

## APLICATIVO GAMMA-GUI: UMA INTERFACE GRÁFICA AMIGÁVEL PARA PLANEJAMENTO DE EXPERIMENTOS NO MATLAB

Diego Galvan<sup>a,\*,#</sup> e Evandro Bona<sup>b,c,#</sup>

<sup>a</sup>Departamento de Química, Universidade Federal de Santa Catarina (UFSC), 88040-900 Florianópolis – SC, Brasil

<sup>b</sup>Universidade Tecnológica Federal do Paraná (UTFPR), 87301-899 Campo Mourão – PR, Brasil

<sup>c</sup>Universidade Tecnológica Federal do Paraná (UTFPR), 80230-901 Curitiba – PR, Brasil

Recebido em 11/09/2023; aceito em 16/10/2023; publicado na web 18/12/2023

GAMMA-GUI APP: A FRIENDLY GRAPHICAL USER INTERFACE FOR THE MULTIVARIATE DESIGN OF EXPERIMENTS IN MATLAB. We present to users unfamiliar with the programming language a user-friendly interface for the multivariate design of experiments (DoE) in Matlab®. The interface of the GAMMA app - *Grupo de Análise Multivariada em Matrizes Alimentares*, in Portuguese-BR, is available for free download on GitHub <https://github.com/appGAMMA>, and has a YouTube channel <https://www.youtube.com/@appGAMMA> with a series of tutorial videos. In this tutorial, some practical demonstrations are made using the DoE interface with real applications: (i) screening design for variable selection; (ii) optimization by Box-Behnken design; (iii) optimization by central composite design; (iv) factorial design for variable selection following by optimization with Doehlert design; (v) optimization by mixture design; (vi) optimization by mixture design with pseudocomponents; (vii) optimization by mixture design with process variable; and (viii) multiple-response optimization; also reproduced in the software Statistica v.13, StatSoft® for comparison. The GAMMA-GUI is constantly developing; suggestions for improvements or new ideas are welcome and can be directed to the developers.

Keywords: response surface methodology; mixture design; process variable optimization; multiple response optimization.

### INTRODUÇÃO

Técnicas de planejamento e otimização de experimentos, em inglês *design of experiments* (DoE), são abordagens estatísticas sistemáticas usadas para planejar, conduzir, analisar e otimizar experimentos.<sup>1,2</sup> Os DoEs aliados a usuários capacitados são ferramentas poderosas que permitem explorar e otimizar sistemas complexos, melhorar a qualidade de produtos ou aumentar a eficiência de processos em diversas áreas.<sup>3</sup>

Nas últimas décadas, a análise multivariada ou modelos quimiométricos, no qual “se varia e avalia tudo ao mesmo tempo” tem se tornado cada vez mais presente no dia a dia dos químicos. Nesse contexto, o emprego de técnicas de planejamento e otimização de experimentos em rotinas de laboratórios descartam as típicas e imprecisas metodologias aleatórias ou não sistemáticas de “tentativa e erro”. Essas ferramentas trazem uma série de benefícios, como obter informações valiosas e confiáveis com uma redução de experimentos, economia de tempo e recursos. Além disso, permitem uma melhor compreensão das relações entre as variáveis, aumentam a capacidade de fazer previsões confiáveis e tomar decisões direcionadas.<sup>1-5</sup>

É fundamental deixar claro que essas técnicas estatísticas não resolvem a todos os nossos problemas, portanto, devem ser utilizadas com parcimônia. Para isso, os usuários precisam ter um bom conhecimento do problema/ferramentas, e é fundamental que o operador realize tomadas de decisões com base em conhecimento científico e bom senso.<sup>3,5</sup> Na literatura existem diversos livros que podem direcionar usuários iniciantes, como: *Statistical Design - Chemometrics*,<sup>1</sup> *Design and Analysis of Experiments*,<sup>2</sup> *Introdução às Técnicas de Planejamento de Experimentos*,<sup>3</sup> *Planejamento de Experimentos e Otimização de Processos*,<sup>4</sup> *Como Fazer Experimentos*,<sup>5</sup> entre outros.

Na revista *Química Nova* alguns tutoriais foram publicados com a finalidade de divulgar a Quimiometria: Planejamento de Experimentos,<sup>6</sup> Análise de Componentes Principais (PCA),<sup>7,8</sup> Regressão por Mínimos Quadrados Parciais (PLS)<sup>9</sup> e Resolução Multivariada de Curvas (MCR).<sup>10</sup> Entretanto, muitos usuários não são familiarizados com linguagem de programação. Neste tutorial apresentamos aos usuários não familiarizados com linguagem de programação, mas com alguns conhecimentos prévios em planejamento e otimização de experimentos, uma interface gráfica e amigável baseada em plataforma Matlab®.

O aplicativo GAMMA (Grupo de Análise Multivariada em Matrizes Alimentares) é fundamentado em programação orientada a objetos OOP (*object-oriented programming*) para Matlab®.<sup>11</sup> Essa programação permite criar interfaces práticas e amigáveis que simplificam a interação entre o usuário e o *software*, tornando-o mais intuitivo.<sup>12</sup> Uma interface gráfica do usuário GUI (*graphical user interface*) permite que usuários novatos e experientes realizem tarefas de forma eficaz e eficiente, mas não descartam que os usuários requeiram conhecimentos prévios em DoE.

Comparado a outros programas, o GAMMA-GUI apresenta alguns diferenciais, uma interface própria para planejamento e otimização de experimentos usando matriz de Doehlert, uma interface para otimização de modelos que combinam variável de processo e mistura, além de permitir a otimização de múltiplas respostas. Além disso, a interface do aplicativo encontra-se em português e conta com vídeos tutoriais demonstrativos no YouTube. Ainda há outras interfaces no aplicativo direcionadas a análise exploratória de dados/métodos multi-bloco, classificação e calibração multivariada.<sup>13</sup> Entretanto, neste tutorial nos limitamos a explorar a interface de técnicas de planejamento e otimização de experimentos utilizando dados reais da literatura.

\*e-mail: [diego.galvan@ufsc.br](mailto:diego.galvan@ufsc.br); [ebona@utfpr.edu.br](mailto:ebona@utfpr.edu.br)

#Grupo de Análise Multivariada em Matrizes Alimentares (GAMMA)

## RECURSOS INTEGRADOS NA INTERFACE DE PLANEJAMENTO DE EXPERIMENTOS

As técnicas de planejamento de experimentos integradas nesta interface do aplicativo GAMMA-GUI podem ser visualizadas na Figura 1. No aplicativo os usuários têm acesso aos planejamentos de triagem (*screening design*), metodologia de superfície de respostas (MSR) e método simplex de otimização. Destaca-se ainda que, o *software* Matlab® possibilita a análise dos mais variados planejamentos, como planejamentos contendo variáveis de processo e de mistura combinadas. Usuários experientes em DoE podem usar o GAMMA-GUI para analisar matrizes de planejamentos diferentes das mais comumente disponíveis em outros *softwares*. Para tanto, basta empregar o modelo desejado através da notação de Wilkinson, que será apresentada nas seções seguintes, que é bastante flexível para a utilização dos mais variados tipos de modelos.

Iniciaremos esse tutorial com uma breve introdução das terminologias fundamentais usadas em planejamento de experimentos.<sup>3,14</sup> Entretanto, incentivamos que usuários iniciantes consultem as referências específicas<sup>1-5</sup> antes de iniciar a utilização do aplicativo.

- Fator: variáveis passíveis de controle em um processo (x), que podem ou não ser controladas, pode ser quantitativa/numéricas ou qualitativa/categóricas (por exemplo: tipo ou concentração do solvente); dependentes ou independentes, isto é, fatores baseados em valores independentes, para o caso das metodologias de superfície de resposta convencionais, ou ainda baseadas em proporções, para o caso do planejamento de misturas;
- Resposta: variáveis quantitativas resultantes de um processo (y) podendo ser única ou de múltiplas respostas;
- Domínio experimental: intervalos aplicados aos fatores durante a otimização, isto é, faixa de estudo investigada;
- Nível: quantas vezes o nível (k) de um fator é modificado na otimização, pelo menos dois níveis são fundamentais;
- Modelo: função matemática passível de interpretação que relaciona as variáveis dependentes (resposta) e independentes (fatores);
- Codificação: coordenadas geométricas que geralmente variam entre (-1) a (+1), porém os valores podem mudar conforme o tipo de planejamento empregado.

A aplicação de planejamento e otimização de experimentos inclui uma série de etapas consecutivas, frequentemente inicia-se com: (i) estudo de triagem ou varredura, no qual uma grande quantidade inicial de fatores suspeitos de influenciar uma ou várias respostas é reduzida, como também investigar os efeitos principais quanto as interações entre fatores, nesse caso são aplicados os *screening*

*design*; (ii) construção da superfície de resposta para fornecer dados experimentais, que podem ser baseados em fatores de processo independentes ou em proporções; (iii) modelagem através de ajuste por método dos mínimos quadrados e validação estatística do modelo obtido; e (iv) otimização uni ou de múltiplas respostas, quando um grande número de respostas está envolvido, sendo a função de desajuste a ferramenta mais utilizada.<sup>15</sup>

A seguir fornecemos uma breve fundamentação teórica sobre os planejamentos de experimentos frequentes na química e áreas correlacionadas que podem ser encontrados no aplicativo GAMMA-GUI.

### Planejamento fatorial completo e fracionário

São planejamentos de experimentos úteis na triagem de variáveis e na avaliação preliminar da influência das variáveis no sistema. Frequentemente, a aplicação dessas ferramentas são as primeiras abordagens a serem aplicadas, pois requerem poucos experimentos e permitem um conhecimento amplo do sistema estudado.<sup>5</sup> Nesses planejamentos são utilizados modelos lineares para avaliar os efeitos principais dos fatores e as interações entre os mesmos. Através da análise estatística dos dados empregando ferramentas gráficas e ANOVA, é possível checar quais fatores tem influência significativa na(s) resposta(s) avaliadas.<sup>15</sup>

Nesses planejamentos, geralmente, somente dois níveis (+1) e (-1) para cada fator são avaliados. Entretanto, uma boa estratégia é adicionar pontos centrais (0) ao planejamento para avaliar se existe falta de ajuste do modelo linear, ou seja, é feita uma análise estatística para avaliar se o valor previsto pelo modelo linear para o ponto central é estatisticamente diferente daquele medido experimentalmente.<sup>6</sup> Quando a falta de ajuste for estatisticamente significativa, é recomendável o uso de planejamentos que permitam o ajuste de modelos quadráticos. No planejamento fatorial completo, o número de experimentos é dado por  $n = 2^k$ , onde o número dois indica os níveis e o k representa número de fatores investigados. Esse planejamento permite avaliar a influência das variáveis, porém, o número de experimentos aumenta com o aumento de k.

Em contrapartida, no planejamento fatorial fracionário, o número de experimentos é dado por  $n = 2^{k-r}$ , onde o r representa a fração de redução de experimentos. Quando  $r = 1$ , o número de experimentos será reduzido pela metade,  $r = 2$  um quarto, e assim por diante. Como vantagem há uma redução significativa no número de experimentos, porém, os efeitos dos fatores estarão “confundidos” dificultando a sua interpretação. O emprego e interpretação dos resultados de

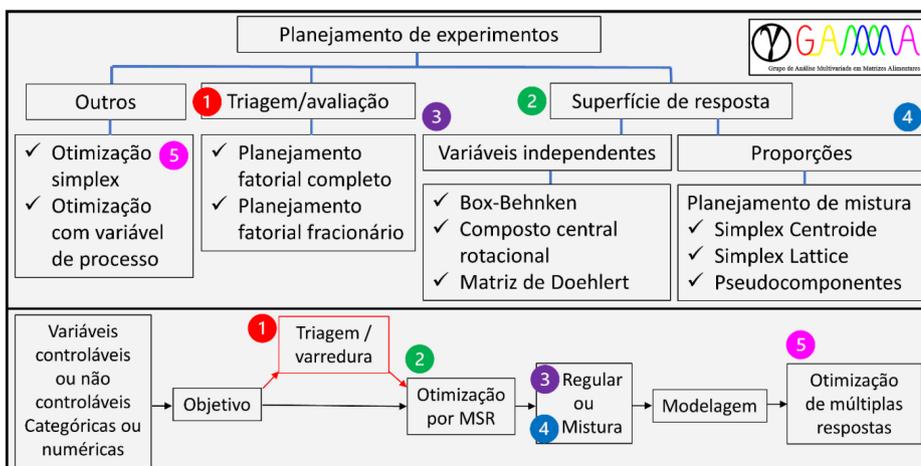


Figura 1. Possibilidades de aplicação do GAMMA-GUI usando a interface de planejamento e otimização de experimentos (superior). Esquema geral para implementação das técnicas de planejamento e otimização de experimentos (inferior) (nota: a etapa (1) é opcional, e depende do número de fatores investigados)

planejamentos fatoriais fracionários requer maior experiência. Entretanto, para usuários com menor experiência é recomendável utilizar fatoriais fracionários com resoluções mínima IV ou V.<sup>1-5</sup> Nos planejamentos de resolução IV os efeitos de segunda ordem estão confundidos entre si, assim esse tipo de planejamento é recomendado apenas para a triagem de fatores. Já na resolução V as interações de segunda ordem se confundem com interações de terceira ordem, esse tipo de planejamento permite uma estimativa mais segura dos efeitos principais e das interações de segunda ordem.

### Metodologia de superfície de resposta (MSR)

O método de regressão múltipla por mínimos quadrados, tem com princípio verificar como os fatores variam em função da resposta:  $y = f(x) + \epsilon$  com o objetivo de otimização de um processo ou produto.<sup>14</sup> Sua utilização consiste na etapa seguinte à triagem, e tem como objetivo modelar uma superfície de resposta aplicando um modelo de segunda ordem.<sup>15</sup> Esses planejamentos podem ser aplicados a fatores baseados em valores independentes ou em proporções. Na abordagem baseada em valores independentes (variáveis de entrada) são tratados como valores independentes/não absolutos. Em outras palavras, o foco está na variação desses valores em relação a uma referência ou ponto central. Enquanto, na abordagem baseada em proporções, a ênfase está nas proporções entre os fatores/componentes avaliados, em vez dos valores absolutos. Esse caso é comumente usado em problemas de misturas ou formulação, onde a composição ou proporção dos componentes é a variável de interesse.<sup>1,2,5</sup>

Embora tais metodologias tenham suas diferenças, ambas têm como principal objetivo a obtenção das condições ótimas através de uma função matemática. As superfícies de resposta podem ser planas (modelo linear), bem como podem ser descritas por um modelo quadrático (Equação 1):

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

onde  $y$  é a variável de resposta,  $\beta_0$  é o termo constante ou intercepto,  $\beta_i$ ,  $\beta_{ii}$  e  $\beta_{ij}$  representam os termos lineares, quadráticos e de interação da equação de regressão,  $x_i$  e  $x_j$  são os valores dos fatores codificados e  $\epsilon$  o termo de erro no modelo. Nesses planejamentos os níveis são  $\geq 3$ , e frequentemente, é recomendável realizar replicatas dos pontos centrais (0) para avaliar a falta de ajuste do modelo e para obter uma medida do erro experimental (*pure error*).

