
Tabela 10 - Médias quadráticas, testes F e coeficientes de determinação para o modelo recalculado no planejamento de misturas.....	35
Tabela 11 - Valores previstos pelo modelo obtido com intervalo de confiança e valores obtidos pelos autores do artigo para cada resposta.....	37

Lista de Figuras

Figura 1 - Pontos Experimentais em 1) Planejamento Fatorial Completo 2) Box- Behnken e 3) Composto Central.....	13
Figura 2 - Gráfico de contorno utilizado no planejamento de misturas.....	15
Figura 3 - a) Porcentagem de Efeitos e b) Probabilidade de Efeitos.....	25
Figura 4 - a) Porcentagem de Efeitos e b) Probabilidade de Efeitos para a resposta 1, diâmetro.....	26
Figura 5 - a) Porcentagem de Efeitos e b) Probabilidade de Efeitos para a resposta 2, distribuição.....	26
Figura 6 - Valores experimentais vs Valores previstos, com intervalo de confiança.....	28
Figura 7 - a) Histograma dos resíduos b) Valores de resíduos vs previsto.....	29
Figura 8 - Coeficientes do modelo de regressão com intervalo de confiança.....	29
Figura 9 - Valores experimentais vs Valores previstos, com intervalo de confiança, para o modelo recalculado.....	31
Figura 10 - a) Histograma dos resíduos b) Valores de resíduos vs previsto, para o modelo recalculado.....	32
Figura 11 - Coeficientes do modelo de regressão recalculado com intervalo de confiança.....	32
Figura 12 - a) Gráfico de contorno e b) Superfície de Resposta para a resposta de intensidade de fluorescência do antimônio.....	33
Figura 13 - Valores experimentais vs Valores previstos, com intervalo de confiança, para o modelo recalculado no planejamento de misturas.....	35
Figura 14 - a) Histograma dos resíduos b) Valores de resíduos vs previsto, para o modelo recalculado para o planejamento de misturas.....	36

Figura 15 - Gráfico de contorno para o modelo obtido utilizando a desejabilidade global.....37

AGRADECIMENTOS

Ao Prof. Dr. Edenir Rodrigues Pereira Filho pela paciência, ajuda e por aceitar orientar este trabalho de conclusão de curso.

Aos meus pais, Silvia e Laércio, por todo apoio e incentivo durante todos esses anos.

À minha família, em especial à minha irmã Ana Silvia, meu cunhado André Luís e minha sobrinha Ana Carolina, pelo apoio.

À Universidade Federal de São Carlos por disponibilizar recursos e a estrutura necessária para minha formação. Em especial ao Departamento de Química e todos docentes e funcionários.

A todos amigos que fiz durante essa etapa, em especial ao André e Pedro pela companhia durante os estudos e por tornar essa etapa mais agradável.

RESUMO

O presente trabalho teve como objetivo adaptar para a linguagem R as rotinas *regression2* e *fabi_efeito* de Octave/Matlab e aplicá-las no contexto do planejamento fatorial em química utilizando alguns exemplos, assim como utilizar pacotes existentes disponíveis para o R na otimização de um experimento e na composição de uma mistura. O planejamento fatorial aplicado à química é utilizado auxiliando no entendimento de sistemas e otimização de processos. Para os cálculos utilizados nestes métodos, comumente são utilizadas planilhas virtuais ou *softwares* capazes de realizar os mesmos, sendo o *software* livre RStudio mais uma opção no tratamento, análise e visualização destes dados, contando com uma vasta biblioteca de pacotes que podem ser aplicadas ao planejamento fatorial em química. Neste trabalho, através do pacote *PacoteEfeitos*, uma adaptação das rotinas *fabi_efeito* e *regression2*, foi possível o cálculo de porcentagem e probabilidade de efeitos, obtenção de modelos de regressão, assim como a geração de gráficos para a visualização dos mesmos e seus parâmetros, além disso, através de pacotes disponíveis para o R, como *NlcOptim*, *rsm* e *mixexp*, foi possível obter gráficos de contorno e superfícies de resposta para um planejamento do tipo Doehlert e gráfico de contorno para um planejamento de misturas, assim como suas regiões de ótimo e respostas previstas.

ABSTRACT

The present work aimed to adapt the `fabi_efeito` and `regression2` scripts from Octave/Matlab to the R language and apply them in the context of factorial design in chemistry using some examples, as well as using existing packages available for R in the optimization of an experiment and in the composition of a mixture. Factorial design applied to chemistry is used to help understand systems and optimize processes. For the calculations used in these methods, virtual spreadsheets or softwares capable of performing it are commonly used, with the free software RStudio being an alternative in the processing, analysis and visualization of these data, with a vast library of packages that can be applied to factorial design in chemistry. In this work, through the `PacoteEfeitos` package, an adaptation of the `fabi_efeito` and `regression2` scripts, it was possible to calculate the percentage and probability of effects, obtain regression models, as well as the plot of graphs to visualize them and their parameters, in addition, through packages available for R, such as `NlcOptim`, `rsm` and `mixexp`, it was possible to obtain contour plots and response surfaces for a Doehlert-type design and contour plot for a mixture design, as well as their optimal regions and predicted responses.

SUMÁRIO

1. Introdução	11
1.1 Planejamento Fatorial Completo	12
1.2 Planejamento Fatorial Fracionário.....	12
1.3 Modelos de regressão e Otimização de experimentos.....	13
1.4 Planejamento de Misturas	15
1.5 Desejabilidade.....	16
2. Objetivos	17
3. Materiais e Métodos	17
3.1 Planejamento Fatorial Completo	17
3.2 Planejamento Fatorial Fracionário.....	19
3.3 Obtenção de um modelo de regressão e otimização.	20
3.4 Planejamento de Misturas	22
4 Resultados e Discussões	25
4.1 Planejamento Fatorial Completo	25
4.2 Planejamento Fatorial Fracionário.....	26
4.3 Obtenção de um modelo de regressão e otimização.	27
4.4 Planejamento de Misturas	33
5 Conclusão	38
6. Referências	39

1. Introdução

O tema principal deste trabalho foi o planejamento fatorial aplicado à química baseado nos trabalhos de PEREIRA e PEREIRA-FILHO (2018) e HILÁRIO et al. (2021), utilizando o *software* livre RStudio na versão 2022.7.1.554 como ferramenta auxiliar na obtenção dos resultados. Assim, foi realizada uma adaptação das rotinas de Octave e Matlab, *fabi_efeito* e *regression2* para a linguagem R.

Através da experimentação, busca-se obter dados que possam elucidar o comportamento de um sistema através da influência de suas variáveis. Entretanto, para que isso ocorra, os experimentos devem ser planejados de modo que as informações obtidas sejam relevantes e confiáveis para que o objetivo do mesmo seja atingido.

O planejamento fatorial, um ramo da estatística utilizado na formulação de experimentos e que pode ser empregado na otimização, na redução da variabilidade de processos e para extrair mais informações acerca do sistema a ser estudado, se destaca por necessitar de menos tempo e menor custo para ser realizado devido ao número reduzido de experimentos em relação ao método OFAT (*One-Factor-at-a-Time*) (J. N. Miller; J. C. Miller, 2005). Em contraste com o método OFAT, em que apenas um fator é alterado por vez, o planejamento fatorial trata-se de uma análise multivariada, ou seja, mais de um fator é alterado simultaneamente, possibilitando também a determinação dos efeitos principais e das interações entre os mesmos (BOX et al., 1978).

Com isso, através do auxílio de programas computacionais, é possível manipular e tratar os dados obtidos através da experimentação a fim de obter modelos de regressão e condições ideais para a otimização das condições experimentais. O R, uma linguagem de programação livre e voltada à manipulação, análise e visualização de dados, pode ser utilizada para essa finalidade. O R, linguagem desenvolvida em 1993 e que em 1995, se tornou um *software* livre, permitindo a colaboração de pessoas para o desenvolvimento de pacotes. Com isso, conta com uma vasta biblioteca de pacotes, pacotes estes que possuem funções adicionais ao *software* para tratamentos e análises mais específicas para determinada área do conhecimento. Alguns desses processos serão detalhados ao longo deste trabalho de conclusão de curso.

1.1 Planejamento Fatorial Completo

O planejamento fatorial completo consiste em realizar todas possíveis combinações de níveis entre os fatores, ou seja, se existem n níveis e k fatores, o número de experimentos será n^k , além de experimentos adicionais a fim de criar réplicas que possibilitem o cálculo do erro experimental. Com apenas 2 níveis, ou seja, um planejamento do tipo 2^k , é possível realizar uma varredura, verificando quais os fatores mais significativos. Para isso, combinações de níveis dos fatores envolvidos são investigadas, verificando a mudança na resposta produzida pela alteração do nível de cada fator. Essa mudança na resposta é denominada efeito de um fator ou efeito principal. Além dos efeitos principais, em que se leva em conta apenas o impacto dos fatores individuais nas mudanças das respostas, existem os efeitos de interação entre estes quando a influência de um fator na resposta varia quando outro fator é alterado (MONTGOMERY, 2017).

1.2 Planejamento Fatorial Fracionário

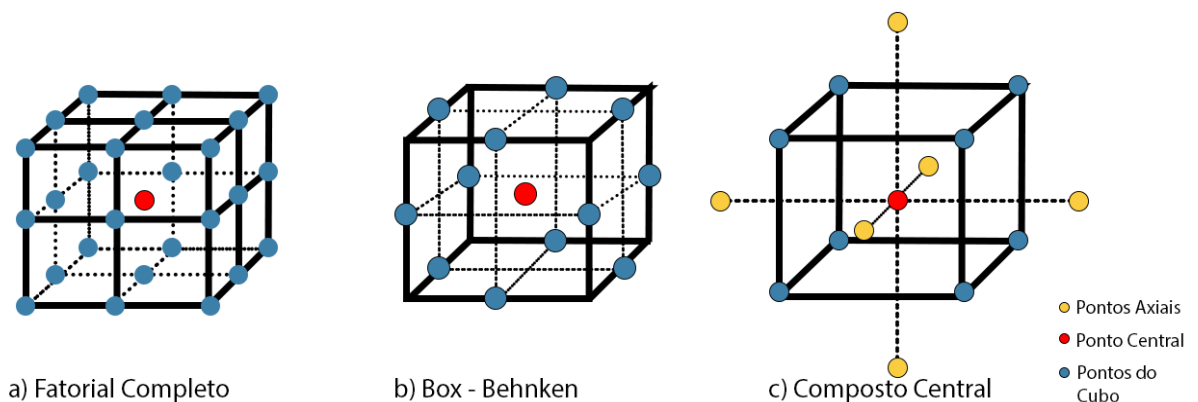
O planejamento fatorial fracionário é utilizado com mais frequência na triagem de experimentos, nos estágios iniciais de um estudo, onde há muitos fatores e pretende-se identificar os mais significativos, ou seja, com maior impacto na resposta (BRERETON, 2003). Isso possibilita a eliminação dos fatores menos importantes e prosseguir com um estudo aprofundado apenas com os mais relevantes, tornando o experimento mais econômico (J. N. Miller; J. C. Miller, 2005).

