



UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E TECNOLÓGICAS



EDENIR RODRIGUES PEREIRA FILHO

**SINERGISMO ENTRE QUÍMICA E
MATEMÁTICA: ENSINO DE
PLANEJAMENTO FATORIAL NOS CURSOS
DE GRADUAÇÃO E PÓS-GRADUAÇÃO EM
QUÍMICA DA UFSCAR**

SÃO CARLOS
ABRIL DE 2022

Edenir Rodrigues Pereira Filho

Sinergismo entre Química e Matemática: Ensino de planejamento fatorial nos cursos de graduação e pós-graduação em Química da UFSCar

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Ensino de Ciências Exatas (PPGECE), da Universidade Federal de São Carlos, como parte dos requisitos para obtenção do título de Mestre em Matemática, sob orientação do Professor Doutor Tomas Edson Barros.

Universidade Federal de São Carlos
Centro de Ciências Exatas e Tecnológicas

Orientador: Tomas Edson Barros

São Carlos
Abril de 2022

Pereira Filho, Edenir Rodrigues

Sinergismo entre Química e Matemática: Ensino de planejamento fatorial nos cursos de graduação e pós-graduação em Química da UFSCar / Edenir Rodrigues Pereira Filho -- 2022.

104f.

Dissertação (Mestrado) - Universidade Federal de São Carlos, campus São Carlos, São Carlos

Orientador (a): Tomas Edson Barros

Banca Examinadora: Daniel Cordeiro de Moraes Filho, Paulo Antonio Silvani Caetano

Bibliografia

1. Planejamento de Misturas. 2. Octave. 3. Pontos Críticos. I. Pereira Filho, Edenir Rodrigues. II. Título.

Ficha catalográfica desenvolvida pela Secretaria Geral de Informática (SIn)

DADOS FORNECIDOS PELO AUTOR

Bibliotecário responsável: Ronildo Santos Prado - CRB/8 7325



UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Ciências Exatas e de Tecnologia
Programa de Pós-Graduação em Ensino de Ciências Exatas

Folha de Aprovação

Defesa de Dissertação de Mestrado do candidato Edenír Rodrigues Pereira Filho, realizada em 29/04/2022.

Comissão Julgadora:

Prof. Dr. Tomas Edson Barros (UFSCar)

Prof. Dr. Daniel Cordeiro de Moraes Filho (UFCG)

Prof. Dr. Paulo Antonio Silvani Caetano (UFSCar)

O Relatório de Defesa assinado pelos membros da Comissão Julgadora encontra-se arquivado junto ao Programa de Pós-Graduação em Ensino de Ciências Exatas.

Esta dissertação de mestrado é dedicada à memória do meu colega, mentor e professor Ronei Jesus Poppi (1961-2020) que me inspirou ao longo dos anos.

Agradecimentos

Agradeço à minha esposa Fabíola Manhas Verbi Pereira pelo amor, apoio, paciência e incentivo durante a realização dessa dissertação. Além disso, enfrentamos juntos os desafios, problemas e as angústias que a pandemia causou.

Agradeço aos professores do Departamento de Matemática da UFSCar pelos ensinamentos oferecidos durante a realização das disciplinas: Grazielle Feliciani Barbosa, José Antonio Salvador, Paulo Antonio Silvani Caetano (incentivador), Pedro Luiz Aparecido Malagutti, Renato José de Moura, Tomas Edson Barros, Luciene Nogueira Bertoncello, Wladimir Seixas (instruções no Latex) e Ivo Machado da Costa.

Agradeço ao Prof. Dr. Tomas Edson Barros pela orientação e discussões sobre matemática e ideias desta dissertação.

Agradeço à Dra. Jeyne Pricylla Castro Castilho pela imprescindível ajuda e discussões na preparação dos algoritmos `mix_target` (para o Matlab®) e `mix_target2` (para o Octave). Além disso, contribuiu com o algoritmo `mydes2` para cálculo de desejabilidade.

Agradeço ao Felipe Ferri Hilário pela preparação do *Template Misturas* no Excel®.

Agradeço aos colegas do Profmat e PPGECE (anos de 2019, 2020 e início de 2021) pelo convívio e troca de experiências sobre o ensino de matemática.

Agradeço à Kelly Schiabelli pela ajuda e prestatividade junto às secretarias do Profmat e PPGECE.

Agradeço aos mais de 200 alunos (graduação e pós-graduação) que acompanharam as disciplinas ministradas em um momento tão conturbado e confuso da nossa história recente.

Agradeço à Dra. Amanda dos Santos Augusto pela correção do *abstract* e aos colegas professores José Mario de Aquino (discussões sobre o ensino remoto) e Alejandro Lopez Castillo (discussões sobre cálculo numérico).

Agradeço à UFSCar e aos Departamentos de Matemática e Química do campus de São Carlos pelas oportunidades oferecidas ao longo dos anos de docência.

Agradeço à banca examinadora pelas contribuições durante a defesa desta dissertação de mestrado profissional.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

*"Se você só fizer o que sabe, nunca será nada além do que já é."
Mestre Shifu*



Meus primeiros cálculos aos 5 ou 6 anos na cidade de Imperatriz no Estado do Maranhão
(início da década de 1980).

Resumo

Sinergismo entre Química e Matemática: Ensino de planejamento fatorial nos cursos de graduação e pós-graduação em Química da UFSCar

Esta dissertação mostra o sinergismo entre duas ciências: a química e a matemática. Na primeira parte da dissertação foi apresentada a aplicação do planejamento de experimentos (*Design of experiments*, DoE) na otimização de misturas. Como exemplos, foi discutido um planejamento de mistura para a otimização de um processo de gravação de semicondutores e foram avaliadas misturas compostas por três ácidos. Duas estratégias foram comparadas e discutidas: (i) o uso de derivada parcial e (ii) a implementação de um algoritmo para determinar a melhor condição experimental. Durante a aplicação da derivada parcial vários aspectos precisam ser avaliados para identificar se o ponto ótimo está localizado no vértice, aresta ou parte interna do triângulo. Por outro lado, o algoritmo proposto empregou aproximações sucessivas evitando a aplicação de várias condições de exceção como demonstrado na derivada parcial. O algoritmo proposto cria uma população inicial de variáveis (v_1, v_2, \dots, v_n) e testa todas elas para identificar a melhor condição. Na segunda parte da dissertação, foi realizada uma comparação entre alunos de graduação e pós-graduação do curso de química. Ambos os grupos de alunos ($n = 200$) tiveram aulas sobre DoE. Essa comparação foi realizada antes e durante a pandemia de Covid-19. Antes do período de pandemia não foi observada diferença entre os dois grupos de alunos. Durante o período de pandemia foram elaborados vídeos sobre diversos temas relacionados ao DoE e os mesmos foram apresentados aos alunos por meio da plataforma YouTube. Além disso, também foram realizados encontros síncronos gravados. O material gravado (≈ 20 horas) foi assistido por mais de 2500 horas. Dois parâmetros apresentaram diferenças estatisticamente significativas ($p - \text{valor} < 0,05$): aspectos sobre as aulas remotas (i) antes e (ii) durante a pandemia. Observou-se que os alunos de pós-graduação estão mais interessados em aulas remotas, do que os de graduação. Aproximadamente 60% dos alunos de pós-graduação estão abertos às estratégias de ensino híbrido em um período pós-pandemia. A experiência com aulas remotas permitiu atingir alunos de diversas partes do Brasil (17 estados), bem como de alguns países da América Latina (Argentina, Colômbia e Peru). Toda a experiência foi muito produtiva e gratificante, sendo possível apresentar conteúdos mais informativos aos alunos. Além disso, abre uma perspectiva para uma situação híbrida na fase pós-pandemia.

Palavras-chave: Planejamento de misturas, Octave, Pontos críticos, Gráfico de contorno, Modelos de regressão, Ensino não presencial, Planejamento fatorial.

Abstract

Chemistry and Mathematics' synergism: Teaching Design of Experiments in undergraduate and graduate courses of Chemistry at UFSCar

The main idea of this dissertation is to highlight the synergism of two sciences: chemistry and mathematics. The first part of the text is dedicated to a chemometric tool known as Design of Experiments (DoE) applied to the optimization of mixtures. The optimization of an etching process was discussed, and mixtures composed of three acids were evaluated for semiconductors surface treatment. Two strategies were assessed: (i) the use of partial derivative and (ii) the implementation of an algorithm to determine the optimal experimental condition. During the application of partial derivative, several aspects need to be analyzed to identify if the optimal point is located at the vertice, edge, or inner part of the mixture (triangle). On the other hand, the proposed algorithm employed successive approximations avoiding the application of several exception conditions as demonstrated in partial derivative strategies. The proposed algorithm creates an initial population of variables combination (v_1, v_2, \dots, v_n) and tests it in order to identify the best condition. The second part of the dissertation, it was performed a comparison between undergraduate and graduate students from chemistry courses. Both groups of students ($n = 200$) had classes about DoE, and two types of data were considered: (i) students' grades obtained during the evaluations and (ii) students' opinions about the course. This comparison was performed before and during the pandemic of Covid-19. During the pandemic period, it was prepared videos about several topics related to DoE and these videos were presented to the students using the YouTube platform. In addition, recorded synchronous meetings were also performed. The recorded material (≈ 20 hours) was watched for more than 2500 hours. Before the pandemic period, it was not observed any difference between both groups of students regarding the grades obtained. On the other hand, the students' opinion pointed differences ($p - value < 0.05$): aspects about the remote classes (i) before and (ii) during the pandemic. It was observed that the graduate students are more interested and open to the strategies of remote classes than the undergraduate. Approximately 60% of the graduate students are open to hybrid classes after the pandemic. The experience with remote classes permitted to reach students from different parts of Brazil (17 states) as well, as some countries from Latin America (Argentina, Colombia, and Peru). The whole experience was very productive and gratifying, being possible to present more informative content to the students. In addition, opens perspectives for a hybrid situation pos-pandemic phase.

Keywords: Mixture design, Octave, Critical points, Contour plot, Regression models, Remote teaching, Design of experiments.

Lista de ilustrações

Figura 1 – Superfície de resposta (a) e gráfico de contorno (b) do modelo apresentado na Função 3.5.	29
Figura 2 – Superfície de resposta (a) e gráfico de contorno (b) do modelo apresentado na Função 3.7, tendo as variáveis v_3 e v_4 fixadas em -0,74 e -1,78, respectivamente.	31
Figura 3 – Sistema ternário envolvendo a preparação de um suco. As imagens de cores branca, azul e amarela indicam as quantidades de açúcar, água e suco concentrado, respectivamente em cada mistura preparada.	33
Figura 4 – Representação dos experimentos executados (círculos) no planejamento fatorial de misturas mostrados na Tabela 1.	35
Figura 5 – Visualização dos coeficientes de regressão e os intervalos de confiança.	41
Figura 6 – Triângulo de vértices $A = (1, 0, 0)$, $B = (0, 1, 0)$ e $C = (0, 0, 1)$	42
Figura 7 – Visualização dos valores calculados do polinômio 3.24 no domínio $[0, 1]$ (a) e identificação dos valores de y_1 (b) e y_2 (c).	44
Figura 8 – Primeira etapa de funcionamento do algoritmo.	53
Figura 9 – Segunda etapa de funcionamento do algoritmo.	54
Figura 10 – Terceira etapa de funcionamento do algoritmo.	55
Figura 11 – Combinações dos valores das variáveis de 1 a 3 (círculos) para valor de <i>Precision</i> igual a 1000.	56
Figura 12 – Combinações dos valores das variáveis de 1 a 3 (círculos) para valor de <i>Precision</i> igual a 5000.	56
Figura 13 – Combinações dos valores das variáveis de 1 a 3 (círculos) para valor de <i>Precision</i> igual a 10000.	57
Figura 14 – Quarta etapa de funcionamento do algoritmo.	58
Figura 15 – Gráficos de contorno obtidos para os modelos hipotéticos das Funções 5.1 (a), 5.2 (b) e 5.3 (c). A região em vermelho representa a maior resposta obtida.	58
Figura 16 – Gráficos de contorno obtidos para os modelos hipotéticos das Funções 5.4 (a), 5.5 (b) e 5.6 (c). A região em vermelho representa a maior resposta obtida.	59
Figura 17 – Quinta etapa de funcionamento do algoritmo.	60
Figura 18 – Visualização dos níveis das variáveis 1 até 3 que apresentaram os cinco melhores resultados (<i>Precision</i> = 1000).	61
Figura 19 – Visualização dos níveis das variáveis 1 até 3 que apresentaram os cinco melhores resultados (<i>Precision</i> = 5000).	62

Figura 20 – Visualização dos níveis das variáveis 1 até 3 que apresentaram os cinco melhores resultados (<i>Precision</i> = 10000).	62
Figura 21 – Visualização das cinco melhores combinações de resultados quando a <i>Precision</i> for igual a 1000 (a), 5000 (b) e 10000 (c).	63
Figura 22 – Sexta etapa de funcionamento do algoritmo.	63
Figura 23 – Visualização dos parâmetros finais das 3 variáveis após o cálculo com valores de <i>Precision</i> de 1000 (a), 5000 (b) e 10000 (c).	64
Figura 24 – Combinações dos valores das variáveis de 1 a 3 (círculos) para valor de <i>Precision</i> igual a 1000 após o refinamento do cálculo.	65
Figura 25 – Ampliação da região ótima de trabalho após a repetição dos cálculos com valores de <i>Precision</i> de 1000 (a), 5000 (b) e 10000 (c).	65
Figura 26 – Gráficos de contorno (a) e de superfície (b) obtidos no <i>Template Misturas</i> do Excel®. A região em vermelho no gráfico de contorno (a) representa a maior resposta obtida.	66
Figura 27 – Notas obtidas pelos alunos de graduação na disciplina Introdução à Quimiometria antes (círculo interno, n = 16) e durante a pandemia (círculo externo, n = 68).	67
Figura 28 – Conceitos obtidos pelos alunos de pós-graduação na disciplina Quimiometria I antes (círculo interno, n = 161) e durante a pandemia (círculo externo, n = 154).	68
Figura 29 – Opinião dos alunos de pós-graduação sobre as questões 1 (a), 2 (b), 3 (c), 4 (d) e 6 (e) no período antes da pandemia (2014 e 2019).	70
Figura 30 – Opinião dos alunos de pós-graduação sobre as questões 1 (a), 2 (b), 3 (c), 4 (d) e 6 (e) no período antes (círculo interno, n = 63) e durante a pandemia (círculo externo, n = 130). Para a questão 6 (e) o número de respondentes para antes da pandemia foi de 40.	72
Figura 31 – Opinião dos alunos de graduação (círculo interno) e pós-graduação (círculo externo) sobre as questões 1 (a), 2 (b), 3 (c), 4 (d), 5 (e) e 6 (f).	74
Figura 32 – Opinião dos alunos de graduação (círculo interno) e pós-graduação (círculo externo) sobre as questões 7 (a), 8 (b) e 10 (c).	77
Figura 33 – Opinião dos alunos de graduação (círculo interno) e pós-graduação (círculo externo) sobre a questão 9 (a) e comparação entre as respostas (b).	78
Figura 34 – Opinião dos alunos de graduação (círculo interno) e pós-graduação (círculo externo) sobre a questão 11.	79
Figura 35 – Comparação entre as respostas dadas pelos alunos de graduação e pós-graduação para a questão 11.	79

Lista de tabelas

Tabela 1	– Experimentos executados para a otimização das condições de gravações na superfície de semicondutores.	34
Tabela 2	– Tabela Anova do modelo proposto para o planejamento de misturas. . .	39
Tabela 3	– Questões apresentadas aos alunos para avaliação das disciplinas de graduação e pós-graduação. Os detalhes sobre as opções de respostas disponíveis para os alunos podem ser visualizados no Apêndice B. . . .	48
Tabela 4	– Conteúdos apresentados nas disciplinas de graduação (Partes 1 a 9) e pós-graduação (Todas as partes).	51
Tabela 5	– Intervalos das variáveis 1 a 3 quando os valores de <i>Precision</i> foram 5000 e 10000	61
Tabela 6	– Características dos resultados após repetição dos cálculos com intervalos mais estreitos.	64
Tabela 7	– Intervalos das variáveis 1 a 3 quando os valores de <i>Precision</i> foram 1000, 5000 e 10000 após a repetição dos cálculos.	64
Tabela 8	– Condições ótimas de trabalho (<i>B</i>) e resposta teórica obtida (<i>Response</i>) para os três valores de <i>Precision</i> empregados	66
Tabela 9	– Alguns parâmetros estatísticos das notas e conceitos obtidos pelos alunos de graduação e pós-graduação antes e durante a pandemia. . . .	69

Lista de abreviaturas e siglas

Anova	<i>Analysis of variance.</i>
Capex	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior.
CCD	<i>Central composite design.</i>
DoE	<i>Design of experiments.</i>
EaD	Ensino à distância.
ENPE	Ensino não presencial emergencial.
GAIA	Grupo de análise instrumental aplicada.
MQEP	Média quadrática do erro puro.
MQFaj	Média quadrática da falta de ajuste.
MQR	Média quadrática da regressão.
MQres	Média quadrática do resíduo.
MQT	Média quadrática total.
OBS	<i>Open Broadcaster Software.</i>
PB	<i>Plackett-Burman.</i>
PC	<i>Personal computer.</i>
PCA	<i>Principal component analysis.</i>
PPGQ	Programa de pós-graduação em Química.
Profmat	Mestrado Profissional em Matemática em Rede Nacional.
Profqui	Programa de Mestrado Profissional em Química em Rede Nacional.
R^2	Coefficiente de determinação.
RSM	<i>Response surface methodology.</i>
SAEd	Serviço de atendimento do Edénir.
SBM	Sociedade Brasileira de Matemática.

SBQ	Sociedade Brasileira de Química.
SQEP	Soma quadrática do erro puro.
SQFaj	Soma quadrática da falta de ajuste.
SQR	Soma quadrática da regressão.
SQres	Soma quadrática do resíduo.
SQT	Soma quadrática total.
TCLE	Termo de Consentimento Livre e Esclarecido.
TVI	Teorema do Valor Intermediário.
UFSCar	Universidade Federal de São Carlos.
t	Variável tempo de pré-concentração.
V	Variável velocidade de varredura.

Sumário

Lista de ilustrações	13
Lista de tabelas	15
PREFÁCIO	21
1 INTRODUÇÃO	23
2 OBJETIVOS E JUSTIFICATIVAS	25
2.1 Objetivos	25
2.2 Justificativas	25
3 REVISÃO BIBLIOGRÁFICA	27
3.1 Proposição de Modelos Empíricos	27
3.2 Identificação dos Pontos Críticos	28
3.3 Proposição de um modelo de regressão em um Planejamento de Misturas	32
3.4 Identificação dos Pontos Críticos em um Planejamento de Misturas	41
4 MATERIAIS E MÉTODOS	47
4.1 Programas Computacionais	47
4.2 O Ensino de Planejamento Fatorial na Pandemia	47
4.3 Coleta de dados e avaliação pelos estudantes	48
4.4 Conteúdos apresentados e características dos vídeos	49
5 RESULTADOS E DISCUSSÃO, ALGORITMO MIX_TARGET2	53
5.1 Preparação e proposição do algoritmo mix_target2	53
5.2 Situações iniciais para o algoritmo mix_target2	57
5.3 Funções mais complexas para o algoritmo mix_target2	60
6 RESULTADOS E DISCUSSÃO, DISCIPLINAS SOBRE PLANEJAMENTO FATORIAL	67
6.1 Desempenho dos alunos nas disciplinas sobre planejamento fatorial	67
6.2 Opinião dos alunos sobre o ensino de planejamento fatorial	69
6.2.1 Cenário antes da pandemia	69
6.2.2 Comparação dos cenários: antes e durante a pandemia	71
6.2.3 Avaliação das percepções dos alunos durante a pandemia	71

6.2.4	Avaliação das percepções dos alunos durante a pandemia: questões que apresentaram diferenças significativas	75
7	CONCLUSÕES, PERSPECTIVAS FUTURAS E CONSIDERAÇÕES FINAIS	81
7.1	Conclusões	81
7.2	Perspectivas Futuras	82
7.3	Considerações Finais	82
	REFERÊNCIAS	85
	APÊNDICE A – INFORMAÇÕES GERADAS PELO ALGORITMO REGRESSION2 NA PROPOSIÇÃO DE MODELOS DE REGRESSÃO	89
	APÊNDICE B – FORMULÁRIO DE AVALIAÇÃO DAS DISCIPLINAS MINISTRADAS	91
	APÊNDICE C – FORMULÁRIO DO SAED	97
	APÊNDICE D – ALGORITMO MIX_TARGET2	99
	APÊNDICE E – ARTIGO CIENTÍFICO PUBLICADO	105

Prefácio

O ensino de ferramentas matemáticas e computacionais voltadas para a aplicação do planejamento fatorial não é uma tarefa trivial. Na área de química, em especial, é necessário que o usuário tenha conhecimentos de matemática, estatística e lógica para atingir os objetivos necessários na otimização de um dado problema ou sistema. A química como uma ciência experimental se beneficia ao lidar com tais ferramentas de otimização. A proposta inicial desta dissertação de mestrado profissional ocorreu em 2018 e tratava da interação entre matemática e química.



Com o surgimento da pandemia de Covid-19 no início de 2020, os desafios se tornaram maiores para o ensino de quaisquer conteúdos, tanto nos cursos de graduação, quanto nos de pós-graduação. Entretanto, oportunidades também foram criadas e, de certa forma, concretizadas. Assim, a proposta definitiva desta dissertação é mostrar a aplicação de diversas ferramentas computacionais, tais como o programa livre Octave, proposição de rotinas computacionais, o programa Microsoft Excel®, vídeos gravados e disponibilizados no YouTube e também encontros síncronos no ensino de planejamento fatorial. No decorrer da parte experimental serão mostrados exemplos relacionados com o planejamento de misturas, onde foi proposto um algoritmo para calcular a região ótima de trabalho (ponto crítico).

Ao longo dos anos de 2020 e 2021 foram aplicados questionários aos alunos de graduação e pós-graduação que assistiram às aulas e o objetivo foi entender e avaliar o impacto das estratégias propostas no ensino remoto. Além disso, algumas perguntas do formulário foram aplicadas no momento pré-pandemia (anos de 2014 e 2019). A dissertação é dividida em diferentes partes que descrevem a proposta do projeto de mestrado profissional, bem como a aplicação de planejamento fatorial e o ensino desta ferramenta fundamental para diversas áreas do conhecimento.



1 Introdução

A quimiometria é uma ciência que utiliza conhecimentos de matemática e estatística em dados químicos [1]. As suas aplicações são amplas e começaram a serem popularizadas na década de 1970 com o surgimento da microeletrônica e dos primeiros computadores acessíveis para empresas, universidades e público em geral [2]. Classicamente, a quimiometria pode ser dividida em técnicas de reconhecimento de padrões, proposição de modelos de regressão multivariada e planejamento fatorial.

O planejamento fatorial (*Design of experiments*, DoE) é uma ferramenta imprescindível para a otimização de diversos problemas nas áreas de química, física e engenharia [3]. A literatura atual conta com diversos tutoriais [4], apresentação de programas computacionais pagos e livres [5] e também algoritmos que permitem um melhor aprendizado e utilização desta ferramenta. Um ponto a ser destacado é o fato de que o ensino de planejamento fatorial é visto como um tópico secundário em diversos cursos de química [6]. Entretanto, em uma ciência predominantemente experimental (como é o caso da Química) tais conhecimentos trazem benefícios extraordinários para reduzir e racionalizar a quantidade de experimentos, além de possibilitar melhor aproveitamento das informações que são obtidas [7]. O uso do planejamento fatorial pode contribuir também com a química verde [8], que possui como um dos princípios a redução do número de experimentos e conseqüente menor geração de resíduos químicos.

O planejamento fatorial pode ser dividido em diferentes partes [9]: (i) planejamento fatorial completo; (ii) planejamento fatorial fracionário; (iii) planejamentos saturados (*Plackett-Burman*, PB); (iv) proposição de modelos empíricos matemáticos com posterior uso da metodologia de superfície de respostas (*Response surface methodology*, RSM) e (v) planejamento de misturas. Este último é bastante útil quando há uma relação constante entre as variáveis, onde o aumento de uma delas acarretará na diminuição das demais. No planejamento fatorial de misturas com três variáveis (v_1 , v_2 e v_3), por exemplo, podem ser propostos modelos cúbicos especiais [3, 10] onde é possível calcular 7 coeficientes (b_i), como mostra a Função 1.1:

$$f(v_1, v_2, v_3) = b_1v_1 + b_2v_2 + b_3v_3 + b_{12}v_1v_2 + b_{13}v_1v_3 + b_{23}v_2v_3 + b_{123}v_1v_2v_3 \quad (1.1)$$

No caso específico de um planejamento fatorial de misturas com 3 variáveis (v_1 até v_3), o modelo proposto será representado por um triângulo. Os vértices do triângulo representam os componentes puros, as arestas compreendem misturas binárias e o interior do triângulo é formado por misturas dos três componentes (misturas ternárias) [11]. A

identificação da região ótima de trabalho (resposta desejada ou melhor resposta) de um planejamento de misturas pode representar esforços consideráveis, principalmente para o caso onde o ponto crítico está no interior do triângulo. No caso de 4 variáveis (v_1 até v_4) a mistura será representada por um tetraedro e a região ótima de trabalho poderá estar em um dos vértices, ou em uma das arestas (no caso de dois componentes serem importantes), ou em uma das faces (no caso de três componentes serem importantes) ou no interior para o caso do ponto ótimo possuir a contribuição dos quatro componentes.

Desta forma, esta dissertação de mestrado profissional visa preencher algumas lacunas referentes ao melhor entendimento do planejamento de misturas pelos alunos de graduação e pós-graduação. Além disso, foram propostas ferramentas que permitem a identificação da região que apresenta a resposta desejada. Os conhecimentos e estratégias de ensino gerados foram propostos e aplicados durante os anos de 2020 até 2022 visando um melhor entendimento dos conteúdos pelos alunos. Tais estratégias estão em contínuo aprimoramento e sendo utilizadas nos semestres subsequentes.

2 Objetivos e Justificativas

2.1 Objetivos

Os objetivos desta dissertação de mestrado profissional estão relacionados com a apresentação de alternativas para a identificação da região ótima de trabalho de um planejamento fatorial de misturas do tipo cúbico especial. Entretanto, as alternativas geradas podem ser adaptadas para outros modelos, como o linear, quadrático e o cúbico completo [3]. Além disso, serão mostradas e discutidas as percepções dos alunos de graduação dos cursos de química da UFSCar e pós-graduação sobre o ensino de planejamento fatorial em anos bastante atípicos como os de 2020 e 2021. Os alunos de pós-graduação envolvidos são majoritariamente do Programa de pós-graduação em Química da Universidade Federal de São Carlos (PPGQ/UFSCar) e também, de outros estados da federação e países da América Latina.

Os seguintes pontos estão relacionados com os objetivos específicos: (i) melhor entendimento das ferramentas matemáticas que podem ser empregadas para o ensino de planejamento fatorial de misturas; (ii) utilização de programa computacional livre com a proposição de algoritmo específico visando a identificação da melhor região de trabalho e (iii) entender as percepções dos alunos de graduação e pós-graduação sobre o ensino de tais ferramentas e identificar similaridades e pontos divergentes na visão dos mesmos.

2.2 Justificativas

Mesmo presente em diversos livros (textos ou *e-books*) de química e química analítica [12, 13, 14, 15], o ensino de estratégias de planejamento fatorial em cursos de graduação e pós-graduação em química é, em muitas situações, colocado como um conteúdo secundário e, quando ministrados, não é dada a devida importância. Em diversas universidades do Brasil e do exterior o planejamento fatorial é tido como uma disciplina optativa e ministrada por professores externos aos da área de química. Em vários casos, os conteúdos de planejamento fatorial são dados *en passant*. Desta forma, na maioria das vezes o diálogo e interpretação correta dos dados, sob o ponto de vista químico, ficam evidentemente comprometidos e incompletos.

3 Revisão Bibliográfica

3.1 Proposição de Modelos Empíricos

Na proposição de modelos empíricos, o experimentador tem como objetivo identificar uma relação causal entre variáveis que podem ser controladas e as respostas que são monitoradas [16]. Na química, por exemplo, essas variáveis que podem ser controladas são configuradas por parâmetros quantitativos, tais como concentrações de reagentes, massas (de amostra e/ou reagentes), temperatura, tempo, vazões, pressões, entre outros. Além disso, podem ser testadas também variáveis qualitativas, onde não é possível atribuir um valor numérico. Entre elas estão, por exemplo, tipos de catalisador, de reator e diferentes operadores [17]. Após a execução dos experimentos é possível obter um modelo matemático por meio de Mínimos Quadrados [18], onde o objetivo é minimizar a distância vertical (erro) entre o valor obtido experimentalmente (y) e aquele previsto pelo modelo (\hat{y}). O método dos Mínimos Quadrados pode ser utilizado para calcular modelos (proposição de polinômios) de quaisquer graus (m), desde que $m < n$, onde n é o número de experimentos distintos executados (x, y) [19].