### Delineamento Box-Behnken

O Box-Behnken consiste em um planejamento fatorial fracionário incompleto de três níveis. No planejamento Box-Behnken, os níveis dos fatores estão nos pontos médios das arestas e no centro de um cubo ou hiper-cubo, sendo que o número de pontos experimentais é dado por  $n = 2k^2 - 2k + PC$ . Onde PC representa a quantidade de experimentos no ponto central. Como vantagem, nesse planejamento não há experimentos somente com os níveis inferiores e superiores, entretanto, requer no mínimo três fatores para sua utilização.<sup>1-3,14</sup>

### Delineamento composto central

É um planejamento estabelecido pela adição de pontos axiais a um planejamento fatorial completo ou fracionário (resolução V ou maior) de dois níveis, resultando em um desenho geométrico com cinco níveis. Os pontos axiais ( $\pm \alpha$ ) variam em tamanho e localização de acordo com o número de fatores considerados no estudo, geralmente variam entre 1 e  $\sqrt{k}$ .<sup>15</sup> Além disso, outras propriedades estatísticas desejadas para o planejamento, como

ortogonalidade e rotacionalidade, podem influenciar na escolha dos valores de  $\alpha$ . Nesse planejamento o número de pontos é dado por  $n = 2^k + 2k + 1 + PC$ . Como vantagem esse planejamento permite uma melhor exploração do domínio experimental, devido ao fato dos pontos axiais adicionados.<sup>1-3,14</sup>

### Matriz de Doehlert

É um eficiente planejamento no qual a distribuição dos pontos ocorre sobre um espaço esférico, de modo que, o número de experimentos é dado por  $n = k^2 + k + PC$ . Esse planejamento possibilita uma investigação minuciosa das variáveis mais importantes, e requer a regra de significância dos fatores, que nada mais é que verificar e separar as variáveis mais significativas utilizando um planejamento fatorial. Quanto, mais importante o fator, mais níveis são atribuídos a essa variável independente, por exemplo, em um estudo com três fatores os níveis variam de 7, 5 e 3 de acordo com sua significância, respectivamente.<sup>1,3,5,14,16</sup>

### Planejamento de misturas

Conforme antecipado, esse tipo de planejamento é baseado em proporções de misturas de dois ou mais componentes, nas quais as características do produto resultante são registradas e as respostas são independentes dos valores absolutos, mas sim da proporção entre os fatores.<sup>17</sup> O fato crucial que torna o planejamento de misturas único, é que, os “fatores” não são mais variáveis independentes, como consequência a matriz é singular impossibilitando a obtenção dos coeficientes  $\beta_j$ . Como alternativa, a “sacada” é utilizar os valores de restrições na expressão do modelo introduzindo o coeficiente  $\beta_0$  em cada termo  $\beta_i$ , onde  $b = \beta_0 + \beta_i$  para  $i = 1, 2, \dots, q$  componentes. Como consequência, as proporções das misturas não podem ser maiores que 1 ou 100%, ou seja,  $\sum_{i=1}^q x_i = 1$ . Exceto esse fato, os modelos são ajustados pelo método dos mínimos quadrados e são estatisticamente validados como os demais planejamentos.<sup>1,2,5,18</sup> Os modelos mais usuais para planejamentos de mistura são dadas pelas Equações 2-5:

$$\text{Linear } y = \sum_{i=1}^q b_i x_i + \epsilon \quad (2)$$

$$\text{Quadrático } y = \sum_{i=1}^q b_i x_i + \sum_{i < j}^{q-1} b_{ij} x_i x_j + \epsilon \quad (3)$$

$$\text{Cúbico especial } y = \sum_{i=1}^q b_i x_i + \sum_{i < j}^{q-1} b_{ij} x_i x_j + \sum_{i < j}^{q-2} \sum_{j < k}^{q-1} b_{ijk} x_i x_j x_k + \epsilon \quad (4)$$

$$\text{Cúbico completo } y = \sum_{i=1}^q b_i x_i + \sum_{i < j}^{q-1} b_{ij} x_i x_j + \sum_{i < j}^{q-1} \sum_j^q \delta_{ij} x_i x_j (x_i - x_j) + \sum_{i < j}^{q-2} \sum_{j < k}^{q-1} b_{ijk} x_i x_j x_k + \epsilon \quad (5)$$

Como resultado da dependência entre os fatores, a superfície de respostas apresenta  $q-1$  dimensões. Por exemplo, uma representação gráfica para três componentes gera um triângulo equilátero, cujos vértices correspondem a 100% de um único componente, as arestas misturas binárias e os pontos internos misturas ternárias.<sup>15</sup> Existem vários tipos de planejamento de misturas, basicamente o que os diferencia é a distribuição, quantidade de pontos e o espaçamento entre eles. Os mais frequentes são os planejamentos do tipo simplex lattice e simplex centroide.

Uma possibilidade interessante para esses planejamentos é a modelagem de misturas baseada em pseudocomponentes, que é empregada quando necessitamos garantir a presença de todos os componentes, ou de no mínimo um deles, para obter um produto

aceitável. Isto é feito impondo limites inferiores não nulos ( $a_i$ ) que simultaneamente impõe limites superiores para as proporções dos componentes.<sup>5</sup> Para um caso geral de  $q$  componentes, as misturas em termos de pseudocomponentes podem ser dadas por  $0 \leq a_i \leq c_i$  e  $\sum_{i=1}^q a_i < 1$ , para  $i = 1, 2, \dots, q$  componentes, onde  $c_i$  representa a proporção do componente  $i$  na mistura. A conversão entre o valor do pseudocomponente ( $x_i$ ) e o valor do componente real ( $c_i$ ) é dada pela Equação 6:<sup>1</sup>

$$x_i = \frac{c_i - a_i}{1 - \sum_{i=1}^q a_i} \quad (6)$$

### Otimização com variável de processo

É um conceito que leva em consideração as variáveis de processo que podem afetar na resposta ou produto almejado. As variáveis de processo podem incluir parâmetros operacionais, propriedades/pureza dos materiais, configurações do equipamento, entre outros. Essas variáveis podem ser contínuas, discretas ou categóricas, e uma vez identificadas, é possível utilizar técnicas de otimização para encontrar os valores ótimos das variáveis de processo. Um papel importante do planejamento de mistura, é no controle de qualidade de processos, uma vez que, os ingredientes podem sofrer variações gerando resultados completamente diferentes dos esperados.<sup>17,19,20</sup> Por exemplo, se todas as variáveis de processo forem independentes, pode-se usar um planejamento fatorial combinado ao de mistura. O modelo de regressão em função das variáveis de mistura e da variável do processo é dado por  $y(x,z)$ :

$$y(x,z) = \sum_{1 \leq i \leq q} \gamma_i^0 x_i + \sum_{1 \leq i \leq q} \gamma_i^1 x_i z + \sum_{1 \leq i \leq j \leq q} \gamma_{ij}^0 x_i x_j + \sum_{1 \leq i \leq j \leq q} \gamma_{ij}^1 x_i x_j z + \gamma_{123}^0 x_1 x_2 x_3 + \gamma_{123}^1 x_1 x_2 x_3 z + \epsilon \quad (7)$$

Os coeficientes de regressão  $\gamma_i^j$  são obtidos pelo método dos mínimos quadrados, através da equação matricial  $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ , onde  $\mathbf{X}$  é a matriz do planejamento incluindo a variável de processo,  $\mathbf{y}$  é o vetor resposta,  $\gamma^0$  os termos sem a variável de processo e  $\gamma^1$  contendo a variável de processo. Em seguida, o modelo é analisado pela ANOVA para a escolha dos modelos, interpretação dos resultados e obtenção da condição de ótimo.<sup>17,19,20</sup>

### Otimização de múltiplas respostas

A otimização de múltiplas respostas é uma forma sistemática de análise que permite a otimização simultânea de forma precisa em um único delineamento experimental. Uma das ferramentas que viabiliza a otimização de múltiplas resposta mais conhecida é a “função de desejabilidade D”, proposta por Derringer & Suich.<sup>21</sup> A função D tem como objetivo ajudar a encontrar uma região comum que satisfaça a todas as respostas dentro do domínio experimental em questão. Ela tem como argumento a transformação de cada variável de resposta estimada  $\hat{y}_i$  ( $i = 1, 2, \dots, k$ ) pelo modelo para um valor de desejabilidade  $d_i$ , onde  $0 \leq d_i \leq 1$ . O valor de  $d_i$  aumenta conforme o “desejo” do correspondente de resposta aumenta. As desejabilidades individuais são então combinadas usando a média geométrica, fornecendo a desejabilidade global (D):

$$D = \sqrt[k]{d_1 \times d_2 \times \dots \times d_k} \quad (8)$$

O intervalo de D estará em um intervalo entre [0, 1] e o valor de D aumentará à medida que o equilíbrio das propriedades se torna mais favorável.<sup>1,21</sup> Uma versatilidade desta função, é que as

desejabilidades individuais podem ser ajustadas separadamente conforme as necessidades do usuário, tendo quatro configurações possíveis: máximo, mínimo, alvo único e faixa de alvo. Outra vantagem é a possibilidade de ajuste de pesos/importâncias para as respostas individuais.

Cabe ainda ressaltar que, a fim de realizar uma otimização de resposta múltipla adequadamente, os valores de  $\hat{y}_i$  devem representar os valores de resposta estimadas para cada variável, isto é, os valores de resposta preditos pelos modelos de superfície de resposta individual estatisticamente validos,<sup>1</sup> e não os valores experimentais, como tem sido vastamente e erroneamente empregada.

### INTERFACE GRÁFICA

Esse tutorial tem como objetivo realizar demonstrações práticas utilizando a interface de planejamento de experimentos do aplicativo GAMMA-GUI. O aplicativo foi desenvolvido utilizando a interface gráfica do usuário (*graphical user interface*, GUI) e implementado usando programação orientada a objetos (*object-oriented programming*, OOP) do *software* Matlab®.<sup>11</sup> Os códigos e algoritmos implementados no aplicativo contém funções nativas do próprio Matlab, *scripts* desenvolvidos pelos próprios autores ou adaptações de *scripts* disponíveis gratuitamente para *download* (créditos devidamente reconhecidos aos desenvolvedores). O aplicativo foi testado nas versões do *software* Matlab R2022b, R2023a e R2023b. A execução de todas as funcionalidades do aplicativo requer funções dos *toolbox* “Global Optimization Toolbox”, “Optimization Toolbox”, “Statistics and Machine Learning Toolbox” e “Signal Processing Toolbox”, entretanto, a interface de planejamento de experimentos requer somente o *toolbox* “Statistics and Machine Learning Toolbox”. O aplicativo GAMMA-GUI e os conjuntos de dados para demonstração usados neste tutorial podem ser visualizados no canal do YouTube <https://www.youtube.com/@appGAMMA>, na página do Instagram [https://www.instagram.com/appgamma\\_gui/](https://www.instagram.com/appgamma_gui/) e o aplicativo está disponível gratuitamente para *download* no GitHub <https://github.com/appGAMMA>. Esse pacote é livremente distribuído, os autores não se responsabilizam pelos resultados obtidos através da livre utilização dessas rotinas, sua comercialização e modificação não está autorizada.

### Instalando e executando o GAMMA-GUI

Após o *download* do aplicativo os usuários terão um arquivo no formato GAMMA.zip. Descompacte o arquivo no local que se deseja realizar a instalação, após isso abra o *software* Matlab e selecione “Home” e então “Set Path”. Uma nova janela se abrirá, selecione “Add with Subfolders”, selecione o diretório onde foi descompactado os arquivos, selecionar pasta com os arquivos, “Save” e “Close”. Veja o esquema de instalação na Figura 2, ou tutorial de instalação no YouTube <https://www.youtube.com/watch?v=c5uRUFJ14Jk>.

Para iniciar o aplicativo os usuários devem digitar na janela “Command Window” do Matlab o comando appGAMMA, e então a tecla ENTER (Figura 3a). Após essa etapa, uma janela semelhante a Figura 3b se abrirá.

Cabe ainda ressaltar que, o GAMMA-GUI utiliza o “ponto” como separador decimal, em vez da “vírgula”. Portanto, é importante que o usuário esteja ciente dessa diferença durante a definição dos *inputs*, como também na interpretação dos *outputs* e saídas gráficas do aplicativo.

### Conhecendo a interface de planejamento de experimentos do GAMMA-GUI

Na interface de “Regressão” os usuários têm a possibilidade

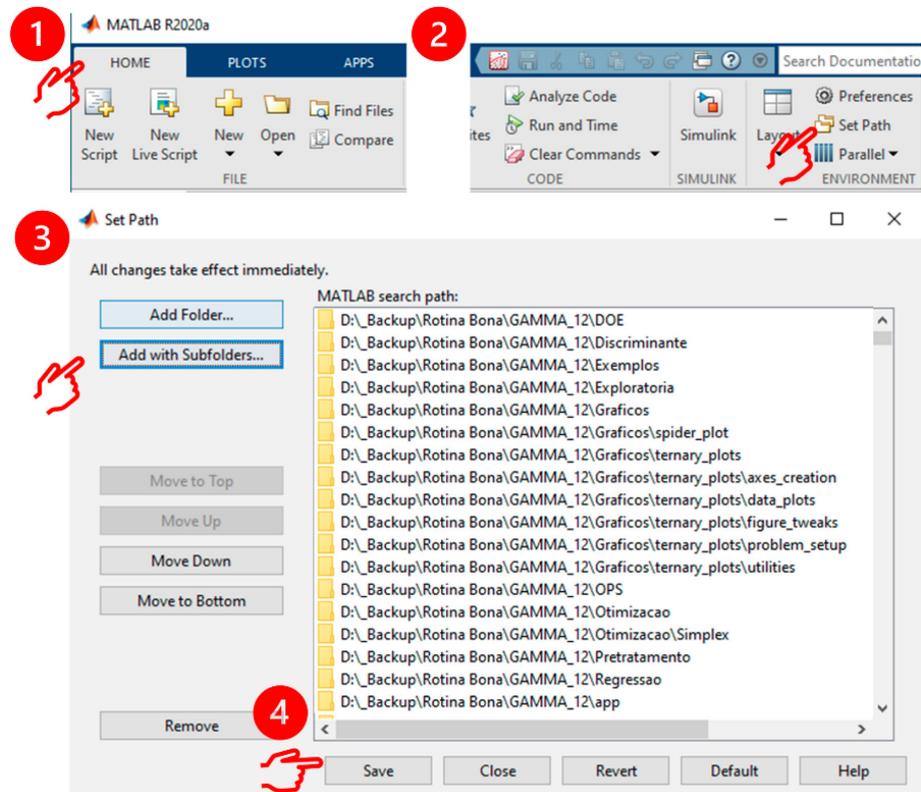


Figura 2. Instalando o aplicativo GAMMA-GUI no Matlab

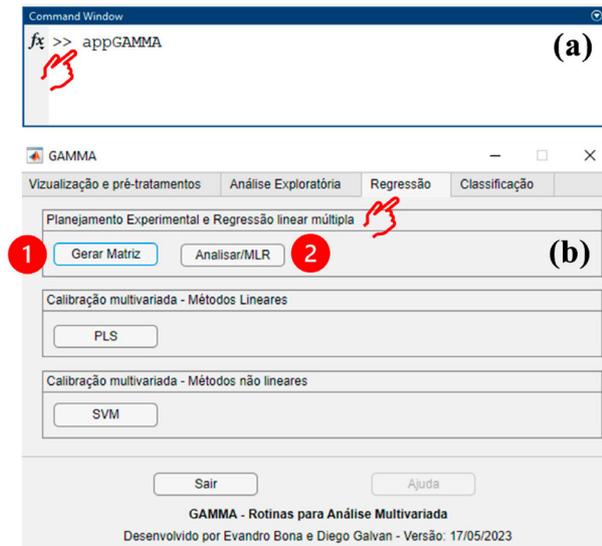


Figura 3. (a) Iniciando o aplicativo na janela "Command Window". (b) Interface inicial do aplicativo GAMMA-GUI. Selecione a janela de "Regressão" que corresponde a interface de planejamento de experimentos