No planejamento fatorial fracionário, uma fração dos experimentos do fatorial completo é feita e é denotado por 2^{k-n} , sendo k e n números inteiros e $1 \leq n < k$ (J. N. Miller; J. C. Miller, 2005). Esse tipo de planejamento também é chamado de $(\frac{1}{2})^p$ frações de 2^k , com isso, são obtidos $2^{(k-p)} - 1$ contrastes, um número reduzido de efeitos em relação ao planejamento fatorial completo. Entretanto, ao fazer a redução de números de experimentos, pode ser perdida informações, pois efeitos e interações são confundidos entre si, ou seja, seus níveis são iguais e para identificá-los, pode-se utilizar a relação do gerador utilizado e multiplicar por todos efeitos, obtendo-se assim todos os contrastes. Para isso, utilizam-se p geradores do experimento, selecionando estes de forma que possibilite o planejamento com a maior resolução possível e não faça com que efeitos de interesse sejam confundidos entre si (MONTGOMERY, 2017).

1.3 Modelos de regressão e Otimização de experimentos

Para a otimização de experimentos, os planejamentos do tipo fatorial completo com três níveis, Box-Behnken e composto central (CCD) são os mais utilizados. Box-Behnken é caracterizado por possuir 3 níveis para cada variável que se encontram equidistantes do ponto central, já o CCD possui além dos pontos presentes no planejamento fatorial completo ou fracionário, pontos axiais e central. (Figura 1). Diferente dos planejamentos fatoriais com apenas dois níveis, em que a superfície de resposta obtida é de 1ª ordem, a utilização de planejamentos com 3 a 5 níveis para cada fator é vantajosa pois permite obter modelos de 2ª ordem, possibilitando a avaliação da curvatura da mesma (FUKUDA et al., 2018).

Figura 1 - Pontos Experimentais em a) Planejamento Fatorial Completo b) Box-Behnken e c) Composto Central



Com o maior número de níveis, mais experimentos serão necessários em comparação com apenas dois níveis, sendo estes tipos de planejamentos fatoriais utilizados após a etapa de varredura e realizados apenas com os fatores mais importantes a serem estudados (FUKUDA et al., 2018).

Os modelos podem ser desde os mais simples, considerando apenas a influência dos fatores de maneira linear (Equação 2), até um modelo completo que leva em conta os fatores, interações e termos quadráticos (Equação 3), em que y é a resposta prevista, β_0 uma constante, β_i os coeficientes lineares, β_{ii} os coeficientes quadráticos e β_{ij} os coeficientes de interação de primeira ordem. A escolha do modelo é baseada na análise de variância (*Analysis of Variance*, ANOVA) a fim de verificar a significância do modelo de regressão, através dos elementos da tabela ANOVA (Tabela 1) em que n_i são o número de repetições no nível i , m o número de níveis

distintos da variável independente, n o número total de observações e p o número de parâmetros do modelo. Com isso, decidir quais coeficientes devem estar presentes no modelo, baseando-se no valor- p de cada um, ou seja, admitindo-se apenas coeficientes que sejam diferentes de 0 de maneira significativa (FUKUDA et al., 2018).

$$y = \beta_0 + \sum_i^k \beta_i \times x_i + \varepsilon \quad (2)$$

$$y = \beta_0 + \sum_i^k \beta_i \times x_i + \sum_{ij}^k \beta_{ij} \times x_i \times x_j + \sum_i^k \beta_{ii} \times x_i^2 + \varepsilon \quad (3)$$

Tabela 1 - Tabela ANOVA

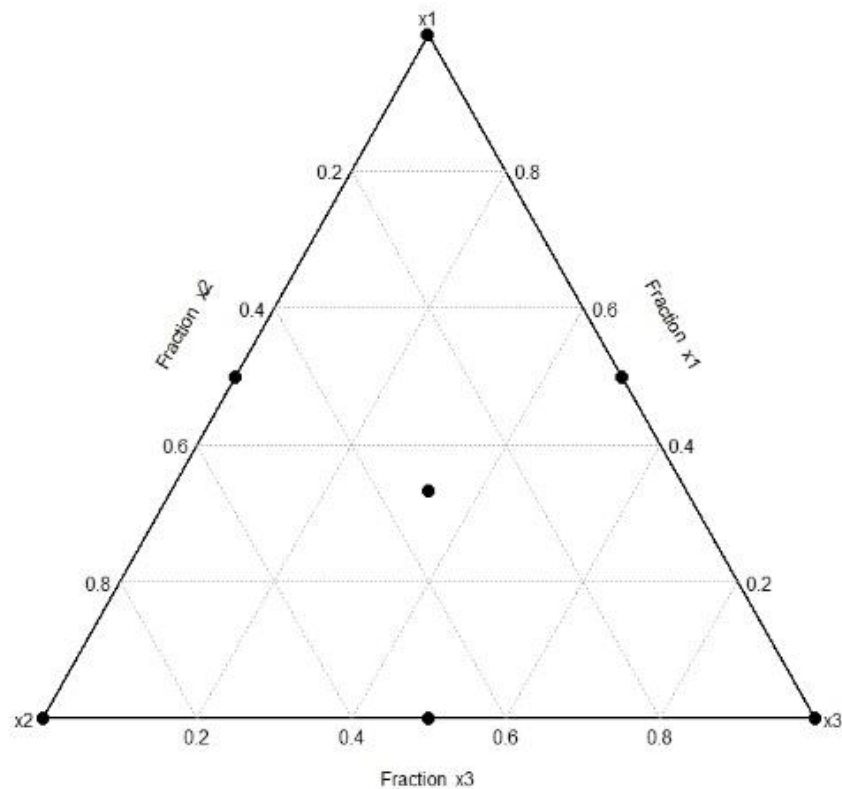
Fonte de variação	Graus de liberdade	Soma Quadrática	Média Quadrática	F calculado
Regressão	$p-1$	$\sum_i^m \sum_j^{n_i} (\hat{y}_i - \bar{y})^2$	$\frac{\text{SQ regressão}}{p-1}$	$\frac{\text{MQ regressão}}{\text{MQ resíduos}}$
Resíduos	$n-p$	$\sum_i^m \sum_j^{n_i} (y_{ij} - \hat{y}_i)^2$	$\frac{\text{SQ resíduos}}{n-p}$	
Falta de Ajuste	$m-p$	$\sum_i^m \sum_j^{n_i} (y_i - \bar{y}_i)^2$	$\frac{\text{SQ falta de ajuste}}{m-p}$	$\frac{\text{MQ falta de ajuste}}{\text{MQ erro puro}}$
Erro Puro	$n-m$	$\sum_i^m \sum_j^{n_i} (y_{ij} - \bar{y}_i)^2$	$\frac{\text{SQ erro puro}}{n-m}$	
Total	$n-1$	MQTot = SQ regressão + SQ resíduos		
% de variação explicada (R^2)	$\frac{\text{SQ regressão}}{\text{SQ total}}$			
% máxima variação explicável	$\frac{\text{SQ total} - \text{SQ erro puro}}{\text{SQ total}}$			

Os coeficientes de determinação R^2 e R^2 máximo devem ser observados, uma vez que são medidas de como o modelo de regressão explica e prevê novas respostas, ambos podem assumir valores entre 0 e 1, quanto mais próximo de 1, maior capacidade do modelo de explicar as mudanças na variável dependente. Além disso, quando existem réplicas no planejamento fatorial, pode-se estimar a falta de ajuste do modelo. Quando a variabilidade pela falta de ajuste do modelo de regressão é maior que o erro puro, trata-se de uma evidência da falta de ajuste do modelo (MONTGOMERY, 2017).

1.4 Planejamento de Misturas

No planejamento de misturas, diferente do planejamento fatorial, existe uma relação entre as variáveis, ou seja, a soma das proporções de todos componentes da mistura deve ser igual à 100%. Com isso, os níveis dos fatores são proporções dos componentes da mistura. Consideremos, por exemplo que 40% de uma mistura é do componente A e 20% do componente B, o componente C fica condicionado a ser de 40% para satisfazer a restrição. Na Figura 2 é possível visualizar um gráfico de contorno para o planejamento de misturas com 3 fatores, em cada vértice encontram-se pontos em que a mistura é composta por somente um fator.

Figura 2 – Gráfico de contorno utilizado no planejamento de misturas



Devido à restrição da soma dos componentes ser sempre igual à 100%, os modelos de misturas se alteram em relação aos modelos quando não há restrição. Para os modelos de mistura ternárias, são utilizados os polinômios como mostrados na Tabela 2 (MONTGOMERY, 2017), sendo o cúbico especial o mais eficiente (BEZERRA et al., 2020).

Tabela 2 – Polinômios utilizados no planejamento de misturas

Linear	$E(y) = \sum_{i=1}^p \beta_i x_i$
Quadrático	$E(y) = \sum_{i=1}^p \beta_i x_i + \sum_{i<j}^p \beta_{ij} x_i x_j$
Cúbico	$E(y) = \sum_{i=1}^p \beta_i x_i + \sum_{i<j}^p \beta_{ij} x_i x_j + \sum_{i<j}^p \delta_{ij} x_i x_j (x_i - x_j) + \sum_{i<j<k}^p \beta_{ijk} x_i x_j x_k$
Cúbico Especial	$E(y) = \sum_{i=1}^p \beta_i x_i + \sum_{i<j}^p \beta_{ij} x_i x_j + \sum_{i<j<k}^p \beta_{ijk} x_i x_j x_k$

1.5 Desejabilidade

A função da desejabilidade surge como uma alternativa ao método de sobrepor gráficos individuais para encontrar a região de ótimo quando trabalhando com múltiplas respostas. Este método consiste em transformar cada resposta em uma desejabilidade individual, esta que por sua vez pode variar de 0 a 1 e é utilizada para maximizar (Equação 4), minimizar (Equação 5) ou atingir um valor específico da resposta (Equação 6). Para o cálculo da desejabilidade individual, L, U, T e r são o limite inferior, superior, alvo e o peso, respectivamente (DERRINGER, SUICH, 1980).

$$d = \begin{cases} 0 & y < L \\ \left(\frac{y-L}{T-L}\right)^r & L \leq y \leq T \\ 1 & y > T \end{cases} \quad (4)$$

$$d = \begin{cases} 1 & y < T \\ \left(\frac{U-y}{U-T}\right)^r & T \leq y \leq U \\ 0 & y > U \end{cases} \quad (5)$$

$$d = \begin{cases} 0 & y < L \\ \left(\frac{y-L}{T-L}\right)^{r1} & L \leq y \leq T \\ \left(\frac{U-y}{U-T}\right)^{r2} & T \leq y \leq U \\ 0 & y > U \end{cases} \quad (6)$$

A partir das m desejabilidades individuais, a desejabilidade global pode ser então calculada (Equação 7) (DERRINGER, SUICH, 1980).

$$D = (d_1, d_2 \dots d_m)^m \quad (7)$$

2. Objetivos

Os objetivos deste trabalho de conclusão de curso estão centrados em adaptar rotinas utilizadas para o cálculo de efeitos e obtenção de modelos de regressão no planejamento fatorial para a linguagem R, além de empregar pacotes existentes na obtenção de gráficos de contorno, superfície de resposta e condições de ótimo em exemplos aplicados à química.