Na área de química analítica, por exemplo, em diversas situações temos que propor uma curva de calibração para a correta utilização de um método de determinação. Essa curva de calibração é formada por pares de pontos: concentração do analito (variável independente) e sinal analítico (variável dependente). Tais pares de pontos fornecem uma função $f(x)$ que está em um sub-espço E . Assim, no método dos Mínimos Quadrados o que se deseja é estabelecer uma função $F(x)$ de E' de tal forma que a distância de $f(x)$ a E' seja mínima[19]. Desta forma, teremos:

$$\text{dist}(f(x), F(x)) = \text{mínima} \quad (3.1)$$

Entretanto, a distância entre $f(x)$ e $F(x)$ é dada por:

$$\text{dist}(f(x), F(x)) = \|f(x) - F(x)\| = [\sqrt{(f - F, f - F)}] \quad (3.2)$$

Onde os símbolos $\| \ \|$ representam a norma (tamanho do vetor). Assim, a ideia central é:

$$Q = \|f(x) - F(x)\|^2 = \text{mínima} \quad (3.3)$$

A literatura específica [20] apresenta uma vasta descrição sobre os cálculos envolvidos e os coeficientes (b_i) podem ser obtidos por meio da Equação 3.4:

$$b_i = (X^t X)^{-1} X^t y \quad (3.4)$$

Na Equação 3.4, a matriz \mathbf{X} contém os parâmetros experimentais dos coeficientes que serão calculados [20]. A dimensão da matriz \mathbf{X} é n (número de experimentos) por m (número de coeficientes que serão calculados). No exemplo mostrado na Função 1.1, a matriz \mathbf{X} é $n \times 7$. Já o vetor y mostra a resposta monitorada e deve possuir dimensão $n \times 1$.

Na avaliação da significância estatística do modelo calculado são utilizadas diversas métricas, tais como o coeficiente de determinação (R^2). Esse parâmetro varia entre 0 e 1 e o desejável são valores mais próximos de 1 possível [9]. Além disso, devem ser também comparadas as médias quadráticas (variâncias) da regressão e do resíduo, e as médias quadráticas da falta de ajuste e do erro puro [21]. Tais parâmetros são avaliados por meio da tabela de Análise de Variância (*Analysis of variance*, Anova) [3, 5, 21]. A literatura científica sobre esses assuntos é bastante rica e diversas delas são voltadas para exemplos da área de química. Assim, a intenção desta parte da dissertação é mostrar de forma breve os parâmetros mais importantes na proposição de um modelo empírico. Além disso, na plataforma do YouTube podem ser encontrados diversos canais com materiais didáticos sobre o assunto, além de rotinas computacionais livres, exemplos didáticos e tutoriais: <https://www.youtube.com/c/EdenirPereiraFilho>.

Um ponto interessante a ser considerado é a identificação da região ótima ou dos pontos críticos. Em um planejamento fatorial completo composto central [16, 21], derivadas parciais juntamente com a regra da potência [22] são excelentes alternativas para calcular os pontos críticos. Além disso, em muitos casos, uma simples inspeção e avaliação visual dos dados e da superfície de resposta obtida pode trazer de forma objetiva e clara a identificação de tais pontos. Nesta parte em diante da dissertação irei me referir ao ponto crítico como a região experimental onde é possível obter a maior resposta de um planejamento fatorial. Isso se deve ao fato de que na maioria dos experimentos executados, os experimentadores estão buscando maximizar uma resposta, tais como rendimento de uma reação, sinal analítico de um analito, entre outras respostas. Caso o contrário ocorra, ou seja, identificação da condição que reflete a menor resposta (redução de custo ou menor limite de detecção, por exemplo), é possível codificar os dados antes de efetuar a proposição do modelo. Na próxima seção serão mostrados e discutidos alguns exemplos.

3.2 Identificação dos Pontos Críticos

Em um planejamento fatorial composto central [9, 21] com duas variáveis, podemos codificar as mesmas entre $-\sqrt{2}$ e $\sqrt{2}$. No trabalho publicado por Vicentini *et al.* [23],

os autores estavam interessados em otimizar a determinação de Ag(I) por um método voltamétrico [12]. As variáveis investigadas foram o tempo de pré-concentração (t , variável 1) e a velocidade de varredura (V , variável 2) e o objetivo era obter sinais altos de Ag(I) (Resposta monitorada, y). Após a execução dos experimentos e validação dos dados, o modelo matemático obtido pode ser visualizado na Função 3.5:

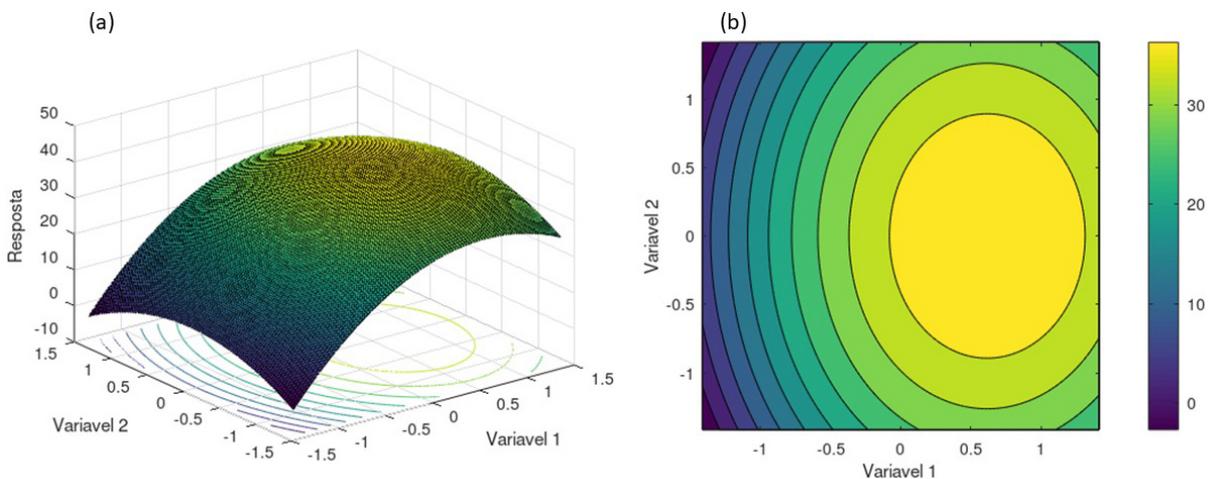
$$\text{Sinal de Ag(I)} = f(t, V) = 32,70 + 9,92t - 8,03t^2 - 4,91V^2 \quad (3.5)$$

A Função 3.5 mostra apenas os coeficientes válidos estatisticamente ao nível de confiança de 95% ($p - \text{valor} < 0,05$). O coeficiente linear ($b_0 = 32,70$) mostra a resposta prevista (\hat{y}) quando as duas variáveis estão no nível codificado 0 (ponto central). É importante notar que o coeficiente linear V e o de interação tV não foram significativos. A Figura 1 mostra a superfície de resposta (Figura 1a) e o gráfico de contorno (Figura 1b) obtidos.

Com o auxílio da Figura 1 é possível visualizar aproximadamente a localização do ponto de máximo do modelo matemático. Analisando essa figura, percebemos que o máximo da função f (Função 3.5) ocorre no interior do domínio $[-\sqrt{2}, \sqrt{2}] \times [-\sqrt{2}, \sqrt{2}]$. Assim, com o auxílio de derivadas parciais [24] é possível identificar com precisão tal ponto:

$$\frac{\partial f}{\partial t}(t, V) = 9,92 - 16,06t \quad \text{e} \quad \frac{\partial f}{\partial V}(t, V) = -9,82V$$

Figura 1 – Superfície de resposta (a) e gráfico de contorno (b) do modelo apresentado na Função 3.5.



Fonte: Compilação do autor.

$$\frac{\partial f}{\partial t}(t, V) = \frac{\partial f}{\partial V}(t, V) = 0 \Leftrightarrow t = \frac{9,92}{16,06} \approx 0,62 \text{ e } V = 0$$

Assim, o ponto $(0,62,0)$ do plano tV é aquele que oferece o maior sinal de Ag(I). O sinal máximo de Ag(I) previsto pelo modelo (\hat{y}) será:

$$\text{Sinal de Ag(I)} = f(t, V) = 32,70 + 9,92 \times 0,62 - 8,03 \times 0,62^2 - 4,91 \times 0^2 \approx 35,8 \quad (3.6)$$

Para casos mais complexos quando temos 3, 4 ou mais variáveis podemos utilizar a Regra de Cramer [25]. No trabalho publicado por Najafpou *et al.* [26], por exemplo, os autores otimizaram um sistema químico para remoção de elementos químicos potencialmente tóxicos de uma bebida denominada Kombucha. Os autores estudaram a influência de quatro variáveis (v_1 a v_4) e o objetivo foi identificar uma condição que permitisse uma alta porcentagem dos íons Hg(II), As(III), Pb(II), Cd(II) e Cr(VI). Para o caso específico da remoção de Hg (II), o modelo matemático obtido é mostrado na Função 3.7:

$$\begin{aligned} Hg(II) = g(v_1, v_2, v_3, v_4) = & 89,4 + 1,1v_1 - 3,6v_2 - 1,1v_3 - 4,4v_4 \\ & - 2,4v_1^2 - 3,6v_2^2 - 1,7v_3^2 - 1,5v_4^2 \\ & + 3,7v_1v_2 - 1,7v_1v_3 - 2,1v_2v_4 \end{aligned} \quad (3.7)$$

Na Função 3.7 são mostrados também apenas os coeficientes válidos ao nível de confiança de 95% ($p - \text{valor} < 0,05$). Após a utilização das derivadas parciais é possível obter 4 equações e 4 incógnitas (v_1 a v_4) que podem ser mostradas pelo sistema:

$$\begin{aligned} -4,8v_1 + 3,7v_2 - 1,7v_3 &= -1,1 \\ 3,7v_1 - 7,2v_2 - 2,4v_4 &= 3,6 \\ -1,7v_1 - 3,4v_3 &= 1,1 \\ -2,1v_2 - 3,0v_4 &= 4,4 \end{aligned}$$

A matriz dos coeficientes do sistema linear anteriormente mostrado é dada por:

$$X = \begin{bmatrix} -4,8 & 3,7 & -1,7 & 0 \\ 3,7 & -7,2 & 0 & -2,1 \\ -1,7 & 0 & -3,4 & 0 \\ 0 & -2,1 & 0 & -3,0 \end{bmatrix}$$

Os valores de v_1 a v_4 são dados pela Equação 3.8:

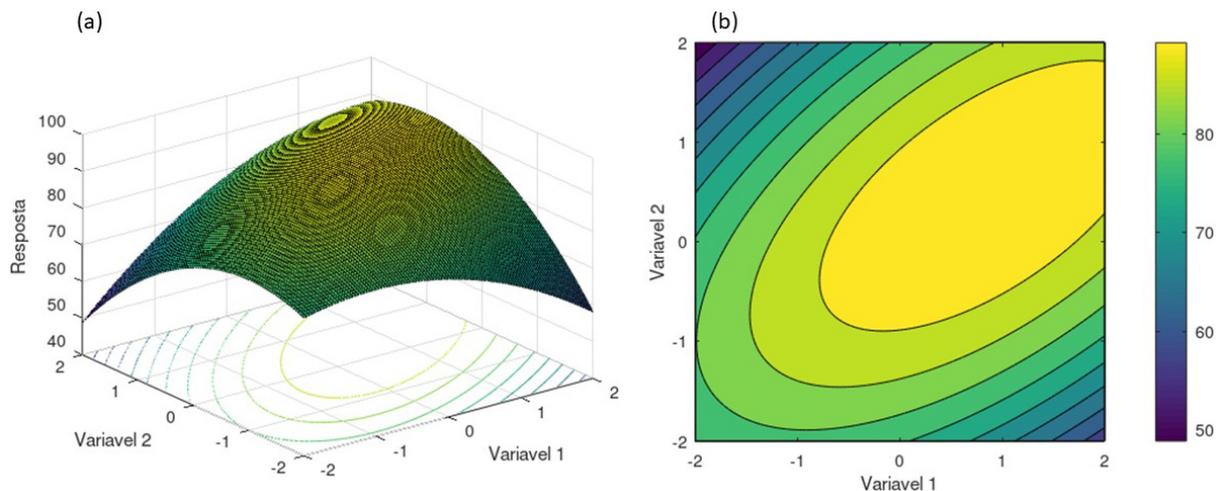
$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} = X^{-1}y \quad (3.8)$$

Onde o vetor y contém as constantes apresentadas no sistema: $y = \begin{bmatrix} -1,1 \\ 3,6 \\ 1,1 \\ 4,4 \end{bmatrix}$

Por fim temos que os valores de v_1 a v_4 que permitem a maior remoção de Hg(II) será dado pelo vetor: $\begin{bmatrix} 0,84 \\ 0,45 \\ -0,74 \\ -1,78 \end{bmatrix}$

A Figura 2 mostra a superfície de resposta (Figura 2a) e o gráfico de contorno (Figura 2b) no domínio $[-2, 2] \times [-2, 2]$ obtidos quando são variadas as variáveis v_1 e v_2 e as variáveis v_3 e v_4 são fixadas em $-0,74$ e $-1,78$, respectivamente. Nas condições ótimas de trabalho, a remoção de Hg(II) é da ordem de $\approx 93\%$.

Figura 2 – Superfície de resposta (a) e gráfico de contorno (b) do modelo apresentado na Função 3.7, tendo as variáveis v_3 e v_4 fixadas em $-0,74$ e $-1,78$, respectivamente.



Fonte: Compilação do autor.

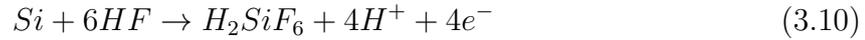
Em planejamentos fatoriais de misturas, a identificação do ponto crítico não é uma tarefa trivial na maioria das vezes. Em diversas situações, o experimentador ao aplicar derivadas parciais, poderá encontrar funções com duas ou mais incógnitas. Além disso, não

é possível a utilização da Regra de Cramer (como ilustrada anteriormente). A dificuldade encontrada no uso do cálculo para determinação dos valores ótimos de cada problema está centrada no grande número de subdivisões executadas. A implementação de diversos casos e situações de contorno trazem dificuldades na programação de um algoritmo para o cálculo desses valores ótimos com auxílio do cálculo diferencial.

3.3 Proposição de um modelo de regressão em um Planejamento de Misturas

O planejamento de misturas apresenta diversas peculiaridades quando comparado com os demais tipos de planejamentos [10]. Uma das características que deve ser apontada é o fato de que há uma relação fixa entre as variáveis envolvidas. Assim, nesta parte da dissertação irei apresentar um exemplo lúdico que mostro durante as aulas de graduação ou pós-graduação. A Figura 3 mostra a preparação de um suco. O suco em questão é constituído por aquelas soluções concentradas comumente encontradas em supermercados. Em casa, o consumidor deve misturar a solução concentrada com água e açúcar de acordo com o seu gosto. Desta forma, a ideia inicial foi preparar um planejamento de misturas do tipo *pseudo*-misturas [3], onde em cada vértice temos uma mistura inicial dos três componentes do suco (suco concentrado + água + açúcar).

Assim, no vértice onde temos 1,00 (ou 100%) de açúcar (variável v_2) puro (açúcar puro não é um suco) teremos uma mistura de 6,25% de água, 6,25% de suco concentrado e 87,5% de uma solução contendo 150g de açúcar em 600 mL de água. Assim, temos o vértice $(0, 1, 0)$ onde uma pessoa poderia dizer que temos um suco muito doce, pois o componente majoritário é açúcar. Nos demais vértices temos sucos muito "aguados" $(0, 0, 1)$ e muito "fortes" $(1, 0, 0)$. As imagens dos componentes nos vértices do triângulo da Figura 3 ajudam na identificação dos mesmos: (i) açúcar com a cor branca, (ii) água e (iii) suco com as cores azul e amarela, respectivamente. Nas arestas do triângulo temos misturas binárias (misturas de dois componentes) e no interior temos misturas ternárias (misturas dos três componentes). Assim, na cozinha da minha casa preparei 22 misturas que contemplassem sucos "aguados", "fortes" e "muito doces", além de sucos que podemos chamar de "normais". As misturas foram colocadas em copos plásticos identificados com um código e oferecidos de forma aleatória (*blind sampling*) a um "*instrumento analítico*" (neste caso, a minha esposa). A minha esposa provou todos os sucos e deu notas de 0 a 10 (ver números no triângulo da Figura 3). Os sucos dos vértices ("aguado", "forte demais", "doce demais") receberam notas baixas que variaram de 0 a 1,5 (ver valores em vermelho na Figura 3). Os sucos das arestas (misturas binárias) tiveram notas de 0,5 até 8,5 (ver valores em azul na Figura 3). Já os sucos das misturas ternárias apresentaram notas de 0,5 a 9,0 (ver valores em verde na Figura 3). Algumas misturas foram realizadas em duplicata ou



Retomando o exemplo que será discutido, temos apresentados na Tabela 1 os experimentos executados.

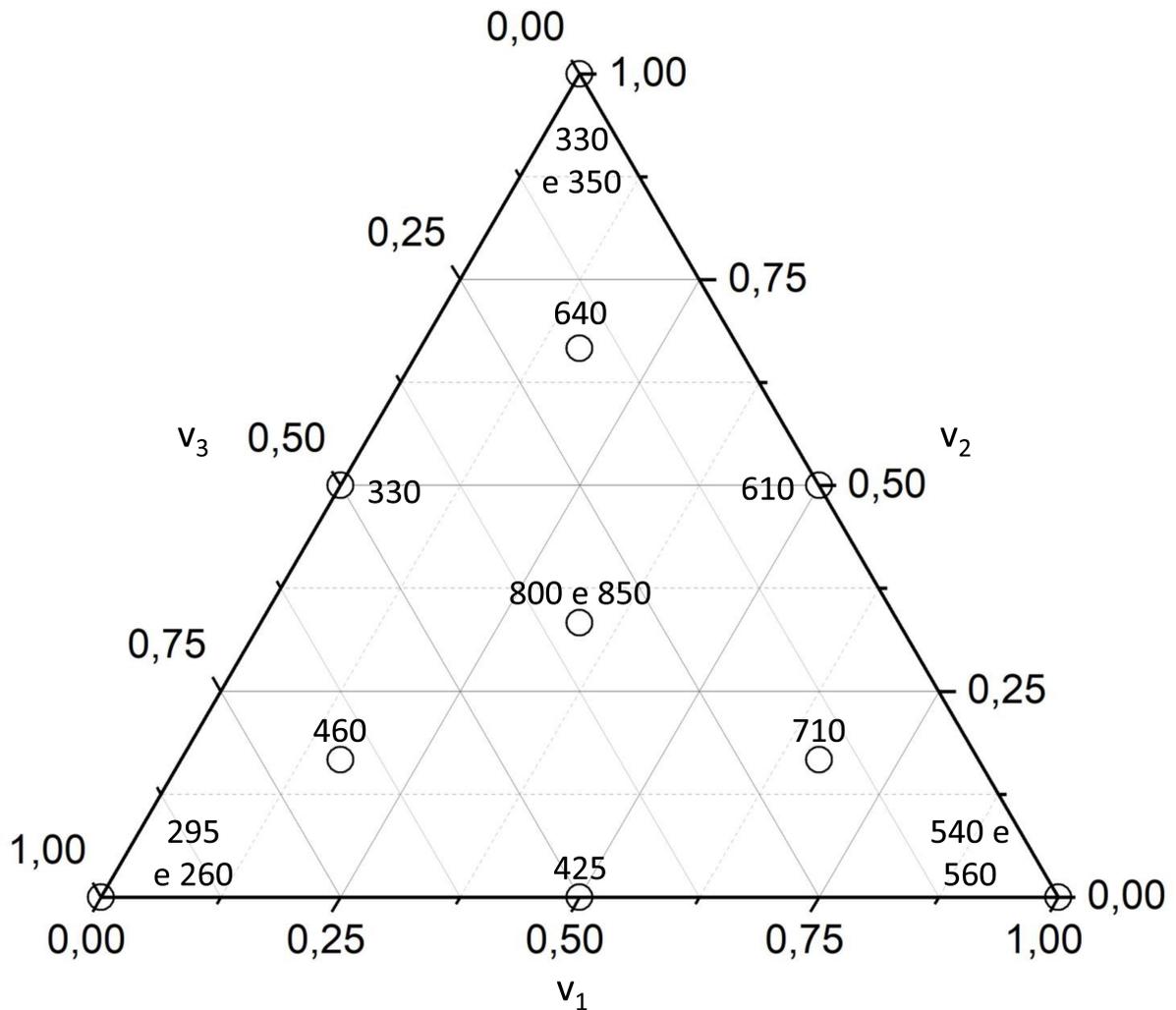
Tabela 1 – Experimentos executados para a otimização das condições de gravações na superfície de semicondutores.

Exp.	v_1 (ácido A)	v_2 (ácido B)	v_3 (ácido C)	Taxa de gravação, y (Å/min)
1 e 2	1	0	0	540 e 560
3 e 4	0	1	0	330 e 350
5 e 6	0	0	1	295 e 260
7	$\frac{1}{2}$	$\frac{1}{2}$	0	610
8	0	$\frac{1}{2}$	$\frac{1}{2}$	330
9	$\frac{1}{2}$	0	$\frac{1}{2}$	425
10	$\frac{2}{3}$	$\frac{1}{6}$	$\frac{1}{6}$	710
11	$\frac{1}{6}$	$\frac{2}{3}$	$\frac{1}{6}$	640
12	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{2}{3}$	460
13 e 14	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	800 e 850

Fonte: elaboração do autor

Os experimentos de 1 a 6 na Tabela 1 representam os vértices do planejamento de misturas e foram realizados em duplicata ($n = 2$). Nas vértices são testados os ácidos (A a C) de forma individual, ou seja, o planejamento não é do tipo *pseudo*-misturas, como o da preparação do suco. Já os experimentos de 7 a 9 representam as arestas do triângulo e os ácidos são misturados para obter soluções binárias 50% entre eles ($A + B$, $B + C$ e $A + C$). Os experimentos 10 até 14 mostram experimentos que foram realizados no interior do domínio experimental (interior do triângulo, $[0,1],[0,1],[0,1]$,) e neste caso temos soluções ternárias entre os ácidos em diferentes proporções. Os experimentos 13 e 14 da solução ternária $v_1=v_2=v_3=\frac{1}{3}$ foram realizados em duplicata e representam o ponto central do domínio experimental. A Figura 4 mostra uma representação dos experimentos executados e o leitor deve ter notado que a soma das três variáveis para um mesmo experimento é sempre igual a 1. Isto indica que há uma relação constante entre as variáveis e o aumento ou diminuição de uma implicará na diminuição ou aumento das demais [3, 10].

Figura 4 – Representação dos experimentos executados (círculos) no planejamento fatorial de misturas mostrados na Tabela 1.



Fonte: Compilação do autor.

Os experimentos realizados em duplicata tem como finalidade verificar a significância do modelo proposto por meio da soma quadrática do erro puro (SQEP). Estes experimentos mostram o erro aleatório (ou indeterminado) inerente aos experimentos e cometido durante a execução dos mesmos no laboratório. Na proposição do modelo matemático podem ser calculados 7 coeficientes como mostrou a Função 1.1.

A matriz de dados \mathbf{X} irá possuir 14 linhas e 7 coeficientes ($b_1, b_2, b_3, b_{12}, b_{13}, b_{23}, b_{123}$):

$$X = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & \frac{1}{4} & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & \frac{1}{4} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & \frac{1}{4} & 0 & 0 \\ \frac{2}{3} & \frac{1}{6} & \frac{1}{6} & \frac{1}{9} & \frac{1}{9} & \frac{1}{36} & \frac{1}{54} \\ \frac{1}{6} & \frac{2}{3} & \frac{1}{6} & \frac{1}{9} & \frac{1}{36} & \frac{1}{9} & \frac{1}{54} \\ \frac{1}{6} & \frac{1}{6} & \frac{2}{3} & \frac{1}{36} & \frac{1}{9} & \frac{1}{9} & \frac{1}{54} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{27} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{27} \end{bmatrix}$$

No cálculo dos coeficientes temos que empregar a Equação 3.4 descrita anteriormente. Assim, a matriz obtida pelas operações $(X^t X)^{-1}$ terá 7 linhas e 7 colunas (7x7) e será:

$$(X^t X)^{-1} = \begin{bmatrix} 0,48 & 0,01 & 0,01 & -1,07 & -1,07 & 0,08 & 1,33 \\ 0,01 & 0,48 & 0,01 & -1,07 & 0,08 & -1,07 & 1,33 \\ 0,01 & 0,01 & 0,48 & 0,08 & -1,07 & -1,07 & 1,33 \\ -1,07 & -1,07 & 0,08 & 19,21 & 2,06 & 2,06 & -56,00 \\ -1,07 & 0,08 & -1,07 & 2,06 & 19,21 & 2,06 & -56,00 \\ 0,08 & -1,07 & -1,07 & 2,06 & 2,06 & 19,21 & -56,00 \\ 1,33 & 1,33 & 1,33 & -56,00 & -56,00 & -56,00 & 792,00 \end{bmatrix}$$

O próximo passo é o cálculo do vetor $X^t y$ (7x1):

$$(X^t y) = \begin{bmatrix} 2824 \\ 2322 \\ 2014 \\ 499 \\ 437 \\ 408 \\ 95 \end{bmatrix}$$

O vetor contendo os coeficientes (7x1), finalmente será:

$$b_i = (X^t X)^{-1} X^t y = \begin{bmatrix} 550 \\ 345 \\ 268 \\ 690 \\ -9 \\ 58 \\ 9243 \end{bmatrix}$$

Antes de prosseguir a discussão sobre o modelo obtido, é importante discutir o significado dos coeficientes calculados para o exemplo em questão. Os coeficientes lineares (b_1 , b_2 e b_3), por exemplo, mostram o efeito individual das variáveis sobre a taxa de gravação. O coeficiente b_1 foi igual a 550 (valor positivo), assim, quando a v_1 está no nível alto (1) é possível obter uma taxa de gravação superior àquela de quando está no nível baixo (0). O mesmo entendimento é válido para os coeficiente b_2 e b_3 e ambos possuem coeficientes positivos, ou seja, contribuem positivamente para a taxa de gravação. A interação das v_1 e v_2 possui um coeficiente positivo ($b_{12} = 690$), assim, quando ambas variáveis tendem a aumentar, a taxa de gravação irá aumentar. A interação das v_1 e v_3 ($b_{13} = -9$) é negativa. Assim, quando essas duas variáveis aumentam simultaneamente (tendem ao nível 1) a taxa de gravação será menor do que aquela obtida quando as duas variáveis estiverem em níveis opostos (uma alta e a outra baixa). O mesmo entendimento pode ser feito para o coeficiente de terciário (b_{123}).