de utilizar diferentes técnicas de planejamento e otimização de experimentos (Figura 3b). Na opção "Gerar matriz" (Figura 4a) os usuários podem escolher o tipo de planejamento a ser utilizado, e gerar sua respectiva matriz de planejamento elegendo o número de fatores ou variáveis de misturas, pontos centrais ou centroides, entre outras opções. Nessa mesma interface os usuários têm acesso ao planejamento fatorial completo e/ou fracionário, metodologia de superfície de resposta como Box-Behnken, central composto rotacional, matriz de Doehlert e planejamento de misturas simplex centroide e simplex lattice. A opção "Analisar/MLR" (Figura 4b), é onde os usuários devem realizar a análise dos dados e visualização dos resultados.

## DEMONSTRAÇÕES USANDO O GAMMA-GUI

Diferentes conjuntos de dados foram selecionados para demonstrar as aplicações e funções do aplicativo GAMMA-GUI usando a interface de "Regressão", seção que corresponde a técnicas de planejamento de experimentos. Para essa seção foram utilizados conjuntos de dados dos próprios autores e outros exemplos da literatura. Neste tutorial fornecemos algumas dicas e truques úteis que podem ajudar os usuários. Entretanto, cabe salientar que, outras dicas e truques podem ser acessadas utilizando recursos visuais através dos vídeos disponibilizados em nosso canal no YouTube. A utilização do GAMMA-GUI por usuários iniciantes também é recomendada e fomentada pelos desenvolvedores, considerando que, eles busquem alguns conhecimentos prévios sobre o assunto antes da sua utilização.

### Screening design

#### Planejamento fatorial completo

Neste exemplo demonstramos passo a passo a construção e aplicação de um planejamento fatorial completo usado para verificar as condições experimentais do ICP-OES para a determinação de urânio em água ([link YouTube: https://www.youtube.com/watch?v=tIDMhYgWz4M](https://www.youtube.com/watch?v=tIDMhYgWz4M)). Foram avaliadas a influência das variáveis: concentração de  $\text{HNO}_3$  ( $\text{mol L}^{-1}$ ) ( $x_1$ ), potência do plasma (kW) ( $x_2$ ) e a vazão do gás do nebulizador ( $\text{L min}^{-1}$ ) ( $x_3$ ) através de um planejamento fatorial completo  $2^3$ , tendo como resposta o sinal de emissão do urânio medidos em diferentes comprimentos de onda ( $y$ ). Para fins didáticos, demonstraremos somente o estudo para o sinal de emissão medido em 367,007 nm. Maiores detalhes podem ser encontrados em Santos *et al.*<sup>22</sup>

A primeira etapa de um planejamento experimental consiste na geração da matriz de planejamento. Para gerar a matriz no GAMMA-GUI, o usuário deve selecionar a opção "Gerar matriz", ajustar em "Fatores independentes" com o botão de rolagem o número três.

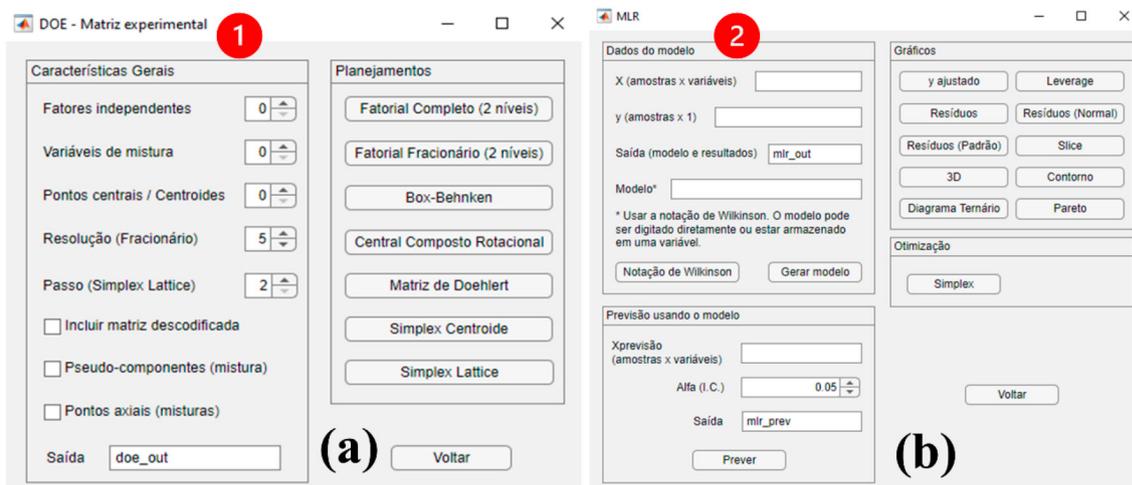


Figura 4. Interface de planejamento de experimentos: (a) “gerar matriz” e (b) “analisar/MLR”

Esses três fatores correspondem a concentração de  $\text{HNO}_3$ , potência do plasma e a vazão do gás do nebulizador. Ainda nessa janela, os usuários devem definir a quantidade de pontos centrais em “Pontos centrais”, que para esse caso foram três, e finalmente apertar o botão “Fatorial Completo (2 níveis)”. Em seguida, uma estrutura de dados “doe\_out” será gerada no *Workspace* do Matlab. Nessa estrutura gerada os usuários têm acesso ao tipo de planejamento e a matriz de planejamento codificada, nesse caso onze experimentos que correspondem a oito pontos experimentais mais três repetições do ponto central. Caso o usuário almeje trabalhar com a matriz de dados descodificados, a opção “Incluir matriz descodificada” deve ser marcada nessa etapa, e os valores descodificados/reais precisam ser digitados na nova aba “Descodificação - DoE” que se abrirá. Ver procedimento na Figura 5. Agora, o usuário deverá apertar o botão “Descodificar”, e as informações da matriz de planejamento descodificada e os níveis reais dos fatores serão inseridos na estrutura de dados “doe\_out” gerada anteriormente.

Após gerar a matriz de planejamento, a próxima etapa consiste em realizar os experimentos para cada condição pré-definida de modo aleatório para evitar erros sistemáticos. Uma etapa crucial do processo, é a criação do vetor  $y$  (variável dependente), que contém os valores obtidos experimentalmente. Os valores de  $y_i$  devem corresponder ao devido experimento de que se trata. Após essa “cautela”, com os dados em mãos, é hora de analisar os resultados utilizando a interface “Analisar/MLR”. Ao apertar o botão

“Analisar/MLR” a interface “MLR” se abrirá, e nela os usuários devem inserir na caixa “Dados modelos” a matriz de dados  $X$  (variáveis independentes) e o vetor  $y$  (variável dependente) a serem analisados, que nesse caso foi o sinal de emissão do urânio medido em 367,007 nm. Optamos nesse exemplo por criar um vetor  $y$  dentro da própria estrutura “doe\_out” usando o comando CTRL + N, renomeando o novo vetor como “y”, e então os respectivos valores experimentais (planilha: Planejamento\_fatorial.xlsx disponibilizada nos conjuntos de dados usados como demonstração) devem ser inseridos neste vetor.

Agora os nomes da matriz de dados  $X$  do planejamento e o vetor  $y$  devem ser digitados nos campos “X (amostras  $\times$  variáveis)” e “vetor y (amostras  $\times$  1)”, respectivamente. Aqui, optamos por “pegar” os valores de  $X$  e  $y$  dentro da própria estrutura “doe\_out” gerada anteriormente, para isso, utilizamos as seguintes notações “doe\_out.matriz” e “doe\_out.y”. Cabe ainda ressaltar que, os usuários têm a possibilidade de utilizar a matriz de dados descodificados. Mas por razões matemáticas e estatísticas, é recomendável que a análise dos dados seja realizada com a matriz de dados codificados, enquanto, as representações gráficas ficam a critério dos usuários utilizar os valores codificados ou não.

Na opção “Modelo”, os usuários podem construir diversas possibilidades de modelos usando a notação de Wilkinson. No botão “Notação de Wilkinson”, os usuários têm alguns exemplos de sua utilização. Mais detalhes podem ainda ser encontrados no “help” do

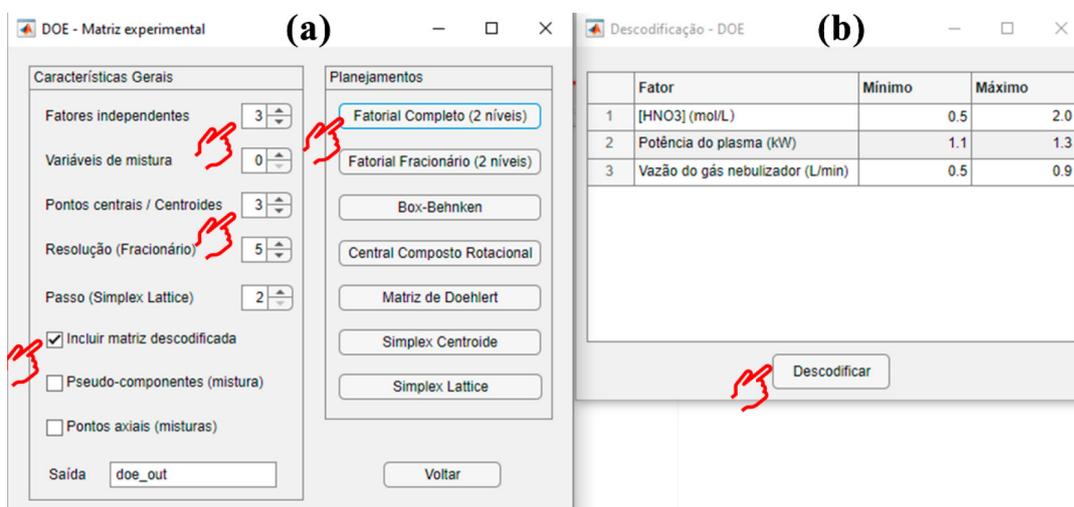


Figura 5. (a) Escolhendo e gerando a matriz de planejamento experimental. (b) Incluindo os valores descodificados na matriz de dados

Matlab e no site <https://www.mathworks.com/help/stats/wilkinson-notation.html>. Nessa demonstração, optamos por avaliar dois modelos “*modelo1*” e “*modelo2*”, que foram previamente escritos usando a notação de Wilkinson e salvos como “*Modelos\_fatorial\_completo.mat*”. Ao abrir esse arquivo, automaticamente os modelos serão carregados para o *Workspace* do Matlab. Ao definir todas as opções, haverá em tela uma imagem semelhante a Figura 6, então o usuário deverá clicar em “*Gerar modelo*”.

O modelo gerado terá como *default* uma estrutura de saída chamada “*mlr\_out*”. Nessa estrutura de dados armazenada no *Workspace*, os usuários têm acesso a ANOVA (análise de variância),  $R^2$  (coeficiente de determinação), MSE (erro quadrático médio), RMSE (raiz quadrada do erro quadrático médio), coeficientes estimados, resíduos, entre outras métricas para avaliação da qualidade e ajuste do modelo com um único clique. Para mais detalhes sobre as métricas e suas interpretações os leitores podem consultar o trabalho de Breitreitz *et al.*<sup>6</sup> Conforme antecipado, nessa demonstração avaliamos a qualidade de ajuste de dois modelos lineares, um com efeitos de interação de primeira e segunda ordem “*modelo1*” e outro com efeitos de interação de primeira, segunda e terceira ordem “*modelo2*”. Para isso, antes da avaliação de cada modelo as saídas foram renomeadas para “*mlr\_out1*” ou “*mlr\_out2*”, que correspondem aos modelos 1 e 2, respectivamente (ver Figura 6b).

Uma etapa importante ao utilizar um planejamento experimental é a avaliação da adequação do modelo construído das respostas obtidas experimentalmente, que ditará sua capacidade de previsão.<sup>6</sup> Ao comparar o ajuste do modelo 1 e 2 pela ANOVA, nota-se que ambos os modelos foram significativos ( $p < 0,05$ ) e não tiveram falta de ajuste ( $p > 0,05$ ). Além disso, os valores de  $R^2$  ( $> 87,9\%$  e  $90,5\%$ ) e RMSE (44,7 e 39,6) para os modelos foram próximos. Nesses casos, os usuários podem checar a significância dos efeitos na tabela “*ANOVA.efeitos*” ou gráfico de Pareto apertando no botão “*Pareto*” usando a estrutura de dados “*mlr\_out2*”, que representa o modelo 2. Essa etapa permite checar se a inclusão do efeito de terceira ordem “ $x_1:x_2:x_3$ ” é significativa ou não em nível de 5%. Nesse caso, a inclusão do efeito de interação de terceira ordem não foi significativa ( $p = 0,24$ ), portanto, seguimos o princípio do “bom senso” e “simplicidade” e optamos por utilizar o modelo 1 (ver Figura 7).