3. Materiais e Métodos

3.1 Planejamento Fatorial Completo

Utilizou-se os pacotes `dplyr`, `matlib`, `ggplot2` e `devtools`, necessários para o funcionamento e instalação do `PacoteEfeitos`, além do próprio `PacoteEfeitos`, que se trata de uma adaptação das rotinas `fabi_efeito` e `regression2` de Octave para o R.

Para a instalação dos pacotes, utilizou-se a função `install.packages`, inserindo entre parênteses o nome do pacote a ser instalado, já para o `PacoteEfeitos`, utilizou-se a função `install_github`, disponível no pacote `devtools` e inseriu-se o endereço do repositório em que o pacote se encontra.

Com os pacotes instalados e carregados, a partir dos dados obtidos por SANTOS et al. (2018), referentes à influência de 3 variáveis na intensidade de fluorescência de antimônio presente em amostras de água (Tabela 3), criou-se a matriz **X** contendo as variáveis codificadas com a função `matrix`, especificando-se o número de linhas e colunas nos argumentos `nrow` e `ncol`, respectivamente. Com a função `cbind` adicionou-se as colunas referentes as interações entre os fatores à matriz **X** e para o vetor `y` contendo os valores das respostas, utilizou-se a função `rbind`.

Tabela 3 - Valores codificados e reais para as variáveis e a resposta.

Experimentos	HCl		NaBH ₄		Tempo de Retenção		IF do Sb
	Codificado	Real (mol/L)	Codificado	Real (%m/v)	Codificado	Real (min)	
1	-1	3	-1	1	-1	10	178,4
2	-1	3	-1	1	1	30	167,5
3	-1	3	1	3	-1	10	225,7
4	-1	3	1	3	1	30	218,1
5	1	5	-1	1	-1	10	86,6
6	1	5	-1	1	1	30	91,0
7	1	5	1	3	-1	10	195,6
8	1	5	1	3	1	30	189,2
9	0	4	0	2	0	20	137,5
10	0	4	0	2	0	20	135,7
11	0	4	0	2	0	20	137,8

Com os dados adicionados ao R, utilizou-se as funções existentes no PacoteEfeitos, `Porcentagem(X,y,k)` e `Probabilidade(X,y,PC,t)`, em que **X** e **y** são a matriz e vetor criados com os dados, **k** o número de fatores, **PC** a combinação de valores no ponto central e **t** o valor de *t* correspondente ao nível de significância desejado, onde o valor de *t* utilizado foi de 0,975, para 95% de significância, ficando dessa forma:

```
library("dplyr")
library("matlib")
library("ggplot2")
library("devtools")

devtools::install_github("ntgj/PacoteEfeitos")

library("PacoteEfeitos")

X <- matrix(c(-1,-1,-1,-1,1,1,1,1,-1,-1,1,1,-1,-1,1,1,-1,1,-1,1,-1,1),
, nrow = 8, ncol = 3)
X <- cbind(X[,1:3],X[,1]*X[,2],X[,1]*X[,3],X[,2]*X[,3])
X <- cbind(X[,1:6],X[,1]*X[,2]*X[,3])

y <- rbind(178.4,167.5,225.7,218.1,86.6,91,195.6,189.2)

Porcentagem(X,y,3)

Probabilidade(X,y,3,c(137.5,135.7,137.8))
```

3.2 Planejamento Fatorial Fracionário

Seguiu-se etapas semelhantes ao do Exemplo 1, criou-se a partir dos dados obtidos por MORAIS et al. (2017), em que o objetivo era de minimizar o diâmetro das partículas e maximizar a distribuição (Tabela 4) a matriz **X** contendo as variáveis codificadas e suas respectivas interações a partir da função matrix, o vetor **y** através da função rbind e executou-se as funções Porcentagem e Probabilidade, inserindo os parâmetros e com nível de significância de 95%. O procedimento foi realizado para as duas respostas, alterando-se apenas o vetor **y** e os valores do ponto central.

Tabela 4 - Valores codificados e reais para as variáveis de planejamento fatorial fracionário e respostas

Experimentos	PEG (1)		Vol. Fase Aq. (2)		PVA (3)		Vol. Fase Aq. (4=123)		Respostas	
	Cod.	Real (mg)	Cod.	Real (mL)	Cod.	Real (%)	Cod.	Real (mL)	1 (µm)	2
1	1	200	-1	1	-1	0,5	1	100	28,4	2,02
2	-1	50	-1	1	1	2	1	100	26,0	1,29
3	1	200	1	3	1	2	1	100	14,1	1,62
4	-1	50	1	3	-1	0,5	1	100	8,1	2,51
5	-1	50	1	3	1	2	-1	30	13,0	4,76
6	-1	50	-1	1	-1	0,5	-1	30	14,7	1,79
7	1	200	-1	1	1	2	-1	30	7,8	3,23
8	1	200	1	3	-1	0,5	-1	30	6,7	3,61
9	0	125	0	2	0	1,25	0	65	20,8	1,56
10	0	125	0	2	0	1,25	0	65	18,6	1,35
11	0	125	0	2	0	1,25	0	65	22,9	1,76

```
X <- matrix(c(1,-1,1,-1,-1,-1,1,1,-1,-1,1,1,1,-1,-1,1,-1,1,1,-1,1,1,-1,1,-1,1,-1),
, nrow = 8, ncol = 3)
X <- cbind(X[,1:3],X[,1]*X[,2],X[,1]*X[,3],X[,2]*X[,3])
X <- cbind(X[,1:6],X[,1]*X[,2]*X[,3])

#Resposta 1
y <- rbind(28.4,26,14.1,8.1,13,14.7,7.8,6.7)

Porcentagem(X,y,3)

Probabilidade(X,y,3,c(20.8,18.6,22.9))
```

```
#Resposta 2
```

```
y <- rbind(2.02,1.29,1.62,2.51,4.76,1.79,3.23,3.61)
```

```
Porcentagem(X,y,3)
```

```
Probabilidade(X,y,3,c(1.56,1.35,1.76))
```

3.3 Obtenção de um modelo de regressão e otimização.

Nesta parte do trabalho foram utilizados novamente os dados obtidos por SANTOS et al. (2018), desta vez num planejamento fatorial do tipo Doehlert a fim de se obter uma condição de ótimo. Para isso, ainda no PacoteEfeitos, utilizou-se a função Regressao (X, y, GL, SQEP), em que **X** é a matriz com as variáveis codificadas, y o vetor contendo as respostas, GL o grau de liberdade do erro puro e SQEP a soma quadrática do erro puro.

Tabela 5 - Valores codificados e reais para as variáveis de planejamento fatorial do tipo Doehlert e respostas

Experimentos	HCl (v1)		NaBH ₄ (v2)		IF do Sb
	Cód.	Real (mol/L)	Cód.	Real (% m/v)	
1	0,866	5	0,5	1,8	367
2	0,866	5	-0,5	1,4	660
3	0	4	-1	1,2	762
4	-0,866	3	-0,5	1,4	787
5	-0,866	3	0,5	1,8	434
6	0	4	1	2	167
7	0	4	0	1,6	651
8	0	4	0	1,6	643
9	0	4	0	1,6	652

Para o funcionamento desta função instalou-se e carregou-se também os pacotes dplyr, matlib. Com isso, criou-se a matriz **X** contendo as variáveis codificadas, os termos de interações e quadráticos, também foi criado o vetor y contendo a resposta. Além disso, calculou-se a SQEP, assim como seu grau de liberdade.

Inseriu-se os parâmetros na função e quando requisitado, inseriu-se a MQres e seu respectivo valor de *t*, por não apresentar falta de ajuste. Verificou-se após isso que o coeficiente b12, identificado pelo número 6 não é significativo.

Com isso, recalculou-se o modelo sem o coeficiente não significativo seguindo os mesmos passos descritos anteriormente e obteve-se os valores das MQs, testes F1, F2 e coeficientes de determinação.

```
library("dplyr")
library("matlib")
library("PacoteEfeitos")

Exemplo_Regressao <- read_excel("C:/Users/NeriGodoy/Desktop/TCC/VídeoR/ex3.xlsx")
X <- Exemplo_Regressao[,1:6]
y <- Exemplo_Regressao[,7]
Regressao(X,y,2,48.7)
```

Além disso, utilizou-se o método da superfície de resposta para a obtenção de uma condição de ótimo. A função rsm foi empregada utilizando as relações entre as variáveis, sendo necessário fornecer os coeficientes de regressão, bastando especificar os lineares (FO), de interação (TWI) e quadráticos (PQ) a serem utilizados, como demonstrado nos comandos.

```
library("rsm")

Exemplo_Regressao <- read_excel("C:/Users/NeriGodoy/Desktop/TCC/Vídeo R/ex3.xlsx")
X <- cbind(superficie_ex3[,2], superficie_ex3[,4])
y <- superficie_ex3[,5]
exemplo3 <- cbind(X,y)
rsm_exe3 <- rsm(y ~FO(HCl, NaBH4) + PQ(HCl, NaBH4), data=exemplo3)
contour(rsm_exe3, ~HCl + NaBH4, image = TRUE)
persp(rsm_exe3, ~HCl + NaBH4, zlab="y", col = cm.colors(100), contours = ("colors"), phi=45, theta=240)
summary(rsm_exe3)
```

A partir da função contour e persp, em que são especificadas as variáveis presentes nos gráficos a serem gerados, além de parâmetros como cores e rotação dos mesmos, obteve-se os gráficos de contorno e da superfície de resposta. Ao utilizar

a função `summary`, obteve-se mais detalhes do modelo calculado pelo pacote `rsm`, sua análise de variância e o ponto estacionário.

3.4 Planejamento de Misturas

Para o exemplo de um planejamento de mistura com múltiplas respostas, utilizou-se os pacotes `mixexp`, `PacoteEfeitos` e `NlcOptim`. Criou-se uma função com a finalidade de se obter o valor da desejabilidade para cada resposta a partir dos valores do valor mínimo, alvo e peso, para isso, utilizou-se a Equação 4.

```
#Maximizar
maxD <- function(y,L,T,r) {
  if(y<L){d <- 0}
  else if(y>T){d <- 1}
  else{d <- ((y-L)/(T-L))^r}
  return(d)}

```

Importou-se a planilha contendo as variáveis e as respostas (Tabela 6) para o R e criou-se a matriz **X** contendo as componentes da mistura, assim como os coeficientes de interação.

