



**Carolina Jesus
Serôdio**

**Processo de Neutralização do Óleo Utilizado na
Produção de Biodiesel: Caracterização e Análise
Estatística**



**Carolina Jesus
Serôdio**

Processo de Neutralização do Óleo Utilizado na Produção de Biodiesel: Caracterização e Análise Estatística

Relatório de estágio apresentado à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Matemática e Aplicações, realizada sob a orientação científica da Professora Doutora Andreia Hall, Professora Associada do Departamento de Matemática da Universidade de Aveiro e coorientação da Professora Doutora Helena Alvelos, Professora Auxiliar do Departamento de Economia, Gestão, Engenharia Industrial e Turismo da Universidade de Aveiro.

Dedico este trabalho em memória da minha avó Judite.

o júri / the jury

presidente / president

Professor Doutor Agostinho Miguel Mendes Agra

Professor Auxiliar da Universidade de Aveiro

vogais / examiners committee

Professora Doutora Ana Raquel Reis Couto Xambre

Professora Auxiliar da Universidade de Aveiro (arguente principal)

Professora Doutora Andreia Oliveira Hall

Professora Associada da Universidade de Aveiro (orientadora)

**agradecimentos /
acknowledgements**

Em primeiro lugar, agradeço profundamente aos meus pais pelo amor, pela força e pelas oportunidades que me dão, não apenas durante o meu percurso académico, mas ao longo da minha vida. Agradeço às minhas irmãs e aos meus avós pelo amor e apoio incondicionais.

Agradeço à Professora Andreia Hall e à Professora Helena Alvelos pela disponibilidade e pelos conhecimentos partilhados.

Agradeço à Engenheira Sónia Henriques pela orientação, empenho, paciência e por todos os conhecimentos que me transmitiu. Agradeço-lhe também pelas palavras de conforto e por acreditar em mim.

Agradeço a todos os colaboradores da Prio Bio, S.A., em particular a todos os chefes de turno, operadores e analistas de laboratório que tanto contribuíram para este trabalho.

Ao Martim, ao Rafael, à Diana, à Rita, à Filipa, ao Hugo Teixeira e ao Hugo Almeida agradeço-lhes em particular pela amizade.

Agradeço sinceramente ao Bernardo, à Cristiana e à Joana por serem os meus “pilares” ao longo destes anos.

Por último, agradeço a todos aqueles que de alguma forma fizeram parte desta caminhada.

Palavras-Chave

Biodiesel, Neutralização do óleo, Óleos alimentares usados, Base, X.

Resumo

Os impactos ambientais são uma crescente preocupação global, pelo que se procura encontrar outras fontes de energia renováveis. Sendo um biocombustível renovável, o biodiesel é uma resposta a esta problemática, uma vez que pode ser produzido através de óleos alimentares usados.

A presente dissertação incide sobre o processo de neutralização do óleo utilizado na Prio Bio, S.A., uma fábrica de produção de biodiesel. O principal objetivo foi a aplicação de algumas técnicas matemáticas a um processo industrial, no sentido de o melhorar.

Numa fase inicial, foi tomado conhecimento sobre todo o processo de neutralização, bem como do processo de produção de biodiesel. Posteriormente, fez-se uma análise exploratória dos dados do processo, recorrendo à análise de correlações e à análise fatorial em componentes principais.

Face às conclusões obtidas, os passos seguintes foram traçados sob a ótica do ciclo *Plan-Do-Check-Act*. Na primeira fase do ciclo- "*Plan*", fez-se um levantamento das principais fragilidades do processo de neutralização, para as quais foram propostas sugestões de melhoria na segunda fase- "*Do*". Na fase "*Check*" e no seguimento das duas fases anteriores foram testados modelos de regressão linear múltipla e foi realizado um estudo de capacidade do processo. Tendo em conta a incerteza dos resultados obtidos com as duas metodologias anteriores, na fase "*Act*", realizou-se uma análise ao sistema de medição para validar o método.

Keywords

Biodiesel, Oil neutralization, Used cooking oils, Lye, X.

Resumo

Nowadays, other sources of renewable energy are sought due to the growing concern about environmental impacts. Being a renewable biofuel, biodiesel is a response to this problem, since it can be produced using used cooking oils.

The present dissertation focuses on the neutralization process of the oil used in Prio Bio, S.A., a biodiesel production plant. The main objective was to apply some mathematical techniques to an industrial process in order to improve it.

At an early stage, the process of neutralization as well as the process of biodiesel production was known. Subsequently, an exploratory analysis of the process data was made, using correlation analysis and factorial analysis in main components.