Ao definir o modelo mais adequado para o conjunto de dados experimentais, os usuários podem verificar os resultados visualmente usando as opções gráficas disponíveis. Nessas opções os usuários podem verificar a qualidade de ajuste do modelo na opção “*y ajustado*”, ou ainda verificar a alavancagem dos pontos experimentais em “*Leverage*”, entre outras opções de resíduos “*Resíduos*”, “*Resíduos (Normal)*” ou “*Resíduo (Padrão)*”. Uma interessante e interativa interface gráfica do GAMMA-GUI é a

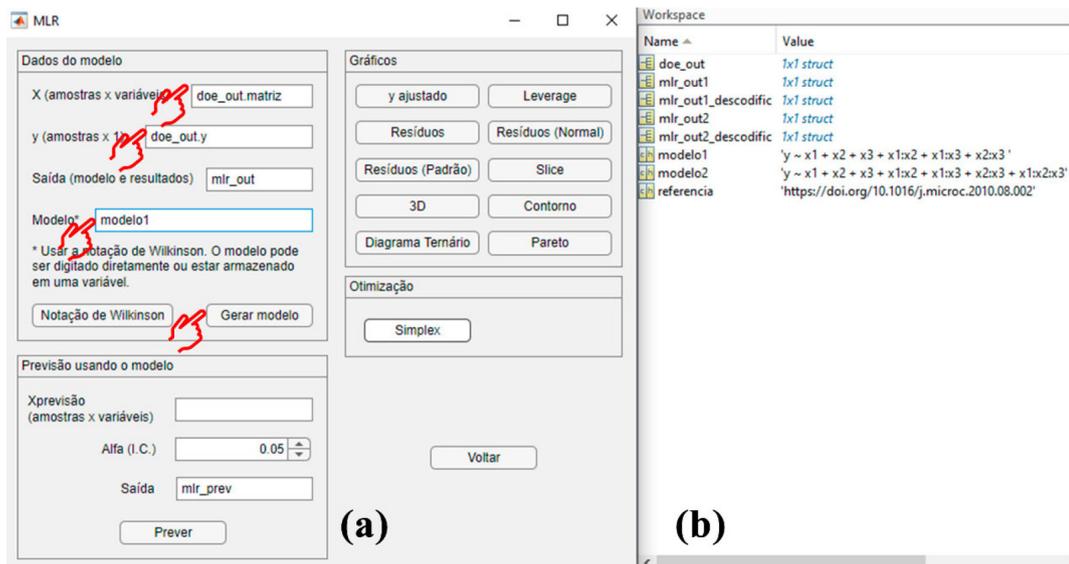


Figura 6. (a) Construindo um modelo. (b) Gerando a estrutura de dados do modelo

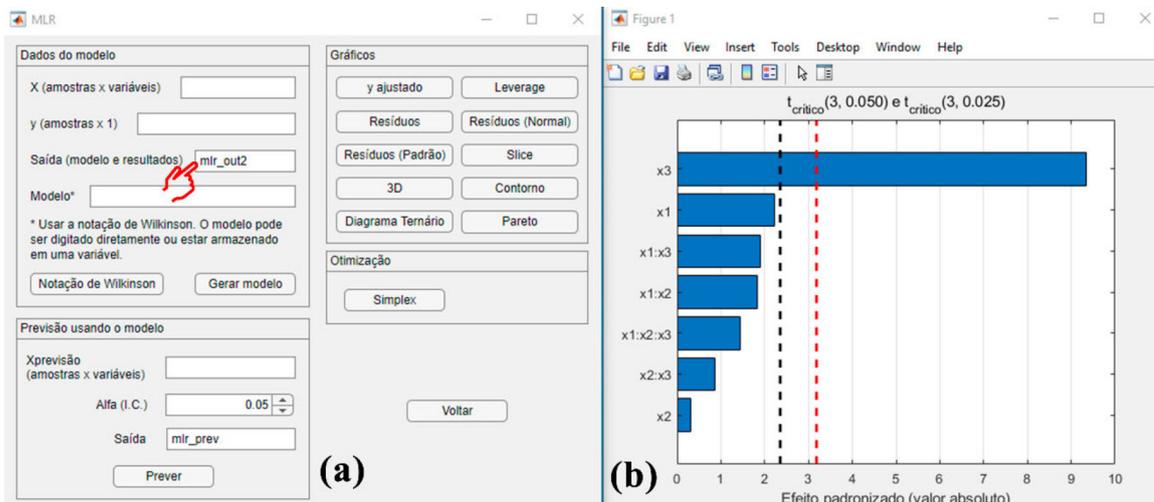


Figura 7. (a) Checando a qualidade de ajuste do modelo 2. (b) Gráfico de Pareto cujo efeito “ $x_1:x_2:x_3$ ” não foi significativo

opção “Slice”. Nessa opção, os usuários podem prever os valores de  $y$  estimados pelo modelo nos diferentes níveis experimentais dos fatores avaliados (Figura 8). Para uma melhor visualização, optamos por usar os valores  $x_1$ ,  $x_2$  e  $x_3$  descodificados nesse gráfico, portanto, um novo modelo com a matriz de dados descodificados precisa ser gerada, ou seja, os campos “ $X$  (amostras  $\times$  variáveis)” e “Modelo\*” devem ser renomeados para “doe\_out\_descodificado” e “mlr\_out\_descodific”, respectivamente (Figura 8a). Na área de cada sub-gráfico há uma linha vertical pontilhada que pode ser direcionada conforme as necessidades do usuário. Essas linhas de  $x_1$ ,  $x_2$  ou  $x_3$  podem ser direcionadas para esquerda ou direita e o valor de  $y$  predito pelo modelo pode ser verificado (Figura 8b). Neste caso, optamos pelo aumento de  $y$ , ou seja, a maximização do sinal de emissão do urânio em 367,007 nm. Para essa resposta a condição de máximo é alcançada quando se utiliza 0,5 mol L<sup>-1</sup> de HNO<sub>3</sub>, uma potência do plasma de 1,1 kW e vazão do gás do nebulizador de 0,9 L min<sup>-1</sup> com um  $y$  de 585,9 U.

Os demais recursos gráficos serão demonstrados nos exemplos adiante. Cabe ainda ressaltar que, nesta demonstração aplicamos um planejamento fatorial completo, porém, o mesmo procedimento pode ser seguido para o planejamento fatorial fracionário, ou superfície de resposta.

### Metodologia de superfície de resposta (MSR)

MSR são ferramentas que fazem o uso de técnicas de regressão múltipla para modelar, analisar e otimizar a relação entre  $X$  (variáveis independentes) e  $y$  (variável dependente). Na MSR um modelo matemático é aproximado por uma função de segunda ordem, que é uma expansão polinomial que inclui termos lineares, quadráticos e de interação entre os fatores que pode ser usado para prever a resposta em diferentes combinações dos fatores e pode ser aplicada na otimização de uma infinidade de processos e produtos. As principais MSR são: planejamento Box-Behnken, planejamento composto central rotacional e matriz de Doehlert.<sup>14</sup>

### Delineamento Box-Behnken

Neste exemplo o processo de hidrólise do amido de arroz para produção de xarope com alto teor de dextrose equivalente (DE) foi otimizado utilizando um planejamento Box-Behnken 3<sup>3</sup> ([link YouTube: https://www.youtube.com/watch?v=3chGfZoGdFg](https://www.youtube.com/watch?v=3chGfZoGdFg)). Nesse estudo foram avaliados os efeitos da temperatura (°C) ( $x_1$ ), relação enzima/substrato (% m m<sup>-1</sup>) ( $x_2$ ) e tempo de liquefação (min) ( $x_3$ ) na %DE alcançada ( $y$ ). Para mais informações, consultar o trabalho de Spinoza *et al.*<sup>23</sup>

Seguindo os mesmos procedimentos descritos no planejamento fatorial completo, os usuários necessitam primeiro definir o número de “Fatores independentes”, “Pontos centrais” e “Planejamento” a ser utilizado. Nesse exemplo, foi empregado um planejamento Box-Behnken com três fatores independentes e três pontos centrais. Nesse planejamento os usuários têm ainda a opção de gerar a matriz de dados descodificada marcando a opção “Incluir matriz descodificada”. Após a definição desses parâmetros, os usuários devem apertar o botão “Box-Behnken”, e então uma estrutura de dados chamada “doe\_out” será gerada.

A etapa seguinte, consiste na aquisição dos valores experimentais de  $y$  para cada experimento. Após as determinações experimentais os valores obtidos devem ser inseridos em um vetor  $y$  que representa a variável dependente, isto é, a resposta em %DE. Uma das opções para a criação do vetor  $y$  é dada na seção anterior. Agora, os usuários devem iniciar a análise dos dados apertando o botão “Analisar/MLR”, e então inserir os dados que correspondem a matriz de dados  $X$  e o vetor  $y$  na interface “MLR” que se abrirá. Optamos por manter os nomes gerados automaticamente pelo aplicativo (“doe\_out.matriz” para  $X$  e “doe\_out.y” para  $y$ ), porém, os usuários podem criar suas próprias nomenclaturas. Nesta mesma interface, os usuários devem ainda definir os modelos a serem utilizados, que podem ser carregados previamente, ou escritos seguindo as notações de Wilkinson. Nesse caso, optamos por carregar o modelo, para isso, os usuários devem carregar o arquivo “Modelo\_Box\_Behnken.mat” disponibilizado nos conjuntos de dados para demonstração. Então, um arquivo chamado “modelo” aparecerá no *Workspace* do Matlab, e esse mesmo nome

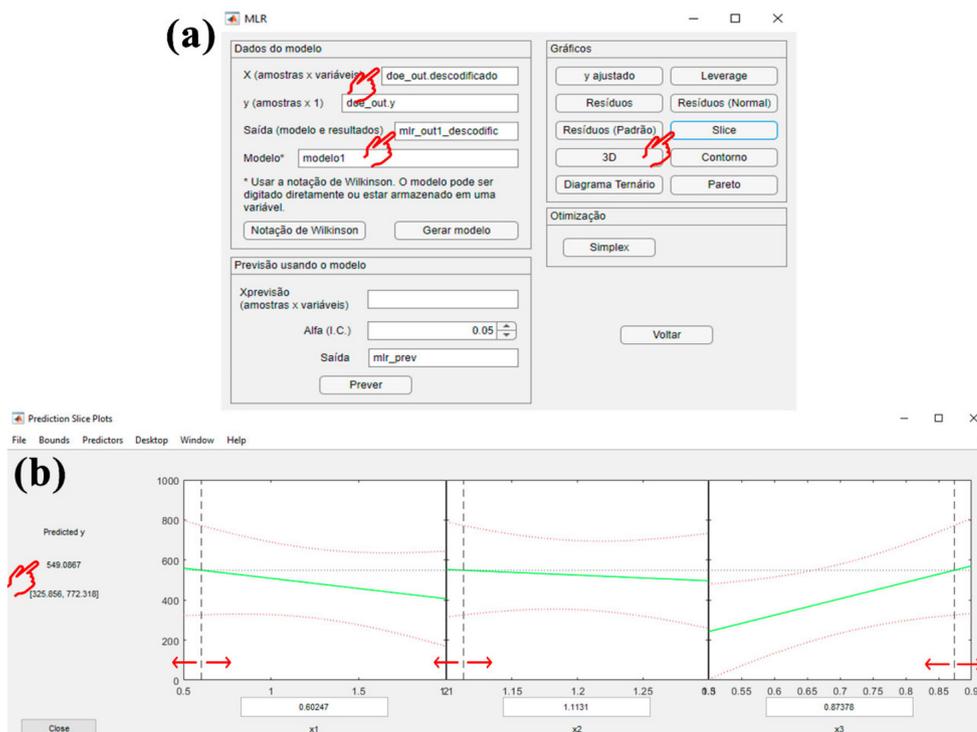


Figura 8. (a) Criando um modelo para os dados descodificados. (b) Gráfico de slice interativo

deverá ser inserido dentro da janela “Modelo\*” na interface “MLR”, e então apertar o botão “Gerar modelo”.

O modelo considerado nesse estudo contém todos os termos individuais, quadráticos e de interação. Na estrutura “mlr\_out” gerada, os parâmetros de qualidade de ajuste do modelo aos dados experimentais podem ser checados. De acordo com a ANOVA, o modelo demonstrou significância, mas falta de ajuste ( $p > 0,05$ ) a nível de 95% de confiança. Embora o modelo tenha falta de ajuste, nesse caso é fundamental checar o valor do quadrado médio do erro experimental.<sup>24</sup> Se o valor do quadrado médio do erro experimental é extremamente baixo, como nesse caso MeanSq de  $1,6 \times 10^{-5}$ , é um indicador positivo de um bom ajuste do modelo, porém, os usuários devem ainda verificar outros parâmetros, como os padrões de distribuição dos resíduos, leverage, entre outras opções disponíveis na interface “MLR”. Somado a esses fatores, os valores de  $R^2$  foram superiores a 94,9% com baixo valor de RMSE. Diante das possibilidades avaliadas e do bom senso, concluímos que, o fato do teste da falta de ajuste significativo pode ser irrelevante neste caso.

O gráfico 3D ou superfície de contorno são fundamentais para visualizar o comportamento das variáveis independentes em função da resposta. Embora tais figuras possam também ser geradas para os planejamentos fatoriais, é nas MSR que elas predominam. Para fins demonstrativos utilizaremos as variáveis descodificadas para melhor interpretação dos resultados. Iniciaremos gerando o gráfico 3D apertando o botão “3D”, em seguida a interface “Gráfico 3D” se abrirá. Ao plotar a superfície 3D somente duas variáveis podem ser apresentadas simultaneamente, portanto, as demais variáveis devem ser fixadas no mesmo valor. Para estudos com somente duas variáveis independentes não há necessidade de ajuste, mas em casos com mais de duas variáveis independentes, como esse em questão, uma variável deve ser fixada. Uma boa opção é usar a variável menos significativa para ser fixada, cuja significância pode ser visualizada na

opção “Pareto”, que neste caso foi a variável  $x_3$  (Figura 9a). Para uma melhor visualização do ótimo experimental, o ideal é fixar a variável menos importante em seu valor ótimo, para isso, os usuários podem usar a opção “Slice”. Usando o gráfico de slice a melhor condição para  $x_3$  foi de 20 min (Figura 9b), dessa forma o usuário deve inserir esse valor como máximo e mínimo para  $x_3$  no gráfico 3D (Figuras 9c e 9d). A superfície de contorno é gerada apertando o botão “Contorno” (Figura 9e), e pode ser obtida da mesma forma que o gráfico 3D.