Tabela 6 - Valores codificados das variáveis e respostas.

Exp.	Variáveis Codificadas			Respostas (%)				
	CLF(x1)	DCM(x2)	ACN(x3)	DIA	DEA	ATZ	DIU	TFB
1	1	0	0	44,6	69,6	81,1	85	56,7
2	0,5	0,5	0	46,6	71,2	87,6	90,7	60,6
3	0	1	0	43,5	65,6	78,5	80,8	53,4
4	0	0,5	0,5	46,1	69,3	86,8	89,8	58,1
5	0	0	1	39,8	62,5	86,1	90,7	57,3
6	0,5	0	0,5	46,3	70,4	91,1	93,4	61,3
7	0,75	0,125	0,125	47,1	72,3	88,3	90,3	63,8
8	0,125	0,75	0,125	44,9	68,4	83,1	85,2	61,6
9	0,125	0,125	0,75	42,7	65,9	86,2	88,4	58,2
10	0,5	0,25	0,25	48,5	74,3	91,9	95,8	61,1
11	0,25	0,5	0,25	41,9	62,8	77,7	80,3	54,7
12	0,25	0,25	0,5	48,8	74	92,7	97,2	61,1
13	0,125	0,44	0,44	45,4	69,4	87,1	90,6	58,8
14	0,44	0,44	0,125	47,2	72	89,2	91,2	73,9
15	0,44	0,125	0,44	48,1	73,4	92,3	96,1	74,1

Com isso, criou-se a matriz contendo a desejabilidade para todas as respostas utilizando a função criada, $\max D$, utilizando-se o valor mínimo para cada resposta, o alvo sendo igual à 100% de recuperação e o peso a princípio igual à 1. Após isso, utilizou-se a função Regressao do PacoteEfeitos para a obtenção de MQ_{Reg} e MQ_r , atribuindo-se pesos às respostas a partir da razão entre MQ_{Reg} e MQ_r . Utilizou-se novamente a função para obtenção da desejabilidade para todas as respostas, desta vez levando em consideração os pesos atribuídos e calculou-se também a desejabilidade global através da Equação 7.

Utilizou-se a função Regressao novamente, utilizando a coluna contendo a desejabilidade global calculada anteriormente, deixando com os valores 0 ainda nos argumentos de SQEP e GL e utilizando o valor de t igual à 1,0 quando requisitado.

Com isso, a partir dos coeficientes significativos obtidos, juntamente com as novas réplicas obtidas e SQEP, utilizou-se novamente a função inserindo estes parâmetros e utilizando somente os coeficientes significativos. Utilizou-se a função lm para armazenar o modelo em uma variável, inserindo os coeficientes significativos obtidos anteriormente. Com o modelo obtido, utilizou-se a função ModelPlot do pacote mixexp para a obtenção do gráfico de contorno, inserindo os parâmetros referentes ao modelo e componentes da mistura.


```

Reg <- X
Reg$D <- Desejabilidade$Dglobal
modelo <- lm(D ~ -1 + b1*b2 + b1*b3, data=Reg)
modelo$coefficients[1:5] <- c(0,0,0,0.22,0.38)
ModelPlot (model = modelo, dimensions = list(v1="b1", v2="b2", v3="b3"), c
ontour = T, fill = T, axislabs = c("CLF","DCM","ACN"), color.palette = cm
.colors, colorkey = T)

```

Para a obtenção da condição de ótimo, criou-se uma variável contendo a função a ser otimizada, neste caso sendo o modelo recalculado, além disso, criou-se também uma variável contendo a restrição presente no planejamento de misturas. Também se criou um ponto de partida, sendo este 1/3 para os três componentes da mistura.

Com isso, utilizando-se a função solnl, especificando-se o ponto de partida, a função a ser otimizada, a restrição e os limites inferiores e superiores que as variáveis podem assumir, ou seja, limite inferior igual à 0 e superior igual à 1. Com isso, obteve-se as proporções de cada componente da mistura na região de ótimo, além de obter o valor esperado nesta região.

```

#Função a ser otimizada
Obj <- function(x){
  y_hat <- predict(modelo,newdata= data.frame(b1 = x[1], b2=x[2], b3=x[3]))
  return(-y_hat)}
#Restrição
con <- function(x){
  f <- NULL
  f <- rbind(x[1]+x[2]+x[3]-1)
  return(list(ceq=f,c=NULL))}
x0 <- c(1/3,1/3,1/3)
Opt <- solnl(X=x0, objfun=obj, confun=con, lb=rep(0,3), ub=rep(1,3))
x_otim <- round(Opt$par, 3)
x_otim

```

Por fim, utilizou-se novamente a Equação 4 para obter-se os valores esperados das respostas na região de ótimo, utilizando-se os coeficientes dos modelos individuais e o intervalo de confiança destes para cada modelo.

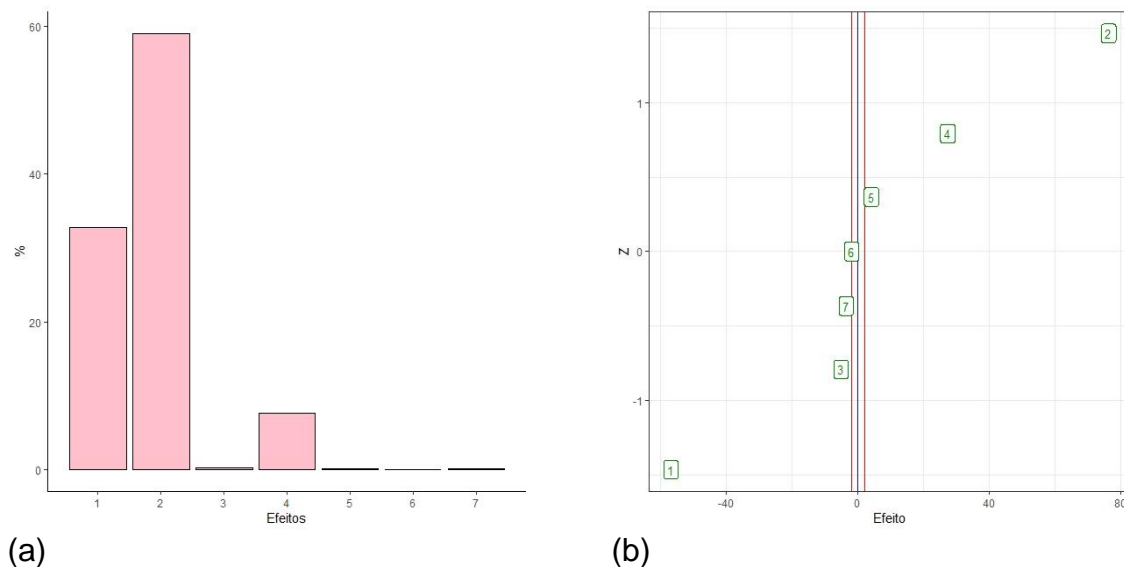
4 Resultados e Discussões

4.1 Planejamento Fatorial Completo

Foi possível obter gráficos ilustrando os efeitos mais importantes e seus impactos nas respostas.

Ao utilizar as funções Porcentagem e Probabilidade, foram gerados o gráfico de porcentagem de efeito (Figura 3a), onde é possível visualizar os fatores ou interações que possuem maior impacto na mudança da resposta e o gráfico de probabilidade de efeito (Figura 3b), em que é possível observar o impacto de cada efeito na resposta juntamente com o intervalo de confiança.

Figura 3 - a) Porcentagem de Efeitos e b) Probabilidade de Efeitos



Com isso, visualiza-se claramente quais efeitos possuem maior impacto e se este é negativo ou positivo na resposta, neste caso, a partir da Figura 3a os fatores 1 e 2 se mostraram mais significativos, tendo o fator 1 e 2 um impacto negativo e positivo na resposta, respectivamente, como visto na Figura 3b. Além disso, a interação entre ambos (efeito 12, designado como 4 na Figura 3) também apresentou importância.

4.2 Planejamento Fatorial Fracionário

Os gráficos de porcentagem e probabilidade de efeito foram obtidos para a primeira resposta, diâmetro (Figura 4a e 4b) e para a segunda resposta, distribuição (Figura 5a e 5b).

Figura 4 - a) Porcentagem de Efeitos e b) Probabilidade de Efeitos para a resposta 1, diâmetro.

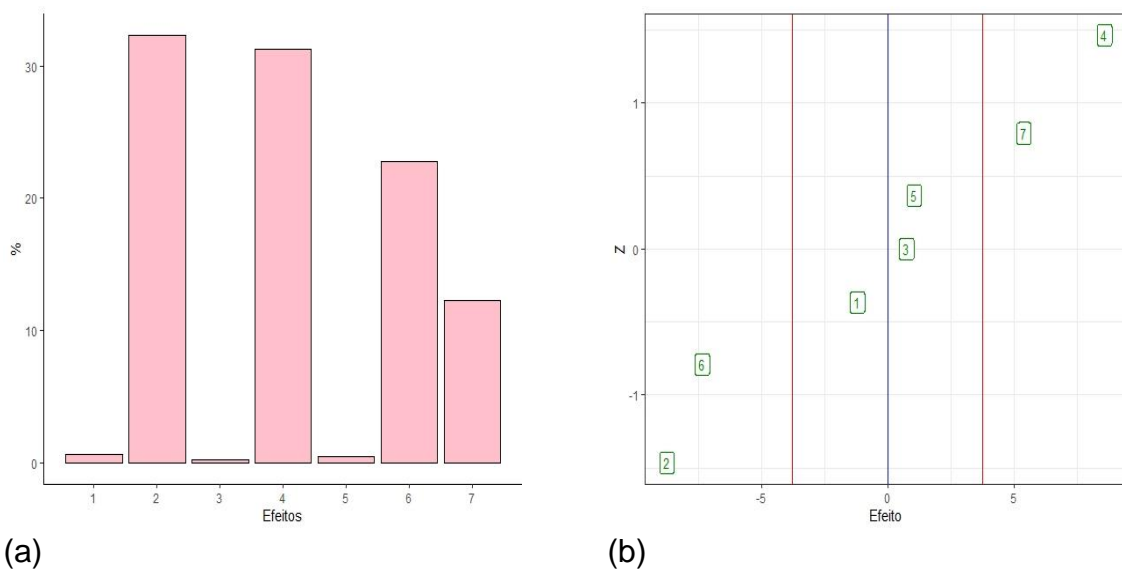
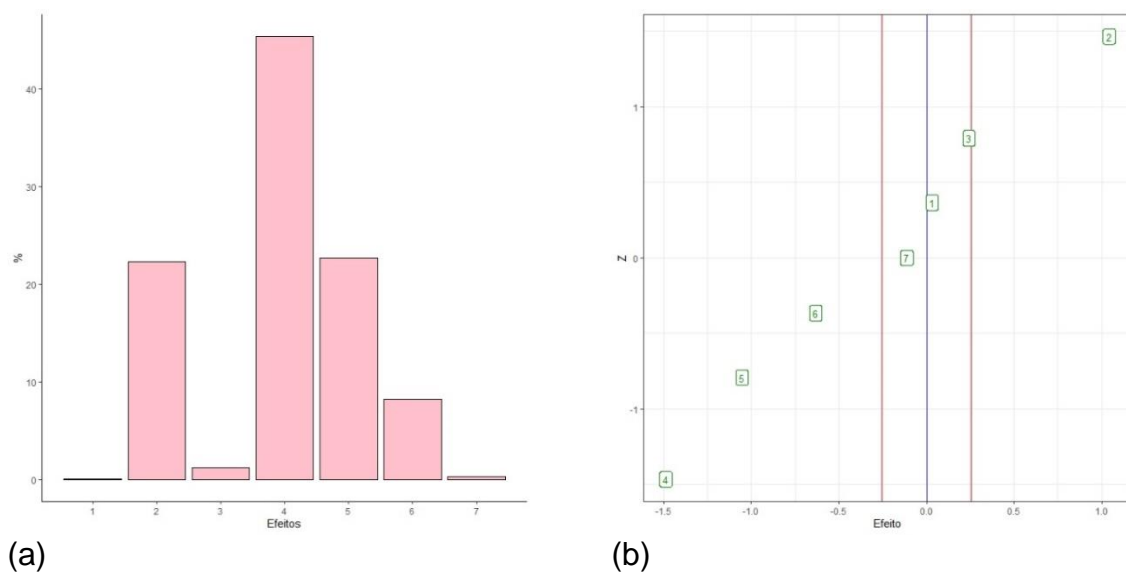