Com o modelo proposto é possível calcular diversos parâmetros e assim verificar a sua significância [2, 5, 9]. O primeiro parâmetro é a Soma quadrática total (SQT), que pode ser definida como a soma quadrática da diferença entre cada valor experimental (y_i) e a média desses valores (\bar{y}), como mostra a Equação 3.11:

$$SQT = \sum_{i=1}^n (y_i - \bar{y})^2 \quad (3.11)$$

A SQT do modelo proposto é 478.821 e esse valor possui $n-1$ graus de liberdade (14-1=13 graus de liberdade), onde n é o número de experimentos realizados. Na sequência é possível calcular a Soma quadrática da regressão (SQR) que possui $p-1$ graus de liberdade. O parâmetro p é o número de coeficientes calculados. No caso do modelo proposto, p será 7 e possui 7-1=6 graus de liberdade. A Equação 3.12 mostra que a SQR está relacionada com a diferença entre os valores previstos (\hat{y}_i) e a média dos valores experimentais (\bar{y}):

$$SQR = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 \quad (3.12)$$

Os valores previstos (\hat{y}_i) pelo modelo podem ser obtidos ao multiplicar a matriz \mathbf{X} (14x7) pelo vetor contendo os coeficientes (b_i , 7x1), como mostra a Equação 3.13:

$$\hat{y}_i = Xb_i \quad (3.13)$$

A SQR do modelo proposto é 470.998. Ao dividir a SQR pela SQT é obtido o R^2 e no caso específico do modelo proposto temos $\frac{470.998}{478.821} = 0,9837$. Esse parâmetro aponta que $\approx 98\%$ da variância dos dados está sendo expressa pelos coeficientes de regressão calculados (b_i) e que $100-98=2\%$ é representado pelo resíduo (erro). A Soma quadrática do resíduo (SQres), ou seja, a parte que não foi modelada pelo modelo proposto é então calculada pela diferença entre a SQT e a SQR. Assim teremos:

$$SQ_{res} = SQT - SQR = 478.821 - 470.988 = 7.824 \quad (3.14)$$

A SQres possui $n - p$ graus de liberdade ($14-7=7$) e pode ser subdividida em duas somas quadráticas: Soma quadrática do erro puro (SQEP, aquela oriunda do erro aleatório dos experimentos) e a Soma quadrática da falta da ajuste (SQFaj). A SQEP trás informações sobre o erro aleatório ou indeterminado [2] cometido durante a execução dos experimentos (veja que esse erro é inevitável) e pode ser calculada pela Equação 3.15:

$$SQEP = \sum_{i=1}^m \sum_{j=1}^{n_j} (y_{ij} - \bar{y}_i)^2 \quad (3.15)$$

A SQEP, no exemplo mostrado, é calculada com os experimentos que possuem réplicas, ou seja, os experimentos 1 e 2, 3 e 4, 5 e 6 e 13 e 14 (ver Tabela 1). Cada réplica de experimentos possui 1 grau de liberdade e parâmetro m da Equação 3.15 será 4. Desta forma, a SQEP será 2.263 com 4 graus de liberdade. Já a SQFaj pode ser computada como sendo:

$$SQFaj = SQ_{res} - SQEP \quad (3.16)$$

Assim, teremos que a SQFaj será $7.824-2.263=5.561$ com $7-4 = 3$ graus de liberdade. Com os parâmetros obtidos é possível então estabelecer a tabela contendo a Análise de variância (Anova). Na tabela Anova são mostrados também os valores de Médias quadráticas (MQ) de cada parâmetro calculado anteriormente (SQR, SQres, SQEP e SQFaj). A Tabela 2 mostra a Anova do modelo proposto.

Tabela 2 – Tabela Anova do modelo proposto para o planejamento de misturas.

Parâmetro	SQ	gl	MQ	Teste F	F tabelado (95%)
SQR	470.998	6	78.500	70,2	3,87 (p -valor = $6,72 \times 10^{-6}$)
SQres	7.824	7	1.118	-	-
SQT	478.821	13	36.832	-	-
SQEP	2.263	4	566	3,28	6,59 (p -valor = 0,141)
SQFaj	5.561	3	1.854	-	-

Fonte: elaboração do autor

Com os valores de MQ é possível verificar a significância estatística do modelo por meio do teste F [9]. Para o caso das Médias quadráticas da regressão e resíduo (MQR e MQres, respectivamente), temos que o valor de F calculado é $\frac{78.500}{1.118} = 70,2$. Tal valor é bastante superior ao valor de F tabelado (3,87) com 95% de confiança e 6 (numerador) e 7 (denominador) graus de liberdade. Além disso, o p -valor é inferior a 0,05. Assim, ambas MQ (MQR e MQres) são estatisticamente diferentes, ou seja, a regressão não se confunde com o resíduo.

Com as informações obtidas anteriormente podemos realizar um teste de hipóteses [30]. As hipóteses são avaliadas em pares que contêm as hipóteses nula (H_0) e a alternativa (H_1). Assim, no cálculo descrito anteriormente temos:

$$H_0 : p\text{-valor} = 0,05$$

$$H_1 : p\text{-valor} < 0,05$$

Com a H_0 assumimos que ambas MQ são iguais, ou seja, p -valor = 0,05. Já no caso da H_1 devemos assumir que as MQ são estatisticamente diferentes (p -valor < 0,05).

Assim, temos que ao comparar MQR e MQres, a hipótese nula (H_0) é rejeitada e a alternativa (H_1) é aceita. Desta forma, é possível afirmar que o modelo é estatisticamente válido ao nível de confiança de 95%. Veja que o resíduo (parte não modelada) não se confunde estatisticamente com a regressão (parte modelada).

Outra comparação importante a ser efetuada é o teste F que avalia a MQFaj e a MQEP: $\frac{1.854}{566} = 3,28$. Como é possível identificar na Tabela 2, o valor de F calculado (3,28) é menor que o tabelado (6,59) com 95% de confiança e 3 (numerador) e 4 (denominador) graus de liberdade. Diante disso, o p -valor é superior a 0,05 (0,141) e, assim, ambas médias quadráticas são estatisticamente iguais (aceita H_0 e rejeita H_1). Assim, o erro aleatório inerente e ocorrido na execução dos experimentos (Erro puro) se confunde com a incapacidade do modelo em se ajustar aos dados experimentais (Falta de ajuste). Como conclusão parcial é possível afirmar que o modelo é estatisticamente válido, pois a MQR

é diferente da MQres e a MQFaj é igual à MQEP [7]. Com isso é possível então calcular a variância, o erro e o intervalo de confiança dos coeficientes propostos.

Para o caso deste modelo, a variância dos coeficientes será dada pela MQres (1.118) multiplicada pela diagonal principal da matriz $(X^tX)^{-1}$. O erro é a raiz quadrada de cada variância calculada e o intervalo de confiança será dado pela multiplicação do erro pelo valor de t correspondente a $n - 1$ graus de liberdade da MQres [2, 5]. No caso específico do modelo proposto, o valor de t é 2,447 (6 graus de liberdade e 95% de confiança). O vetor apresentado a seguir mostra o cálculo envolvido na obtenção dos intervalos de confiança:

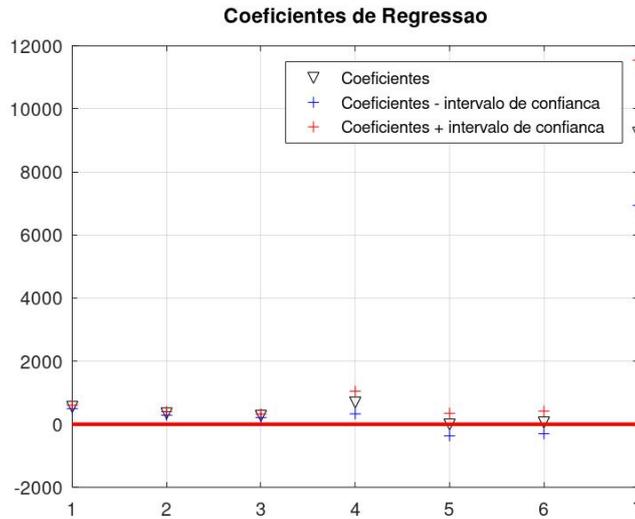
$$\text{Vetor com os intervalos de confiança} = \begin{bmatrix} \sqrt{1.118 \times 0,48} \times 2,447 = 56,83 \\ \sqrt{1.118 \times 0,48} \times 2,447 = 56,83 \\ \sqrt{1.118 \times 0,48} \times 2,447 = 56,83 \\ \sqrt{1.118 \times 19,21} \times 2,447 = 358,51 \\ \sqrt{1.118 \times 19,21} \times 2,447 = 358,51 \\ \sqrt{1.118 \times 19,21} \times 2,447 = 358,51 \\ \sqrt{1.118 \times 792,00} \times 2,447 = 2302,19 \end{bmatrix}$$

Com os valores de intervalo de confiança calculados é possível identificar que os coeficientes b_{13} ($-9 \pm 358,51$) e b_{23} ($58 \pm 358,51$) são estatisticamente insignificantes. Valores de b_{13} e b_{23} estão no intervalo $[-367,349]$ e $[-300, 416]$, respectivamente e podem assumir o valor 0. Pereira Filho [21] propôs um algoritmo (Regression2) para a realização dos cálculos descritos anteriormente e a Figura 5 mostra a avaliação dos coeficientes de regressão. Os números (1 a 7) apresentados no eixo x da Figura 5 representam os 7 coeficientes calculados ($b_1, b_2, b_3, b_{12}, b_{13}, b_{23}, b_{123}$, nesta ordem). O Apêndice A mostra as demais informações geradas pelo algoritmo Regression2 para a proposição de modelos de regressão e a sua avaliação por meio da tabela Anova. O modelo final proposto é:

$$\text{Taxa de gravação} = f(v_1, v_2, v_3) = 550v_1 + 345v_2 + 268v_3 + 690v_1v_2 + 9243v_1v_2v_3 \quad (3.17)$$

A próxima seção discutirá como identificar o ponto máximo do modelo da Função 3.17 por meio da utilização de derivadas parciais.

Figura 5 – Visualização dos coeficientes de regressão e os intervalos de confiança.



Fonte: Compilação do autor.

3.4 Identificação dos Pontos Críticos em um Planejamento de Misturas

Na seção anterior foi proposto um modelo matemático para descrever como três variáveis (ácidos) podem impactar a taxa de gravação na superfície de semicondutores. O modelo pode ser descrito pela Função 3.18:

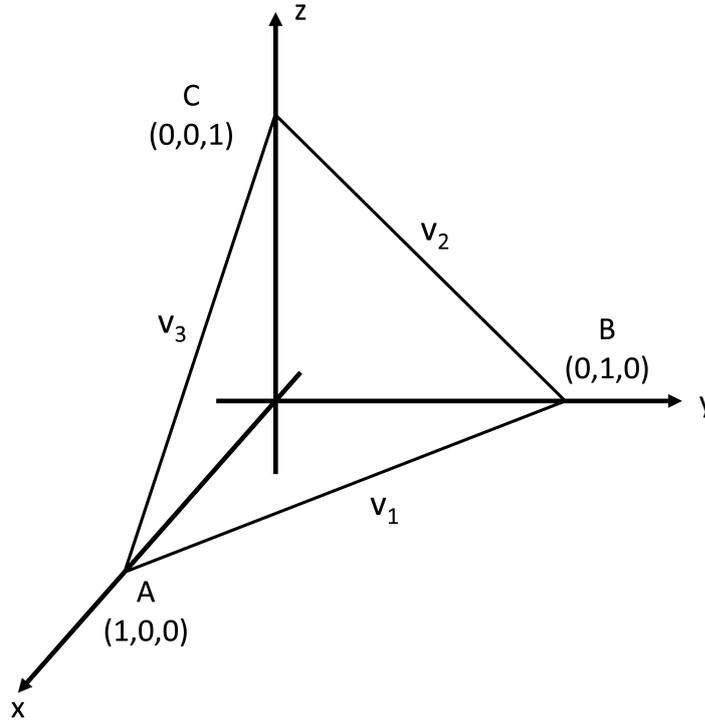
$$\text{Taxa de gravação} = f(v_1, v_2, v_3) = 550v_1 + 345v_2 + 268v_3 + 690v_1v_2 + 9243v_1v_2v_3 \quad (3.18)$$

A partir deste ponto irei referir às variáveis v_1 , v_2 e v_3 como x , y e z , respectivamente. Assim, as variáveis x , y e z satisfazem a relação, $x + y + z = 1$. Além disso, $x \geq 0$, $y \geq 0$ e $z \geq 0$. Assim, o objetivo é determinar o ponto máximo da Função 3.18 que satisfaça às essas condições. A região determinada pelas condições $x + y + z = 1$, $x \geq 0$, $y \geq 0$ e $z \geq 0$ é representada pela região triangular delimitada pelo triângulo (T) de vértices $A = (1, 0, 0)$, $B = (0, 1, 0)$ e $C = (0, 0, 1)$, como mostra a Figura 6.

Temos que, se $x + y + z = 1$ então, $z = 1 - x - y$. Substituindo esse valor de z na expressão que define a Função 3.18, obtemos:

$$\begin{aligned} f(x, y, 1 - x - y) &= 550x + 345y + 268(1 - x - y) + 690xy + 9243xy(1 - x - y) = \\ &= 550x + 345y + 268 - 268x - 268y + 690xy + 9243xy - 9243x^2y - 9243xy^2 = \\ &= 268 + 282x + 77y + 9933xy - 9243x^2y - 9243x^2y \end{aligned} \quad (3.19)$$

Figura 6 – Triângulo de vértices $A = (1, 0, 0)$, $B = (0, 1, 0)$ e $C = (0, 0, 1)$.



Fonte: Compilação do autor.

Assim, se $x + y + z = 1$ temos que $f(x, y, z) = F(x, y)$, sendo, $F : T \rightarrow \mathbb{R}$ definida por,

$$F(x, y) = 268 + 282x + 77y + 9933xy - 9243xy^2 - 9243x^2y \quad (3.20)$$

$$\text{e } T = \{(x, y) \in \mathbb{R}^2: x \geq 0, y \geq 0 \text{ e } x + y \leq 1\}.$$

Assim, o problema inicial de determinar o ponto de máximo de f sujeito às condições $x + y + z = 1$, $x \geq 0$, $y \geq 0$ e $z \geq 0$ é equivalente à determinação do ponto máximo de F . Temos que T (Figura 6) é a região triangular com vértices em $(0,0)$, $(1,0)$ e $(0,1)$. Portanto T é compacto. O termo "compacto" indica que T é um subconjunto não-vazio de \mathbb{R} , sendo fechado e limitado (limitado com os vértices mencionados anteriormente) [31].

Para determinar o ponto máximo de F primeiramente devemos determinar os pontos críticos de F :

$$\frac{\partial F}{\partial x}(x, y) = 0 \iff 282 + 9933y - 9243y^2 - 18486xy = 0 \quad (3.21)$$

$$\frac{\partial F}{\partial y}(x, y) = 0 \iff 77 + 9933x - 18486xy - 9243x^2 = 0 \quad (3.22)$$

Isolando x em 3.21 obtemos,

$$x = \frac{282 + 9933y - 9243y^2}{18486y} = \frac{94 + 3311y - 3081y^2}{6162y} \quad (3.23)$$

Substituindo esse valor de x em 3.22 obtemos,

$$\frac{10699809y^2 - 4080476y^3 + 28477683y^4 - 8836}{4108y^2} = 0 \quad (3.24)$$

A identificação das raízes do polinômio mostrado em 3.24 exige consideráveis esforços e uma das maneiras para realizar tal procedimento é por meio de aproximações sucessivas e o Teorema do Valor Intermediário (TVI) [22]. Inicialmente é necessário compreender que não precisamos identificar todas as raízes do polinômio e sim aquelas que estão no domínio ou intervalo fechado $[0, 1]$ (as variáveis variam entre 0 e 1). Assim, as raízes (L_i) serão números que $f(a) \leq L_i \leq f(b)$ ou $f(b) \leq L_i \leq f(a)$. Assim, existe pelo menos um valor de $x \in [a, b]$ tal que $f(x_i) = L_i$.

Para o caso particular do numerador ($10699809y^2 - 4080476y^3 + 28477683y^4 - 8836$) do polinômio mostrado em 3.24, as raízes serão denominadas de y_i e no intervalo $[0, 1]$ foram identificadas duas raízes (y_1 e y_2). Finalmente, as raízes do polinômio 3.24, contidas no intervalo $[0, 1]$, são dadas por,

$$y_1 \approx 0,0305 \text{ e } y_2 \approx 0,342$$

Para efeito de ilustração, a Figura 7 mostra os valores calculados de y (Figura 7a) do polinômio de 3.24 e a identificação de y_1 (Figura 7b) e y_2 (Figura 7c).

Substituindo os valores de y_1 e y_2 em 3.23 obtemos, respectivamente,

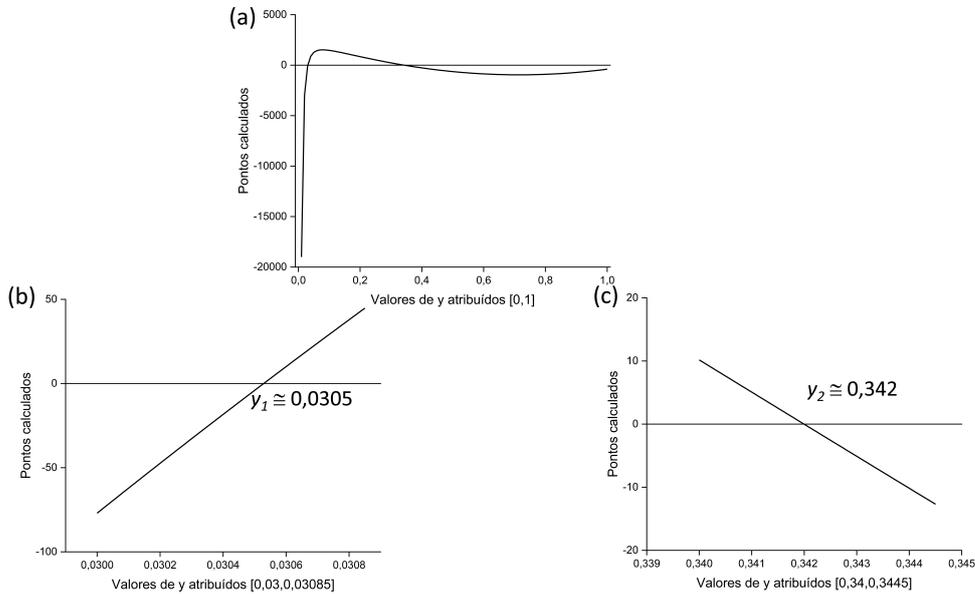
$$x_1 \approx 1,0217 \text{ e } x_2 \approx 0,411$$

Como x_1 e y_1 não pertencem a T (veja que $x_1 > 1$), segue que $P_1 = (x_2, y_2) = (0,411, 0,342)$ é o único ponto crítico de F .

Outros pontos de que devem ser avaliados são os pontos do bordo de T . O bordo de T é formado pelos segmentos:

- $T_1 = \{(x, y) \in T : y = 0, 0 \leq x \leq 1\}$
- $T_2 = \{(x, y) \in T : x = 0, 0 \leq y \leq 1\}$ e
- $T_3 = \{(x, y) \in T : x \geq 0, y \geq 0 \text{ e } x + y = 1\}$.

Figura 7 – Visualização dos valores calculados do polinômio 3.24 no domínio $[0, 1]$ (a) e identificação dos valores de y_1 (b) e y_2 (c).



Fonte: Compilação do autor.

Em $T_1 = \{(x, y) \in T: y = 0, 0 \leq x \leq 1\}$:

Temos que, se $(x, y) \in T_1$ então,

$$F(x, y) = F(x, 0) = 268 + 282x = g_1(x) \text{ e } g_1: [0, 1] \rightarrow \mathbb{R}$$

Assim, temos que g_1 não possui pontos críticos, pois $g_1'(x) = 282 \neq 0$ para todo $x \in [0, 1]$. Logo, os candidatos a ponto de máximo ou mínimo de g_1 são, $x = 0$ e $x = 1$. Esses valores são os extremos do intervalo de definição $[0, 1]$ de g_1 . Desta forma, $P_2 = (0, 0)$ e $P_3 = (1, 0)$ são também candidatos a pontos de máximo ou mínimo de F .

Em $T_2 = \{(x, y) \in T: x = 0, 0 \leq y \leq 1\}$:

Temos que, se $(x, y) \in T_2$ então,

$$F(x, y) = F(0, y) = 268 + 77y = g_2(y) \text{ e } g_2: [0, 1] \rightarrow \mathbb{R}$$

Assim, temos que g_2 não possui pontos críticos, pois $g_2'(y) = 77 \neq 0$ para todo $y \in [0, 1]$. Logo, os candidatos a ponto de máximo ou mínimo de g_2 são, $y = 0$ e $y = 1$. Esses valores são os extremos do intervalo de definição $[0, 1]$ de g_2 . Esses pontos reproduzem $P_2 = (0, 0)$ e $P_4 = (0, 1)$ como candidatos a pontos de máximo ou mínimo de F .

Em $T_3 = \{(x, y) \in T: x \geq 0, y \geq 0 \text{ e } x + y = 1\}$:

Temos que, se $(x, y) \in T_3$ então,

$$F(x, y) = F(x, 1-x) = 268 + 282x + 77(1-x) + 9933x(1-x) - 9243x(1-x)^2 - 9243x^2(1-x) \\ = 345 + 895x - 690x^2 = g_3(x) \text{ e } g_3 : [0, 1] \rightarrow \mathbb{R}$$

Como $g_3'(x) = 0 \iff 895 - 1380x = 0 \iff x = \frac{1380}{895} > 1$ temos que g_3 não possui pontos críticos. Logo, os candidatos a ponto de máximo ou mínimo de g_3 são, $x = 0$ e $x = 1$ os quais reproduzem $P_3 = (1, 0)$ e $P_4 = (0, 1)$ como candidatos a pontos de máximo ou mínimo de F .

Desta forma, os candidatos a ponto de máximo (absoluto) de F são P_1, P_2, P_3 e P_4 . Assim, basta testar esses pontos, calculando o valor de F nos mesmos:

$$\bullet F(P_1) = F(0, 411, 0, 342) \approx 828, 13$$

$$\bullet F(P_2) = F(0, 0) = 268$$

$$\bullet F(P_3) = F(1, 0) = 268 + 282 = 550$$

$$\text{e } \bullet F(P_4) = F(0, 1) = 268 + 77 = 345$$

Temos finalmente que $F(P_2) < F(P_4) < F(P_3) < F(P_1)$. Assim, $P_2 = (0, 0)$ é ponto mínimo absoluto de F e $P_1 = (0, 411, 0, 342)$ é ponto de máximo absoluto de F . Neste momento é importante recordar que $z = 1 - x - y$ e teremos: $(0, 411, 0, 342, 0, 247)$ o ponto de máximo absoluto de f sujeito às condições, $x \geq 0, y \geq 0, z \geq 0$ e $x + y + z = 1$.

O leitor deve ter percebido que a identificação do ponto de máximo de um planejamento de misturas não é uma tarefa trivial, envolvendo uma análise criteriosa e atenta das diversas situações de contorno que foram exemplificadas anteriormente. Além disso, temos também polinômios de grau 4 que apresentam dificuldades consideráveis para a determinação das suas raízes. A implementação de tais estratégias de contorno em um algoritmo computacional é uma tarefa que envolve considerável esforço. Desta forma, nas próximas seções desta dissertação de mestrado profissional serão mostradas alternativas para a identificação do ponto de máximo de um planejamento de misturas com a utilização de um algoritmo especialmente desenhado para tal atividade. Além disso, foi desenvolvido um arquivo no Microsoft Excel® para visualização dos modelos empíricos gerados.

4 Materiais e Métodos

4.1 Programas Computacionais

No desenvolvimento da parte experimental foram empregados diversos computadores do tipo *desktop* e *notebook* para o desenvolvimento e teste do algoritmo. Os computadores não apresentavam nenhuma especificidade e as suas características físicas (memória RAM, velocidade e capacidade de armazenamento) são as comumente encontradas em computadores pessoais (*Personal computer*, PC). Mais detalhes sobre a configuração do computador serão discutidos na sequência.

Os programas computacionais empregados foram GNU Octave, versão 6.1.0 (<https://www.gnu.org/software/octave/>). O Octave é um programa livre de programação numérica. Além disso, foi empregado o Microsoft Excel® (versão Microsoft 365®). No Octave foi proposto o algoritmo `mix_target2` para a identificação dos pontos críticos de um modelo de um planejamento de misturas. No caso do Microsoft Excel®, foi preparado um arquivo denominado *Template Misturas* que conta com a aplicação da ferramenta Solver® para a identificação dos pontos críticos de um modelo.

4.2 O Ensino de Planejamento Fatorial na Pandemia

Como discutido anteriormente, as atividades de ensino foram realizadas em plena pandemia de Covid-19 e foram, também avaliadas as percepções dos alunos sobre os procedimentos adotados. Com a chegada e o avanço da Pandemia de Covid-19 no Brasil no final de fevereiro de 2020, diversas mudanças foram deflagradas no sentido de dar continuidade ao ensino de forma remota. As dificuldades iniciais foram imensas, onde universidades, professores e alunos tiveram que se adaptar à nova realidade de forma bastante rápida [32]. Além disso, alunos e professores tiveram que identificar e conhecer as novas ferramentas para o ensino, além da proposição de avaliações sem a supervisão do professor [33]. No ano de 2020 ofereci a disciplina Quimiometria 1 (10 créditos) em três momentos e todas de forma remota: (i) primeiro semestre, (ii) segundo semestre e (iii) na primeira edição do Ensino Remoto não Presencial (ENPE1) para os alunos de graduação da UFSCar. A Quimiometria 1 é oferecida de forma regular no PPGQ/UFSCar desde 2007. No ENPE1, parte do conteúdo de Quimiometria 1 foi ministrado para alunos de graduação e a disciplina é optativa sendo denominada de Introdução à Quimiometria (60 horas/aula e 4 créditos). A primeira edição do ENPE (ENPE1) foi ministrada a partir do mês de agosto/2020. Nos anos de 2021 tivemos os ENPE2 e 3 e em 2022 temos o ENPE4. No ano de 2021 as disciplinas Quimiometria 1 e Introdução à Quimiometria

foram novamente oferecidas no primeiro e segundo semestres, respectivamente.

4.3 Coleta de dados e avaliação pelos estudantes

Na avaliação das percepções dos alunos foram encaminhados formulários com questões nos anos de 2020 (ENPE1) e 2021 (ENPE2 e ENPE3). Ao todo foram recolhidas respostas de 32 e 130 alunos de graduação e pós-graduação, respectivamente. A porcentagem de alunos que responderam variou de 46% a 86% para a graduação e pós-graduação, respectivamente.

Para os cursos de graduação em Química da UFSCar (bacharelado e licenciatura) foi oferecida em duas oportunidades a disciplina optativa Introdução à Quimiometria (60 horas/aula, 4 créditos). Já para o PPGQ/UFSCar foi oferecida a disciplina Quimiometria 1 (10 créditos) em três oportunidades (duas vezes em 2020 e uma em 2021). Em todos os casos, as disciplinas contaram com aulas online gravadas e também encontros síncronos gravados. A Tabela 3 mostra as questões apresentadas para avaliação das disciplinas oferecidas. Os formulários eram enviados sempre ao final do semestre e foi dado um intervalo de tempo de duas semanas para que os alunos respondessem aos questionamentos.

Tabela 3 – Questões apresentadas aos alunos para avaliação das disciplinas de graduação e pós-graduação. Os detalhes sobre as opções de respostas disponíveis para os alunos podem ser visualizados no Apêndice B.

Questão	Descrição
1	Qual é a sua opinião geral sobre a disciplina que cursou?
2	O que você achou da didática de forma geral?
3	O que você achou da teoria de forma geral?
4	O que você achou dos exemplos dados?
5	O que você achou sobre a utilização do Google Forms para responder as atividades (trabalhos, testes e provas)?
6	O que você achou dos vídeos?
7	Sobre a duração dos vídeos, o que você achou?
8	Como ficou o seu nível de aprendizado?
9	O que você achou do formato da disciplina (vídeos gravados e disponibilizados para os alunos)?
10	Você faria outra disciplina como essa (100% com aulas gravadas e encontros síncronos gravados)?
11	Pensando em um cenário pós-pandemia, onde a vida voltaria ao "normal."Quais das opções abaixo são mais atraentes? Marque quantas opções sejam necessárias.
12	Dúvidas ou comentários/sugestões?

Fonte: elaboração do autor

As questões de 1 a 11 eram obrigatórias e a 12, opcional. O Apêndice B mostra o formulário apresentado e o mesmo contou com um Termo de Consentimento Livre e Esclarecido (TCLE), que permite a utilização das informações sem a identificação dos alunos que responderam. Algumas questões tinham cinco níveis de opção para a resposta: Regular, Bom ou Boa, Muito boa ou Muito bom, Ótimo e Excelente.

Antes do início da pandemia foram realizadas coletas de dados apenas com os alunos da pós-graduação (disciplina Quimiometria 1) e o número de alunos que responderam foi de 63 que correspondeu a 86% do total. Entretanto, antes da pandemia as questões apresentadas aos alunos foram apenas as 1, 2, 3, 4 e 6 (ver detalhes na Tabela 3 e no Apêndice B).

Além da opinião dos alunos sobre as disciplinas, foram efetuadas também uma comparação das notas (graduação) e conceitos (pós-graduação) obtidos pelos estudantes antes e durante a pandemia. Entretanto, é importante mencionar que os resultados dessa comparação devem ser analisados com cautela, pois as avaliações foram feitas de maneiras diferentes. Antes da pandemia tanto para a graduação quanto para a pós-graduação era dada uma prova onde o aluno tinha cerca de 30 dias para concluir. Nessa prova era apresentado um conjunto de dados da literatura e o aluno deveria: (i) identificar qual foi o tipo de planejamento executado; (ii) identificar possíveis réplicas e calcular a SQEP e; (iii) propor um modelo utilizando as ferramentas computacionais mostradas nas aulas presenciais. Entretanto, o aluno deveria também ir até a sala do professor e mostrar o entendimento dele sobre o conjunto de dados. Como é possível verificar, tal tipo de avaliação não poderia ser realizada no período pandêmico. Assim, com o ensino remoto a avaliação descrita anteriormente foi substituída por testes (5 para a graduação e 7 para a pós-graduação).