O objetivo principal em um estudo de MSR é a obtenção dos valores otimizados. No aplicativo GAMMA-GUI a condição de ótimo do modelo é obtida através da otimização simplex. Esse método já é consagrado para resolver problemas de otimização baseado em uma figura geométrica irregular que “caminha” pela superfície de respostas afim de encontrar a solução ótima do problema. Para mais detalhes, consultar o trabalho de Bona *et al.*<sup>25</sup> Ao apertar o botão “Simplex” a janela “Otimização Simplex” se abrirá. Nessa janela os usuários têm possibilidade de realizar a otimização de uma ou múltiplas respostas usando a função D. Nesse exemplo otimizamos somente a %DE, portanto, o usuário deve inserir na caixa do modelo 1 a estrutura a ser otimizada. Optamos por utilizar as variáveis descodificadas, então no campo “Estrutura” para o “modelo 1” o usuário deverá digitar “mlr\_out.descodific” e apertar o botão “Continuar”. A janela “Função de desajustabilidade” se abrirá, e nela os usuários têm a possibilidade de otimização de máximo, mínimo ou de alvo. Além disso, para casos de otimização multirespostas pesos podem ser atribuídos aos modelos. Nesse caso, queremos maximizar a %DE, então devemos inserir no valor de alvo o valor de y máximo. Após essas definições apertar o botão “Otimizar”. Durante a otimização uma série de gráficos são apresentados, e ao término do processo uma estrutura de dados chamada “otm\_out” será gerada no Workspace. Os usuários podem acompanhar o processo iterativo de otimização nos gráficos, mas na estrutura “otm\_out.resultado.potm” é onde os valores

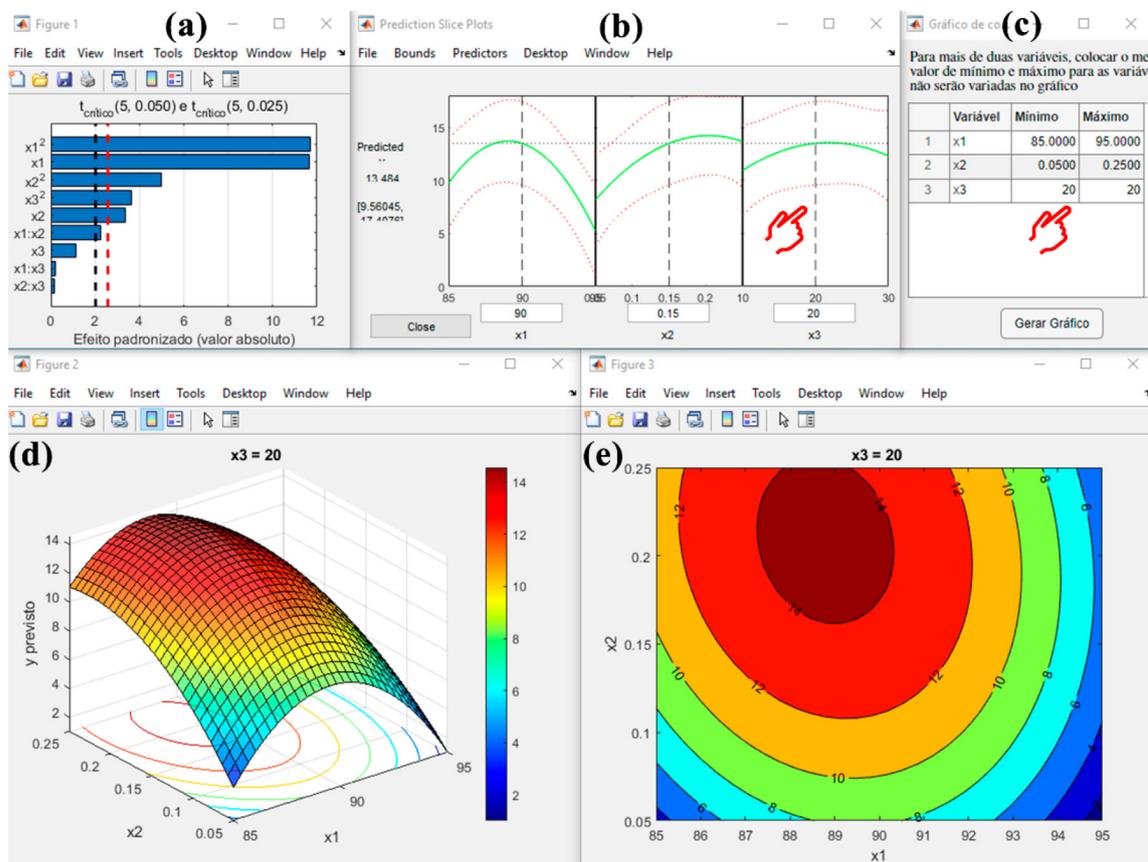


Figura 9. (a) Gráfico de Pareto. (b) Gráfico de slice. (c) Interface para fixação das variáveis. (d) Gráfico 3D. (e) Superfície de contorno

de  $y$ ,  $x_1$ ,  $x_2$  e  $x_3$  otimizados são encontrados. Nessa demonstração, a condição de ótimo obtida foi de 13,8 %DE a uma temperatura de 89,2 °C, razão enzima/substrato de 0,25% m m<sup>-1</sup> e tempo de liquefação de 17,6 min.

#### Delineamento composto central

Nessa demonstração, a codigestão anaeróbica da composição do substrato de resíduos agrícolas foi otimizada usando planejamento composto central rotacional para maximizar a produção de metano ([link YouTube: https://www.youtube.com/watch?v=wq-kl3dc6ro](https://www.youtube.com/watch?v=wq-kl3dc6ro)). Diferentes frações mássicas ( $\omega$ ) ( $x_1$ ) e relação de carbono/nitrogênio (R) ( $x_2$ ) dos resíduos agrícolas de raiz de beterraba, esterco de vaca e esterco de aves foram selecionados como variáveis independentes tendo a produção de metano como resposta  $y$ . Consultar o trabalho de Dima *et al.*<sup>26</sup> para mais detalhes.

Esse estudo contém duas variáveis independentes ( $\omega$  e R), que totalizam doze execuções experimentais, dos quais quatro são repetições do ponto central. A matriz de planejamento pode ser gerada na interface “Gerar matriz” no botão “Composto Central Rotacional”, selecionando o número de fatores, pontos centrais, e incluindo ou não a matriz de dados decodificada, conforme descrito anteriormente. Os usuários podem carregar os dados experimentais no arquivo “Planejamento\_CCD\_digestao.mat”, disponível no material de demonstração. Ao carregar o arquivo uma estrutura de dados “doe\_out” aparecerá no *Workspace*. Nessa estrutura os usuários têm acesso a matriz de dados codificada e decodificada, como também os valores experimentais de  $y$  obtidos para cada execução experimental.

A próxima etapa consiste na escolha e avaliação do modelo que melhor se ajusta aos dados experimentais. O modelo completo foi escolhido para essa demonstração, e consiste na inclusão dos termos individuais, quadráticos e de interação. O modelo pode ser carregado usando o arquivo “Modelo\_CCD.mat”, então uma variável do tipo “char” chamado modelo aparecerá no *Workspace* do Matlab. Agora na interface “MLR”, os usuários devem digitar nos campos a matriz de dados  $X$ , vetor  $y$  e modelo, conforme demonstramos nos exemplos anteriores. Ao apertar “Gerar modelo” uma estrutura “mlr\_out” será gerada, e nela todos os parâmetros de qualidade de ajuste do modelo podem ser verificados. Para esse caso, o modelo foi significativo ( $p < 0,05$ ), não apresentou falta de ajuste ( $p > 0,05$ ) e teve valores de  $R^2$  superior a 93,9%.

O gráfico 3D e a superfície de contorno gerada utilizando as variáveis decodificadas demonstram a condição de ótimo do processo. Nesse caso, um novo modelo foi construído

alterando os campos “ $X$  (amostras  $\times$  variáveis)” e “Modelo” para “doe\_out.descodificado” e “mlr\_out\_descodific”, respectivamente. Para obter os valores exatos da condição de ótimo o usuário deve apertar o botão “Simplex”, e seguir os mesmos passos descritos anteriormente no planejamento Box-Behnken. Na estrutura “otm\_out” gerada após o processo de otimização, os usuários podem obter a condição de ótimo, que nesse caso foi  $\omega$  de 0,36 e R de 26,7 tendo um rendimento máximo de metano de 349,8 mL g<sup>-1</sup> VS<sup>-1</sup>.

#### Matriz de Doehlert

A matriz de Doehlert é frequentemente utilizada associada ao planejamento fatorial para a seleção prévia dos fatores mais importantes a serem otimizados. Neste exemplo, foi desenvolvido um procedimento de digestão de Kjeldahl em micro-ondas para determinação de nitrogênio em feijão ([link YouTube: https://www.youtube.com/watch?v=vQrS6OCd\\_iQ&t=1s](https://www.youtube.com/watch?v=vQrS6OCd_iQ&t=1s)). Inicialmente, um planejamento fatorial completo 2<sup>4</sup> foi utilizado para verificar a influência da temperatura de decomposição ( $x_1$ ), volume de H<sub>2</sub>SO<sub>4</sub> (mL) ( $x_2$ ), de H<sub>2</sub>O<sub>2</sub> (mL) ( $x_3$ ) e massa de K<sub>2</sub>SO<sub>4</sub> (g) ( $x_4$ ) na % (m m<sup>-1</sup>) de proteína (%PTN) ( $y$ ). Após a etapa de varredura os fatores estatisticamente significativos foram otimizados utilizando a matriz de Doehlert. Para mais detalhes consultar o trabalho de Korn *et al.*<sup>27</sup>

Nos dados disponíveis para *download* os usuários devem carregar o arquivo “Planejamento\_fatorial\_completo\_com\_Matriz\_Doehlert.mat”, ao carregar o arquivo uma série de estruturas aparecerão no *Workspace* do Matlab. Iniciaremos nossa demonstração com o planejamento fatorial completo 2<sup>4</sup>, selecionando a estrutura nomeada como “doe\_out\_fatorial\_completo”. Nessa estrutura, os usuários têm acesso a matriz de dados codificados e decodificados, níveis e valores de  $y$  determinados experimentalmente.

Na interface “MLR” os valores de  $X$ ,  $y$  e modelo devem ser especificados nas devidas caixas de texto, inserindo os seguintes comandos “doe\_out\_fatorial\_completo.matriz”, “doe\_out\_fatorial\_completo.y” e “modelo1\_fatorial”. Nesse caso, construiremos um modelo para o planejamento fatorial e outro para a matriz de Doehlert, portanto, na janela de saída do modelo “Saída modelo e resultados” chamaremos o modelo para o planejamento fatorial de “mlr\_out1”, e então apertar o botão “Gerar modelo”. Após essa etapa a estrutura de dados que corresponde aos resultados “mlr\_out1” aparecerá no *Workspace*. Nessa estrutura o usuário pode checar a qualidade de ajuste do modelo entre outras opções de saídas gráficas. Neste caso o modelo foi significativo, sem falta de ajuste com um  $R^2$  superior a 67,0%.

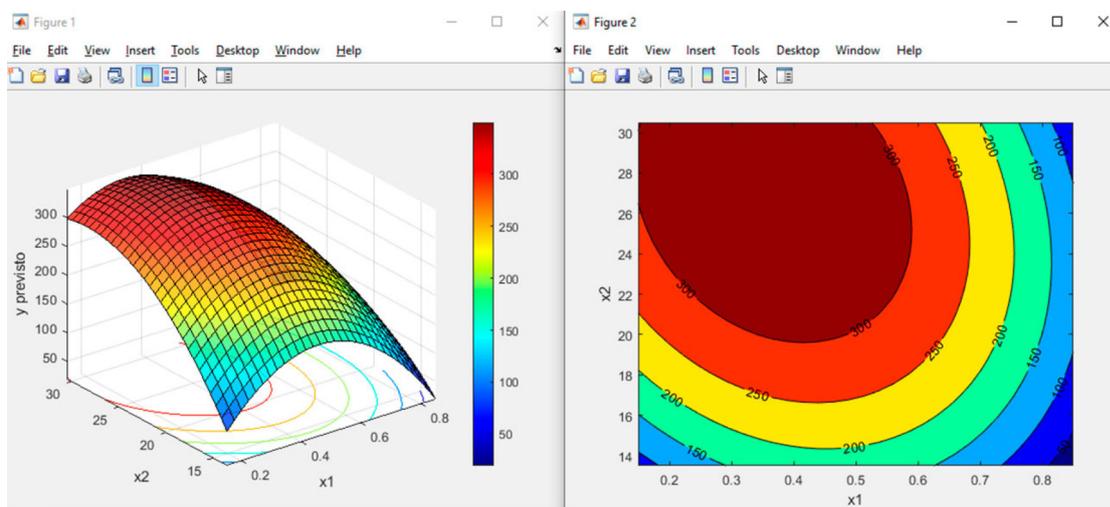


Figura 10. Condição de ótimo experimental

Para a seleção dos fatores mais importantes e seus níveis a serem otimizados, os usuários podem utilizar a opção “Pareto” combinada com o gráfico “Slice”. Com o gráfico de Pareto (Figura 11a), nota-se que as variáveis  $x_1$  (temperatura) e  $x_2$  (volume de  $H_2SO_4$ ) foram estatisticamente significativas, enquanto o gráfico de slice (Figura 11b) ajuda na definição dos níveis que maximizam a resposta, que nesta demonstração corresponde a menores temperaturas e maiores volumes de  $H_2SO_4$ .

O planejamento fatorial auxiliou na seleção das variáveis e níveis que serão otimizados pelo planejamento de Doehlert, neste caso as duas variáveis mais importantes. Neste exemplo, os autores optaram por avaliar a temperatura de decomposição entre 270 e 310 °C e o volume de ácido sulfúrico de 8 a 12 mL, respectivamente. A matriz de planejamento e os dados experimentais dessa etapa podem ser visualizados na estrutura “doe\_out\_doehlert”. Com os dados em mãos, a próxima etapa consiste em carregar os dados nas suas respectivas caixas, matriz  $X$ , vetor  $y$  e modelo, usando as seguintes notações “doe\_out\_doehlert.matriz”, “doe\_out\_doehlert.y” e “modelo2\_doehlert”, respectivamente. Precisamos ainda, na janela de saída do modelo “Saída modelo e resultados” renomear o modelo do planejamento Doehlert como “mlr\_out2”, e então apertar o botão “Gerar modelo”. Em seguida, checamos a qualidade de ajuste do modelo acessando a estrutura “mlr\_out2”. Neste caso, os valores de  $R^2$  foram superiores a 79,0%, porém o modelo não foi significativo ( $p = 0,069$ ) e teve falta de ajuste ( $p = 0,018$ ). Embora esse modelo tenha certa limitação estatística, uma alternativa seria realizar uma nova otimização empregando outra faixa de volume de  $H_2SO_4$  (ver Figura 12). Caso contrário, o princípio do bom senso deve prevalecer, e os prós e contras do uso desse modelo devem ser levantados. Nesse caso, o modelo foi considerado, então os autores justificam os motivos

de suas escolhas, além disso, cabe ainda ressaltar que a condição de ótimo foi validada experimentalmente.

Na Figura 12 temos uma superfície de resposta com ponto de sela, ou seja, existe um platô em relação à variável  $x_2$ , indicando que a sua variação dentro dos níveis estudados não afeta o sistema modelado. O ponto ótimo para essa condição pode ser definido usando a opção “Simplex”, e corresponde à temperatura de decomposição de 279,0 °C e volume de  $H_2SO_4$  de 9,0 mL para um valor máximo de 22,8 %PTN.