Figura 5 - a) Porcentagem de Efeitos e b) Probabilidade de Efeitos para a resposta 2, distribuição.



Verifica-se que considerando as duas respostas, a partir da Figura 5a, os contrastes mais significativos nas respostas são os identificados pelos números 2

(2+134), 4 (4+123), 5 (12+34), 6 (13+24) e 7 (14+23). Para minimizar o diâmetro das partículas, a variável 2 deve estar em nível alto, 4 em nível baixo e por fim as variáveis 1 e 3 devem estar em níveis opostos devido ao impacto do contraste 13+24. Para maximizar a distribuição é necessário que as variáveis 1, 2 e 4 estejam nos níveis baixo, alto e baixo, respectivamente e a variável 3, também devido ao impacto do contraste 13+24 necessita estar no nível alto. Além disso, o contraste 13+24, impacta ambas respostas negativamente, para que uma seja minimizada e outra maximizada, um nível intermediário das variáveis 1 e 3 pode ser o ideal para se atingir este objetivo.

4.3 Obtenção de um modelo de regressão e otimização.

Com os dados do primeiro exemplo na forma de um planejamento fatorial do tipo Doehlert, obteve-se a SQEP de 48,67 e também utilizando a função Regressao, obteve-se a princípio um modelo considerando todos coeficientes.

Neste exemplo, por não apresentar evidências de falta de ajuste no modelo, foi inserida a MSres e seu respectivo valor de t de 4,30 quando requisitado, obtendo-se então as médias quadráticas, o teste F1 e F2, os coeficientes de determinação (Tabela 7) e os coeficientes do modelo de regressão e seus respectivos intervalos de confiança (Equação 8).

$$y = 648,7_{\pm 31,5} - 56,0x_{1\pm 31,5} - 306,0x_{2\pm 31,47} - 54,2x_{1^2\pm 49,8} - 184,2x_{2^2\pm 49,8} + 34,6x_{12\pm 62,9} \quad (8)$$

Tabela 7 - Médias quadráticas, testes F e coeficientes de determinação para o modelo obtido.

	Média Quadrática	F Calculado/F Tabelado
Regressão	66.459	45,88
Resíduos	161	
Falta de Ajuste	433	0,96
Erro Puro	24,35	
% de variação explicada (R ²)		0,9986
% máxima variação explicável		0,9999

Foram obtidos através da função os gráficos referentes aos valores experimentais vs valores previstos, considerando o intervalo de confiança no nível alto e nível baixo (Figura 6), o histograma dos resíduos, este que deve apresentar uma distribuição ao redor de 0 (Figura 7a).

Também se obteve os gráficos de resíduos em função dos valores previstos, devendo idealmente apresentar uma distribuição aleatória (Figura 7b) e o gráfico apresentando os coeficientes juntamente com o intervalo de confiança (Figura 8), valores estes que foram obtidos também em uma tabela através do console.

Figura 6 - Valores experimentais vs Valores previstos, com intervalo de confiança.

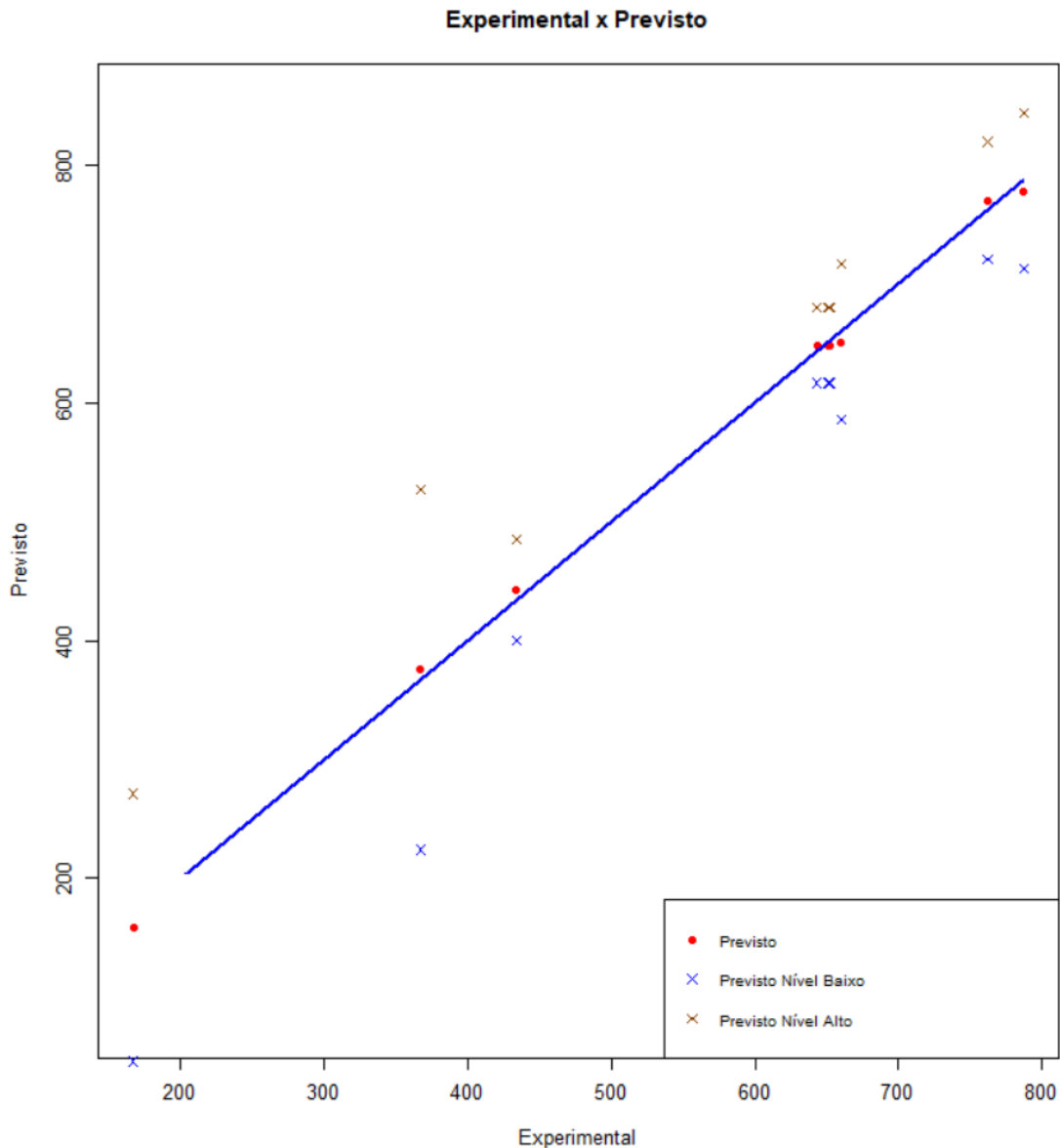


Figura 7 – a) Histograma dos resíduos b) Valores de resíduos vs previsto.