Esses testes também versaram sobre conjuntos de dados da literatura, mas o aluno tinha que apresentar as respostas por meio de formulários eletrônicos. As questões eram, na sua maioria, de múltiplas escolhas, mas haviam campos onde os alunos deveriam redigir uma resposta para determinada pergunta.

4.4 Conteúdos apresentados e características dos vídeos

Como afirmado anteriormente, as disciplinas ofertadas na pós-graduação e graduação eram denominadas de Quimiometria 1 e Introdução à Quimiometria, respectivamente. A Tabela 4 mostra os conteúdos mostrados aos alunos. Os conteúdos foram apresentados por meio de vídeos previamente gravados e também em encontros síncronos (que eram gravados também). Os vídeos previamente gravados apresentaram uma duração média de cerca de 10 minutos.

Os vídeos preparados foram gravados, na sua grande maioria, na minha residência

(em 2020) ou na minha sala no Departamento de Química (em 2021) e tiveram como característica serem curtos (de 10 a 15 minutos no máximo). Na gravação foram empregados dois programas: Zoom® e o aTube Catcher®. Existem outras ferramentas disponíveis, tais como o *Open Broadcaster Software* (OBS) e o Loom®. O computador (*laptop*) utilizado foi um dispositivo convencional e sem qualquer acessório ou configuração diferente do comumente encontrado. Para obter uma melhor qualidade do som, foi empregado um microfone Samson® (modelo C01UPRO). Os vídeos dos encontros síncronos foram gravados por meio do *Google Meet* da plataforma G Suite (Google®) oferecida pela UFSCar.

As avaliações foram realizadas ao utilizar *Google forms*. As questões eram compostas por testes de múltipla escolha ou campos onde os alunos poderiam escrever uma resposta. As 15 partes ministradas para os alunos da pós-graduação (ver Tabela 4) corresponderam à 10 horas de material gravado divididos em 41 vídeos. Além do conteúdo teórico, foram apresentados também 8 práticas utilizando o Excel® e o Octave para os alunos da pós-graduação. No caso da graduação (partes de 1 até 9, Tabela 4), o tempo de gravação correspondeu a 6 horas (27 vídeos) e 5 práticas. Além do conteúdo gravado, foram realizados encontros síncronos que corresponderam, em média, à adicionais 10 e 11 horas para a pós-graduação e graduação, respectivamente.

Durante a realização das aulas remotas foi utilizado um formulário eletrônico onde os alunos poderiam encaminhar dúvidas sobre os conteúdos apresentados. Esse formulário foi chamado de Serviço de atendimento do Edenir (SAEd) e os alunos não precisavam necessariamente se identificarem (ver Apêndice C). A proposta de ensino estava centrada no fato de que os alunos deveriam assistir aos vídeos previamente gravados, apresentar dúvidas utilizando o SAEd e os questionamentos seriam discutidos e respondidos durante os encontros síncronos que foram realizados. Entre os anos de 2020 e 2022 (até abril) o SAEd recebeu mais de 520 consultas dos alunos.

Tabela 4 – Conteúdos apresentados nas disciplinas de graduação (Partes 1 a 9) e pós-graduação (Todas as partes).

Parte	Assuntos abordados
1	Introdução e contextualização
2	Planejamento fatorial completo
3	Planejamento fatorial completo, refinando
4	Planejamento fatorial fracionário
5	Planejamentos saturados
6	Proposição de modelos
7	Proposição de modelos, continuação
8	Proposição de modelos em um planejamento fatorial
9	Proposição de modelos, um exemplo
10	Uso do simplex
11	Planejamento Doehlert e Box-Behnken
12	Problemas com múltiplas respostas
13	Planejamento fatorial + <i>Principal component analysis</i> , PCA
14	Planejamento de misturas
15	Tópicos finais

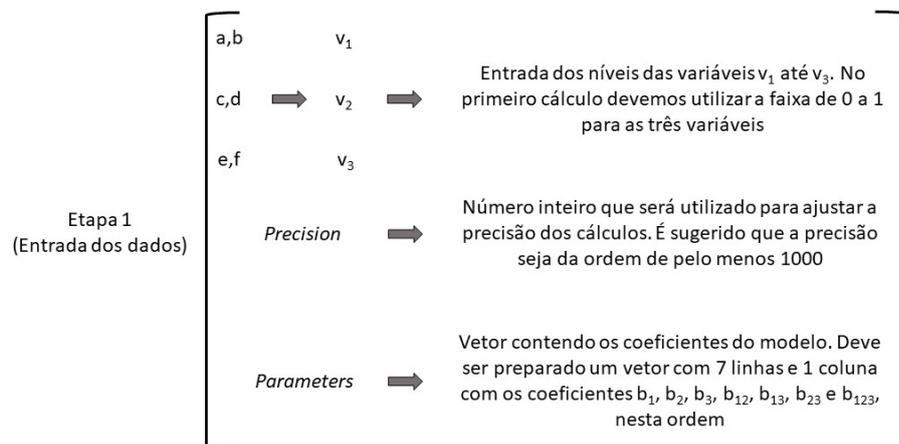
Fonte: elaboração do autor

5 Resultados e discussão, algoritmo mix_target2

5.1 Preparação e proposição do algoritmo mix_target2

A partir desta seção será apresentado um algoritmo desenhado para a identificação do ponto de máximo de um modelo obtido em um planejamento de misturas. A linha de comando do algoritmo `mix_target2` possui a seguinte configuração: $[B, Response] = \text{mix_target2}(a, b, c, d, e, f, Precision, Parameters)$. As informações de entrada (*input*) são formadas pelos níveis das variáveis v_1 (a e b), v_2 (c e d) e v_3 (e e f), a *Precision* e *Parameters*. Os níveis de a até f são valores que inicialmente devem ser 0 (menor nível) e 1 (maior nível). Desta forma, para a variável 1, por exemplo, os valores de a e b devem ser 0 e 1, respectivamente (domínio $[0,1]$). É importante lembrar que a soma dos valores das três variáveis deve ser 1 [3]. A *Precision* é correspondente ao tamanho da população de dados que será gerada. Para conciliar velocidade de processamento e precisão dos cálculos, é recomendado que a *Precision* seja pelo menos igual a 1000. Já a entrada *Parameters* trás um vetor coluna com os coeficientes válidos do modelo ($b_1, b_2, b_3, b_{12}, b_{13}, b_{23}$ e b_{123} , nesta ordem). A Figura 8 mostra o início do funcionamento do algoritmo `mix_target2`.

Figura 8 – Primeira etapa de funcionamento do algoritmo.

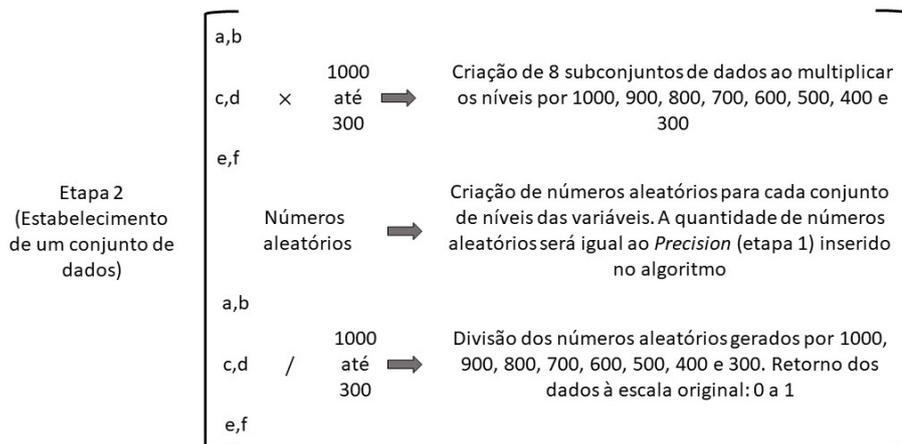


Fonte: Compilação do autor.

Coefficientes insignificantes devem ser substituídos por 0. As informações de saída (*output*) são B e *Response*. O *output* B é um vetor que contém as condições das variáveis v_1, v_2 e v_3 que representam a maior resposta do modelo. A maior resposta é apresentada no objeto (1x1) *Response*. Na sequência à primeira etapa, são criados 8 subconjuntos de dados. Esses subconjuntos são estabelecidos ao multiplicar os níveis das 3 variáveis por 1000, 900, 800, 700, 600, 500, 400 e 300. Os novos níveis da variável 1, por exemplo, serão

0 e 1000, 0 e 900, 0 e 800, 0 e 700, 0 e 600, 0 e 500, 0 e 400 e 0 e 300 para os subconjuntos de 1 a 8, respectivamente. Dentro de cada um dos subconjuntos são estabelecidas uma quantidade de números aleatórios correspondentes ao valor *Precision* inserido na Etapa 1. A Figura 9 mostra um resumo dos processos executados na Etapa 2.

Figura 9 – Segunda etapa de funcionamento do algoritmo.



Fonte: Compilação do autor.

É importante ressaltar que após a multiplicação dos dados de entrada por valores de 1000 até 300 é aplicada a função *round* no Octave. Esta função permite a obtenção de valores inteiros. Na sequência é utilizada a função *randi* que possibilita a geração de uma quantidade pré-estabelecida (*Precision*) de números aleatórios entre os intervalos de valores fornecidos: 0 a 1000 até 0 a 300. Assim, com a multiplicação dos valores de entrada (a até f) por números que variam de 1000 a 300, é possível obter uma melhor amplitude de números aleatórios. Com um valor de *Precision* igual a 1000 são gerados 8000 números para cada uma das variáveis. Após a geração dos números aleatórios, os mesmos são divididos por 1000 até 300 para retornar à escala inicial de 0 até 1. Desta forma, temos o último procedimento da etapa 2 (Figura 9). Outro ponto importante a ser mencionado é que os subconjuntos gerados foram estabelecidos até a multiplicação e divisão por 300. Esta configuração permite a obtenção de combinações de níveis das variáveis 1 até 3 nas imediações do ponto central do planejamento de misturas: $v_1 = \frac{1}{3}$, $v_2 = \frac{1}{3}$ e $v_3 = \frac{1}{3}$. O objetivo deste processo é gerar uma população de dados que ocupe toda o domínio experimental de um planejamento de misturas, como aquele mostrado na Figura 11.

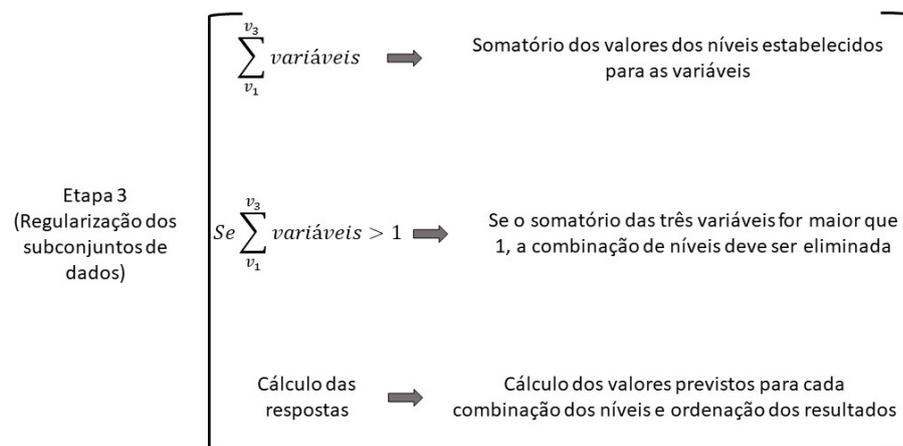
Na Etapa 3 do algoritmo é calculada inicialmente o somatório dos níveis das três variáveis. Se o valor obtido for > 1 , a combinação de níveis é eliminada, sendo mantida apenas combinações onde a soma dos níveis é de no máximo 1. Novamente, vale lembrar que em um planejamento de misturas, há uma relação constante entre as três variáveis e o aumento de uma acarretará na diminuição de outras. No final da Etapa 3 são calculadas

as respostas previstas para todas combinações úteis:

$$v_1 + v_2 + v_3 \leq 1 \text{ e } v_1 \geq 0; v_2 \geq 0; v_3 \geq 0$$

A Figura 10 mostra um resumo dos processos executados na Etapa 3. A Figura 11 mostra as combinações geradas para as variáveis de 1 a 3 quando o valor da *Precision* foi igual a 1000.

Figura 10 – Terceira etapa de funcionamento do algoritmo.



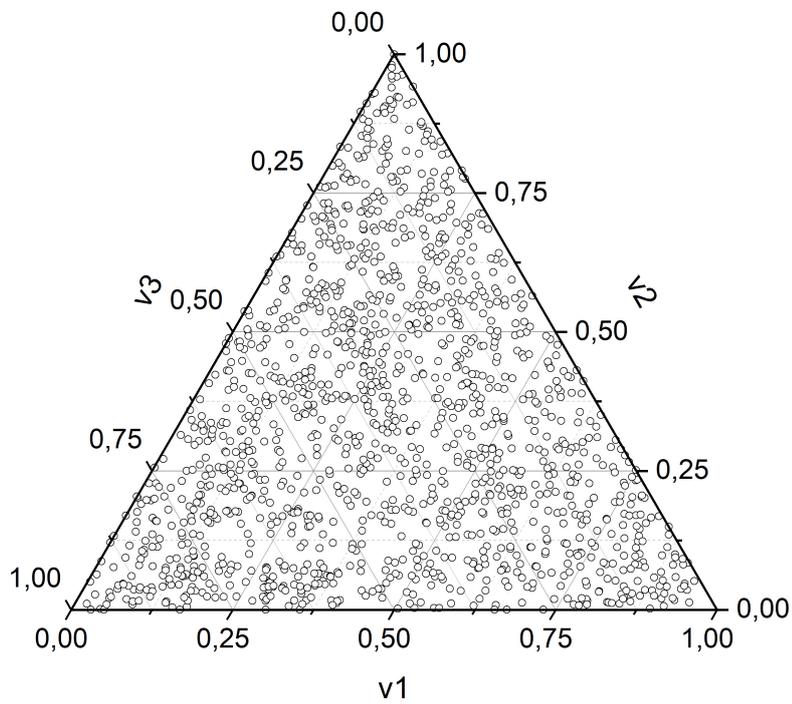
Fonte: Compilação do autor.

É importante notar que todo espaço amostral foi preenchido. No caso específico do valor de *Precision* igual a 1000, foram obtidos no final dos cálculos cerca de 1400 valores úteis dos intervalos das três variáveis ($v_1 + v_2 + v_3 \leq 1$).

As Figuras 12 e 13 mostram os espaços amostrais quando os valores de *Precision* foram 5000 e 10000, respectivamente. Para um valor de *Precision* de 5000, o número de combinações úteis foi da ordem de 6500 (de 40000 possíveis) e para *Precision* de 10000, esse valor saltou para 13382 (de 80000 possíveis).

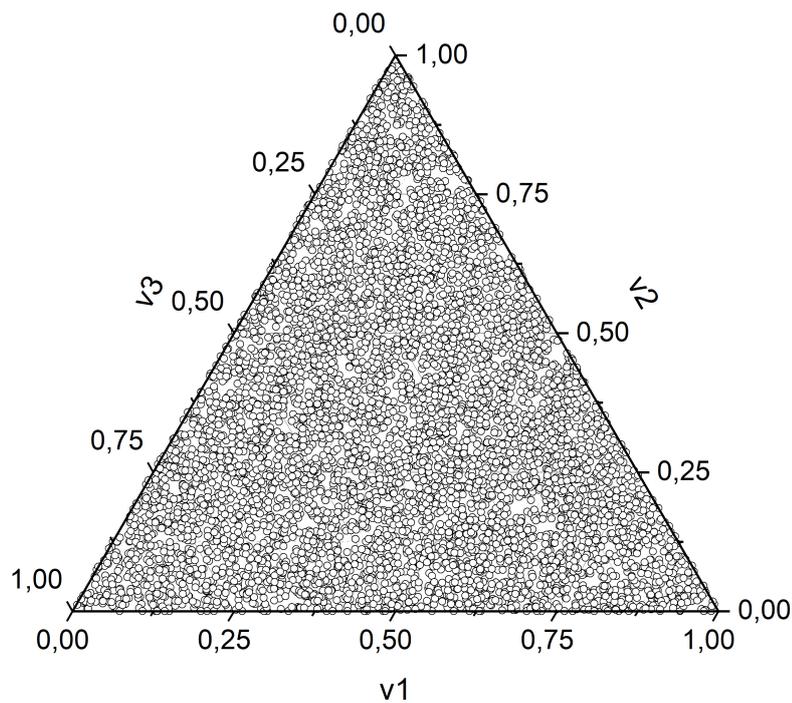
Nitidamente o espaço amostral fica melhor preenchido quando os valores de *Precision* são altos (> 1000). Entretanto, tal aumento não comprometeu de forma significativa a velocidade de processamento dos cálculos no computador utilizado. Para valores de *Precision* de 1000, 5000 e 10000, o tempo de processamento foi de 3, 10 e 32 segundos, respectivamente, para um PC do tipo *desktop* com Processador Intel(R) Core(TM) i7-3770 CPU @ 3,40 GHz e 12,0 GB de RAM instalada.

Figura 11 – Combinações dos valores das variáveis de 1 a 3 (círculos) para valor de *Precision* igual a 1000.



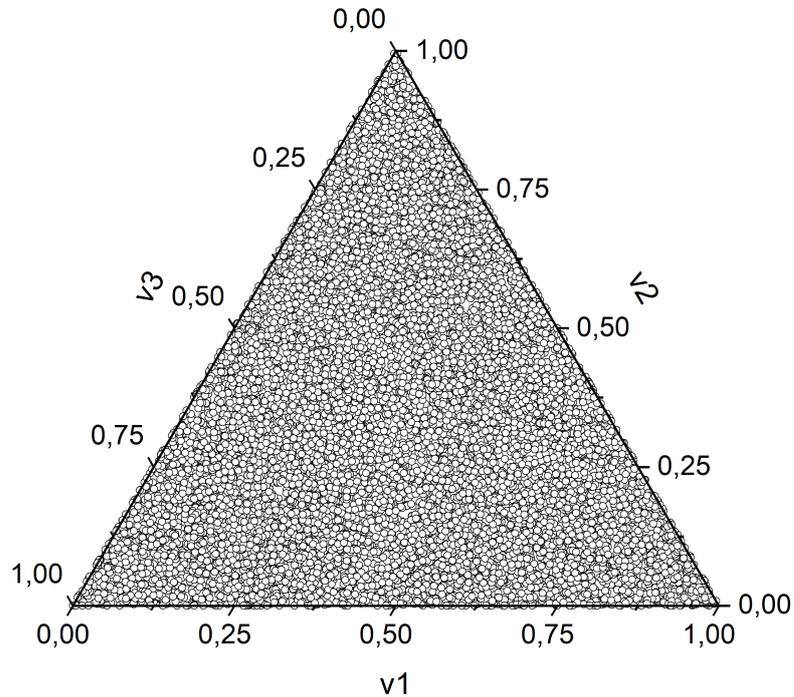
Fonte: Compilação do autor.

Figura 12 – Combinações dos valores das variáveis de 1 a 3 (círculos) para valor de *Precision* igual a 5000.



Fonte: Compilação do autor.

Figura 13 – Combinações dos valores das variáveis de 1 a 3 (círculos) para valor de *Precision* igual a 10000.



Fonte: Compilação do autor.

5.2 Situações iniciais para o algoritmo *mix_target2*

Após o estabelecimento do espaço amostral, é necessário verificar a melhor condição de trabalho, ou seja, a combinação de variáveis 1 a 3 que permite a obtenção dos maiores valores de respostas. Assim, na Etapa 4 (Figura 14) temos 3 situações que podem ser observadas: (i) Somente um dos coeficientes lineares (b_1 , b_2 e b_3) é significativo, (ii) Somente um dos coeficientes quadráticos é significativo (b_{12} , b_{13} e b_{23}) e (iii) somente o coeficiente terciário é significativo (b_{123}).

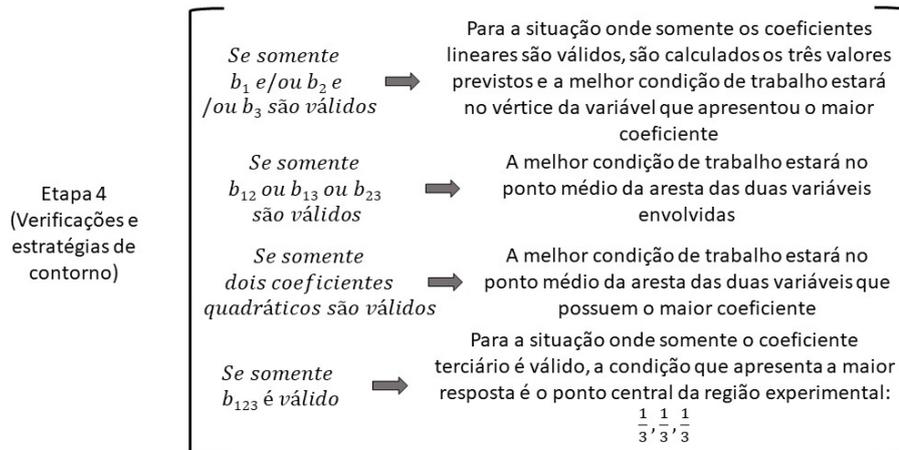
No caso onde somente um dos coeficientes lineares é significativo, por exemplo, b_1 , como mostra o exemplo hipotético da Função 5.1:

$$f(v_1, v_2, v_3) = 1000v_1 \quad (5.1)$$

Teremos que a região ótima de trabalho é a aresta onde $v_1 = 1$ e $v_2=v_3=0$ (1, 0, 0). No exemplo da Função 5.2, onde o maior coeficiente é o b_2 , teremos a região ótima de trabalho na aresta onde $v_2 = 1$ e $v_1=v_3=0$:

$$f(v_1, v_2, v_3) = 1000v_1 + 5000v_2 \quad (5.2)$$

Figura 14 – Quarta etapa de funcionamento do algoritmo.



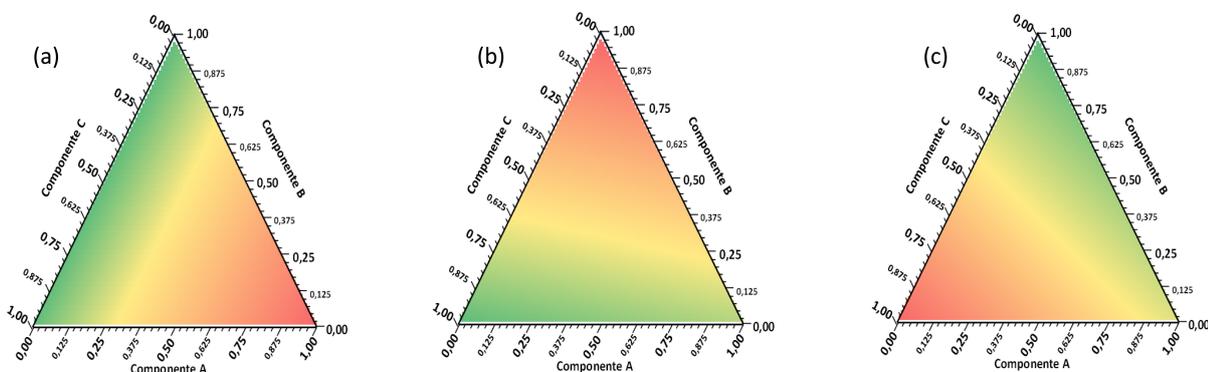
Fonte: Compilação do autor.

Para a situação onde temos 3 coeficientes lineares significativos (Função 5.3) também teremos como região ótima de trabalho a aresta relativa ao maior coeficiente. A Função 5.3 ilustra esta situação e o máximo é observado quando temos $v_3 = 1$ e $v_1=v_2=0$.

$$f(v_1, v_2, v_3) = 1000v_1 + 500v_2 + 2000v_3 \quad (5.3)$$

A Figura 15 exemplifica por meio de gráficos de contorno as situações mostradas nas Funções 5.1 (Figura 15a), 5.2 (Figura 15b) e 5.3 (Figura 15c). Estas figuras foram obtidas utilizando o arquivo *Template Misturas* preparado no Microsoft Excel® e a área em vermelho mostra a região com maior resposta prevista pelo modelo. Mais detalhes sobre o *Template Misturas* serão apresentados nas seções subsequentes desta dissertação.

Figura 15 – Gráficos de contorno obtidos para os modelos hipotéticos das Funções 5.1 (a), 5.2 (b) e 5.3 (c). A região em vermelho representa a maior resposta obtida.



Fonte: Compilação do autor.

Quando um coeficiente quadrático é válido (b_{12} ou b_{13} ou b_{23}), a região ótima de

trabalho estará localizada no ponto médio da aresta das duas variáveis envolvidas. Para o caso da Função 5.4, por exemplo, temos que a região que apresentará a maior resposta é a aresta que está no ponto médio das variáveis 1 e 2 ($v_1 = v_2 = \frac{1}{2}$ e $v_3 = 0$).

$$f(v_1, v_2, v_3) = 1000v_1v_2 \quad (5.4)$$

Para situações onde somente dois coeficientes quadráticos são válidos (b_{12} e b_{13} ou b_{12} e b_{23} ou b_{13} e b_{23}) teremos que a melhor condição de trabalho estará localizada também no ponto médio da aresta das duas variáveis que apresentarem o maior coeficiente. Para o exemplo da Função 5.5 teremos que a região onde $v_1 = v_3 = \frac{1}{2}$ e $v_2 = 0$, é a que apresenta a maior resposta:

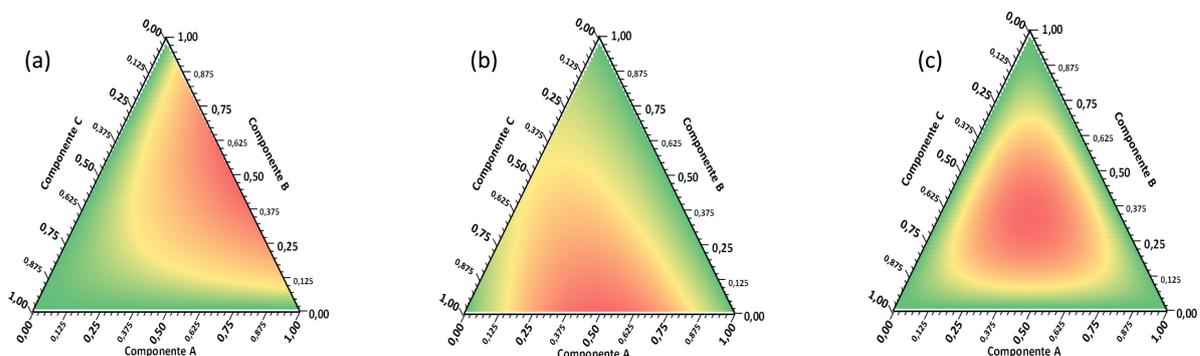
$$f(v_1, v_2, v_3) = 1000v_1v_3 + 500v_2v_3 \quad (5.5)$$

Quando somente o coeficiente terciário (b_{123}) for válido (ver exemplo na Função 5.6), a maior resposta é o ponto central do planejamento de misturas, ou seja, $v_1 = v_2 = v_3 = \frac{1}{3}$.

$$f(v_1, v_2, v_3) = 1000v_1v_2v_3 \quad (5.6)$$

A Figura 16 exemplifica por meio de gráficos de contorno as situações mostradas nas Funções 5.4 (Figura 16a), 5.5 (Figura 16b) e 5.6 (Figura 16c) e a área em vermelho mostra a região com maior resposta prevista pelo modelo.

Figura 16 – Gráficos de contorno obtidos para os modelos hipotéticos das Funções 5.4 (a), 5.5 (b) e 5.6 (c). A região em vermelho representa a maior resposta obtida.



Fonte: Compilação do autor.