## Planejamento de misturas

### Simplex centroide

Neste exemplo foi realizada a otimização da energia de ativação ( $E_a$ ) do processo de oxidação do biodiesel utilizando planejamento de misturas simplex centroide ([link YouTube: https://www.youtube.com/watch?v=NUXCDeuiyks&t=41s](https://www.youtube.com/watch?v=NUXCDeuiyks&t=41s)). Esse estudo tinha como objetivo verificar os efeitos sinérgicos na inibição da reação de oxidação do biodiesel utilizando três antioxidantes sintéticos: *terc*-butil-hidroquinona (TBHQ) ( $x_1$ ), butil-hidroxitolueno (BHT) ( $x_2$ ) e butil-hidroxianisol (BHA) ( $x_3$ ), isto é, maximização da  $E_a$  ( $y$ ). Para mais detalhes consultar o trabalho de Galvan *et al.*<sup>28</sup>

A primeira etapa consiste em gerar a matriz que corresponde ao planejamento de mistura simplex centroide. Para isso, o usuário deve selecionar a opção “Gerar matriz”, e em “Variáveis de mistura” inserir o número três, que corresponde aos antioxidantes BHA, BHT e TBHQ, e então definir o número de centroides, nesse caso três, e por fim apertar o botão “Simplex centroide”. Após essa etapa, uma estrutura “doe\_out” será gerada no *Workspace*. Crie e insira no vetor  $y$  os valores experimentais, seguindo os passos descritos nas seções anteriores. Na interface “Analisar/MLR”, defina a matriz de dados  $X$ , vetor  $y$

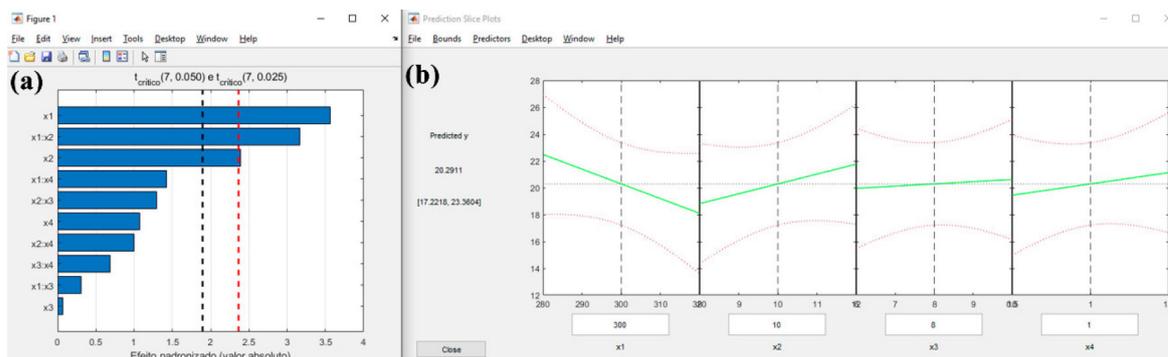


Figura 11. (a) Gráfico de Pareto do planejamento fatorial completo  $2^4$ . (b) Gráfico de slice indicando os níveis dos fatores a serem otimizados

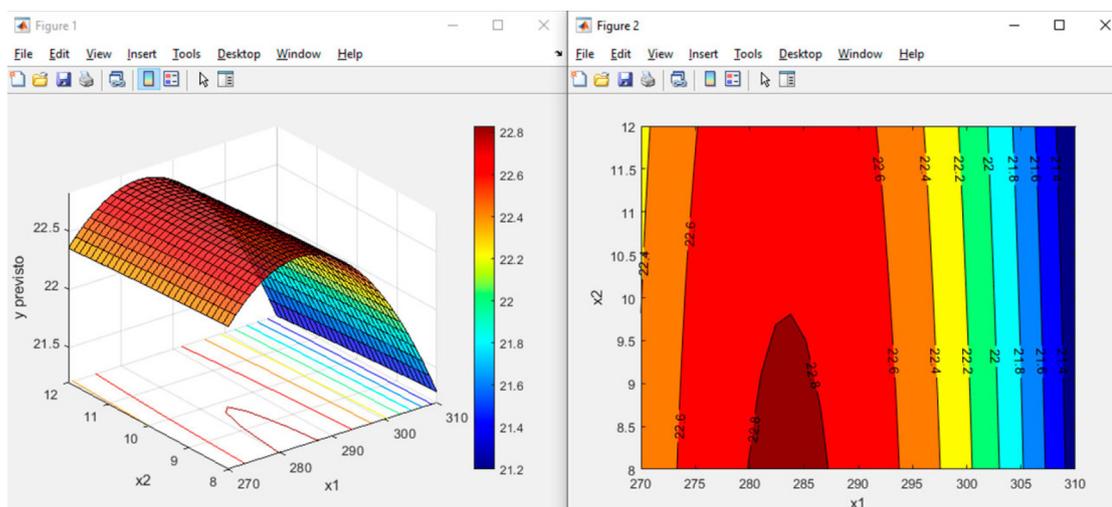


Figura 12. Condição de ótimo experimental para o planejamento de Doehlert

como “*doe\_out.matriz*”, “*doe\_out.y*”, respectivamente. Na opção “*Modelo*”, testaremos o ajuste de três possibilidades: modelo linear “*modelo1*”, quadrático “*modelo2*” e cúbico especial “*modelo3*” que podem ser carregados em “*Modelos\_experimentos\_de\_mistura.mat*”. É importante ressaltar que, no planejamento de misturas, não há um termo constante (intercepto), pois a ideia aqui, é estudar o efeito das proporções dos componentes da mistura e não o efeito absoluto de cada componente individualmente. Além disso, o valor do intercepto é incorporado aos demais para adequar o modelo à restrição de misturas.<sup>17</sup> Portanto, ao construir um modelo usando a notação de Wilkinson os usuários devem retirar o intercepto usando a notação “-1”, conforme mostrado neste exemplo para o modelo linear: “ $y \sim -1 + x_1 + x_2 + x_3$ ”. Destaca-se que o gráfico de slice disponibilizado no GAMMA-GUI não pode ser aplicado para planejamento de misturas, pois ele não obedece a restrição de mistura.

A etapa seguinte consiste em checar o modelo que mais se ajusta aos dados experimentais. Para isso, os usuários devem informar na caixa “*Modelo\**”, o modelo a ser avaliado, simultaneamente em “*Saída de modelo*”, as saídas devem ser renomeadas. Nessa demonstração adotaremos “*modelo1*” e saída “*mlr\_out1*”; “*modelo2*” e saída “*mlr\_out2*”; e “*modelo3*” saída “*mlr\_out3*”. Ao checar os parâmetros e ajuste dos modelos, selecionamos o quadrático (efeito de interação  $x_1:x_2:x_3$  não foi significativo) (ver Figuras 13a, 13b, 13c), cujos parâmetros de ajuste e validação estão armazenados dentro da estrutura “*mlr\_out2*”. Ao utilizar o modelo quadrático os valores  $R^2$  foram superiores a 95,4%. Além disso, a ANOVA mostrou que o modelo foi significativo ( $p = 0,0076$ ) e não teve falta de ajuste significativa ( $p = 0,1067$ ).

Após eleger o modelo adequado e verificar a qualidade de ajuste, os usuários podem gerar o diagrama de mistura apertando

o botão “*Diagrama Ternário*” (Figura 13d), e então obter as proporções otimizadas na opção “*Simplex*” inserido “*mlr\_out2*” na aba “*Otimização Simplex*” que se abrirá. Na interface “*função de otimização*”, os usuários não podem esquecer de inserir o número de variáveis de misturas na caixa “*Variáveis de misturas*”, três para esse caso e então apertar o botão “*Otimizar*”. Na estrutura de dados “*otm\_out2*” gerada no *Workspace* que corresponde aos valores otimizados para o modelo quadrático a condição otimizada foi de 66,6% TBHQ e 16,7% de BHT e BHA, que corresponde a um valor  $y$  de 104,6 kJ mol<sup>-1</sup>.

#### Pseudocomponentes

Em muitos casos de modelagem de misturas necessitamos da presença de todos os componentes para obter um produto apropriado. Nessa demonstração, temos a formulação de uma membrana de revestimento para um eletrodo seletivo com o objetivo de maximizar o sinal analítico ( $y$ ) de substratos biológicos (*link* YouTube: <https://www.youtube.com/watch?v=Rd9vOI0ZGdo>). A formulação desta membrana só é possível se as soluções de pirrol ( $x_1$ ), KCl ( $x_2$ ) e  $K_4[Fe(CN)_6]$  ( $x_3$ ) estiverem todas presentes na mistura. Esse é um caso de planejamento de misturas com pseudocomponentes, isto é, variáveis “fictícias” são criadas para representar as proporções dos componentes reais da mistura. Nesse delineamento os limites inferiores dos componentes nas misturas foram definidos em 10%, que simultaneamente restringe o limite superior para 80%. Para mais detalhes consultar o trabalho de Barros Neto *et al.*<sup>5</sup>

Iniciaremos gerando a matriz de experimentos na interface “*Gerar Matriz*”. O planejamento de mistura empregado nesse estudo foi o simplex aumentado, que contém pontos adicionais entre o ponto central e os vértices. O usuário deverá ajustar para três variáveis de

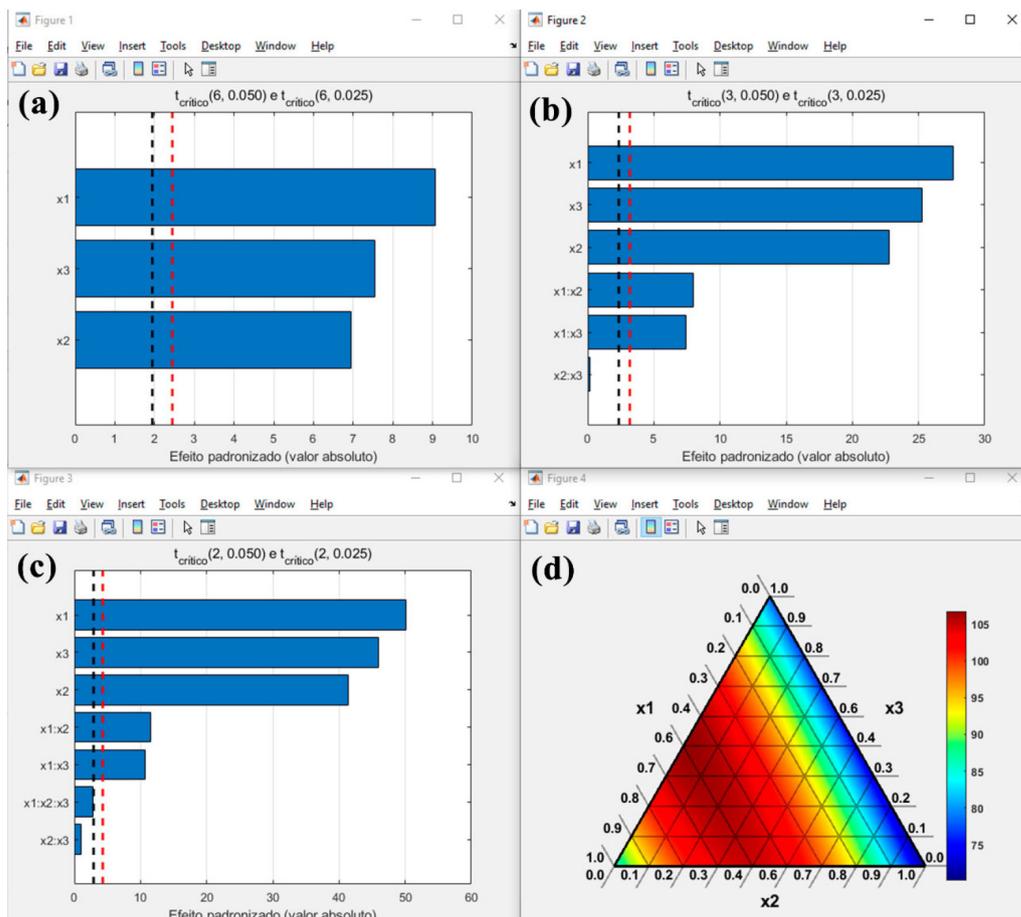


Figura 13. (a) Modelo linear. (b) Modelo quadrático. (c) Modelo cúbico especial. (d) Diagrama ternário para o planejamento de misturas simplex centrado

misturas a opção “Variáveis de mistura” e incluir dois pontos centrais em “Centroides”. Além disso, as opções “Pontos axiais (mistura)” e “Pseudo-componentes (misturas)” devem ser marcadas, e por fim apertar o botão “Simplex Centroide” (Figura 14a). Uma nova interface “Pseudocomponentes - DOE” se abrirá, e nesta janela os valores mínimos de cada componente devem ser definidos, que nesse caso foi 10% para cada, que corresponde a proporção de 0,1 (Figura 14b). Em seguida, aperte o botão “Descodificar” gerando a estrutura de dados “doe\_out”, que contém as informações das restrições aplicadas, a matriz de experimentos codificada/pseudocomponentes e descodificada/real. Esse planejamento é composto por onze experimentos, dos quais dois são repetições no ponto central e três pontos axiais. Entretanto, alguns experimentos foram realizados em duplicata e outros em triplicata, totalizando vinte experimentos presentes na estrutura “doe\_out\_replicatas” que pode ser carregada em “Pseudocomponente\_eletrodo\_livro.mat” disponível no material de demonstração.

Para a análise dos dados na interface “MLR” os usuários precisam definir a matriz de dados  $X$ , vetor  $y$  e modelo usando as notações “doe\_out\_replicatas.matriz”, “doe\_out\_replicatas.y” e “modelo”, respectivamente. Na estrutura “mlr\_out” os usuários podem checar a qualidade do modelo, conforme vimos nas seções anteriores. Uma etapa fundamental ao utilizar esse tipo de experimento, é de

fato saber qual a proporção real de cada componente na mistura, e como apresentar os resultados. Nesse caso, podemos apresentar o diagrama de mistura através dos pseudocomponentes (Figura 15a) ou das concentrações dos componentes puros (Figura 15b) construindo um novo modelo com a variáveis descodificadas, alterando “doe\_out\_replicatas.matriz”, para “doe\_out\_replicatas.descodificado” e “mlr\_out” para “mlr\_out\_descodific”, respectivamente. As representações são equivalentes, entretanto, para o caso das representações dos componentes puros, temos que ter a consciência que não é aceitável prever condições fora do domínio experimental avaliado, isto é, fora do triângulo interno (Figura 15b).