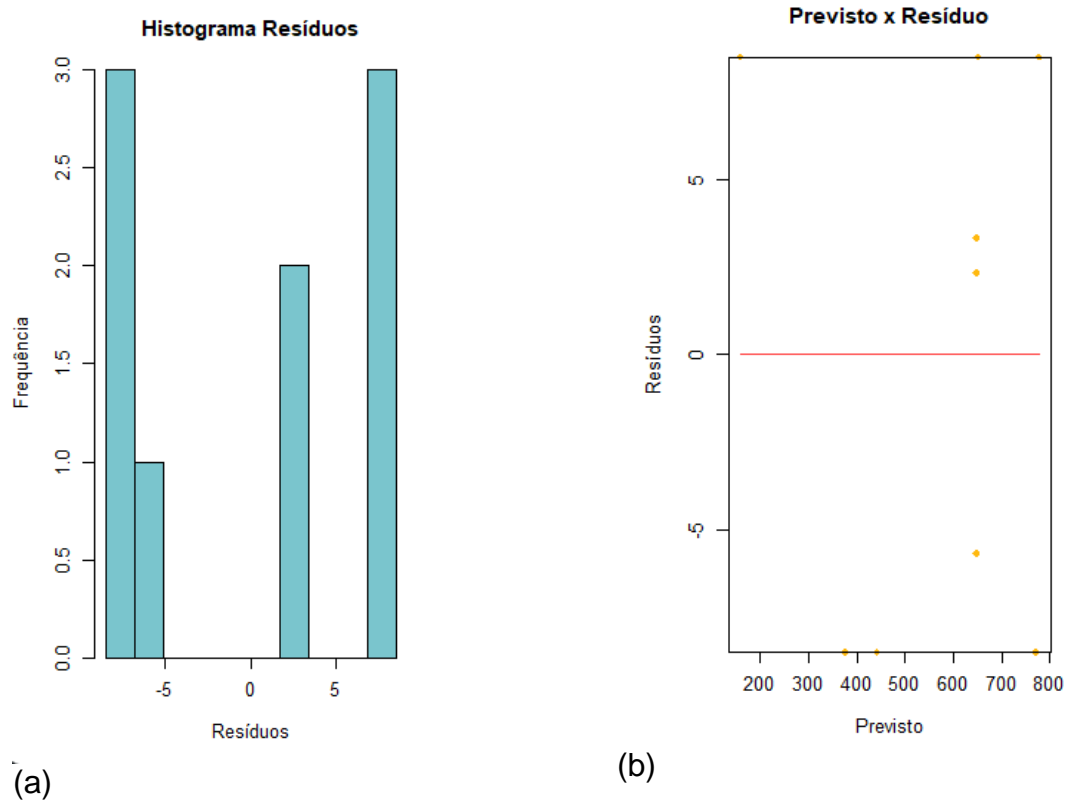
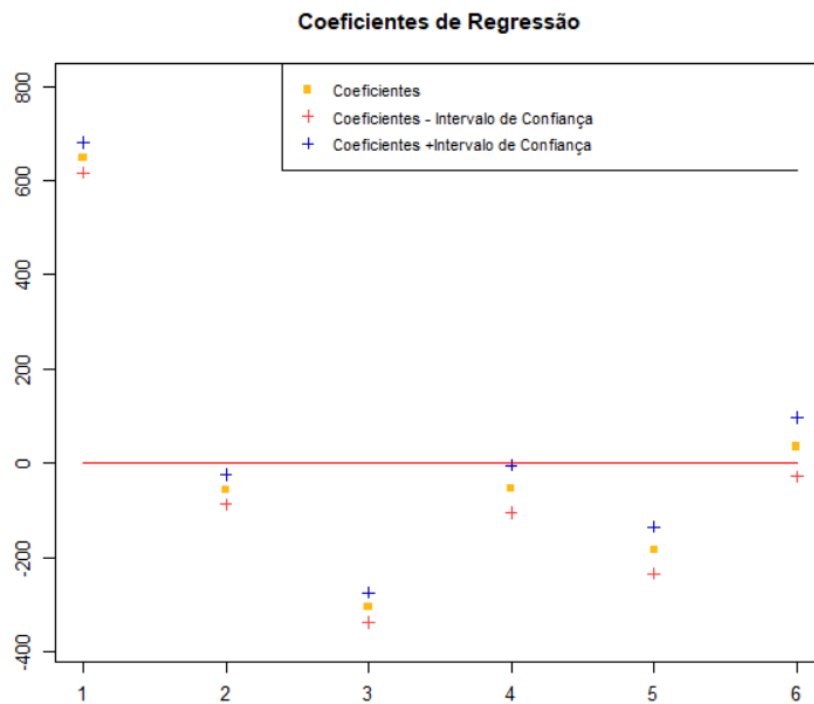


Figura 8 - Coeficientes do modelo de regressão com intervalo de confiança.



Porém, o coeficiente b12, número 6 na Figura 8, não é significativo, havendo necessidade de recalculando o modelo, o excluindo. Recalculando o modelo, obteve-se os valores das MQs, testes F1, F2 e coeficientes de determinação (Tabela 8).

Tabela 8 - Médias quadráticas, testes F e coeficientes de determinação para o modelo recalculado.

	Média Quadrática	F Calculado/F Tabelado
Regressão	82.848	37,53
Resíduos	345	
Falta de Ajuste	666	1,44
Erro Puro	24,35	
% de variação explicada (R ²)		0,9958
% máxima variação explicável		0,9999

O modelo recalculado apresentou evidência de falta de ajuste, porém utilizando a MQFaj, apenas os coeficientes 1 e 3 serão significativos. A partir da MQres, apesar da evidência da falta de ajuste, obteve-se os gráficos de experimental vs previsto (Figura 9), histograma dos resíduos, previsto vs resíduos (Figuras 10a e 10b), além dos 5 coeficientes válidos que podem ser visualizados na Equação 9 e Figura 11.

$$y = 648,7_{\pm 34,1} - 56,0x_{1\pm 34,1} - 306,0x_{2\pm 34,1} - 54,2x_{1^2\pm 53,9} - 184,2x_{2^2\pm 53,9} \quad (9)$$

Figura 9 - Valores experimentais vs Valores previstos, com intervalo de confiança, para o modelo recalculado.

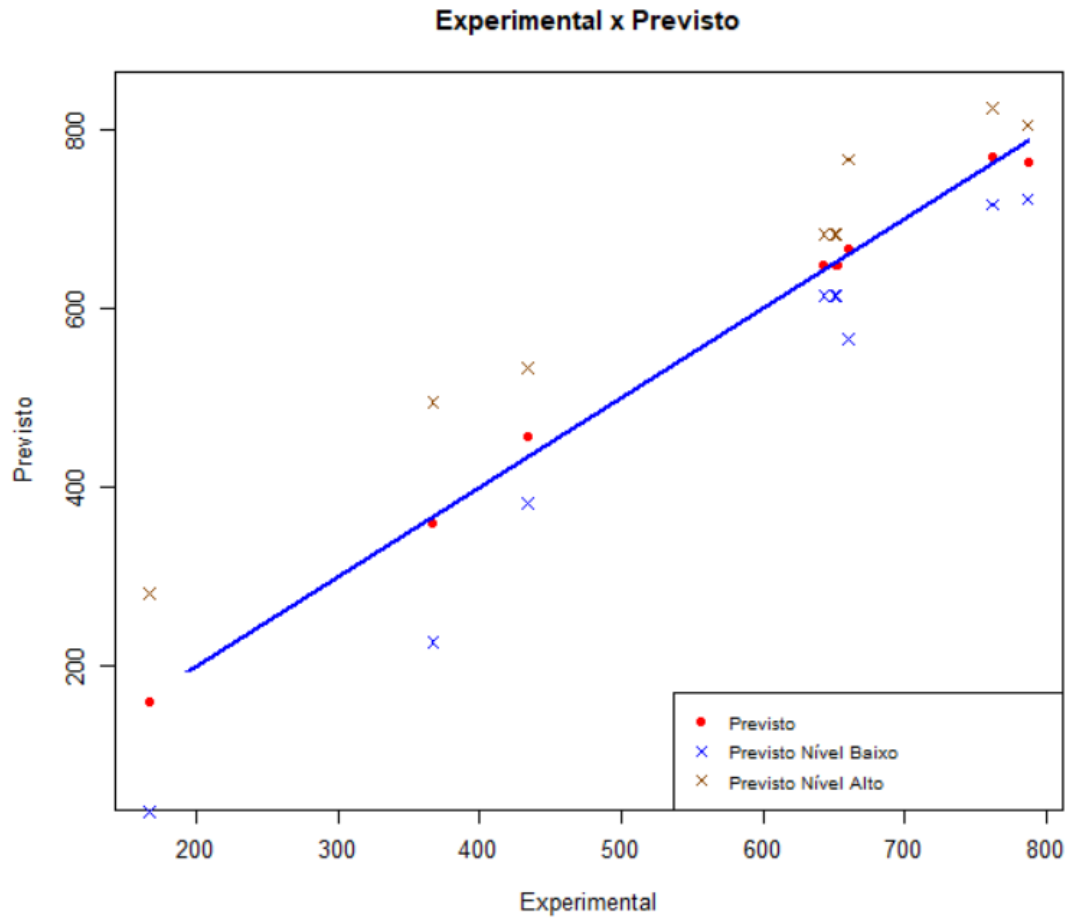


Figura 10 - a) Histograma dos resíduos b) Valores de resíduos vs previsto, para o modelo recalculado.

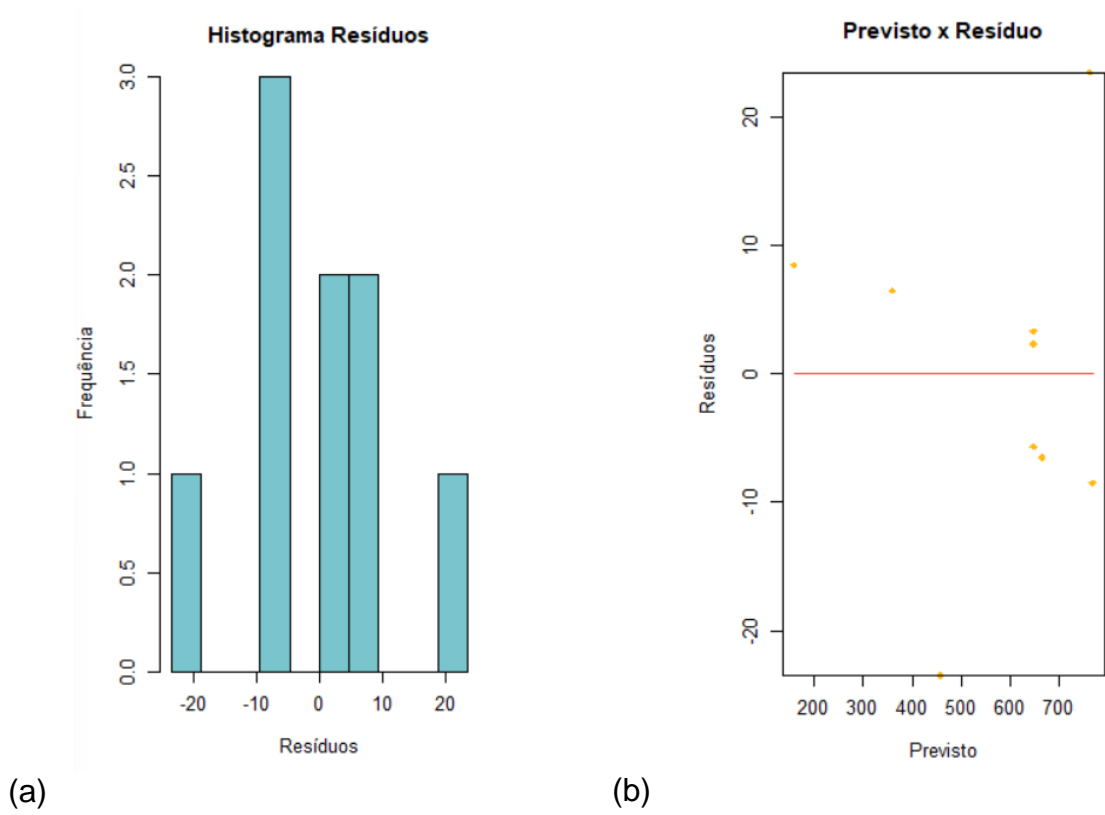
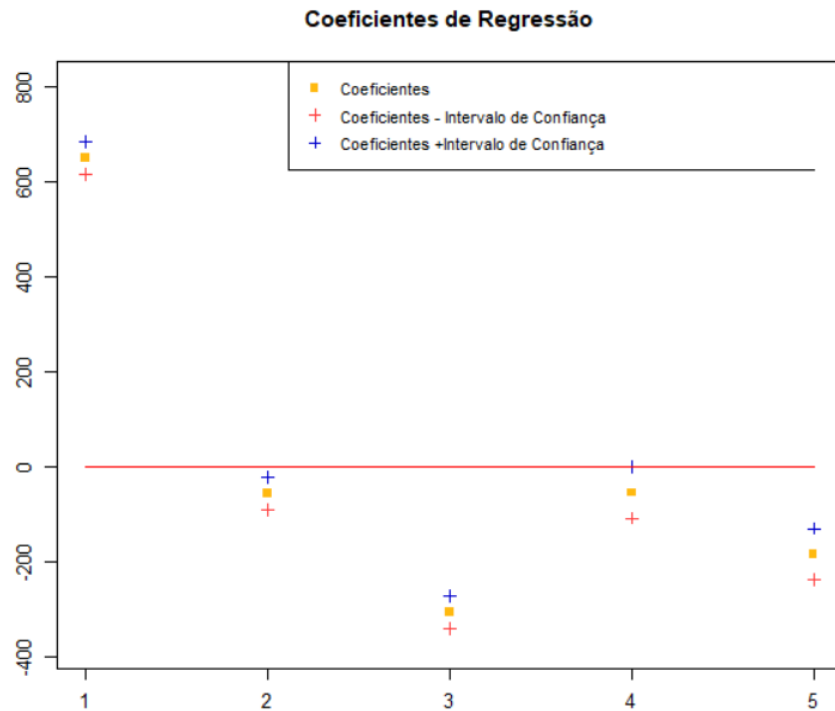