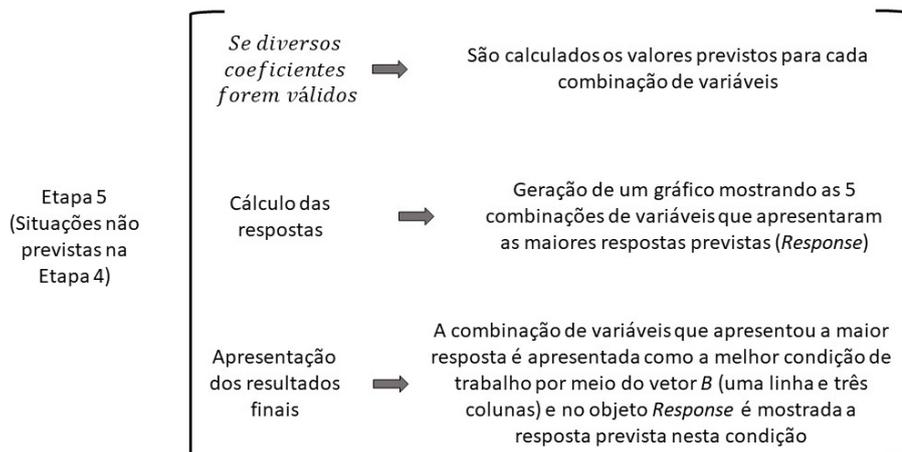
5.3 Funções mais complexas para o algoritmo *mix_target2*

Na maioria dos modelos gerados a identificação da melhor condição de trabalho (maior resposta) não é óbvia e de fácil identificação como exemplificado nas situações anteriores (Funções de 5.1 até 5.6). Assim, a Etapa 5 do algoritmo mostra uma situação onde a melhor região de trabalho está no interior do gráfico da mistura ternária. A Função 5.7 mostra novamente o exemplo que foi ilustrado anteriormente (ver Tabela 1 e Função 3.17):

$$f(v_1, v_2, v_3) = 550v_1 + 345v_2 + 268v_3 + 690v_1v_2 + 9243v_1v_2v_3 \quad (5.7)$$

Com o modelo proposto são calculados os valores de respostas previstas com os dados gerados na Figura 11 onde a *Precision* foi 1000 (ver círculos no interior do triângulo). A combinação de níveis que gerar a maior resposta é mostrada na saída *B*, como sendo a melhor condição de trabalho. Além disso, a resposta prevista é mostrada no objeto *Response*. A Figura 17 mostra a descrição dos processos que ocorrem na Etapa 5. Os procedimentos descritos até aqui evitam a análise mostrada na seção 3.4 *Identificação dos Pontos Críticos em um Planejamento de Misturas*.

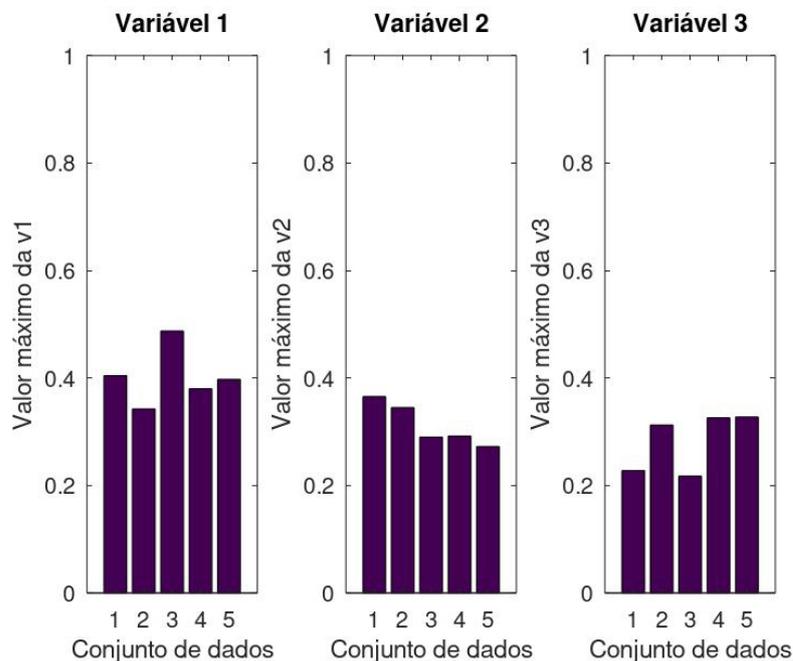
Figura 17 – Quinta etapa de funcionamento do algoritmo.



Fonte: Compilação do autor.

Ao final do algoritmo é gerada uma figura que mostra as cinco melhores combinações, ou seja, as combinações das 3 variáveis investigadas que apresentaram as maiores respostas. Estas combinações são apresentadas na Figura 18 e o objetivo é auxiliar o operador na identificação dos intervalos dos valores das variáveis 1 a 3 que oferecem os melhores resultados.

Figura 18 – Visualização dos níveis das variáveis 1 até 3 que apresentaram os cinco melhores resultados (*Precision* = 1000).



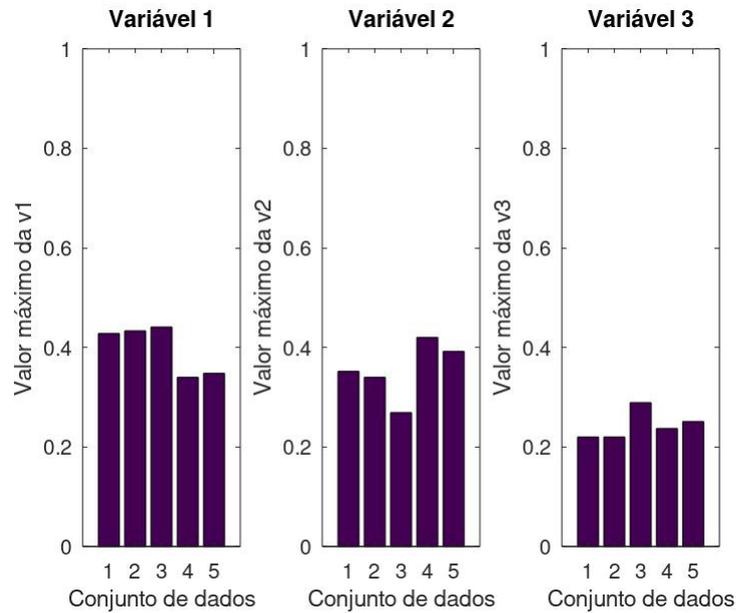
Fonte: Compilação do autor.

Com as informações da Figura 18 nota-se que os 5 maiores valores para a variável 1 variaram de 0,343 até 0,488. Para as variáveis 2 e 3 esses intervalos foram de 0,273 até 0,366 e 0,218 até 0,328, respectivamente. Ao repetir o mesmo cálculo para valores de *Precision* de 5000 (Figura 12) e 10000 (Figura 13), percebe-se que os intervalos de valores não diferiram muito como mostra a Tabela 5 e as Figuras 19 (*Precision* de 5000) e 20 (*Precision* de 10000).

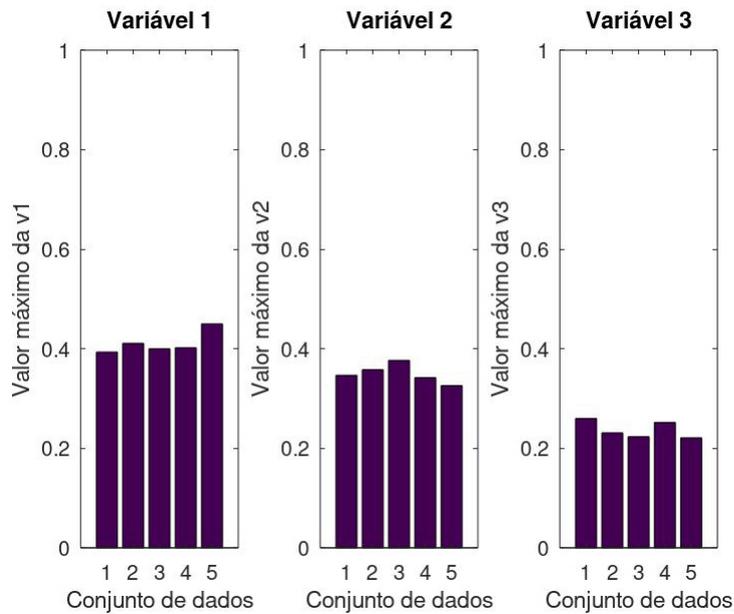
Tabela 5 – Intervalos das variáveis 1 a 3 quando os valores de *Precision* foram 5000 e 10000

<i>Precision</i>	Intervalos da v_1	Intervalos da v_2	Intervalos da v_3
5000	0,340 a 0,441	0,269 a 0,420	0,220 a 0,289
10000	0,393 a 0,450	0,326 a 0,377	0,221 a 0,260

Fonte: elaboração do autor

Figura 19 – Visualização dos níveis das variáveis 1 até 3 que apresentaram os cinco melhores resultados (*Precision* = 5000).

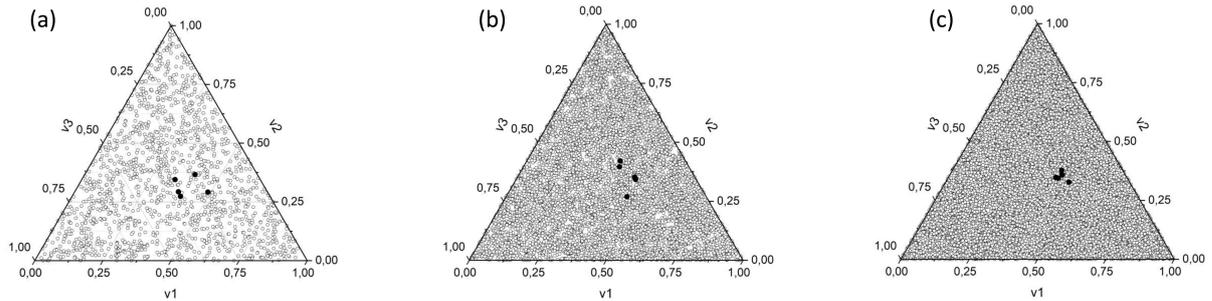
Fonte: Compilação do autor.

Figura 20 – Visualização dos níveis das variáveis 1 até 3 que apresentaram os cinco melhores resultados (*Precision* = 10000).

Fonte: Compilação do autor.

A Figura 21 mostra uma ilustração da posição das 5 melhores condições experimentais (ver círculos preenchidos) das 3 variáveis estudadas quando os valores de *Precision* foram 1000 (Figura 21a), 5000 (Figura 21b) e 10000 (Figura 21c). É possível notar que os resultados mais agrupados são obtidos quando a *Precision* está ajustada para 10000.

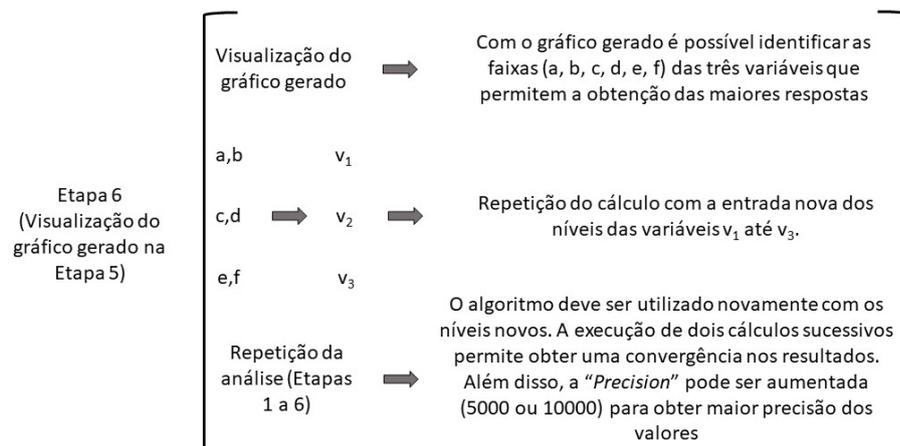
Figura 21 – Visualização das cinco melhores combinações de resultados quando a *Precision* for igual a 1000 (a), 5000 (b) e 10000 (c).



Fonte: Compilação do autor.

Como mencionado anteriormente, com as Figuras 18, 19 e 20 é possível observar os intervalos das variáveis de 1 a 3 que apresentam os melhores resultados. Desta forma, na Etapa 6 é possível identificar esse intervalo e preparar novamente o algoritmo *mix_target2* para uma repetição dos cálculos. Essas informações podem ser vistas na Figura 22.

Figura 22 – Sexta etapa de funcionamento do algoritmo.



Fonte: Compilação do autor.

Os cálculos foram repetidos com as informações dos intervalos descritos anteriormente (ver Tabela 5) e mais detalhes podem ser visualizados na Tabela 6. Nota-se que o tempo de processamento variou de 5 a 157s para valores de *Precision* de 1000 e 10000, respectivamente. Além disso, foi possível obter mais combinações úteis das variáveis ($v_1 + v_2 + v_3 \leq 1$) ao empregar um domínio menor para as 3 variáveis.

Tabela 6 – Características dos resultados após repetição dos cálculos com intervalos mais estreitos.

<i>Precision</i>	Tempo de processamento (s)	Número de combinações úteis ($v_1+v_2+v_3 \leq 1$)
1000	5	3591
5000	81	23746
10000	157	30085

Fonte: elaboração do autor

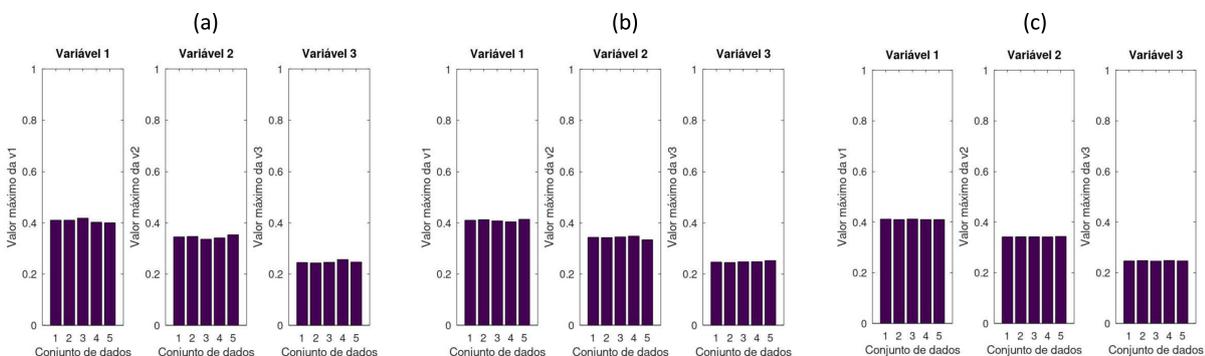
Com a utilização de intervalos mais estreitos das 3 variáveis foi possível obter condições ótimas mais definidas para os três valores de *Precision* empregados. A Tabela 7 mostra um resumo dos dados, sendo possível observar uma concordância entre eles.

Tabela 7 – Intervalos das variáveis 1 a 3 quando os valores de *Precision* foram 1000, 5000 e 10000 após a repetição dos cálculos.

<i>Precision</i>	Intervalos da v_1	Intervalos da v_2	Intervalos da v_3
1000	0,400 a 0,418	0,336 a 0,353	0,243 a 0,257
5000	0,406 a 0,415	0,340 a 0,344	0,241 a 0,251
10000	0,410 a 0,412	0,342 a 0,343	0,246 a 0,248

Fonte: elaboração do autor

A Figura 23 mostra os gráficos gerados após o término dos cálculos para os valores de *Precision* de 1000 (Figura 23a), 5000 (Figura 23b) e 10000 (Figura 23c). Os resultados apontados na Figura 23 mostram também que o aumento do valor de *Precision* não contribui de forma significativa para a precisão dos resultados finais.

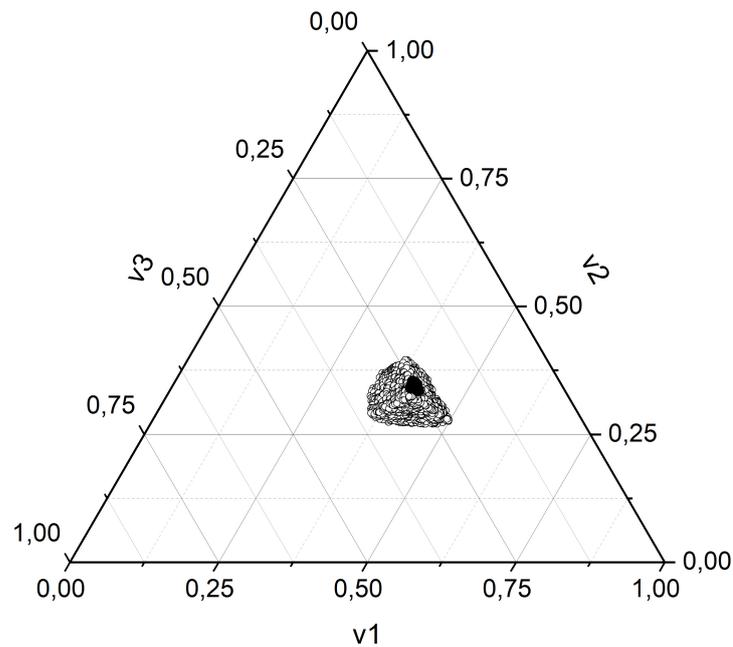
Figura 23 – Visualização dos parâmetros finais das 3 variáveis após o cálculo com valores de *Precision* de 1000 (a), 5000 (b) e 10000 (c).

Fonte: Compilação do autor.

A Figura 24 mostra o agrupamento de combinações de valores das 3 variáveis após repetição quando a *Precision* foi igual a 1000. É possível notar um agrupamento dos

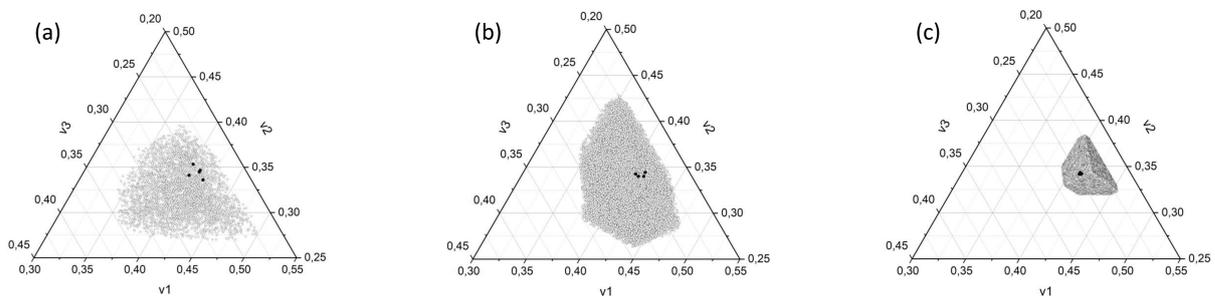
valores das 3 variáveis nas imediações dos níveis apontados na Tabela 7 (ver *Precision* 1000). A Figura 25 mostra uma ampliação da região onde está a melhor condição de trabalho (Figura 25a). O mesmo cálculo foi executado para as *Precision* 5000 e 10000 e as Figuras 25b e 25c mostram, respectivamente os resultados. Os eixos das variáveis (v_1 , v_2 e v_3) foram limitados aos intervalos $[0, 30, 0, 55]$ (v_1), $[0, 25, 0, 50]$ (v_2) e $[0, 20, 0, 45]$ (v_3) para facilitar a comparação dos resultados.

Figura 24 – Combinações dos valores das variáveis de 1 a 3 (círculos) para valor de *Precision* igual a 1000 após o refinamento do cálculo.



Fonte: Compilação do autor.

Figura 25 – Ampliação da região ótima de trabalho após a repetição dos cálculos com valores de *Precision* de 1000 (a), 5000 (b) e 10000 (c).



Fonte: Compilação do autor.

Com a finalização da repetição dos cálculos temos condições de identificar a melhor condição de trabalho. A Tabela 8 mostra a condição obtida (valores de B no algoritmo), bem como a resposta (*Response*) teórica calculada para os três valores de *Precision*.

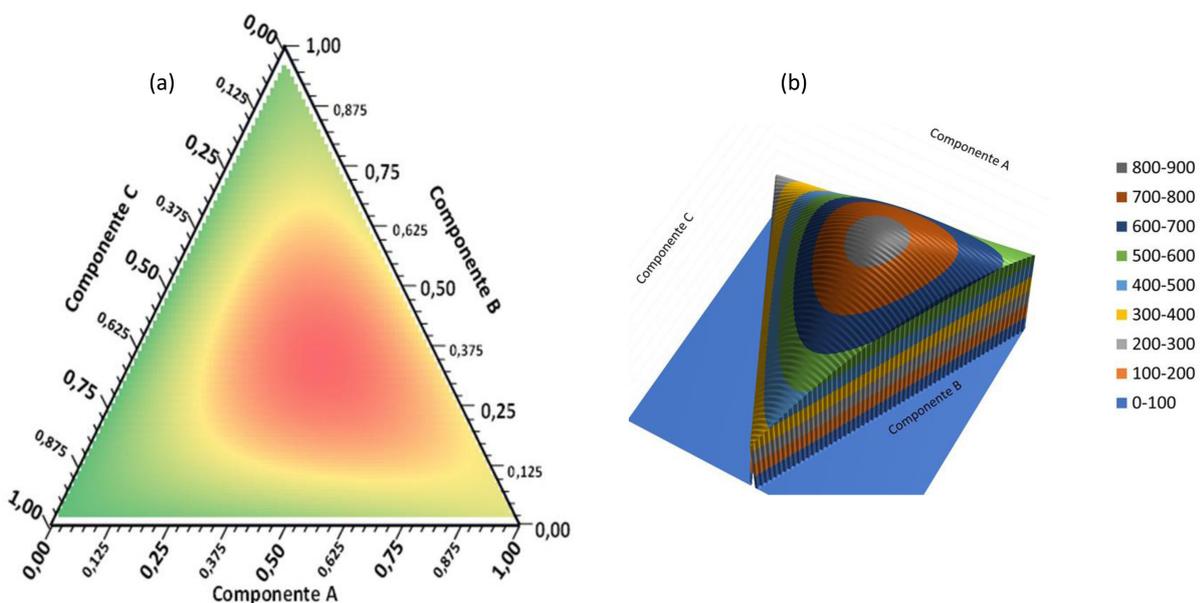
Tabela 8 – Condições ótimas de trabalho (B) e resposta teórica obtida ($Response$) para os três valores de $Precision$ empregados

$Precision$	Condição ótima de trabalho (vetor B)	Resposta teórica obtida ($Response$)
1000	0,410; 0,345; 0,245	828,104
5000	0,410; 0,340; 0,250	828,105
10000	0,412; 0,342; 0,247	828,144
Derivada parcial	0,411; 0,342; 0,247	$\approx 828,13$

Fonte: elaboração do autor

Os resultados obtidos (Tabela 8) são claramente semelhantes. A Tabela 8 mostra também o resultado obtido pelo uso de derivadas parciais (seção 3.4) e os valores são próximos. A mesma análise foi efetuada no arquivo *Template Misturas*, sendo possível observar uma condição ótima de trabalho de 0,411, 0,342 e 0,247 para v_1 , v_2 e v_3 , respectivamente. A resposta teórica (\hat{y}) obtida foi de 828,131 e o *Template Misturas* utiliza a função Solver® do Excel® para identificar o máximo dos modelos de regressão calculados. A descrição completa, comandos e cálculos do algoritmo *mix_target2* podem ser observados no Apêndice D desta dissertação. A Figura 26 mostra o gráfico de contorno (Figura 26a) e a superfície de resposta obtidos (Figura 26b).

Figura 26 – Gráficos de contorno (a) e de superfície (b) obtidos no *Template Misturas* do Excel®. A região em vermelho no gráfico de contorno (a) representa a maior resposta obtida.



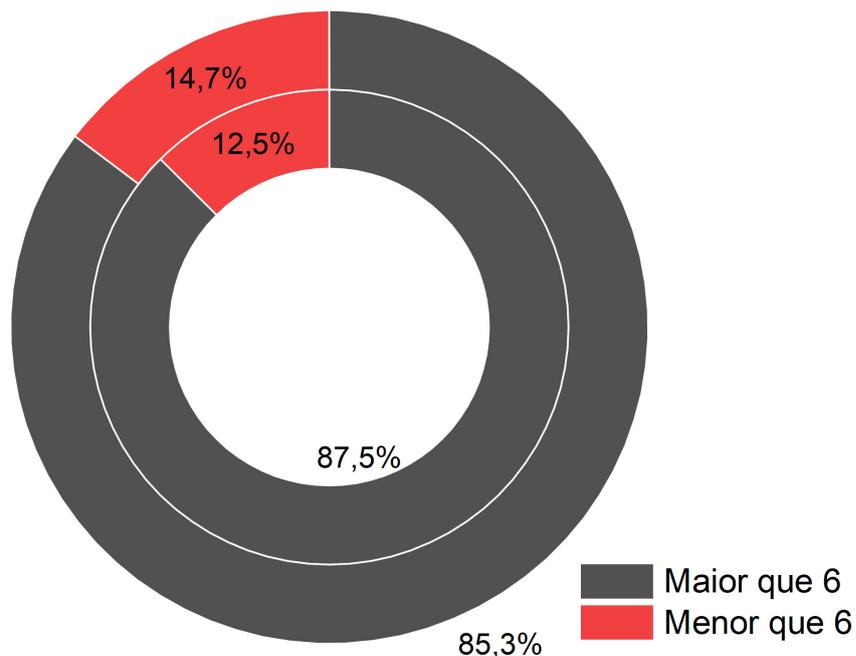
Fonte: Compilação do autor.

6 Resultados e discussão, Disciplinas sobre planejamento fatorial

6.1 Desempenho dos alunos nas disciplinas sobre planejamento fatorial

Nesta parte da dissertação serão mostrados os resultados relacionados com a experiência didática no ensino de planejamento fatorial antes e durante a pandemia de Covid-19. Durante a discussão será apresentado um paralelo entre as percepções dos alunos de graduação e pós-graduação, bem como uma análise estatística dos resultados acumulados ao longo dos anos. Antes do período pandêmico a disciplina Introdução à Quimiometria para a graduação havia sido oferecida apenas uma vez em 2019. Nesse período o número de alunos regularmente matriculados foi 16. Durante a pandemia de Covid-19 a mesma disciplina foi oferecida duas vezes de forma remota (2020 e 2021) e o número total de alunos matriculados foi de 68. A Figura 27 mostra o panorama geral das notas obtidas pelos alunos de graduação.

Figura 27 – Notas obtidas pelos alunos de graduação na disciplina Introdução à Quimiometria antes (círculo interno, n = 16) e durante a pandemia (círculo externo, n = 68).



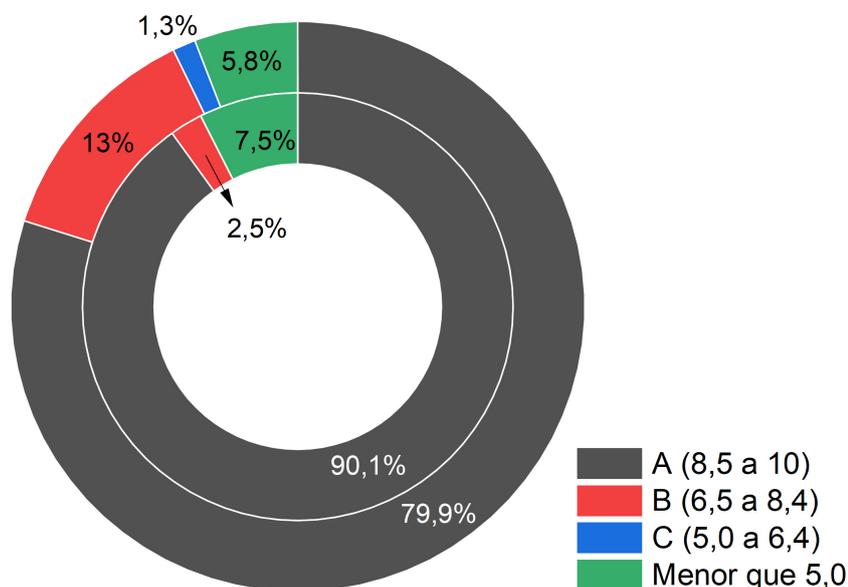
Fonte: Compilação do autor.

Nos cursos de graduação da UFSCar é necessário obter uma nota final $\geq 6,0$

para ser aprovado. Além disso, é necessário 75% de frequência. Com as informações da Figura 27 nota-se que a porcentagem de alunos aprovados foi superior a 85% em ambos períodos. Com os dados apresentados foi executada uma Anova e não foi observada diferença significativa entre ambos períodos ($p - valor = 0,71$). Este teste estatístico permite verificar se a diferença entre os dois grupos (antes e durante) é maior que a diferença dentro dos grupos [34].