Finalmente, é hora de obter a condição de ótimo da formulação apertando o botão “Simplex”. Cabe ressaltar que, a obtenção da condição de ótimo fica a critério do usuário, que pode obter as condições baseados em pseudocomponentes ou reais. Nesse caso, optamos por obter as condições reais usando a estrutura “mlr\_out\_descodific”. O processo de otimização segue os mesmos passos descritos na demonstração do simplex centroide. Na estrutura “otm\_out\_descodific” gerada, os usuários podem checar a condição de ótimo real que corresponde a 55,4% de solução de pirrol, 10,0% de KCl e 34,6% e  $K_4[Fe(CN)_6]$  tendo como sinal analítico um valor máximo de 4,3. Se o usuário optar pela otimização usando a saída “mlr\_out”, os resultados obtidos serão

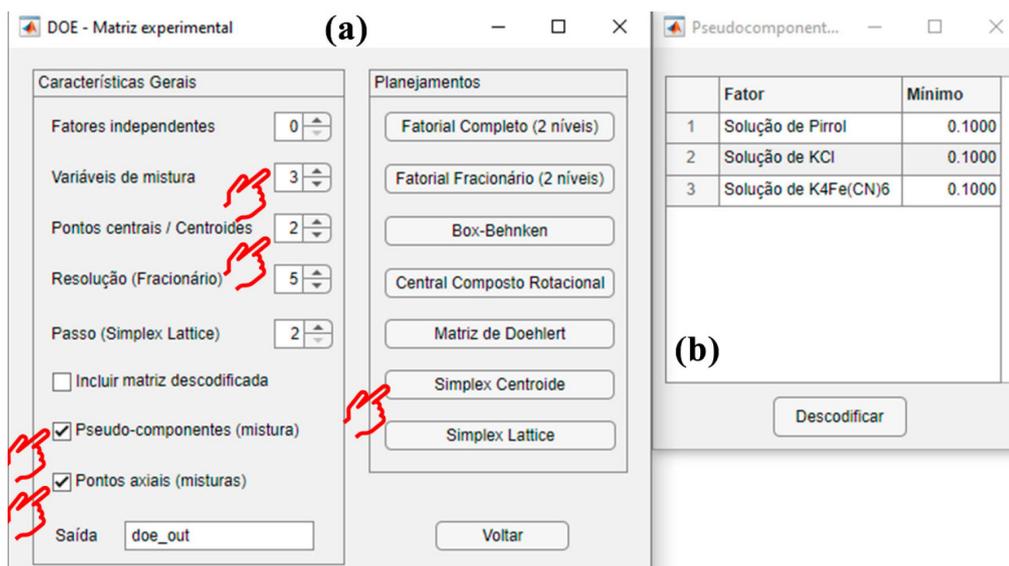


Figura 14. (a) Aplicando a opção de pseudocomponentes. (b) Definindo os limites dos pseudocomponentes

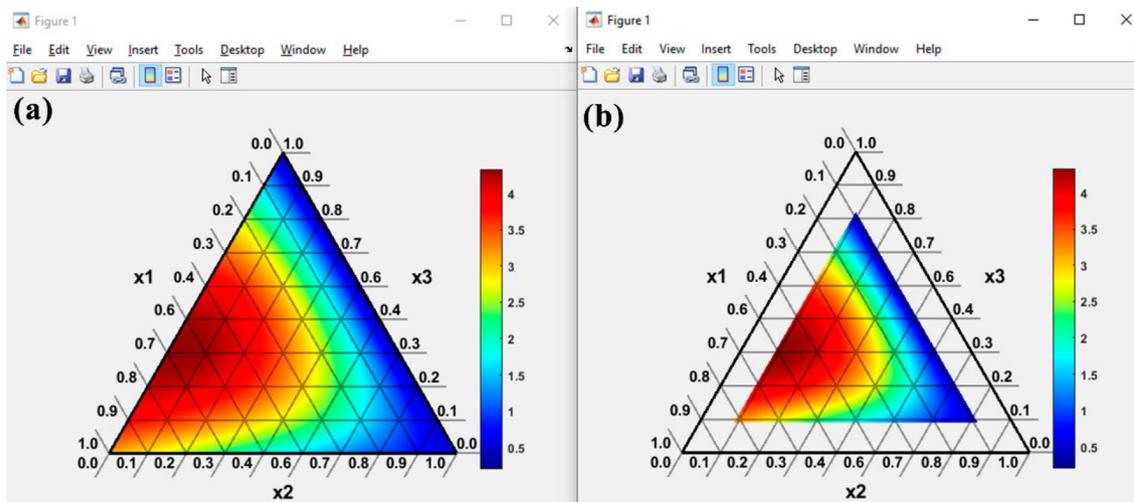


Figura 15. (a) Diagrama de misturas em pseudocomponentes e (b) região de domínio experimental avaliado

expressos em pseudocomponentes e a descodificação será requerida. Ver expressão de descodificação na página 338 do trabalho de Barros Neto *et al.*<sup>5</sup>

### Otimização com variável de processo

Neste estudo foi desenvolvida uma formulação de bolo de chocolate com substituição parcial da farinha de trigo pela farinha de yacon e maca ([link YouTube: https://www.youtube.com/watch?v=rjxVbRmfogQ&t=1s](https://www.youtube.com/watch?v=rjxVbRmfogQ&t=1s)). Testes preliminares mostraram que uma proporção mínima de farinha de trigo de 0,35 kg kg<sup>-1</sup> era necessária para produzir uma massa aceitável, portanto, os componentes originais da mistura foram descritos em termos de pseudocomponentes. As proporções da farinha de trigo ( $x_1$ ), yacon ( $x_2$ ) e maca ( $x_3$ ) foram variadas por planejamento de mistura simplex centroide e a quantidade de água foi incluída como variável de processo em três níveis: 120 mL (-1), 150 mL (0) e 180 mL (+1), tendo como resposta  $y$  a aceitabilidade do produto. Mais informações são fornecidas no trabalho de Tormena *et al.*<sup>20</sup>

Nesse estudo sete formulações de bolo foram feitas para as três quantidades de água avaliadas, totalizando vinte e um experimentos. O planejamento experimental e os resultados desse estudo podem ser carregados em "Variavel\_de\_processo\_bolo.mat". Para gerar a matriz do planejamento, os usuários devem construir a matriz manualmente, uma vez que varia de caso a caso. Ao carregar o conjunto de dados deste exemplo, os usuários terão no *Workspace* acesso as formulações, variáveis de misturas, matriz do planejamento baseada em pseudocomponentes, valores de resposta e o modelo  $y$  incluindo a variável de processo. Ao construir a equação do modelo usando a notação de Wilkinson para casos que incluem variável de processo, o  $z$  precisa ser expresso pelo caractere  $x$ , pois só é possível representar um modelo com a notação de Wilkinson em função dos caracteres  $x$  e  $y$ . Nesse exemplo a variável de processo  $z$  é representada pelo  $x_4$ : " $y \sim -1 + x_1 + x_2 + x_3 + x_1:x_2 + x_1:x_3 + x_2:x_3 + x_1:x_2:x_3$

$+ x_1:x_4 + x_2:x_4 + x_1:x_3:x_4 + x_2:x_3:x_4 + x_1:x_2:x_3:x_4 + x_2:x_4:x_4 + x_3:x_4:x_4 + x_1:x_2:x_4:x_4 + x_1:x_3:x_4:x_4 + x_2:x_3:x_4:x_4$ ".

Nesse tipo de otimização, as variáveis de mistura devem sempre ser as primeiras do modelo, nesse caso,  $x_1$ ,  $x_2$  e  $x_3$ , seguidas pela variável de processo  $x_4$ , ou seja, não é possível inverter, por exemplo, chamar a variável de processo de  $x_1$  e as de mistura  $x_2$ ,  $x_3$  e  $x_4$  pois a programação não considera essa hipótese. Cabe ainda ressaltar que, neste exemplo optamos pela remoção de alguns efeitos de ordem superior, pois são menos prováveis de serem estatisticamente significativos. Essa abordagem permite ainda uma redução do número de grau de liberdade possibilitando analisar a significância estatística dos outros coeficientes e o ajuste do modelo.

A etapa seguinte consiste em definir na interface "MLR" a matriz de planejamento em " $X$  (*amostras*  $\times$  *variáveis*)", as respostas em " $y$  (*amostras*  $\times$   $1$ )" e o modelo em "*Modelo\**", digitando "*planejamento\_codific*", " $y$ " e "*modelo\_artigo*", respectivamente, e então apertar "*Gerar modelo*". Na estrutura de dados "*mlr\_out*" gerada os usuários podem conferir a qualidade de ajuste do modelo, que nesse caso teve um  $R^2$  de 98,9%, modelo significativo ( $p < 0,05$ ) e falta de ajuste não significativa ( $p > 0,05$ ). Devido à complexidade dos termos (variável de processo associada a codificação para pseudocomponentes), a avaliação direta dos efeitos não é trivial. Dessa forma, as superfícies de resposta são ótimas alternativas para verificação dos resultados, e podem ser obtidas na opção "*Diagrama Ternário*". Geramos para essa demonstração as três possibilidades usando os níveis de água ( $x_4$ ) fixados em (-1), (0) e (+1), respectivamente (ver Figura 16).

De acordo com a Figura 16, nota-se que a quantidade de água influencia na aceitabilidade do produto. Portanto, a condição de ótimo pode ser obtida usando o algoritmo simplex no botão "*Simplex*", que buscará a aceitabilidade geral otimizada respeitando-se a restrição de mistura. Na aba "*Simplex otimização*" que se abrirá, a estrutura "*mlr\_out*" deverá ser fornecida e então o usuário deve apertar "*Continuar*". Na aba "*Função de desejabilidade*" que se abrirá, o

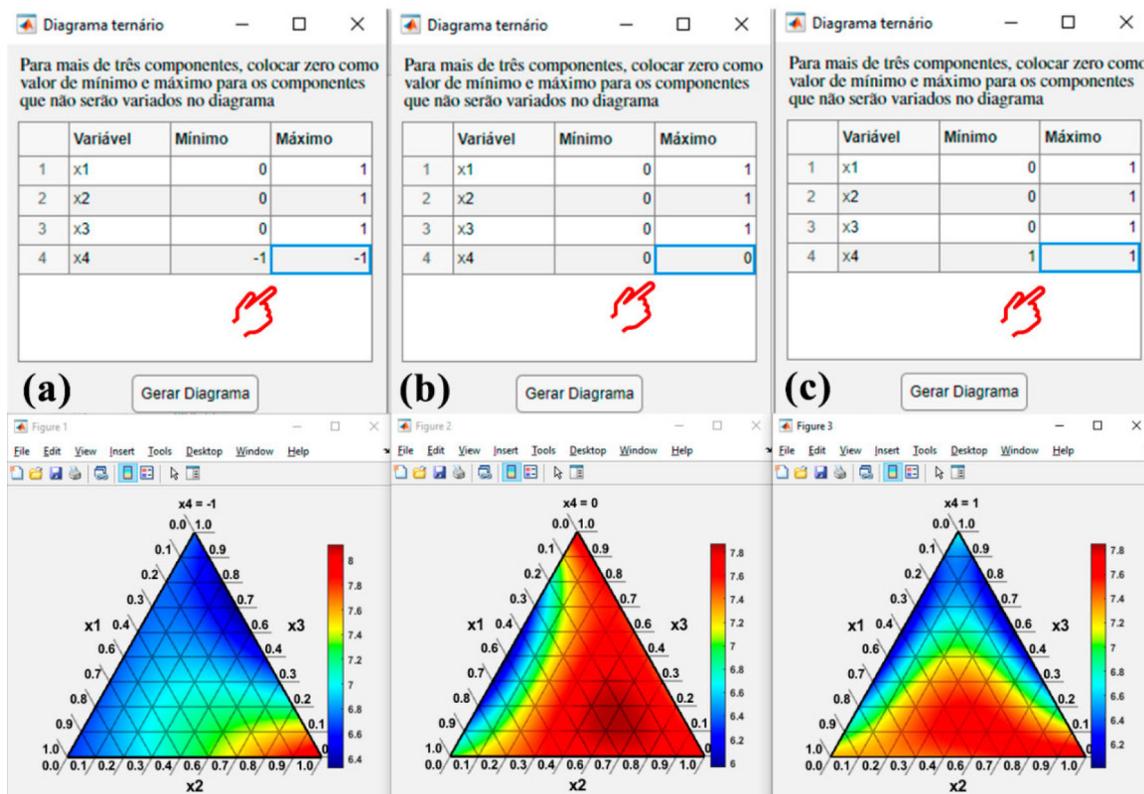


Figura 16. Diagrama de mistura em termos de pseudocomponentes fixando a variável de processo  $x_4$  em (a) 120 mL (-1), (b) 150 mL (0) e (c) 180 mL (+1)

usuário deverá definir para três o número de “Variáveis de mistura”, seguindo os mesmos procedimentos descritos na demonstração do simplex centroide anteriormente.

De acordo com a estrutura “*otm\_out*” gerada no *Workspace* o resultado da formulação otimizada baseado em pseudocomponentes foi de 0,0% de farinha de trigo, 100% de yacon, 0,0% de maca e (-1) para a água, que correspondem a valores reais de 0,35 kg de farinha de trigo/kg, 0,65 kg de farinha de yacon/kg, 0,00 kg de farinha de maca/kg e 120,0 mL de água, alcançando uma aceitabilidade de 8,12.

### Otimização de múltiplas respostas

Nessa seção demonstraremos uma otimização de múltiplas respostas para a formulação de um biodiesel que atende as especificações da Agência Nacional de Petróleo (ANP) e que tenha um menor custo de produção (*link* YouTube: <https://www.youtube.com/watch?v=0oCbNGNym8E&t=5s>). Nesse exemplo, misturas de óleo de soja ( $x_1$ ), sebo bovino ( $x_2$ ) e gordura de frango ( $x_3$ ) foram avaliadas utilizando planejamento de mistura simplex centroide tendo como resposta o rendimento da reação ( $y_1$ ), período de indução ( $y_2$ ), ponto de entupimento a frio ( $y_3$ ) e custo da matéria-prima por tonelada ( $y_4$ ). Esse é um caso interessante de otimização, algumas respostas devem ser maximizadas e outras minimizadas, além disso, os valores de algumas respostas devem atender as normas estabelecidas pela ANP. Para mais detalhes os leitores devem consultar o trabalho de Orives *et al.*<sup>29</sup>

Nessa demonstração ocultaremos as etapas prévias, como a construção da matriz de experimentos e a avaliação dos modelos. Nesse caso, iniciaremos no processo de otimização de múltiplas respostas. Para reprodução desse exemplo, o usuário precisa carregar o arquivo “Mutiresposta\_planejamento\_mistura\_Biodiesel.mat”, disponível no material de demonstração. Ao carregar o arquivo as estruturas de dados “*mlr\_out1*”, “*mlr\_out2*”, “*mlr\_out3*” e

“*mlr\_out4*” geradas através do modelo quadrático para  $y_1$ ,  $y_2$ ,  $y_3$  e  $y_4$ , respectivamente, aparecerão no *Workspace* do Matlab. Em seguida, os usuários podem visualizar os modelos usando a opção “*Diagrama Ternário*”. Para gerar o digrama o usuário precisa definir na caixa “*Saída (modelo e resultados)*” os dados que deseja plotar “*mlr\_out1*”, “*mlr\_out2*”, “*mlr\_out3*” e “*mlr\_out4*”, esse procedimento deve ser repetido para cada uma das quatro estruturas. Na Figura 17, demonstramos o diagrama ternário para cada resposta. É possível observar que encontrar visualmente uma região de ótimo que satisfaça a todas as respostas não é uma tarefa fácil. Porém, com a opção de otimização “*Simplex*” do GAMMA-GUI os usuários têm a possibilidade de realizar otimização de múltiplas respostas através da versátil função de desejabilidade.