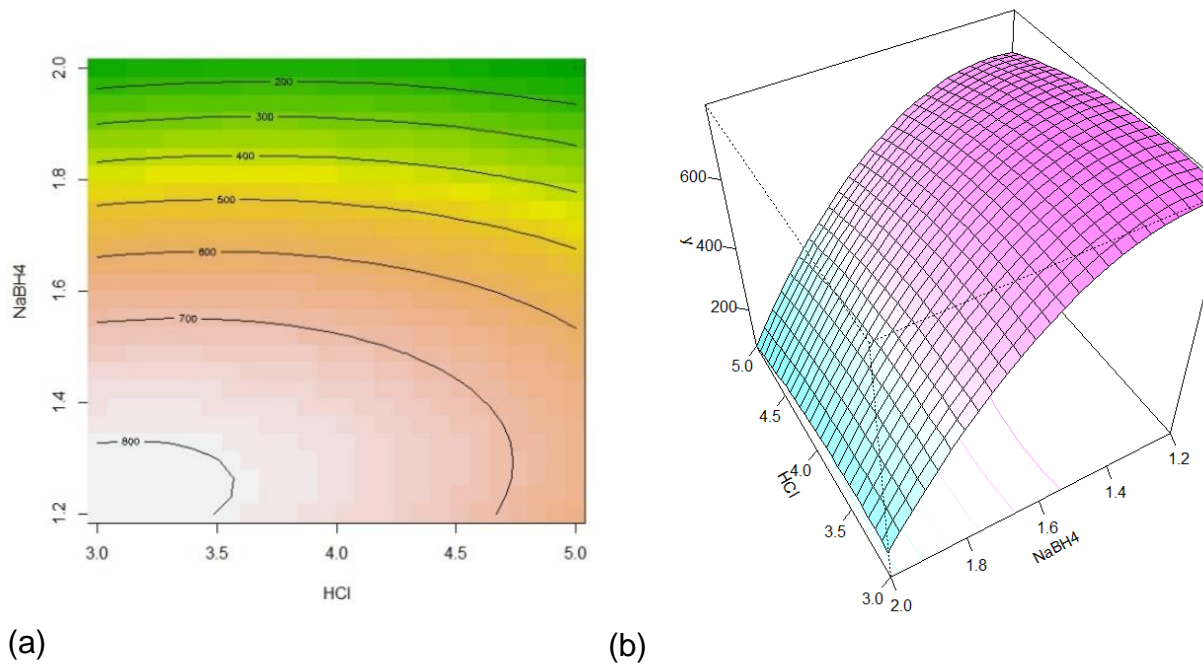
Figura 11 - Coeficientes do modelo de regressão recalculado com intervalo de confiança.



Com a obtenção de um modelo de regressão, foi possível a construção dos gráficos de contorno (Figura 12a), em que é possível ver a relação entre as variáveis independentes e a resposta, e sua superfície de resposta (Figura 12b), sendo o ponto estacionário desta última, um ponto de máximo.

A partir do pacote rsm e dos gráficos, obteve-se, utilizando o modelo obtido na Equação 5, a condição de ótimo de 3,40 mol/L de HCl e 1,27% m/v para o NaBH₄, valores estes próximos aos obtidos por SANTOS et al. (2018), de 3,07 mol/L para o HCl e 1,24% m/v para o NaBH₄.

Figura 12 - a) Gráfico de contorno e b) Superfície de Resposta para a resposta de intensidade de fluorescência do antimônio.



4.4 Planejamento de Misturas

Neste exemplo, ESTEVÃO, PERALTA-ZAMORA e NAGATA (2018) verificaram a influência de 3 variáveis, Clorofórmio (CLF) (x1), Diclorometano (DCM) (x2) e Acetonitrila (ACN) (x3) na recuperação dos pesticidas Diuron (DIU), teflubenzuron (TFB), atrazina (ATZ), desisopropilatrizona (DIA) e desetilatrizona (DEA) a partir de amostras de água natural com a finalidade de identificar a região de ótimo para maximizar esta recuperação, analisando de forma separada cada resposta.

A partir das três componentes da mistura, é construída uma matriz contendo colunas com os coeficientes b_1 , b_2 , b_3 , b_{12} , b_{13} , b_{23} , b_{123} e as respostas. Sendo b_1 , b_2 e b_3 , os níveis codificados das componentes v_1 , v_2 e v_3 . B_{12} é obtido através da multiplicação entre as colunas contendo os níveis codificados dos componentes 1 e 2, seguindo de forma semelhante para os demais coeficientes.

Com a função criada a partir da Equação 4, obteve-se a planilha contendo a desejabilidade calculada para todas respostas, considerando peso 1 para todas. A partir disso, com a função Regressao e utilizando a desejabilidade calculada como resposta, obteve-se os coeficientes para os modelos individuais e obteve-se a MQReg e MQres para estes modelos, permitindo a atribuição de pesos para cada resposta como pode ser visto na Tabela 9.

Tabela 9 – Pesos atribuídos à desejabilidade de resposta.

Resposta	Peso
Diuron	3
Teflubenzuron	2
Atrazina	3
Desisopropilatrazina	2
Desetilatrazine	1

Com os pesos atribuídos e a desejabilidade global calculada, a partir da função Regressao obteve-se os coeficientes válidos para o modelo, com isso obteve-se novas réplicas e a SQEP pôde ser calculada, tendo valor de $8,62 \cdot 10^{-4}$ e grau de liberdade igual à 3. A função Regressao foi utilizada novamente, desta vez utilizando somente os coeficientes significativos, a desejabilidade global, SQEP e grau de liberdade, obtendo-se então os parâmetros do modelo (Tabela 10), os gráficos de experimental vs previsto (Figura 13), histograma dos resíduos, previsto vs resíduos (Figuras 14a e 14b), além dos coeficientes válidos (Equação 10).

$$y = 0,22x_1x_{2\pm 0,16} + 0,38x_1x_{3\pm 0,16} \quad (10)$$

Tabela 10 - Médias quadráticas, testes F e coeficientes de determinação para o modelo recalculado no planejamento de misturas.

	Média Quadrática	F Calculado/F Tabelado
Regressão	$1,46 \cdot 10^{-2}$	4,62
Resíduos	$7,0 \cdot 10^{-4}$	
Falta de Ajuste	$8,0 \cdot 10^{-4}$	0,31
Erro Puro	$3,0 \cdot 10^{-4}$	
% de variação explicada (R^2)		0,6237
% máxima variação explicável		0,9632

Figura 13 - Valores experimentais vs Valores previstos, com intervalo de confiança, para o modelo recalculado no planejamento de misturas.