No caso da disciplina Quimiometria I (pós-graduação), no período pré-pandêmico (presencial) foram coletadas informações de 6 edições da disciplina (intervalo de tempo de 2014 a 2019). A quantidade de alunos regularmente matriculados nesse período foi de 161. No período pandêmico (remoto) a mesma disciplina foi oferecida 3 vezes (duas em 2020 e uma vez em 2021) e o total de alunos matriculados foi de 154. Para a pós-graduação são utilizados conceitos para avaliar os alunos. Os conceitos A (notas entre 8,5 a 10), B (6,5 a 8,4) e C (5,0 e 6,4) permitem o aproveitamento dos créditos. A Figura 28 mostra os resultados computados para ambos os períodos.

Figura 28 – Conceitos obtidos pelos alunos de pós-graduação na disciplina Quimiometria I antes (círculo interno, $n = 161$) e durante a pandemia (círculo externo, $n = 154$).



Fonte: Compilação do autor.

As porcentagens de alunos com conceito A foram da ordem de 80 a 90% para os períodos antes e durante a pandemia, respectivamente. Na Figura 28 percebe-se que o estrato de conceito B para o período pandêmico (13%) é maior quando comparado com o pré-pandêmico (2,5%). Mesmo com essas distinções, não foi observada uma diferença significativa entre ambos os períodos ($p - valor = 0,29$). É importante notar também que, tanto na graduação e na pós-graduação, durante a pandemia de Covid-19 foram realizados mais atividades de avaliação (entre 5 e 7 testes) e isso pode ter possibilitado um

maior refinamento das notas com melhor distribuição das pontuações entre os diferentes conceitos. A Tabela 9 mostra uma visão geral dos dados (média \pm desvio padrão, mediana e moda) dos dados referentes às avaliações.

Tabela 9 – Alguns parâmetros estatísticos das notas e conceitos obtidos pelos alunos de graduação e pós-graduação antes e durante a pandemia.

Parâmetro	Graduação		Pós-graduação	
	Antes	Durante	Antes	Durante
Média \pm desvio padrão	8,5 \pm 3,3	8,2 \pm 2,6	9,0 \pm 2,6	8,7 \pm 1,8
Mediana	10	9,4	10	9,3
Moda	10	9,8	10	10
n	16	68	161	154

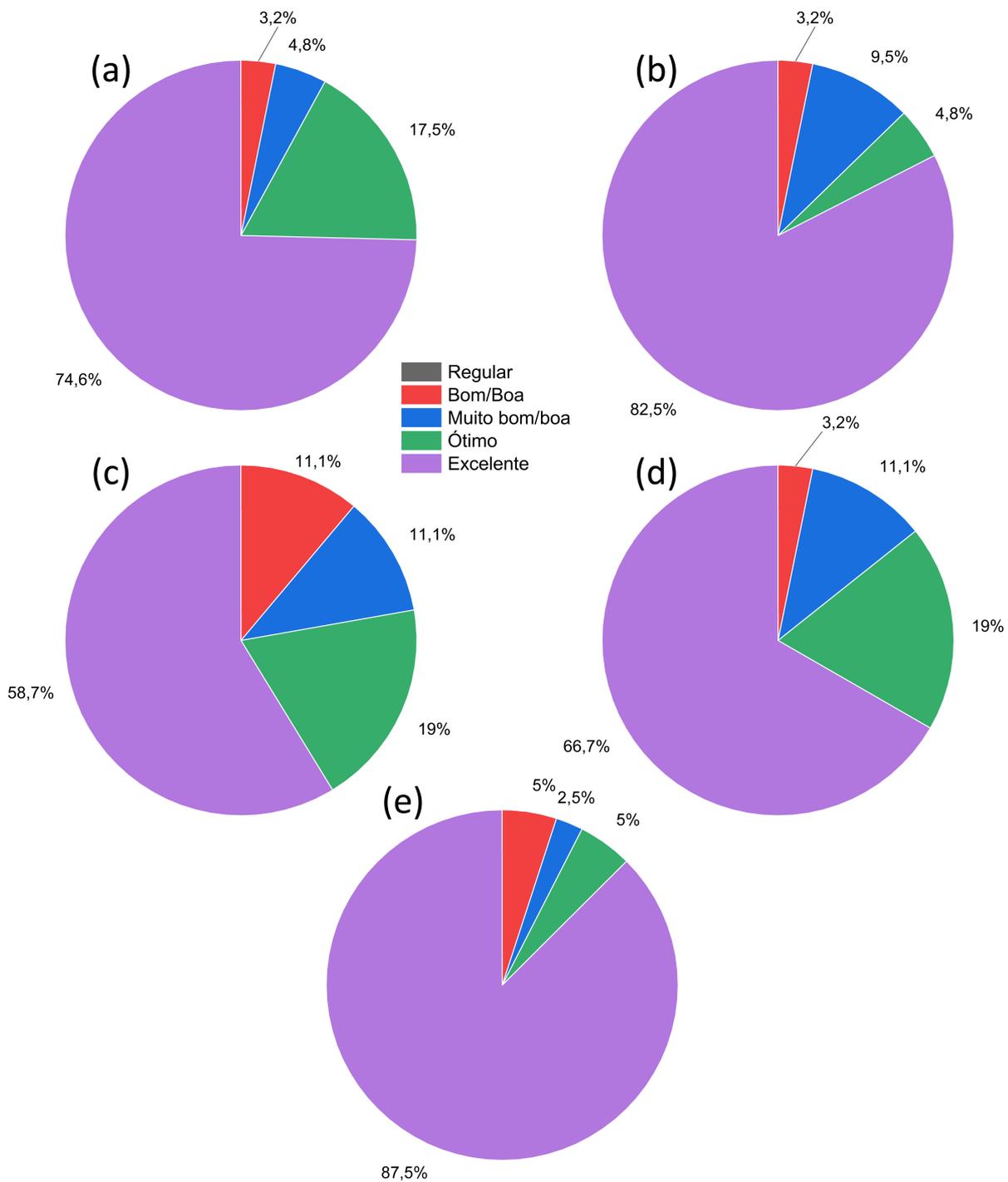
Fonte: elaboração do autor

6.2 Opinião dos alunos sobre o ensino de planejamento fatorial

6.2.1 Cenário antes da pandemia

Como afirmado anteriormente, a disciplina Quimiometria 1 (pós-graduação) é oferecida de forma regular desde 2007. Nesta parte da dissertação serão mostrados resultados apenas da pós-graduação, pois a partir de 2019 foi iniciada a disciplina Introdução à Quimiometria para os alunos de graduação. Antes do desencadeamento da pandemia de Covid-19, a avaliação da disciplina pelos alunos foi realizada ao final do semestre letivo. A primeira avaliação foi em 2014 e a última em 2019. Das questões apresentadas aos alunos, apenas cinco delas (1, 2, 3, 4 e 6, ver Tabela 3) puderam ser utilizadas para comparar com as questões apresentadas durante a pandemia. A Figura 29 mostra um perfil geral da opinião dos alunos. É possível notar que para todas as questões, as avaliações foram bastante positivas. Vale destacar a questão 6 (Figura 29e) que reflete a opinião dos alunos sobre os vídeos gravados e disponibilizados. Os vídeos discutidos aqui são aqueles presentes no canal do YouTube e também alguns que foram gravados para mostrar tópicos pontuais da disciplina. As questões de 1 (Figura 29a) até 4 (Figura 29d) contaram com 63 alunos respondentes. Já a questão 6 (Figura 29e) foi respondida por 40 alunos do ano de 2019. Nenhum aluno apontou a opção *regular* como resposta a estas questões.

Figura 29 – Opinião dos alunos de pós-graduação sobre as questões 1 (a), 2 (b), 3 (c), 4 (d) e 6 (e) no período antes da pandemia (2014 e 2019).



Fonte: Compilação do autor.

6.2.2 Comparação dos cenários: antes e durante a pandemia

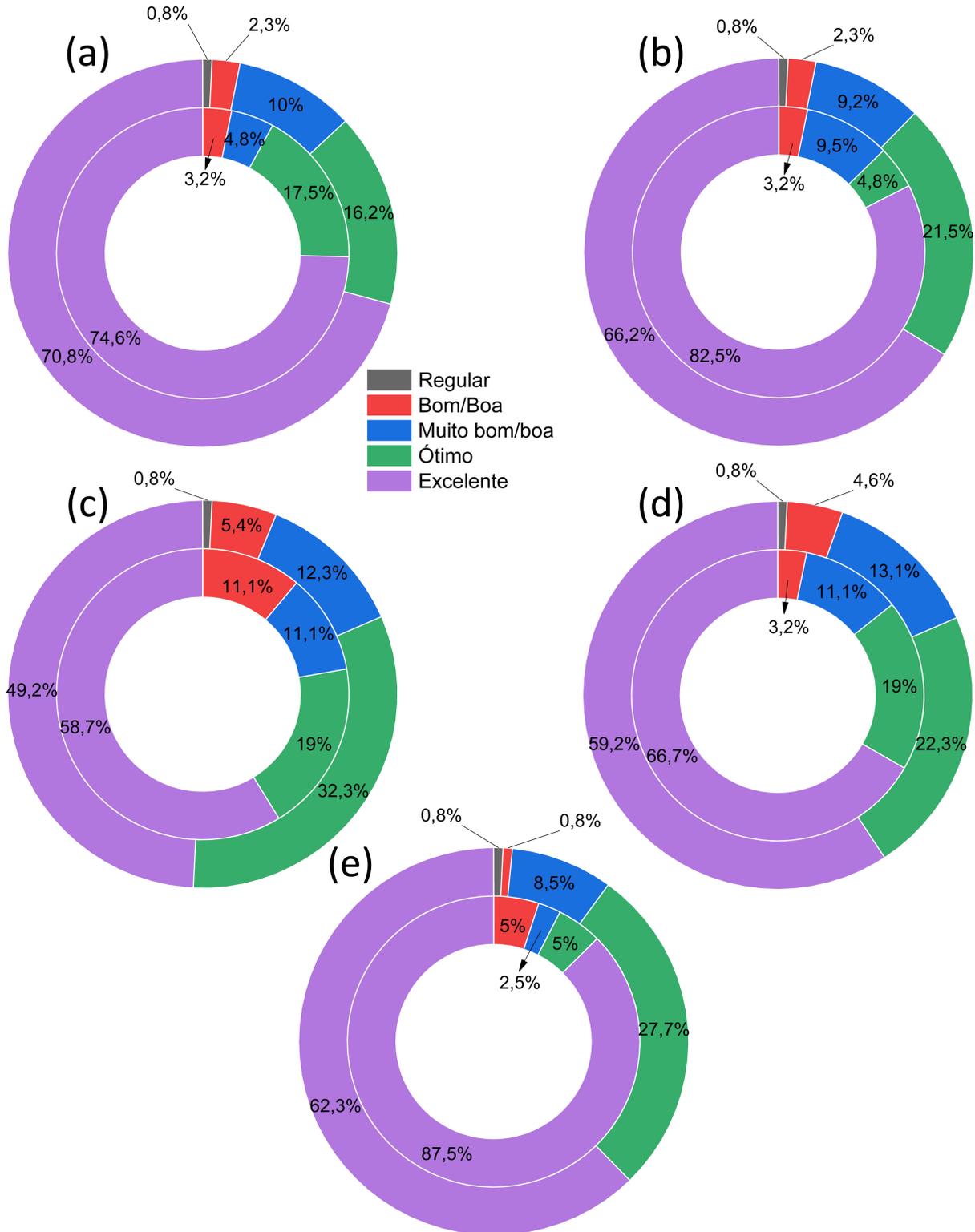
As cinco questões mostradas na Figura 29 (antes da Pandemia) foram comparadas com o período pandêmico (2020 até 2021). Durante os anos de 2020 e 2021 as questões foram respondidas por 193 alunos da pós-graduação. A Figura 30 mostra os resultados coletados. Os dados apresentados nos círculos internos e externos da Figura 30 se referem aos períodos antes e durante a pandemia, respectivamente. Os resultados são bastante similares e no caso da questão 1 (Figura 30a), por exemplo, cerca de 74% (antes) e 70% (durante) dos alunos opinaram que de forma geral a disciplina foi excelente. O mesmo panorama foi observado para as demais questões com uma tendência de mudança para o estrato "Ótimo". Com os dados coletados foi realizada uma análise de variância para verificar se a diferença entre os dois grupos (antes e durante) é maior que a diferença dentro dos grupos [34]. Em todos os casos não foi observada diferenças estatisticamente significativas ao nível de confiança de 95% ($p - valor = 0,05$). Apenas na questão 6 (Figura 30e) foi notada uma leve diferença ($p - valor = 0,07$) devido ao aumento do estrato "Ótimo". Para os demais casos, os valores de $p - valor$ variaram de 0,18 (questão 2, Figura 30b) até 0,92 (questão 3, Figura 30c).

6.2.3 Avaliação das percepções dos alunos durante a pandemia

Antes de apresentar a avaliação das percepções dos alunos é importante mencionar que os vídeos preparados tiveram cerca de 30000 visualizações entre os anos de 2020 e 2021. Esse valor correspondeu a cerca de 2500 horas (aproximadamente 100 dias de exibição). Além disso, foi questionado aos alunos quantas vezes eles assistiam (ou pausavam) os vídeos para reter os conteúdos apresentados. De forma, geral a maioria dos alunos (85%) necessitavam de um tempo de duas à três vezes maior do conteúdo inicialmente gravado.

Durante a pandemia de Covid-19 as disciplinas de quimiometria para a graduação e pós-graduação foram oferecidas em 2 e 3 semestres, respectivamente. Além disso, nestes cinco oferecimentos foi aplicado o questionário mostrado no Apêndice B. A Figura 31 mostra os resultados para as seis primeiras questões (ver Tabela 3). Estas questões possuem 5 opções possíveis de respostas (*regular*, *bom/boa*, *muito bom/boa*, *ótimo* e *excelente*). Os dados apresentados nos círculos interno e externo são referentes aos alunos de graduação ($n = 32$) e pós-graduação ($n = 130$), respectivamente. É possível notar que as percepções dos alunos são bastante parecidas para essas 6 questões. Tal observação também foi comprovada por meio de uma análise de variância [34]. Os $p - valores$ para estas 6 questões variaram de 0,11 para a questão 4 até 0,84 para a questão 2. Nestas 6 primeiras questões, em todos os casos a opção "Excelente" foi escolhida por mais de 50% dos alunos. Outro ponto de destaque é que nenhum aluno de graduação (círculos internos) apontou a opção "Regular" para as questões 1 (Figura 31a), 2 (Figura 31b), 4 (Figura 31d) e 6 (Figura 31f).

Figura 30 – Opinião dos alunos de pós-graduação sobre as questões 1 (a), 2 (b), 3 (c), 4 (d) e 6 (e) no período antes (círculo interno, n = 63) e durante a pandemia (círculo externo, n = 130). Para a questão 6 (e) o número de respondentes para antes da pandemia foi de 40.



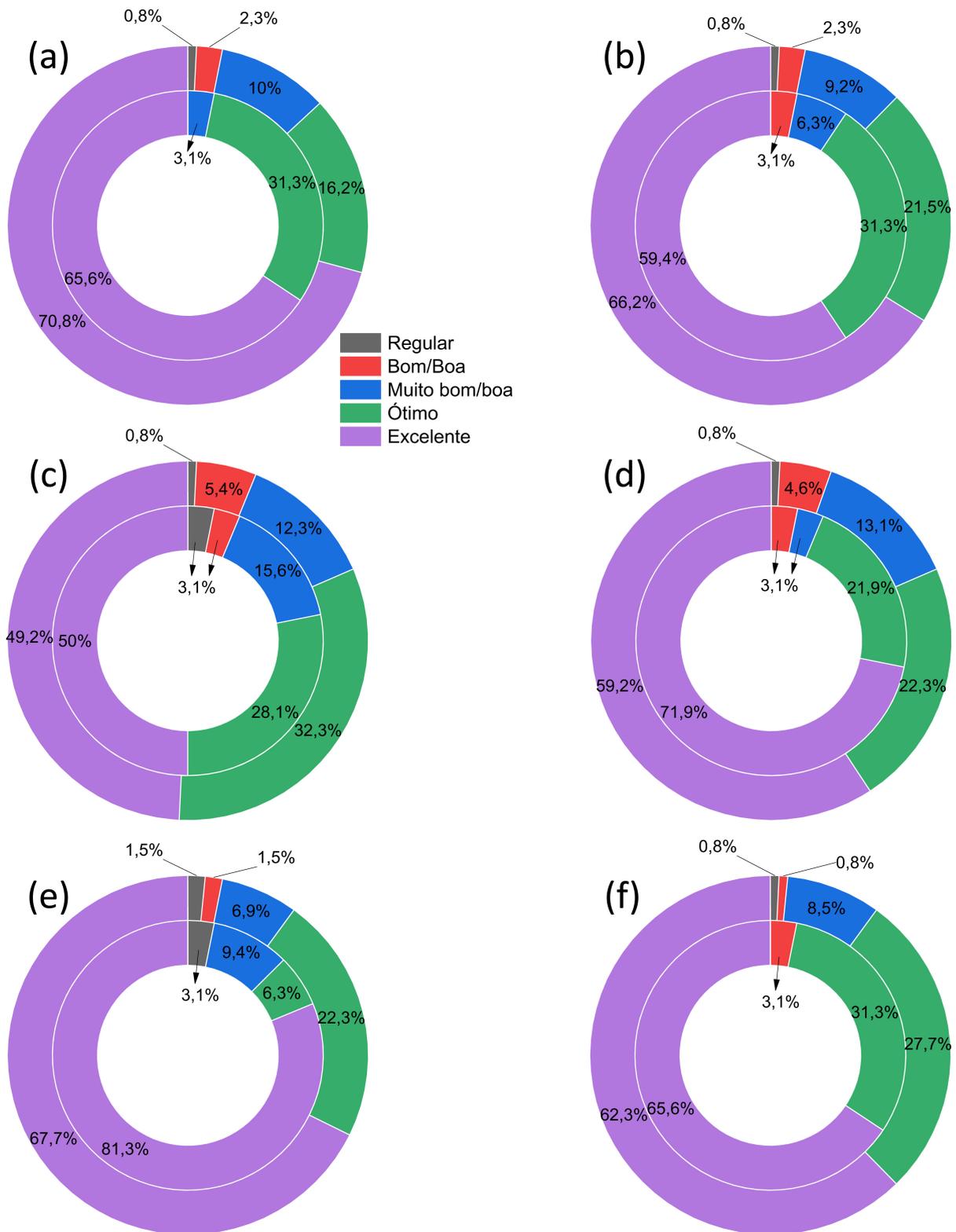
Fonte: Compilação do autor.

Além das disciplinas de planejamento fatorial, foram oferecidas também outras, tais como Química Analítica Geral (graduação) e Tópicos Especiais (para pós-graduação). As mesmas avaliações (Tabela 3, Apêndice B) dos alunos foram coletadas e os resultados finais são distintos. Tais dados de outras disciplinas não serão discutidos aqui, mas podem ser acompanhados no seguinte link: https://www.dropbox.com/s/ftwsho6q3hrqcw2/relatorio_enpe.pdf?dl=0.

Para o caso específico da questão 7 (Figura 32a) é possível notar que a grande maioria dos alunos opinou que a duração dos vídeos é adequada. Menos de 10% dos pós-graduandos acharam os vídeos muito curtos ou muito longos. As percepções dos alunos de graduação e pós foram semelhantes estatisticamente ($p - valor = 0,68$). A duração média dos vídeos gravados foi da ordem de 10 minutos. Já nos encontros síncronos, a duração média dos vídeos foi da ordem de 30 a 50 minutos. Geralmente os encontros síncronos duravam cerca de 90 a 120 minutos e eram constituídos por 2 a 4 vídeos. Esses vídeos eram agrupados por assuntos: (i) dúvidas dos alunos, (ii) resolução dos testes e (iii) tópicos futuros.

A questão 8 trás uma das informações mais importantes sobre a avaliação efetuada pelos alunos. Esta questão é sobre a percepção do nível de aprendizado dos alunos (Tabela 3, Apêndice B). Para os alunos de graduação (círculo interno na Figura 32b) nota-se que nenhum aluno apontou que o aprendizado ficou abaixo da expectativa. De forma geral, cerca de 50% dos alunos informaram que o nível de aprendizado ficou dentro da expectativa. Para os alunos de pós-graduação, menos de 2% afirmaram que o nível de aprendizado ficou abaixo da expectativa e os demais optaram por escolher dentro (50%) ou acima (48,5%) da expectativa. Novamente não foi notada diferença significativa entre as opiniões dos alunos de graduação e pós-graduação ($p - valor = 0,76$).

Figura 31 – Opinião dos alunos de graduação (círculo interno) e pós-graduação (círculo externo) sobre as questões 1 (a), 2 (b), 3 (c), 4 (d), 5 (e) e 6 (f).



Fonte: Compilação do autor.

A questão 10 indaga se os alunos fariam outras disciplinas remotas como as que estão sendo oferecidas durante a pandemia. É possível notar que os alunos de pós-graduação (círculo externo na Figura 32c) são mais receptivos à essa proposta e cerca de 85% afirmaram que sim. Neste ponto é importante afirmar que diversos alunos de 17 estados da federação (AL, BA, CE, ES, GO, MA, MG, MT, PA, PB, PE, PI, PR, RJ, RN, RO, RS, SP e TO) e alguns de outros países da América do Sul (Argentina, Colômbia e Peru) também se matricularam na disciplina da pós-graduação. Outro ponto a ser destacado é que no caso dos alunos de graduação da UFSCar, tivemos matrículas dos cursos de Química do campus de Sorocaba-SP. Esses aspectos são importantes, pois levanta a possibilidade de oferecer disciplinas inter-campi para o caso específico da Introdução à Quimiometria para a graduação. Assim, alunos dos campus de Araras, Buri e Sorocaba poderiam cursar diferentes disciplinas optativas e, assim, ampliar as opções nas suas grades curriculares. Já no caso da pós-graduação, poderiam ser ofertadas disciplinas nacionais com a participação de diversos programas de pós-graduação e também da Sociedade Brasileira de Química (SBQ). Tais pontos podem contribuir para a descentralização dos cursos de pós-graduação, diminuição das assimetrias e fomentar a interiorização da pesquisa em áreas remotas do país (regiões Norte e Nordeste, por exemplo).

Ainda para os alunos de graduação, 25% afirmaram que só fariam outra disciplina remota em uma situação como a imposta pela pandemia (círculo interno na Figura 32c). Um ponto a ser destacado é um certo antagonismo nas percepções dos alunos de pós-graduação, onde cerca de 3% afirmaram que *não* fariam esse tipo de disciplina após a pandemia e 85% afirmaram que *sim*. Ao comparar os dois grupos de alunos não foi notada uma diferença estatística significativa entre eles ($p - valor = 0,42$).

6.2.4 Avaliação das percepções dos alunos durante a pandemia: questões que apresentaram diferenças significativas

Na questão 9 (ver Tabela 3) foi identificada uma diferença significativa ($p - valor = 0,009$) entre as respostas dos alunos de graduação e pós-graduação. Nota-se mais uma vez que os alunos de pós-graduação são mais receptivos às aulas oferecidas de forma remota (ver Figura 33a, círculo externo). Cerca de 48% dos alunos de pós-graduação afirmaram que as aulas deveriam ser sempre remotas (vídeos gravados e disponibilizados). No caso dos alunos de graduação (Figura 33a, círculo interno), 28,1% afirmaram que preferem as aulas convencionais (presenciais). Essa observação é condizente com o grau de amadurecimento do aluno, onde os de graduação necessitam da figura do professor presente em sala de aula e a vivência universitária com os demais membros da comunidade (alunos e professores). A Figura 33b também mostra a média (quadrado interno) \pm desvio padrão. O desvio padrão é representado pelas linhas horizontais inferiores e superiores da caixa de cada nível de aluno (graduação e pós-graduação). A linha horizontal no interior

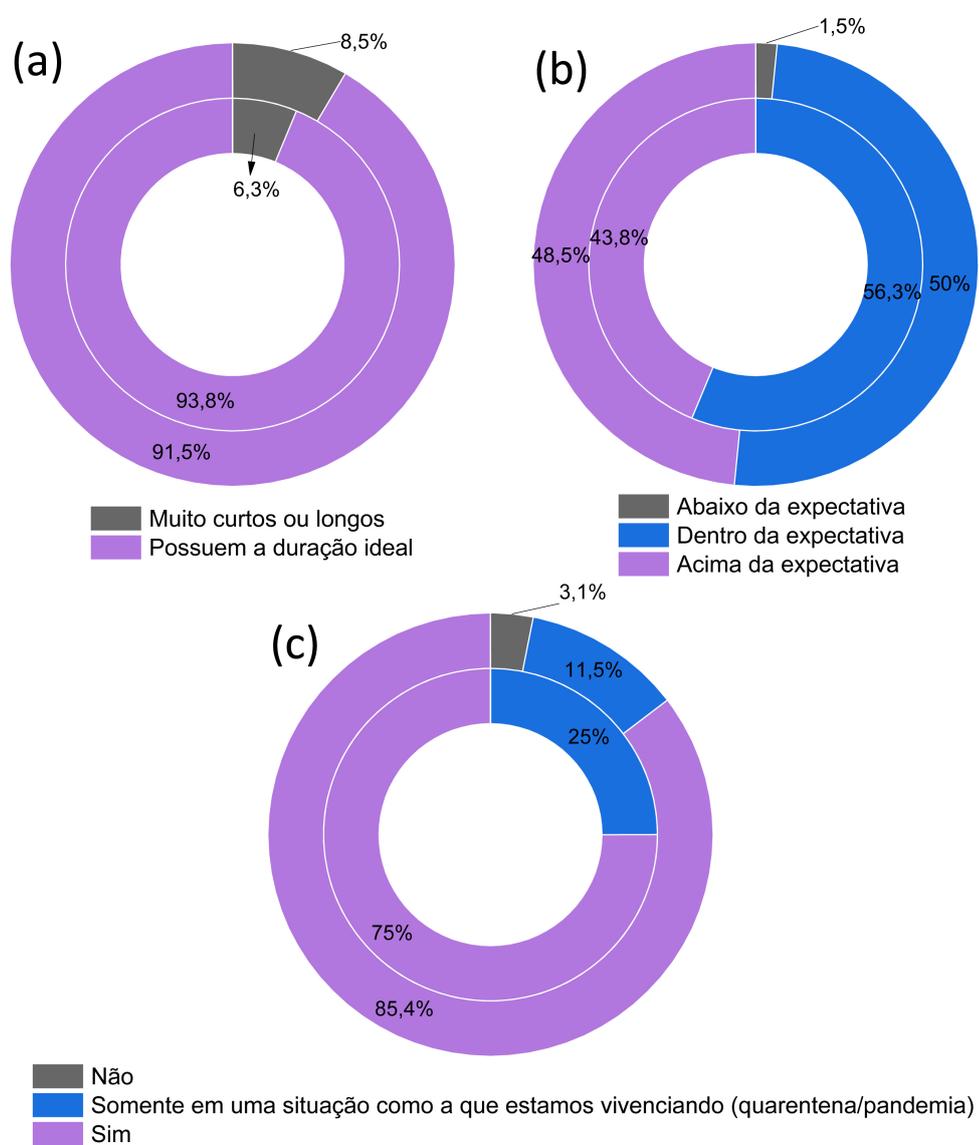
da caixa representa a mediana dos resultados. Além disso, temos a distribuição de todas as respostas dos alunos de graduação e pós-graduação. Para os alunos de graduação ($n = 32$) a média e a mediana das respostas (quadrado e linha horizontal, respectivamente) ficaram consolidadas como "Poderia ser adotado um meio termo (50% presencial e 50% aulas gravadas)".

As respostas dos alunos de pós-graduação ($n = 130$) tenderam ao "Deveria ser sempre assim"(remotas, Figura 33b) e a média é diferente da mediana.

Com as informações da Figura 33b é possível discutir duas observações importantes: assimetria e curtose. A assimetria representa o quão deslocados estão os dados em relação à média. Assimetria igual a 0 indica que a média, a moda (valor que mais se repete) e a mediana são iguais. Valores de assimetria negativos ou positivos indicam que os dados estão mais deslocados para a esquerda ou para a direita da média, respectivamente. Valores de curtose positivos indicam distribuição em cume e os valores negativos correspondem à uma distribuição achatada [35]. Para o caso específico da questão 9 (ver Figuras 33a e 33b) os valores de assimetria para os dados da graduação e pós-graduação foram -0,06 (praticamente centralizados) e -0,59 (deslocados para a esquerda da média), respectivamente. Os valores de curtose foram negativos para ambas classes de alunos e variaram de -1,33 (graduação) e -0,63 (pós-graduação). Os dados da graduação correspondem à uma curva mais achatada do que da pós-graduação.