Para iniciar o processo de otimização o usuário deverá apertar o botão “*Simplex*”, então a janela “*Otimização Simplex*” se abrirá. Nessa janela, o nome de cada estrutura “*mlr\_out1*”, “*mlr\_out2*”, “*mlr\_out3*” e “*mlr\_out4*” deve ser digitado em uma caixa, isto é, modelo 1, 2, 3 e 4, respectivamente, e então “*Continuar*”, ver Figura 18a. Após, outra janela “*função de desejabilidade*” se abrirá, nessa janela os usuários têm a possibilidade de definir uma condição de máximo, mínimo ou de alvo, como também atribuir diferentes pesos a cada resposta. Neste caso, desejamos maximizar o rendimento ( $y_1$ ) e o período de indução ( $y_2$ ), entretanto, a ANP na ocasião requeria um valor mínimo de 6 h, portanto, o valor mínimo para  $y_2$  foi ajustado para 6. Em contrapartida, o ponto de entupimento a frio ( $y_3$ ) e o custo de matéria prima por tonelada ( $y_4$ ) devem ser minimizados. Na ocasião o valor máximo de ponto de entupimento considerado pela ANP era de 5 °C, portanto, o valor máximo de  $y_3$  foi ajustado para 5, ver detalhes na Figura 18b. Para finalizar essa etapa, antes de apertar o botão “*Otimizar*”, os valores de “*Variáveis de mistura*” devem ser ajustados para três, que correspondem aos três tipos de matéria-prima avaliadas. Nota importante, a opção “*Variáveis de mistura*” só será usada para casos de planejamento de misturas.

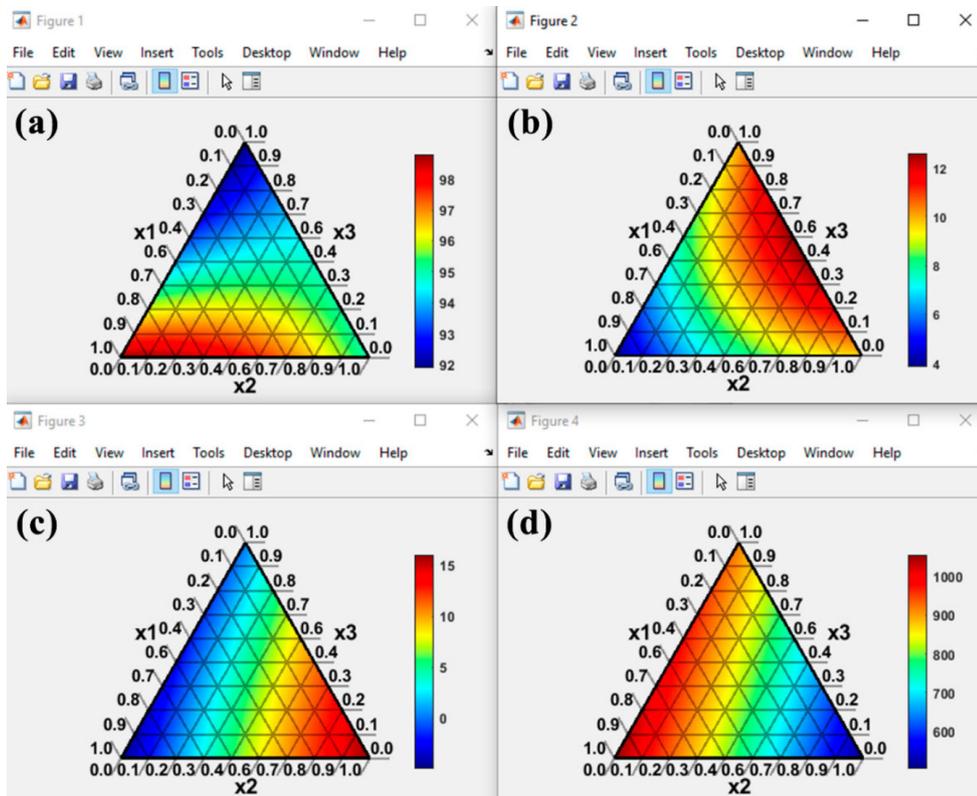


Figura 17. Diagrama de mistura para cada resposta: (a) rendimento da reação ( $y_1$ ), (b) período de indução ( $y_2$ ), (c) ponto de entupimento a frio ( $y_3$ ), (d) custo da matéria-prima por tonelada ( $y_4$ )

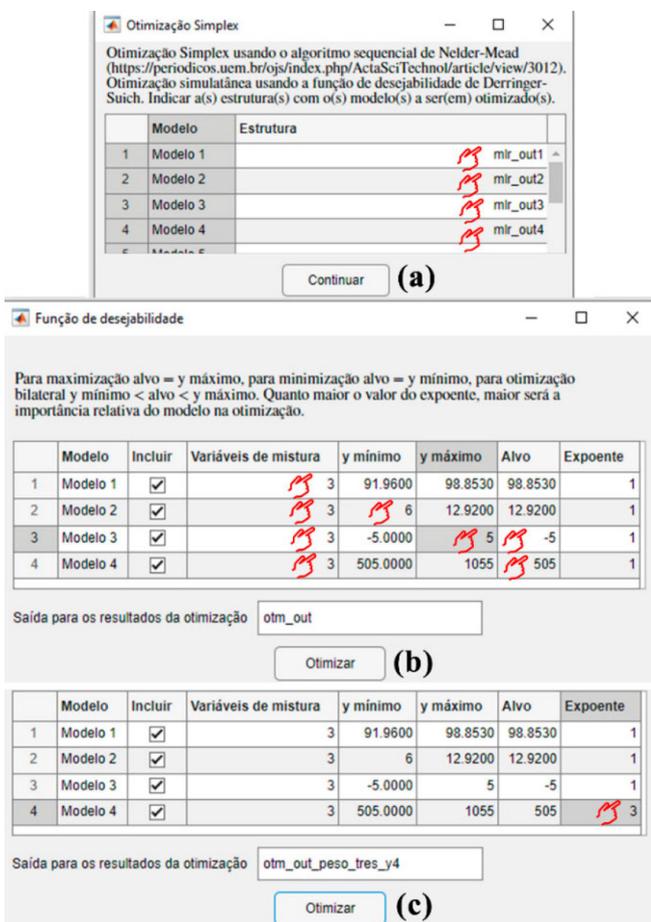


Figura 18. (a) Janela de otimização múltiplas respostas, (b) ajustando as desejeabilidades individuais e (c) atribuindo peso ao modelo de custo ( $y_4$ )

A formulação ótima que atende a todos esses requisitos exigidos corresponde a uma mistura nas proporções de 48,8% de óleo de soja, 19,4% de sebo bovino e 31,8% de gordura de frango, chegando a um rendimento de 95,5%, período de indução de 8,5 h, ponto de entupimento a frio de 1,9 °C e um custo de US\$ 904,4 por tonelada, ver estrutura de dados “otm\_out”. Considerando dessa vez que, almejamos reduzir o custo de produção ainda mais, podemos atribuir um maior peso ao “Modelo 4”, mudando o “Expoente” para três, ver Figura 18c. Essa nova formulação elaborada consiste em 44,4% de óleo de soja, 24,6% de sebo bovino e 31,0% de gordura de frango com rendimento de 95,6%, período de indução de 8,9 h, ponto de entupimento a frio de 3,1 °C e custo de US\$ 877,3 por tonelada, ver “otm\_out\_peso\_tres\_y4”.

## CONCLUSÃO

Neste tutorial demonstramos exemplos de uso do aplicativo GAMMA-GUI baseado em uma interface gráfica e amigável de planejamento e otimização de experimentos desenvolvido na graphical user interface (GUI) e implementado usando programação orientada a objetos *object-oriented programming* (OOP) do software Matlab®.<sup>11</sup> Diferentes aplicações práticas foram retiradas da literatura e reproduzidas sem a necessidade de conhecimentos prévios em linguagem de programação. Simultaneamente os exemplos foram reproduzidos no software Statística v.13, StatSoft®<sup>30</sup> para comparação dos resultados.

O aplicativo se encontra em português e todos os conjuntos de dados de demonstração juntamente com o aplicativo estão disponíveis gratuitamente para download no GitHub. O aplicativo

conta ainda com um canal no YouTube com uma série de vídeos tutoriais. O GAMMA-GUI está em constante desenvolvimento, quaisquer sugestões de melhorias ou novas ideias são incentivadas e bem-vindas e podem ser encaminhadas aos e-mails dos autores correspondentes.

Dando segmento aos nossos tutoriais com o aplicativo GAMMA-GUI, o próximo tutorial será direcionado a demonstrações utilizando a interface de métodos de reconhecimento de padrões não supervisionados por Análise de Componentes Principais (PCA), Análise de Agrupamento Hierárquico (HCA), Mapas Auto-organizáveis (SOM), Agrupamento  $k$ -means ( $k$ -means) e Análise por Dimensões Comuns (ComDim).

## AGRADECIMENTOS

Os autores agradecem aos usuários do aplicativo por todas as contribuições, sugestões de melhorias e ideias de implementação fornecidas. Agradecemos também aos órgãos de fomento à pesquisa: CNPq (processo 312595/2021-2) e CAPES.

## REFERÊNCIAS

- Bruns, R. E.; Scarminio, I. S.; de Barros Neto, B.; *Statistical Design-Chemometrics*, 1<sup>st</sup> ed.; Elsevier: Netherlands, 2006.
- Montgomery, D. C.; *Design and Analysis of Experiments*, 8<sup>th</sup> ed.; John Wiley: Danvers, 2012.
- Ferreira, S. L. C.; *Introdução às Técnicas de Planejamento de Experimentos*, 1<sup>a</sup> ed.; Vento Leste: Salvador, 2015.
- Rodrigues, M. I.; Iemma, A. F.; *Planejamento de Experimentos e Otimização de Processos*, 3<sup>a</sup> ed.; Cárita: Campinas, 2014.
- Barros Neto, B.; Scarminio, I. S.; Bruns, R. E.; *Como Fazer Experimentos: Pesquisa e Desenvolvimento na Ciência e na Indústria*, 4<sup>a</sup> ed.; Bookman: Porto Alegre, 2010.
- Breitkreitz, M. C.; de Souza, A. M.; Poppi, R. J.; *Quim. Nova* **2014**, *37*, 564. [Crossref]
- de Souza, A. M.; Poppi, R. J.; *Quim. Nova* **2012**, *35*, 223. [Crossref]
- Lyra, W. S.; da Silva, E. C.; de Araújo, M. C. U.; Frago, W. D.; Veras, G.; *Quim. Nova* **2010**, *33*, 1594. [Crossref]
- de Souza, A. M.; Breitkreitz, M. C.; Filgueiras, P. R.; Rohwedder, J. J. R.; Poppi, R. J.; *Quim. Nova* **2013**, *36*, 1057. [Crossref]
- Da-Col, J. A.; Dantas, W.; Poppi, R.; *Quim. Nova* **2017**, *41*, 345. [Crossref]
- Matlab®, R2023a; The MathWorks Inc., Natick, MA, USA, 2023.
- The MathWorks Inc.; <https://www.mathworks.com/products/matlab/app-designer.html>, acessado em Dezembro 2023.
- Galvan, D.; de Aguiar, L. M.; Bona, E.; Marini, F.; Killner, M. H. M.; *Anal. Chim. Acta* **2023**, *1273*, 341495. [Crossref]
- Bezerra, M. A.; Santelli, R. E.; Oliveira, E. P.; Villar, L. S.; Escalera, L. A.; *Talanta* **2008**, *76*, 965. [Crossref]
- Azcarate, S. M.; Pinto, L.; Goicoechea, H. C.; *J. Chemom.* **2020**, *34*, 1. [Crossref]
- Ferreira, S. L. C.; dos Santos, W. N. L.; Quintella, C. M.; de Barros Neto, B.; Bosque-Sendra, J. M.; *Talanta* **2004**, *63*, 1061. [Crossref]
- Cornell, J. A.; *A Primer on Experiments with Mixtures*, 1<sup>st</sup> ed.; Wiley: New Jersey, 2011.
- Galvan, D.; Effting, L.; Cremasco, H.; Conte Junior, C. A.; *Foods* **2021**, *10*, 1941. [Crossref]
- Cini, J. R. M.; Borsato, D.; Guedes, C. L. B.; Silva, H. C.; Coppo, R. L.; *Quim. Nova* **2013**, *36*, 79. [Crossref]
- Tormena, M. M. L.; de Medeiros, L. T.; de Lima, P. C.; Possebon, G.; Fuchs, R. H. B.; Bona, E.; *J. Sci. Food Agric.* **2017**, *97*, 3559. [Crossref]
- Derringer, G.; Suich, R.; *Journal of Quality Technology* **1980**, *12*, 214. [Crossref]

22. Santos, J. S.; Teixeira, L. S. G.; Araújo, R. G. O.; Fernandes, A. P.; Korn, M. G. A.; Ferreira, S. L. C.; *Microchem. J.* **2011**, *97*, 113. [Crossref]
23. Spinosa, W. A.; dos Santos Júnior, V.; Galvan, D.; Fiorio, J. L.; Castro Gomez, R. J. H.; *Acta Sci., Technol.* **2016**, *38*, 13. [Crossref]
24. Waszczynskyj, N.; Rao, C. S.; da Silva, R. S. F.; *Cereal Chem.* **1981**, *58*, 264. [Link] acessado em Dezembro 2023
25. Bona, E.; Borsato, D.; Sérgio, R.; Herrera, P.; *Acta Sci.* **2000**, *22*, 1201. [Link] acessado em Dezembro 2023
26. Dima, A. D.; Pârvulescu, O. C.; Mateescu, C.; Dobre, T.; *Biomass Bioenergy* **2020**, *138*, 105602. [Crossref]
27. Korn, M. D. G. A.; dos Santos, W. P. C.; Korn, M.; Ferreira, S. L. C.; *Talanta* **2005**, *65*, 710. [Crossref]
28. Galvan, D.; Orives, J. R.; Coppo, R. L.; Rodrigues, C. H. F.; Spacino, K. R.; Pinto, J. P.; Borsato, D.; *Quim. Nova* **2014**, *37*, 244. [Crossref]
29. Orives, J. R.; Galvan, D.; Coppo, R. L.; Rodrigues, C. H. F.; Angilelli, K. G.; Borsato, D.; *Energy Convers. Manage.* **2014**, *79*, 398. [Crossref]
30. *Statistica*, v.13; StatSoft Inc., Tulsa, OK, USA, 2018.