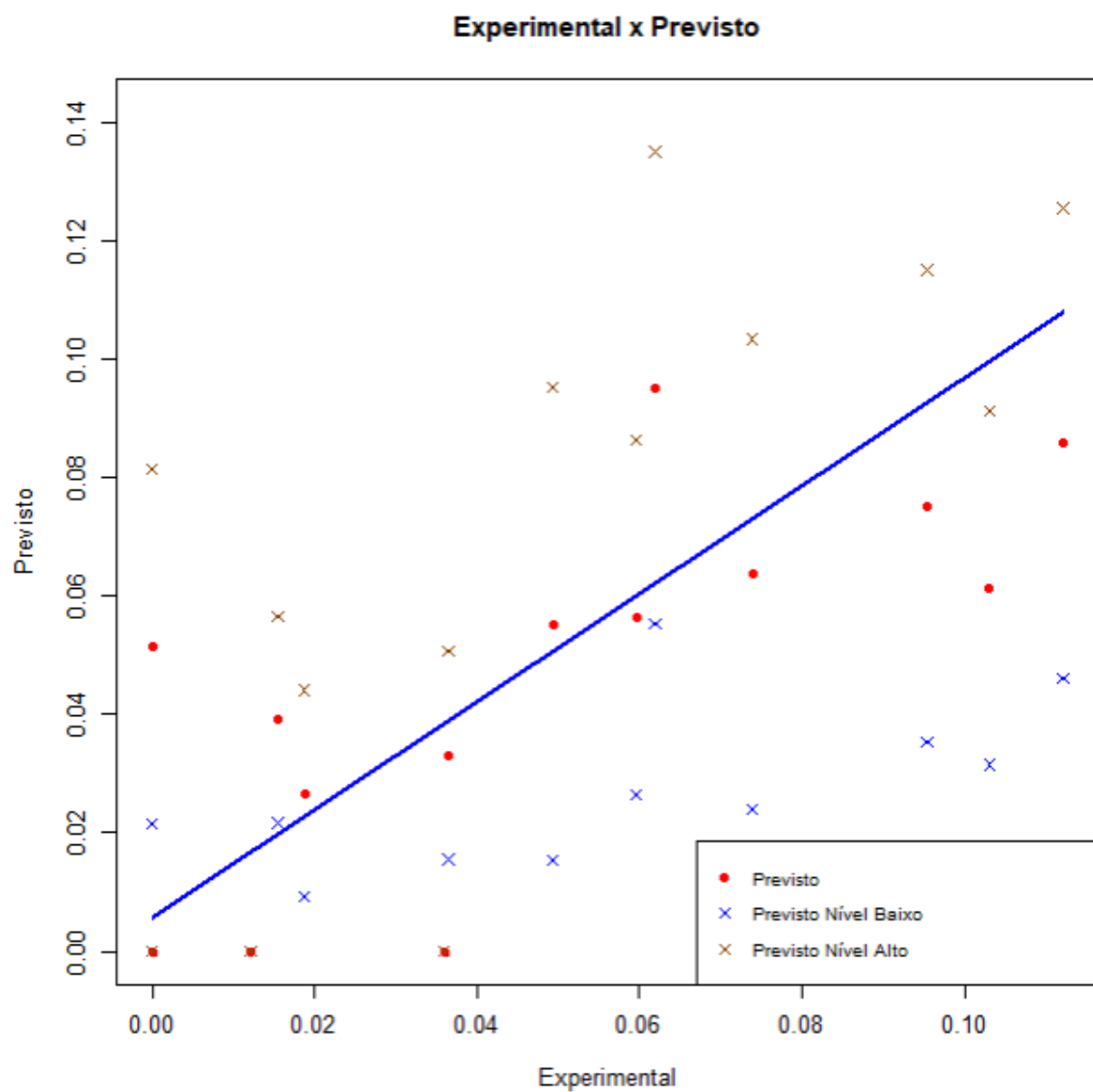
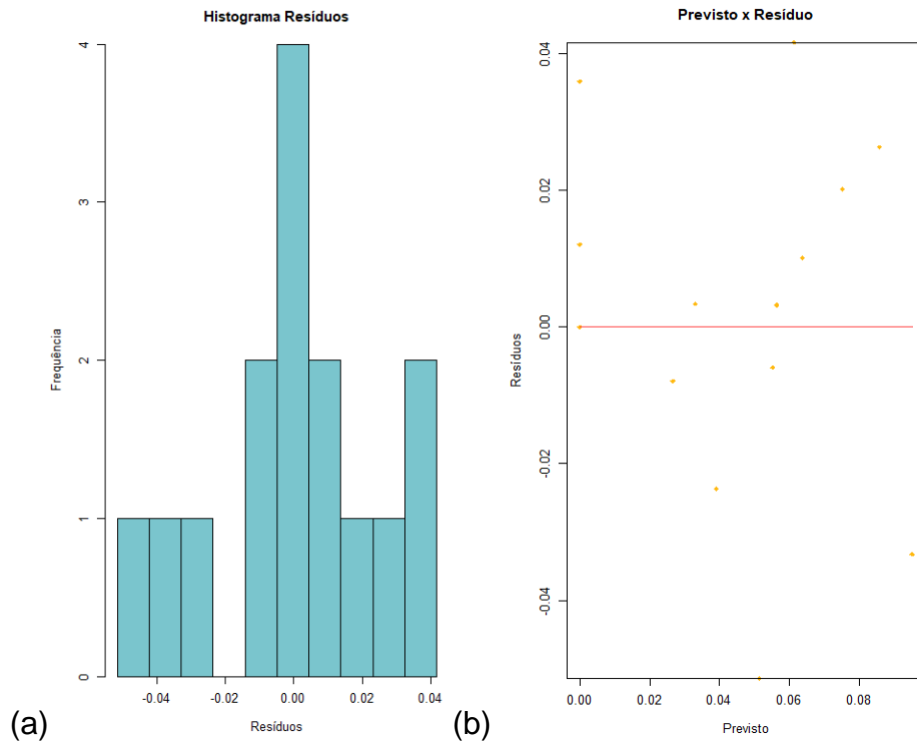
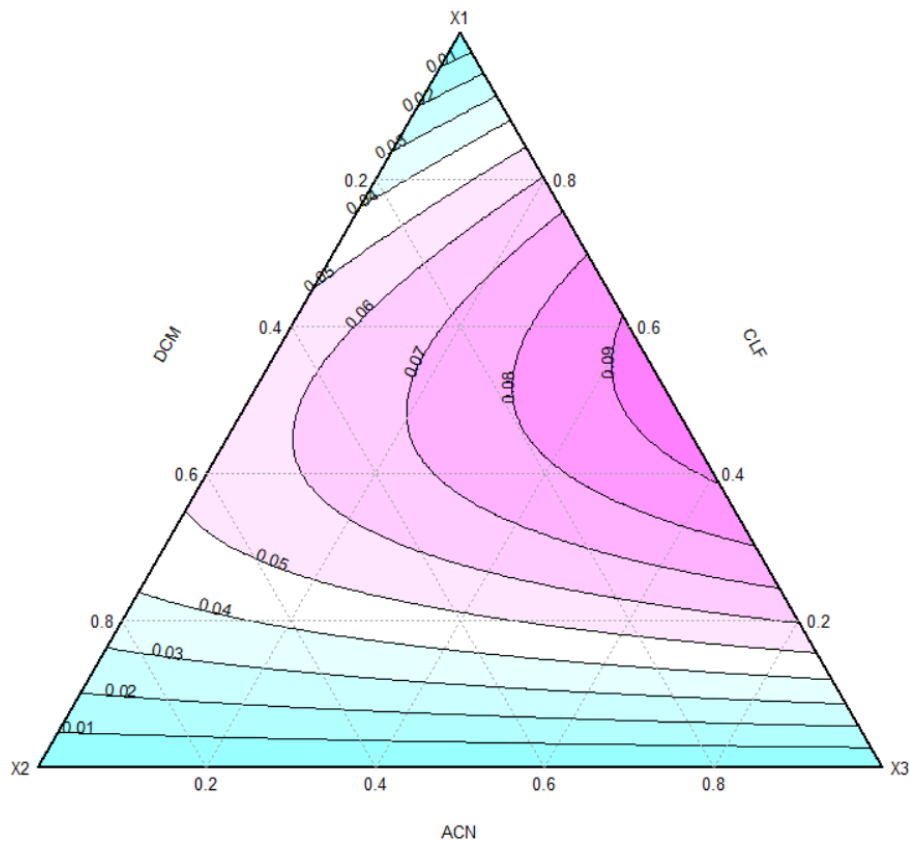


Figura 14 - a) Histograma dos resíduos b) Valores de resíduos vs previsto, para o modelo recalculado para o planejamento de misturas.



Após isso, obteve-se com a função ModelPlot, presente no pacote mixexp o gráfico de contorno (Figura 15), em que é possível observar a maximização da resposta na região de cor magenta, especificamente na área em que ACN e CLF se encontram próximos à 50%.

Figura 15 – Gráfico de contorno para o modelo obtido utilizando a desejabilidade global.



Com a criação da função a ser otimizada, utilizando o modelo obtido através da função Im , a restrição do planejamento de misturas e o ponto de partida, através da função $solnl$ presente no pacote $NlcOptim$, obteve-se as proporções das variáveis na região de ótimo, sendo de 50% para o CLF e 50% para a ACN. Além disso, obteve-se as respostas previstas pelo modelo nestas condições através da Equação 4 utilizando os coeficientes dos modelos individuais para cada resposta, juntamente com o intervalo de confiança (Tabela 11).

Tabela 11 – Valores previstos pelo modelo obtido com intervalo de confiança e valores obtidos pelos autores do artigo para cada resposta.

Resposta	Valor Previsto (%)	Valor obtido pelos autores do artigo (%)
Diuron	45,02 \pm 2,38	49,40
Teflubenzuron	68,63 \pm 3,79	73,20
Atrazina	86,99 \pm 4,45	95,10
Desisopropilatrazina	88,08 \pm 5,05	94,70
Desetilatrazine	61,42 \pm 6,98	74,90

Os valores obtidos são semelhantes aos obtidos por ESTEVÃO, PERALTA-ZAMORA e NAGATA (2018), de 47,5% para o CLF, 47,5% para a ACN e 5% para o DCM, apesar da metodologia diferente ao analisar os dados.

5. Conclusão

Com este trabalho de conclusão de curso, foi possível o tratamento de dados para o planejamento fatorial aplicado à química. A partir das funções Porcentagem, Probabilidade e Regressao do PacoteEfeitos, adaptados para a linguagem R das rotinas `fabi_efeito` e `regression2` de Octave e Matlab, foi possível analisar os efeitos nos exemplos de planejamento fatorial completo e fracionário, assim como obter um modelo de regressão e seus parâmetros para a análise do mesmo no exemplo de obtenção de modelo de regressão e otimização.

Além disso, utilizando pacotes existentes na biblioteca do R como o `NlcOptim`, `rsm` e `mixexp`, foi possível obter gráficos de contorno e a superfície de resposta no exemplo de otimização e o gráfico de contorno no exemplo de planejamento de misturas, assim como calcular as condições ideais para ambos exemplos.

Dito isso, o *software* livre RStudio mostrou-se eficiente e mais uma alternativa no tratamento de dados para o planejamento fatorial em química, além de possuir uma vasta biblioteca de pacotes que expandem suas funcionalidades. Os cálculos apresentados neste trabalho de conclusão de curso fazem parte de uma playlist que foi preparada para o Canal do Prof. Edenir (<https://www.youtube.com/c/EdenirPereiraFilho>). A playlist conta com 12 vídeos e pode ser acessada por meio de (https://www.youtube.com/playlist?list=PL4CuftF4l_fCL_NicWd5Ax5xWIEUn7Zd3).

6. Referências

- 1- SANTOS, Gerffeson S. dos et al. Analytical strategies for determination and environmental impact assessment of inorganic antimony species in natural waters using hydride generation atomic fluorescence spectrometry (HG AFS). **Journal of the Brazilian Chemical Society**, v. 29, 2018.
- 2- PEREIRA, Fabíola Manhas Verbi; PEREIRA-FILHO, Edenir Rodrigues. Aplicação de programa computacional livre em planejamento de experimentos: um tutorial. **Química Nova**, v. 41, 2018.
- 3- ESTEVÃO, Priscila LS; PERALTA-ZAMORA, Patricio; NAGATA, Noemi. Binary solvent dispersive liquid-liquid microextraction for the determination of pesticides in natural water samples. **Journal of the Brazilian Chemical Society**, v. 29, 2018.
- 4- FUKUDA, Isa Martins et al. Design of experiments (DoE) applied to pharmaceutical and analytical quality by design (QbD). **Brazilian Journal of Pharmaceutical Sciences**, v. 54, 2018.
- 5- BRERETON, Richard G. **Chemometrics: data analysis for the laboratory and chemical plant**. John Wiley & Sons, 2003.
- 6- J. N. Miller, J. C. Miller. **Statistics and Chemometrics for Analytical Chemistry**, 5th Ed, Pearson Education Limited, 2005.

- 7- MONTGOMERY, Douglas C. **Design and analysis of experiments**. John Wiley & sons, 2017.
- 8- MORAIS, Waldenice A. et al. Coencapsulation of trans-Dehydrocrotonin and trans-Dehydrocrotonin: hydroxypropyl- β -cyclodextrin into Microparticles. **Journal of the Brazilian Chemical Society**, v. 28, 2017.
- 9- BEZERRA, Marcos Almeida et al. Application of mixture design in analytical chemistry. **Microchemical Journal**, v. 152, 2020.
- 10- DERRINGER, George; SUICH, Ronald. Simultaneous optimization of several response variables. **Journal of quality technology**, v. 12, n. 4, 1980.
- 11- BOX, George EP et al. **Statistics for experimenters**. New York: John Wiley and sons, 1978.
- 12- HILÁRIO, Felipe Ferri et al. Planejamento de misturas e visualização da região ótima com planilhas no Excel: um tutorial. **Química Nova**, v. 44, 2021.