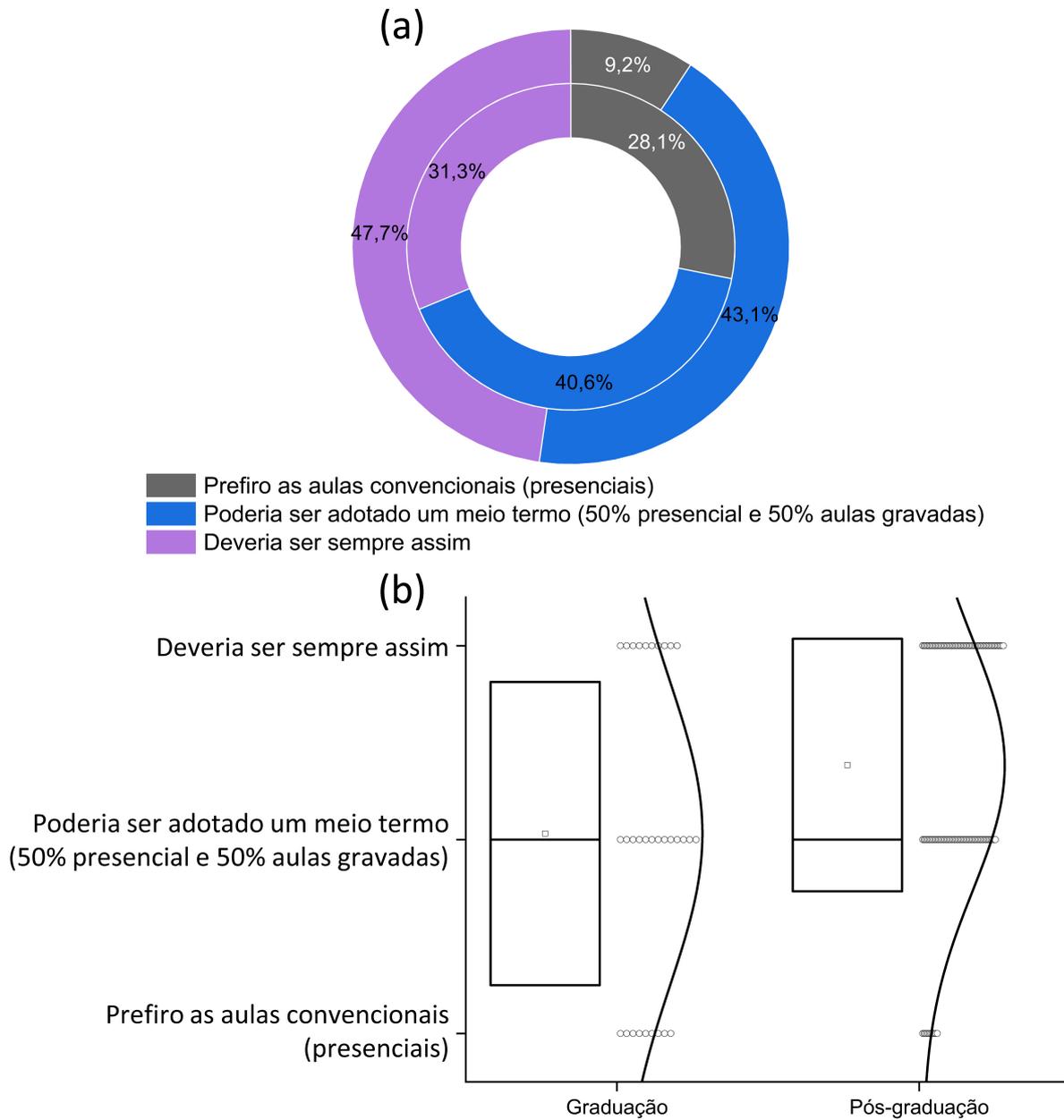
No caso específico da questão 11 (Figuras 34 e 35) também foi observada uma diferença significativa ($p - valor = 0,00004$) entre as opiniões dos alunos de graduação e pós-graduação. Nesta questão, os alunos poderiam selecionar mais de uma opção de resposta. Nenhum aluno de graduação, por exemplo, selecionou a opção "As disciplinas deveriam ser somente à distância". A maioria dos alunos de graduação (40%) selecionaram a opção "Poderíamos ter um meio termo: 25% à distância e 75% presencial". Para 20% dos alunos de pós-graduação, as aulas deveriam ser somente à distância com alguns momentos para reuniões síncronas não presenciais. Além disso, os alunos de pós-graduação ficaram divididos entre as opções que ofereciam uma condição intermediária entre ensino à distância e presencial (Figura 35). Cerca de 66% (20+25+20) dos alunos de pós-graduação optaram pelas opções que ofereciam uma condição meio termo entre ensino à distância e remoto. Os valores de assimetria e curtose foram todos negativos e variaram de -0,57 (curtose dos alunos da graduação) até -0,22 (assimetria dos alunos da graduação).

Figura 32 – Opinião dos alunos de graduação (círculo interno) e pós-graduação (círculo externo) sobre as questões 7 (a), 8 (b) e 10 (c).



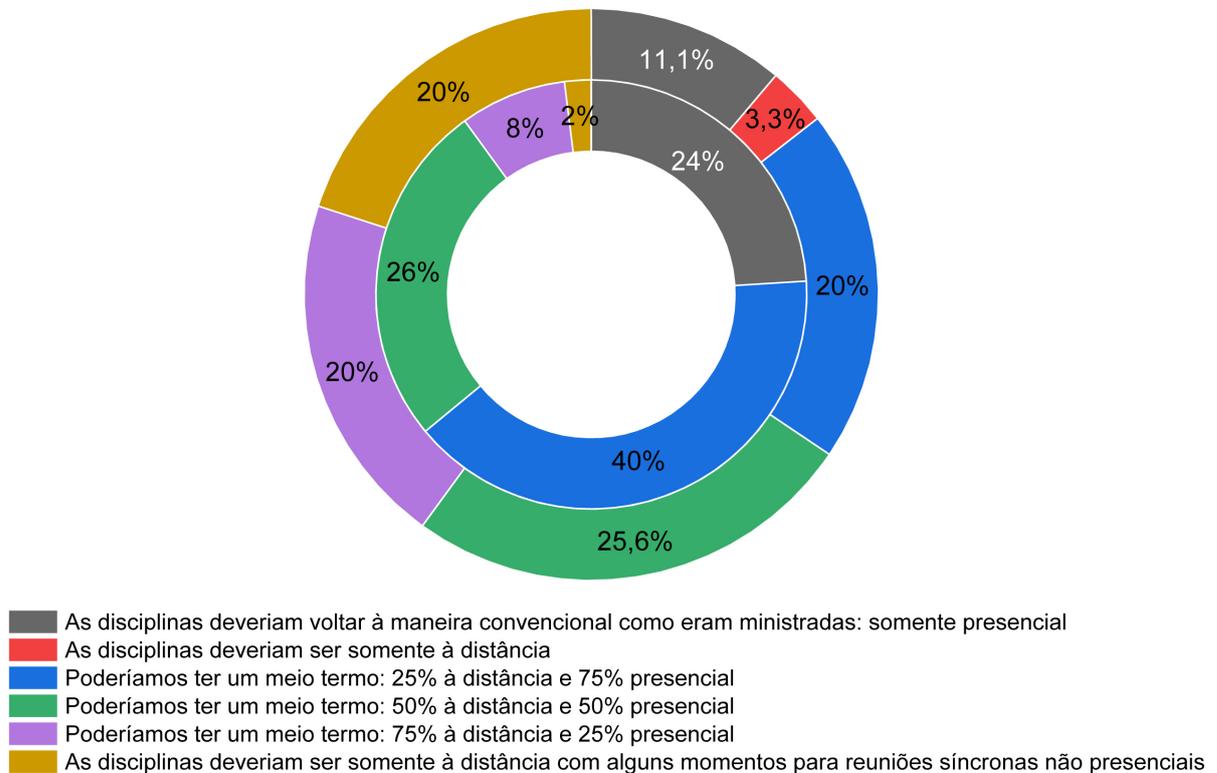
Fonte: Compilação do autor.

Figura 33 – Opinião dos alunos de graduação (círculo interno) e pós-graduação (círculo externo) sobre a questão 9 (a) e comparação entre as respostas (b).



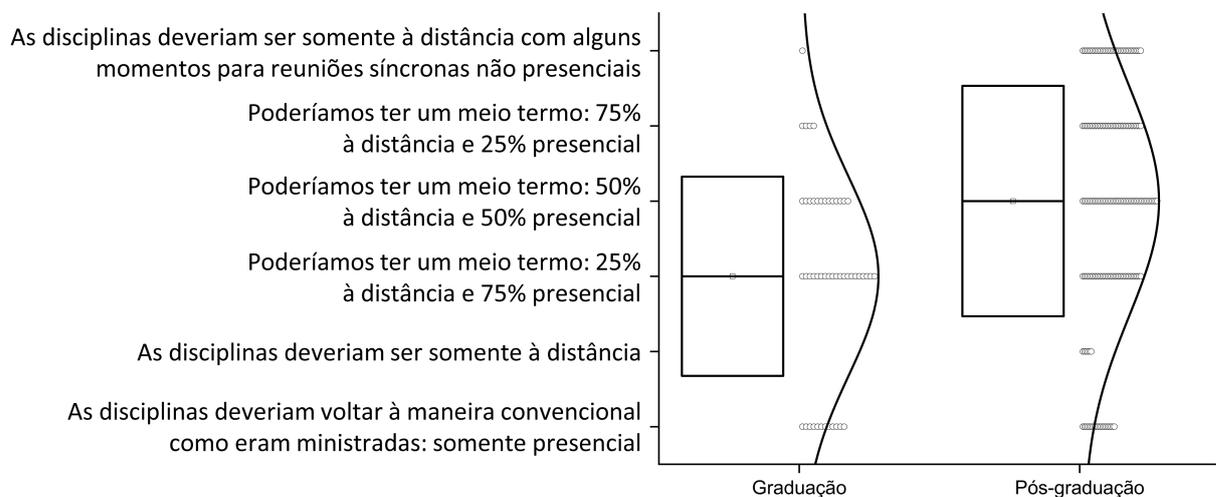
Fonte: Compilação do autor.

Figura 34 – Opinião dos alunos de graduação (círculo interno) e pós-graduação (círculo externo) sobre a questão 11.



Fonte: Compilação do autor.

Figura 35 – Comparação entre as respostas dadas pelos alunos de graduação e pós-graduação para a questão 11.



Fonte: Compilação do autor.

7 Conclusões, Perspectivas Futuras e Considerações Finais

7.1 Conclusões

A proposição, execução e discussão dos dados dessa dissertação de mestrado trouxeram reflexões importantes sobre como ciências imprescindíveis para o desenvolvimento humano podem impactar de forma considerável o ensino para os cursos de química (graduação e pós-graduação). O conhecimento de tais informações é importante, pois permite um maior embasamento teórico sobre diferentes assuntos: álgebra linear, geometria analítica, cálculo numérico, operações com matrizes e uso de recursos computacionais em sala de aula.

O algoritmo desenvolvido para determinar a região ótima de trabalho de um planejamento de misturas foi bastante eficaz e rápido. Além disso, é de fácil operação e entendimento. O mesmo foi utilizado em sala de aula e a aceitação dos alunos foi excelente. Além disso, o texto apresentado trouxe um paralelo sobre o uso de derivadas parciais para a identificação da região ótima de trabalho. As etapas apontadas para a identificação do máximo são muitas e exigem atenção do analista.

A avaliação dos conteúdos apresentados pelos alunos de graduação e pós-graduação mostraram que os mesmos entenderam a proposta de ensino em um momento delicado da nossa história recente. O uso de vídeos gravados, encontros síncronos e recursos computacionais diversos mostraram avanços e a percepção de que atividades positivas podem ser realizadas e implementadas. Além disso, com o ensino remoto foi possível apresentar mais conteúdos aos alunos (situação improvável em tempos pré-pandêmicos). De forma geral as percepções de ambos grupos de alunos (graduação e pós-graduação) foi bastante positiva. Entretanto, foram observadas diferenças significativas para o conjunto de questões que abordaram sobre a continuidade do ensino remoto ou se deveria ser adotado uma situação intermediária. A aceitação dos alunos de pós-graduação pela modalidade remota é nítida e compreensível. Além disso, as aulas remotas e os encontros síncronos aumentaram a capilaridade das disciplinas oferecidas na graduação e pós-graduação. No caso da graduação foi possível oferecer a mesma disciplina para alunos de diferentes campi da UFSCar. Para a pós-graduação foram atingidos alunos de diferentes estados da federação. Posso considerar que os efeitos observados foram positivos, principalmente diante das muitas manifestações de apoio recebidas por e-mail ou por meio da questão 12 do questionário (ver Apêndice B).

7.2 Perspectivas Futuras

Como perspectivas futuras pode ser apontada a alta capilaridade que as disciplinas alcançaram. Assim, essa experiência poderia ser mantida e replicada para outras disciplinas. A ideia central aqui é propor, por exemplo, disciplinas nacionais que seriam oferecidas por programas de pós-graduação conceitos 6 e/ou 7 da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) para diversos cursos de pós-graduação espalhados pelo Brasil. Tal iniciativa já é utilizada por muitos cursos que funcionam em rede, como o Mestrado Profissional em Matemática em Rede Nacional (Profmat) e o Programa de Mestrado Profissional em Química em Rede Nacional (Profqui). Ambos contam com o apoio e reconhecimento da Capes, além do apoio das respectivas sociedades: Sociedade Brasileira de Matemática (SBM) e SBQ. A perspectiva levantada seria de difundir essas experiências para outros programas além daqueles que funcionam em rede ou de forma à distância (Ensino à distância, EaD). Além disso, vários professores poderiam dividir as aulas e docentes do exterior poderiam contribuir também. Quanto à avaliação, poderiam ser adotados testes *online* ou avaliações presenciais (pensando em um período pós-pandêmico) na universidade onde o aluno está matriculado.

Tais iniciativas requerem um maior engajamento das sociedades (SBQ no caso da Química) e dos programas de pós-graduação. Outra deficiência que poderia ser minimizada é a concentração de programas de pós-graduação em Química nas regiões Sul e Sudeste do Brasil. As regiões Centro-Oeste, Norte e Nordeste do Brasil não possuem, por exemplo, programas de pós-graduação em Química com conceito 7 (conceito mais alto) da Capes. No que se refere à graduação, universidades multicampi como a UFSCar, poderiam oferecer simultaneamente as mesmas disciplinas para os alunos espalhados nos seus 4 campi (São Carlos, Araras, Sorocaba e Buri). Os temas relacionados ao planejamento fatorial, por exemplo, são importantes para os cursos de engenharia de alimentos que está localizado no campus Lagoa do Sino na cidade de Buri-SP.

7.3 Considerações Finais

Uma das principais lições na execução dessa dissertação de mestrado profissional é o fato de que mesmo em situações adversas e absurdas como as da pandemia e suas consequências, é possível realizar atividades que trazem benefícios para a sociedade. Os anos de 2020, 2021 e 2022 foram difíceis, mas também trouxeram alegrias no sentido de mais uma vez mostrar para a sociedade a importância da universidade e do ensino formal de qualidade. Outro ponto a ser destacado é o papel de ferramentas *online* para a difusão do conhecimento, tais como o YouTube. Além disso, ao longo desses anos contei com o apoio de diversas pessoas: professores da UFSCar (Departamentos de Química e Matemática), de outras instituições de ensino superior e de outros países (Dinamarca, Espanha, Estados

Unidos da América e Moçambique), alunos de graduação e pós-graduação, e profissionais egressos do Grupo de Análise Instrumental Aplicada (GAIA, www.gaia.ufscar.br). Isso mostra que devemos nos unir e atividades positivas só são realizadas com a participação de diferentes pessoas e profissionais. O Apêndice E mostra um artigo tutorial produzido referente aos dados gerados na seção 5 desta dissertação. O artigo apresentado é uma continuação do tutorial publicado em 2018 [5] e contou com a participação dos então alunos Felipe F. Hilário (iniciação científica) e Jeyne P. Castro Castilho (doutorado), bem como do Prof. Dr. Tomas E Barros. O tutorial [36] apresenta diversos exemplos sobre planejamento de misturas, vídeos sobre os cálculos envolvidos e também material suplementar (dados originais e algoritmos).

Referências

- [1] Brown, Steven; Tauler, Romá; Walczak, Beata. **Comprehensive chemometrics**. Elsevier: Amsterdã. 2009.
- [2] Barros Neto, Benício; Scarmínio, Ieda; Bruns, Roy. **25 anos de quimiometria no Brasil**. Química Nova, 29, 1401-1406, 2006. <https://doi.org/10.1590/S0100-40422006000600042>
- [3] Myers, Raymond; Montgomery, Douglas; Anderson-Cook, Christine. **Response Surface Methodology Process and Product Optimization Using Designed Experiments**. Hoboken: John Wiley Sons, Inc. 2009.
- [4] Leardi, Riccardo. **Experimental design in chemistry: A tutorial**. Analytica Chimica Acta, 652, 161-172, 2009. <https://doi.org/10.1016/j.aca.2009.06.015>
- [5] Pereira, Fabiola; Pereira-Filho, Edenir. **Aplicação de programa computacional livre em planejamento de experimentos: um tutorial**. Química Nova, 41, 1061-1071, 2018. <http://dx.doi.org/10.21577/0100-4042.20170254>
- [6] Lang, Patricia; Miller, Benjamin, Nowak, Abigail. **Introduction to the design and optimization of experiments using response surface methodology**. Journal of Chemical Education, 83, 280-282, 2006. <https://doi.org/10.1021/ed083p280>
- [7] Box, George; Hunter, Stuart; Hunter, William. **Statistics for Experimenters Design, Innovation, and Discovery**. Hoboken: John Wiley Sons, Inc. 2005.
- [8] Ferreira, Sergio *et al.*. **Application of multivariate techniques in optimization of spectroanalytical methods**. Applied Spectroscopy Reviews, 42, 475-491, 2007. <https://doi.org/10.1080/05704920701551506>
- [9] Barros Neto, Benício; Scarminio, Ieda; Bruns, Roy. **Como fazer experimentos**. Porto Alegre: Bookman. 2010.
- [10] Novaes, Cleber; *et al.*. **Otimização de Métodos Analíticos Usando Metodologia de Superfícies de Respostas - Parte II: Variáveis de Mistura**. Revista Virtual de Química, 10, 393-420, 2018. <http://dx.doi.org/10.21577/1984-6835.20180030>
- [11] Gozálviz, J.; García-Díaz, J.. **Mixture design experiments applied to the formulation of colorant solution**. Journal of Chemical Education, 83, 647-650, 2006. <https://doi.org/10.1021/ed083p647>

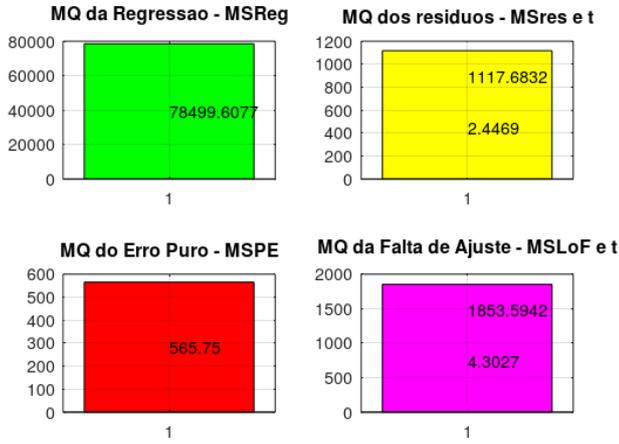
- [12] Harvey, David. **Analytical Chemistry 2.0**. Open Educational Resource (OER) - UNM, 2016. <http://solr.bccampus.ca:8001/bcc/file/ef2dfb89-d60c-41df-a235-d551346adf3f/1/AnalChem2.0.pdf>. Os direitos autorais desta obra foram transferidos para o autor em 2008.
- [13] Mendham, J. *et al.*. **Análise Química Quantitativa**. LTC Editora: Rio de Janeiro. 2002.
- [14] Harris, Daniel. **Análise Química Quantitativa**. LTC Editora: Rio de Janeiro. 2012.
- [15] Kellner, R. *et al.* **Analytical Chemistry**. Wiley-VCH Verlag GmbH Co KGaA: Weinheim. 2004.
- [16] Breitzkreitz, Márcia; Souza, André; Poppi, Ronei. **Experimento didático de quimiometria para planejamento de experimentos: avaliação das condições experimentais na determinação espectrofotométrica de ferro II com o-fenantrolina. Um tutorial, Parte III**. Química Nova, 37, 564-573, 2014. <http://dx.doi.org/10.5935/0100-4042.20140092>
- [17] Teófilo, Reinaldo; Ferreira, Márcia. **Quimiometria II: Planilhas eletrônicas para cálculos de planejamentos experimentais, um tutorial**. Química Nova, 29, 338-350, 2006. <https://doi.org/10.1590/S0100-40422006000200026>
- [18] Hoffmann, Rodolfo. **Análise estatística de relações lineares e não lineares**. Piracicaba: ESALQ-USP. 2016. <https://doi.org/10.11606/9788592105716>
- [19] Franco, Neide Bertoldi. **Cálculo numérico**. Pearson Prentice Hall: São Paulo. 2006.
- [20] Heumann, Christian; Shalabh; Schomaker, Michael. **Introduction to statistics and data analysis**. Suíça: Springer. 2016. <https://doi.org/10.1007/978-3-319-46162-5>
- [21] Pereira Filho, Edénir. **Planejamento Fatorial em Química Maximizando a Obtenção de Resultados**. São Carlos: Edufscar. 2015.
- [22] Patrão, Mauro. **Cálculo 1 Derivada e Integral em uma variável**. Brasília: Editor UnB. 2011.
- [23] Vicentini, Fernando *et al.*. **Planejamento fatorial e superfície de resposta: otimização de um método voltamétrico para a determinação de Ag(I) empregando um eletrodo de pasta de nanotubos de carbono**. Química Nova, 34, 825-830, 2011. <https://doi.org/10.1590/S0100-40422011000500018>
- [24] Muniz Neto, Antonio. **Fundamentos de Cálculo**. Rio de Janeiro: Sociedade Brasileira de Matemática. 2015.

- [25] Delgado, Jorge; Frensel, Katia; Crissaff, Lhaylla. **Geometria Analítica**. Rio de Janeiro: Sociedade Brasileira de Matemática. 2017.
- [26] Najafpour, Azimeh *et al.*. **Study of heavy metals biosorption by tea fungus in Kombucha drink using central composite design**. Journal of Food Composition and Analysis, 86, 103359, 2020. <https://doi.org/10.1016/j.jfca.2019.103359>
- [27] Miranda, Cláudia. *et al.*. **Morphological and Optical Characteristics of Porous Silicon Produced by Anodization Process in HF-Acetonitrile and HF-Ethanol Solutions**. Journal of the Brazilian Chemical Society, 19, 769-774, 2008. <https://doi.org/10.1590/S0103-50532008000400022>
- [28] Silva, José. *et al.*. **Retention of Copper(II) Metal Ions in a Silicon-Glass Microfluidic Device**. Journal of the Brazilian Chemical Society, 18, 1531-1536, 2007. <https://doi.org/10.1590/S0103-50532007000800013>
- [29] Bisi, O. *et al.*. **Porous silicon: a quantum sponge structure for silicon based optoelectronics**. Surface Science Reports, 38, 1-126, 2000. [https://doi.org/10.1016/S0167-5729\(99\)00012-6](https://doi.org/10.1016/S0167-5729(99)00012-6)
- [30] Kaslik, Pete. **Foundations in statistical reasoning**. <https://sites.google.com/site/offthebeatenmathpath/foundations-in-statistical-reasoning-second-edition>
- [31] Moreira, Cassio Neri; Cabral, Marco Aurélio Palumbo. **Curso de Análise Real**. Rio de Janeiro: Editora Instituto de Matemática. 2021. www.im.ufrj.br
- [32] Holme, Thomas. **Chemistry education in times of disruption and the times that lie beyond**. Journal of Chemical Education, 97, 1219-1220, 2020. <http://orcid.org/0000-0003-0590-5848>
- [33] Clark, Ted *et al.*. **Testing in the time of covid-19: a suden transition to unprocted online exams**. Journal of Chemical Education, 97, 3413-3417, 2020. <https://dx.doi.org/10.1021/acs.jchemed.0c00546>
- [34] Montgomery, Douglas. **Introdução ao controle estatístico da qualidade**. Rio de Janeiro: LTC - Livros Técnicos e Científicos Editora Ltda. 2016.
- [35] Ribeiro Júnior, José. **Análises Estatísticas no SAEG**. Universidade Federal de Viçosa: Viçosa. 2001.
- [36] Hilário, F. F. *et al.*. **Planejamento de misturas e visualização da região ótima com planilhas no excel: um tutorial**. Química Nova, 44, 874-881, 2021. <http://dx.doi.org/10.21577/0100-4042.20170750>

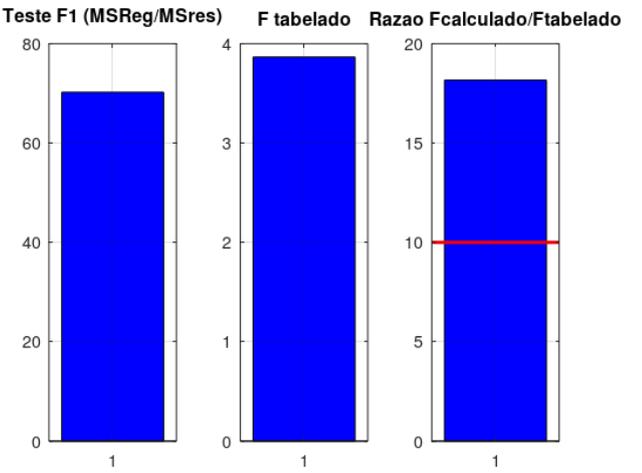
APÊNDICE A – Informações geradas pelo algoritmo Regression2 na proposição de modelos de regressão

Gráficos gerados pelo algoritmo Regression2 [5] na proposição de modelos de regressão. O primeiro gráfico (a) mostra as MQR, MQres e seu respectivo valor de t , MQEP e MQFaj com seu respectivo valor de t . O segundo e terceiro gráficos (b e c) mostram as avaliações, por meio de testes F 's, das MQR e MQres (b) e MQFaj e MQEP (c). O gráfico (d) mostra os valores de R^2 , o gráfico (e) o valor experimental *versus* previsto e o último gráfico (f) o valor previsto *versus* Resíduo e o histograma de distribuição dos resíduos.

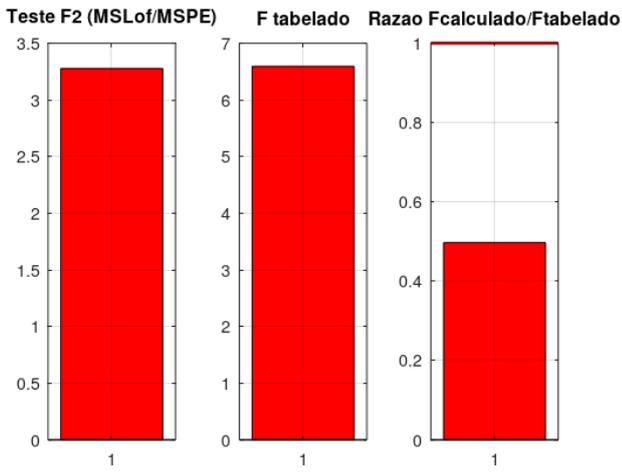
(a)



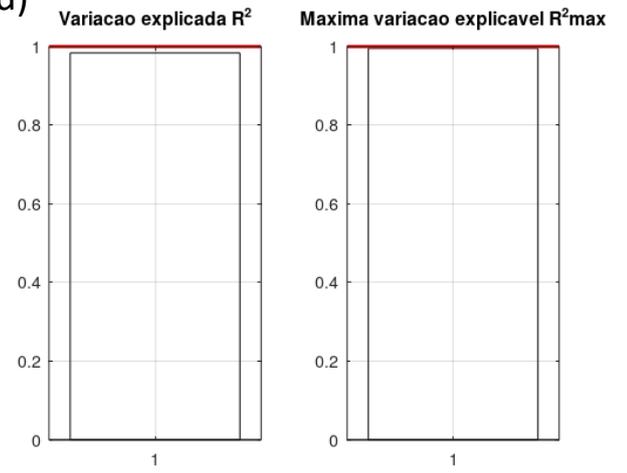
(b)



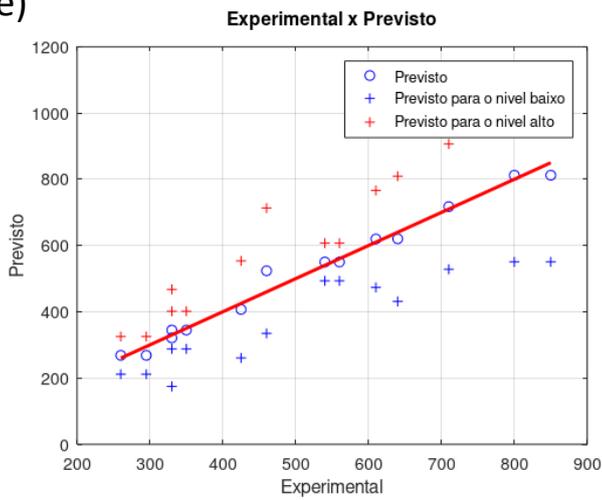
(c)



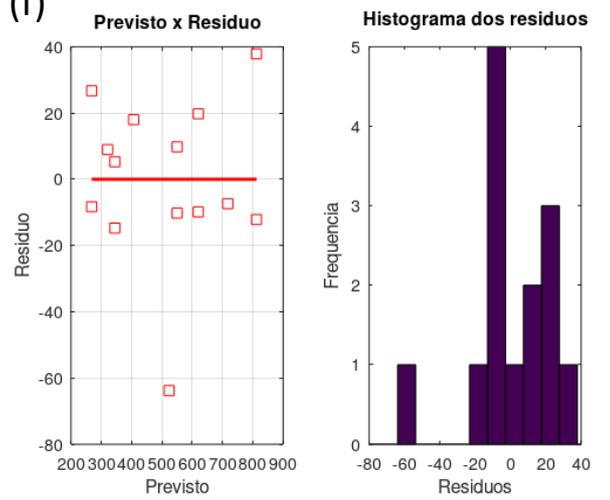
(d)



(e)



(f)



APÊNDICE B – Formulário de avaliação das disciplinas ministradas

Neste apêndice é mostrado o formulário de avaliação das disciplinas de graduação e pós-graduação ministradas nos anos de 2019 até 2021. As questões 1, 2, 3, 4 e 6 foram também respondidas pelos alunos antes da pandemia (cursos presenciais de pós-graduação) nos anos de 2014 (questões 1, 2, 3 e 4) e 2019 (questões 1, 2, 3, 4 e 6).

Formulário de avaliação das disciplinas Quimiometria (pós-graduação) e Introdução à Quimiometria (graduação)

 erpf@ufscar.br (não compartilhado) [Alternar conta](#)



*Obrigatório

Termo de Consentimento Livre e Esclarecido

Você está sendo convidado(a) a participar desta pesquisa. Informamos que sua participação é fundamental para o êxito da pesquisa, mas você é livre para participar ou não. O tempo estimado para resposta é inferior a cinco minutos. Na publicação dos resultados desta pesquisa, as identidades dos participantes serão mantidas no mais rigoroso sigilo. A participação neste estudo é voluntária e se você decidir não participar ou quiser desistir de continuar a qualquer momento, tem absoluta liberdade de fazê-lo.

Faça conexões entre o que você vivência na universidade com a sua vida!



1- Qual é a sua opinião geral sobre a disciplina que cursou? *

- Regular
- Boa
- Muito boa
- Ótima
- Excelente

2- O que você achou da didática de forma geral? *

- Regular
- Boa
- Muito boa
- Ótima
- Excelente

3- O que você achou da teoria de forma geral? *

- Regular
- Boa
- Muito boa
- Ótima
- Excelente

4- O que você achou dos exemplos dados? *

- Regulares
- Bons
- Muito bons
- Ótimos
- Excelentes



5 - O que você achou sobre a utilização do Google Forms para responder as atividades (trabalhos, teste e provas)? *

- Regular
- Boa
- Muito boa
- Ótima
- Excelente

6- O que você achou dos vídeos? *

- Regulares
- Bons
- Muito bons
- Ótimos
- Excelentes

7- Sobre a duração dos vídeos, o que você achou? *

- Muito curtos
- Muito longos
- Possuem a duração ideal

8- Como ficou o seu nível de aprendizado? *

- Abaixo da expectativa
- Dentro da expectativa
- Acima da expectativa



9- O que você achou do formato da disciplina (vídeos gravados e disponibilizados para os alunos)? *

- Deveria ser sempre assim
- Prefiro as aulas convencionais (presenciais)
- Poderia ser adotado um meio termo (50% presencial e 50% aulas gravadas)

10- Você faria outra disciplina como essa (100% om aulas gravadas e encontros síncronos gravados)? *

- Não
- Sim
- Somente em uma situação como a que estamos vivenciando (quarentena/pandemia)

11- Pensando em um cenário pós-pandemia, onde a vida voltaria ao "normal:" Quais das opções abaixo são mais atraentes? Marque quantas opções sejam necessárias. *

- As disciplinas deveriam voltar à maneira convencional como eram ministradas: somente presencial
- As disciplinas deveriam ser somente à distância
- Poderíamos ter um meio termo: 25% à distância e 75% presencial
- Poderíamos ter um meio termo: 50% à distância e 50% presencial
- Poderíamos ter um meio termo: 75% à distância e 25% presencial
- As disciplinas deveriam ser somente à distância com alguns momentos para reuniões síncronas não presenciais

12- Dúvidas ou comentários/sugestões?

Sua resposta

Página 2 de 2

[Voltar](#)

[Enviar](#)

[Limpar formulário](#)

Nunca envie senhas pelo Formulários Google.

Este formulário foi criado em Universidade Federal de São Carlos. [Denunciar abuso](#)

Google Formulários



APÊNDICE C – Formulário do SAEd

Neste apêndice é mostrado o formulário do SAEd que era utilizado pelos alunos na apresentação de dúvidas e comentários durante a execução da disciplina de Quimiometria 1.

Dúvidas, comentários e sugestões para a disciplina Quimiometria 1

erpf@ufscar.br [Alternar conta](#)



A foto e o nome associados à sua Conta do Google serão registrados quando você fizer upload de arquivos e enviar este formulário.. Seu e-mail não faz parte da resposta.

***Obrigatório**

Nome (opcional)

Sua resposta

e-mail (opcional)

Sua resposta

Dúvidas, comentários e sugestões! *

Sua resposta

Se deseja, vc pode anexar um arquivo!

 [Adicionar arquivo](#)

Enviar

[Limpar formulário](#)

Nunca envie senhas pelo Formulários Google.

Este formulário foi criado em Universidade Federal de São Carlos. [Denunciar abuso](#)



Google Formulários



APÊNDICE D – Algoritmo `mix_target2`

Neste apêndice é mostrado o algoritmo `mix_target2`. No início e no final do algoritmo foi colocado o comando *tic* (no começo) e *toc* (no final) que tem como finalidade contar o tempo de processamento dos dados.

```

1  function [B,Response]=mix_target2(a,b,c,d,e,f,Precision,Parameters);
2  %[B,Response]=mix_target2(a,b,c,d,e,f,Precision,Parameters);
3  %Essa função foi criada para verificar o ponto máximo em um planejamento de
4  %misturas.
5  %Precisa ser fornecido a faixa dos números aleatórios:
6  %a e b: referentes a variável 1;
7  %c e d: referentes a variável 2;
8  %e e f: referentes a variável 3.
9  % Para a primeira tentativa, recomenda-se colocar de 0 a 1.
10 % Outro input é a Precision, que seria o número de combinações que a rotina
11 % irá fazer. Quanto maior esse número mais lento fica o cálculo.
12 % Recomenda-se o mínimo de 1000.
13 %Além disso, deve ser fornecido um vetor (7x1) com os coeficientes (b1, b2,
14 %b3, b12, b13, b23, b123). Esse vetor precisa conter essa ordem, não deve
15 %mudar os coeficientes de lugar.
16
17 %Como resposta, temos B com os valores ótimos das variáveis individuais
18 %(v1, v2 e v3) e a Response com a resposta prevista do planejamento.
19 % Deve-se iniciar com 0 e 1 (faixa para cada variável).
20 % Tem-se diferentes casos:
21 % Caso 1: apenas para os coeficientes lineares; quando um ou dois ou os
22 % três coeficientes lineares são válidos.
23 % Caso 2: apenas para os coeficientes quadráticos; quando um deles ou dois
24 % deles são válidos.
25 % Caso 3: apenas o coeficiente terciário é válido.
26 % Caso 4: mais complexo. Quando tem-se coeficientes lineares e quadráticos
27 % válidos. Lineares, quadráticos e terciários válidos. Todos os quadráticos
28 % válidos. Quadráticos e terciários válidos. Nesse caso, é necessário um
29 % cálculo para achar os coeficientes. Assim, aparecerá um gráfico de barras
30 % com os 5 maiores resultados para cada variável. Você verá se chegou na
31 % resposta exata quando as barras estiverem uniformes (praticamente
32 % iguais). Quando se faz uma vez o cálculo variando de 0 a 1 cada variável,
33 % geralmente o resultado será bem próximo do resultado exato. Aconselha-se
34 % realizar novamente a rotina com um intervalo mais curto de acordo com o
35 % gráfico de barras. No intervalo recomenda-se colocar números redondos.
36
37 %% PRIMEIRA RODADA
38 % Criando o conjunto de dados com inúmeras alternativas
39 tic
40 a2=a*1000;b2=b*1000;c2=c*1000;d2=d*1000;e2=e*1000;f2=f*1000;
41 a2=round(a2);
42 b2=round(b2);
43 c2=round(c2);
44 d2=round(d2);
45 e2=round(e2);
46 f2=round(f2);
47 v1_1 = randi([a2 b2],Precision,1);v1_1=v1_1./1000;
48 v2_1 = randi([c2 d2],Precision,1);v2_1=v2_1./1000;
49 v3_1 = randi([e2 f2],Precision,1);v3_1=v3_1./1000;
50
51 a2=a*900;b2=b*900;c2=c*900;d2=d*900;e2=e*900;f2=f*900;
52 a2=round(a2);
53 b2=round(b2);
54 c2=round(c2);
55 d2=round(d2);
56 e2=round(e2);
57 f2=round(f2);
58 v1_2 = randi([a2 b2],Precision,1);v1_2=v1_2./900;
59 v2_2 = randi([c2 d2],Precision,1);v2_2=v2_2./900;
60 v3_2 = randi([e2 f2],Precision,1);v3_2=v3_2./900;
61
62 a2=a*800;b2=b*800;c2=c*800;d2=d*800;e2=e*800;f2=f*800;
63 a2=round(a2);
64 b2=round(b2);
65 c2=round(c2);
66 d2=round(d2);
67 e2=round(e2);
68 f2=round(f2);
69 v1_3 = randi([a2 b2],Precision,1);v1_3=v1_3./800;
70 v2_3 = randi([c2 d2],Precision,1);v2_3=v2_3./800;
71 v3_3 = randi([e2 f2],Precision,1);v3_3=v3_3./800;
72
73 a2=a*700;b2=b*700;c2=c*700;d2=d*700;e2=e*700;f2=f*700;

```

```

74 a2=round(a2);
75 b2=round(b2);
76 c2=round(c2);
77 d2=round(d2);
78 e2=round(e2);
79 f2=round(f2);
80 v1_4 = randi([a2 b2],Precision,1);v1_4=v1_4./700;
81 v2_4 = randi([c2 d2],Precision,1);v2_4=v2_4./700;
82 v3_4 = randi([e2 f2],Precision,1);v3_4=v3_4./700;
83
84 a2=a*600;b2=b*600;c2=c*600;d2=d*600;e2=e*600;f2=f*600;
85 a2=round(a2);
86 b2=round(b2);
87 c2=round(c2);
88 d2=round(d2);
89 e2=round(e2);
90 f2=round(f2);
91 v1_5 = randi([a2 b2],Precision,1);v1_5=v1_5./600;
92 v2_5 = randi([c2 d2],Precision,1);v2_5=v2_5./600;
93 v3_5 = randi([e2 f2],Precision,1);v3_5=v3_5./600;
94
95 a2=a*500;b2=b*500;c2=c*500;d2=d*500;e2=e*500;f2=f*500;
96 a2=round(a2);
97 b2=round(b2);
98 c2=round(c2);
99 d2=round(d2);
100 e2=round(e2);
101 f2=round(f2);
102 v1_6 = randi([a2 b2],Precision,1);v1_6=v1_6./500;
103 v2_6 = randi([c2 d2],Precision,1);v2_6=v2_6./500;
104 v3_6 = randi([e2 f2],Precision,1);v3_6=v3_6./500;
105
106 a2=a*400;b2=b*400;c2=c*400;d2=d*400;e2=e*400;f2=f*400;
107 a2=round(a2);
108 b2=round(b2);
109 c2=round(c2);
110 d2=round(d2);
111 e2=round(e2);
112 f2=round(f2);
113 v1_7 = randi([a2 b2],Precision,1);v1_7=v1_7./400;
114 v2_7 = randi([c2 d2],Precision,1);v2_7=v2_7./400;
115 v3_7 = randi([e2 f2],Precision,1);v3_7=v3_7./400;
116
117 a2=a*300;b2=b*300;c2=c*300;d2=d*300;e2=e*300;f2=f*300;
118 a2=round(a2);
119 b2=round(b2);
120 c2=round(c2);
121 d2=round(d2);
122 e2=round(e2);
123 f2=round(f2);
124 v1_8 = randi([a2 b2],Precision,1);v1_8=v1_8./300;
125 v2_8 = randi([c2 d2],Precision,1);v2_8=v2_8./300;
126 v3_8 = randi([e2 f2],Precision,1);v3_8=v3_8./300;
127
128 v1_T=[v1_1;v1_2;v1_3;v1_4;v1_5;v1_6;v1_7;v1_8];
129 v2_T=[v2_1;v2_2;v2_3;v2_4;v2_5;v2_6;v2_7;v2_8];
130 v3_T=[v3_1;v3_2;v3_3;v3_4;v3_5;v3_6;v3_7;v3_8];
131
132
133 vT = [v1_T v2_T v3_T];
134 vT(:,4)=sum(vT,2);
135
136 [m1,n1]=size(vT);
137
138 % A soma das três variáveis precisa resultar em 1. Se a soma é maior que 1,
139 % deve ser substituído por 0.
140
141 for i=1:m1
142 if vT(i,4)>1
143     vT(i,1:4) = 0;
144 end
145 end
146

```

```

147 vT = sortrows(vT,-4);
148 positive = find(vT(:,4) > 0);
149 vT=vT(positive,:);
150
151
152 % Caso 1: Criando as condições para quando apenas os coeficientes lineares são válidos
153 while Parameters(4:7,1) == 0
154     if (Parameters(1,1) > Parameters(2:3,1))
155         B=[1,0,0];
156     elseif (Parameters(2,1) > Parameters([1,3],1))
157         B = [0,1,0];
158     elseif (Parameters(3,1) > Parameters(1:2,1))
159         B = [0,0,1];
160
161     end
162     Response =
163         (B(1,1)*Parameters(1,1)+B(1,2)*Parameters(2,1)+B(1,3)*Parameters(3,1)+B(1,1)*B(1,2)*Pa
164         rameters(4,1)+B(1,1)*B(1,3)*Parameters(5,1)+B(1,2)*B(1,3)*Parameters(6,1)+B(1,1)*B(1,2
165         )*B(1,3)*Parameters(7,1));
166
167     break
168     end
169
170 % Caso 2: Criando as condições para quando apenas os quadráticos são válidos
171 %Para b12 e b13
172 while Parameters([1:3,6:7],1) == 0
173     if (Parameters(4,1) > Parameters(5,1))
174         B = [0.5,0.5,0];
175     else
176         B = [0.5,0,0.5];
177     end
178     Response =
179         (B(1,1)*Parameters(1,1)+B(1,2)*Parameters(2,1)+B(1,3)*Parameters(3,1)+B(1,1)*B(1,2
180         )*Parameters(4,1)+B(1,1)*B(1,3)*Parameters(5,1)+B(1,2)*B(1,3)*Parameters(6,1)+B(1,
181         1)*B(1,2)*B(1,3)*Parameters(7,1));
182     break
183     end
184
185 %Para b12 e b23
186 while Parameters([1:3,5,7],1) == 0
187     if (Parameters(4,1) > Parameters(6,1))
188         B = [0.5,0.5,0];
189     else
190         B = [0,0.5,0.5];
191     end
192     Response =
193         (B(1,1)*Parameters(1,1)+B(1,2)*Parameters(2,1)+B(1,3)*Parameters(3,1)+B(1,1)*B(1,2
194         )*Parameters(4,1)+B(1,1)*B(1,3)*Parameters(5,1)+B(1,2)*B(1,3)*Parameters(6,1)+B(1,
195         1)*B(1,2)*B(1,3)*Parameters(7,1));
196     break
197     end
198
199 %Para b13 e b23
200 while Parameters([1:4,7],1) == 0
201     if (Parameters(5,1) > Parameters(6,1))
202         B = [0.5,0,0.5];
203     else
204         B = [0,0.5,0.5];
205     end
206     Response =
207         (B(1,1)*Parameters(1,1)+B(1,2)*Parameters(2,1)+B(1,3)*Parameters(3,1)+B(1,1)*B(1,2
208         )*Parameters(4,1)+B(1,1)*B(1,3)*Parameters(5,1)+B(1,2)*B(1,3)*Parameters(6,1)+B(1,
209         1)*B(1,2)*B(1,3)*Parameters(7,1));
210     break
211     end
212
213 % Caso 3: somente o coeficiente terciário é válido
214 while Parameters([1:6],1) == 0
215     if (Parameters(7,1) ~= 0)
216         B = [0.33,0.33,0.33];
217     end
218     Response =

```

```

    (B(1,1)*Parameters(1,1)+B(1,2)*Parameters(2,1)+B(1,3)*Parameters(3,1)+B(1,1)*B(1,2)
    )*Parameters(4,1)+B(1,1)*B(1,3)*Parameters(5,1)+B(1,2)*B(1,3)*Parameters(6,1)+B(1,
    1)*B(1,2)*B(1,3)*Parameters(7,1));
208     break
209 end
210
211 % Caso 4: Para as demais situações: multiplicação pelos coeficientes:
212 % Para quando algum coeficiente linear, quadrático e terciário é válido:
213
214 [m2,n2]=size(vT);
215
216 while sum(Parameters(1:3,1))~= 0
217     if sum(Parameters(4:7,1))~= 0
218     for i=1:m2
219
220         vT(i,5)=(vT(i,1)*Parameters(1,1)+vT(i,2)*Parameters(2,1)+vT(i,3)*Parameters(3,1)+vT
221         T(i,1)*vT(i,2)*Parameters(4,1)+vT(i,1)*vT(i,3)*Parameters(5,1)+vT(i,2)*vT(i,3)*Par
222         ameters(6,1)+vT(i,1)*vT(i,2)*vT(i,3)*Parameters(7,1));
223         C = sortrows(vT,-5);
224     B = C(1,1:3);
225     Response = C(1,5);
226
227 end
228     subplot(1,3,1)
229     bar(C(1:5,1));
230     xlabel('Conjunto de dados')
231     ylabel('Valor máximo da v1')
232     ylim([0 1])
233     title ('Variável 1');
234     subplot(1,3,2)
235     bar(C(1:5,2));
236     xlabel('Conjunto de dados')
237     ylabel('Valor máximo da v2')
238     ylim([0 1])
239     title ('Variável 2');
240     subplot(1,3,3)
241     bar(C(1:5,3));
242     xlabel('Conjunto de dados')
243     ylabel('Valor máximo da v3')
244     ylim([0 1])
245     title ('Variável 3');
246     end
247     break
248 end
249
250 % Para quando algum coeficiente, quadrático e terciário é válido:
251
252 [m3,n3]=size(vT);
253
254 while sum(Parameters(4:6,1))~= 0
255     if Parameters(7,1)~= 0
256     for i=1:m3
257
258         vT(i,5)=(vT(i,1)*Parameters(1,1)+vT(i,2)*Parameters(2,1)+vT(i,3)*Parameters(3,1)+vT
259         T(i,1)*vT(i,2)*Parameters(4,1)+vT(i,1)*vT(i,3)*Parameters(5,1)+vT(i,2)*vT(i,3)*Par
260         ameters(6,1)+vT(i,1)*vT(i,2)*vT(i,3)*Parameters(7,1));
261         C = sortrows(vT,-5);
262     B = C(1,1:3);
263     Response = C(1,5);
264
265 end
266     subplot(1,3,1)
267     bar(C(1:5,1));
268     xlabel('Conjunto de dados')
269     ylabel('Valor máximo da v1')
270     ylim([0 1])
271     title ('Variável 1');
272     subplot(1,3,2)
273     bar(C(1:5,2));
274     xlabel('Conjunto de dados')
275     ylabel('Valor máximo da v2')
276     ylim([0 1])
277     title ('Variável 2');
278     subplot(1,3,3)
279     bar(C(1:5,3));
280     xlabel('Conjunto de dados')
281     ylabel('Valor máximo da v3')
282     ylim([0 1])
283     title ('Variável 3');
284     end
285     break
286 end

```

```

272 bar(C(1:5,3));
273 xlabel('Conjunto de dados')
274 ylabel('Valor máximo da v3')
275 ylim([0 1])
276 title ('Variável 3');
277     end
278         break
279 end
280
281 % Para quando todos os coeficientes quadráticos são válidos:
282
283 [m4,n4]=size(vT);
284
285 while Parameters(4:6,1)~= 0
286
287     for i=1:m4
288
289         vT(i,5)=(vT(i,1)*Parameters(1,1)+vT(i,2)*Parameters(2,1)+vT(i,3)*Parameters(3,1)+vT
290         T(i,1)*vT(i,2)*Parameters(4,1)+vT(i,1)*vT(i,3)*Parameters(5,1)+vT(i,2)*vT(i,3)*Par
291         ameters(6,1)+vT(i,1)*vT(i,2)*vT(i,3)*Parameters(7,1));
292         C = sortrows(vT,-5);
293     B = C(1,1:3);
294     Response = C(1,5);
295
296     end
297     subplot(1,3,1)
298     bar(C(1:5,1));
299     xlabel('Conjunto de dados')
300     ylabel('Valor máximo da v1')
301     ylim([0 1])
302     title ('Variável 1');
303     subplot(1,3,2)
304     bar(C(1:5,2));
305     xlabel('Conjunto de dados')
306     ylabel('Valor máximo da v2')
307     ylim([0 1])
308     title ('Variável 2');
309     subplot(1,3,3)
310     bar(C(1:5,3));
311     xlabel('Conjunto de dados')
312     ylabel('Valor máximo da v3')
313     ylim([0 1])
314     title ('Variável 3');
315     break
316     end
317     toc
318

```

APÊNDICE E – Artigo científico publicado

Neste apêndice é mostrada primeira página do artigo científico [36] publicado no periódico Química Nova no ano de 2021.

PLANEJAMENTO DE MISTURAS E VISUALIZAÇÃO DA REGIÃO ÓTIMA COM PLANILHAS NO EXCEL: UM TUTORIAL[†]Felipe Ferri Hilário^a, Jeyne Pricylla Castro^a, Tomas Edson Barros^b e Edenir Rodrigues Pereira-Filho^{a,b,*} ^aDepartamento de Química, Universidade Federal de São Carlos, 13565-905 São Carlos – SP, Brasil^bDepartamento de Matemática, Universidade Federal de São Carlos, 13565-905 São Carlos – SP, Brasil

Recebido em 03/02/2021; aceito em 18/03/2021; publicado na web em 14/04/2021

MIXTURE DESIGN AND OPTIMUM REGION VISUALIZATION WITH SPREADSHEETS AT EXCEL: A TUTORIAL. This tutorial shows how to perform the data treatment for mixture design with two and multiple responses. Two data sets from the literature were used to exemplify the calculations. A template in Excel was proposed to prepare contour plot and response surface to visualize the optimum region of the mixture, which usually is made with commercial computational programs. The desirability function was used in example 2 (optimization of 5 responses simultaneously), and a code also available for download was proposed. Besides, 21 videos were prepared to show all the details for the readers and, it is available on YouTube. With this tutorial, it is possible to learn in a practical fashion how to handle the data from mixture design and, the authors hope to contribute with the researchers in this area.

Keywords: Octave; mixture design; analysis of variance; regression model; template mixtures; partial derivatives.

INTRODUÇÃO

O planejamento de misturas é bastante empregado em otimizações nas quais é fixa a relação entre as variáveis.¹ Um exemplo típico é a otimização da composição da fase móvel em um método de separação cromatográfico.² Nesse tipo de planejamento, os níveis das variáveis não podem ser alterados de forma independente e sempre é mantida uma relação entre as mesmas. A literatura científica apresenta diversos textos excelentes sobre este tipo de planejamento.³⁻⁷ Entretanto, a visualização da superfície de resposta e gráfico de contorno está restrita a alguns programas computacionais pagos e específicos.

Dessa forma, o objetivo deste tutorial é mostrar a aplicação de uma planilha no Microsoft Excel para a visualização do gráfico de contorno e superfície de resposta. Além disso, é empregado o Solver (Microsoft Excel) para o cálculo da melhor condição de trabalho. Os cálculos para obtenção do modelo de regressão podem ser efetuados em programa livre (Octave) e todos os procedimentos para execução dos cálculos já foram descritos no tutorial anterior publicado na Química Nova.⁸ Um dos objetivos deste tutorial é também apresentar uma análise criteriosa e com rigor matemático sobre a identificação da região ótima de um dado modelo.

Portanto, com este tutorial é possível apresentar ao público interessado uma alternativa viável e de fácil acesso para a utilização de planejamento fatorial no contexto do planejamento de misturas. Além da descrição dos conjuntos de dados e cálculos efetuados, foram preparados vídeos explicativos que estão disponíveis no Canal do prof. Edenir Pereira Filho no YouTube (<https://www.youtube.com/c/EdenirPereiraFilho>, ver Playlist 15).

PARTE EXPERIMENTAL**Descrição dos conjuntos de dados empregados**

Para exemplificar as ferramentas deste tutorial, foram selecionados dois conjuntos de dados a partir de trabalhos publicados no periódico *Journal of the Brazilian Chemical Society* (JBCS). A utilização

desses estudos, bem como o conjunto de dados neles publicados, foi gentilmente cedida pela Editora da Sociedade Brasileira de Química (PubliSBQ). Tais artigos foram selecionados por utilizarem planejamento de misturas objetivando a resolução de diferentes problemas analíticos.

Exemplo 1

No primeiro exemplo foi utilizado o trabalho publicado por Damasceno *et al.*,⁹ no qual os autores utilizaram um planejamento de misturas do tipo *simplex centroide* para visualizar a relação do chumbo presente em descarte de material eletrônico (placas de circuito impresso) com diferentes tipos de solos. Nesse caso, cada pseudocomponente foi avaliado em 6 níveis. A Tabela 1 mostra em detalhes o planejamento experimental utilizado pelos autores com três variáveis, OM (v_1 - matéria orgânica), IO (v_2 - óxido de ferro) e KA (v_3 - caulinita), além de duas respostas que serão estudadas neste tutorial, F2 (concentração de Pb extraído na fração orgânica) e F3 (concentração de Pb extraído na fração residual). No experimento 1, por exemplo, da Tabela 1 temos uma mistura composta por 80 g (codificado em 1) de matéria orgânica (OM) e 10 g de cada um dos demais componentes: óxido de ferro (IO, codificado em 0) e caulinita (KA, codificado em 0).

Exemplo 2

Para o exemplo 2, foi utilizado o trabalho publicado por Estevão *et al.*¹⁰ e o objetivo foi abordar um estudo com múltiplas respostas. Nesse caso, os autores otimizaram um procedimento de extração utilizando a técnica de microextração líquido-líquido dispersivo em solvente binário para a determinação de pesticidas em amostras de água natural. Na parte experimental foram realizados dois planejamentos de misturas com três componentes. O primeiro planejamento abordou uma faixa muito ampla da composição da mistura dos componentes extratores, sendo necessário um segundo com uma faixa mais restrita. Nesse tutorial abordamos apenas o segundo planejamento.

A mistura foi realizada com dois extratores (diclorometano, DCM e clorofórmio, CLF) e um solvente dispersor (acetonitrila, ACN). O objetivo foi obter a melhor relação entre os três componentes que proporcionasse recuperações adequadas (da ordem de 100%) dos seguintes pesticidas nas amostras de água natural: *Diuron* (DIU),

*e-mail: erpf@ufscar.br

[†]Esse trabalho é dedicado à memória do Prof. Dr. Ronei Jesus Poppi