In view of the conclusions obtained, the following steps were drawn from the perspective of the *Plan-Do-Check-Act* cycle. In the first phase of the "*Plan*" cycle, a survey was made of the main fragilities of the neutralization process, for which suggestions for improvement were proposed in the second phase- "*Do*". In the "*Check*" phase and following the previous two phases, multiple linear regression models were tested and a process capability study was performed. Considering the uncertainty of the results obtained with the two previous methodologies, in the "*Act*" phase, an analysis was made to the measurement system to validate the method.

Conteúdo

Conteúdo	i
Lista de Figuras	iii
Lista de Tabelas	v
Nomenclatura	vii
1 Introdução	1
1.1 Enquadramento	1
1.2 Objetivos	2
1.3 Estrutura	3
2 Apresentação da Prio Energy, S.A.	5
2.1 Biodiesel	6
2.2 Produção de biodiesel na Prio Bio, S.A.	7
2.3 Neutralização do óleo	9
2.3.1 Dados Operacionais	13
2.3.2 <i>Check List em Gemba</i> (chão de fábrica)	13
2.3.3 Análise Laboratorial	13
3 Análise Exploratória	15
4 Análise Estatística da variável X	27
4.1 1ª Fase do ciclo PDCA- <i>Plan</i> (Planeamento)	27
4.2 2ª Fase do ciclo PDCA – <i>Do</i> (Fazer)	28
4.3 3ª Fase do ciclo PDCA- <i>Check</i> (Verificar)	30
4.3.1 Regressão Linear Múltipla	30
4.3.2 Estudo de Capacidade do Processo	40
4.4 4ª Fase do ciclo PDCA – <i>Act</i> (Atuar)	44
4.4.1 Análise do Sistema de Medição	44

5	Conclusões e Oportunidades Futuras	55
	Bibliografia	59
A	Análise do processo de neutralização	61
B	Regressão Linear Múltipla	63

Lista de Figuras

2.1	Áreas de negócio da Prio Energy, S.A.	5
2.2	Reação de transesterificação	6
2.3	Processo de produção de biodiesel na Prio Bio, S.A.	8
2.4	Reação química principal da neutralização do óleo	9
2.5	Tanques de armazenamento	10
2.6	Processo de neutralização do óleo	11
3.1	Caixa de bigodes das variáveis XA, XC e XF1, respetivamente	16
3.2	Histograma da variável XF1 para a formulação A + C	17
3.3	Caixa de bigodes e histograma para XD3, respetivamente (formulação A + C)	17
3.4	Caixa de bigodes das variáveis XB, XC e XF1, respetivamente	18
3.5	Histograma da variável XF1 para a formulação B + C	18
3.6	Caixa de bigodes e histograma para XD3, respetivamente (formulação B + C)	19
3.7	Série temporal para a formulação A + C	19
3.8	Série temporal para a formulação B + C	20
3.9	<i>Biplot</i>	22
3.10	Teste KMO e Bartlett	23
3.11	Matriz anti-imagem	24
3.12	<i>Scree plot</i>	24
3.13	Variância total explicada	25
3.14	Comunalidades	25
3.15	Matriz de componentes sem rotação e com rotação <i>Varimax</i> , respetivamente .	26
4.1	Ciclo <i>PDCA</i>	27
4.2	Gráficos de dispersão de XD3	28
4.3	Resumo do modelo 1 para XD3	32
4.4	Esquema de variáveis relacionadas com G	33
4.5	Resumo do modelo 2	33
4.6	Coefficientes do modelo 2	34
4.7	Estatísticas dos resíduos do modelo 2	35

4.8	QQ <i>plot</i> dos resíduos do modelo 2	36
4.9	Teste de normalidade para o modelo 2	36
4.10	Gráfico de dispersão do modelo 2	36
4.11	Gráfico de dispersão	37
4.12	Resumo do modelo 3	38
4.13	Estudo de capacidade para XD3 (01-10-2017 a 08-03-2018)	42
4.14	Estudo de capacidade de XD3 (08-03-2018 a 21-05-2018)	43
4.15	Procedimentos MSA	45
4.16	Variação total de um processo [17]	46
4.17	Estudo da Linearidade e do <i>Bias</i> para XD3	48
4.18	Resultados do estudo Gage R&R para XD3	48
4.19	Resultados do estudo Gage R&R (Nested) para XD3	49
4.20	Gage Run Chart of Measurement by Part, Operator para XD3	50
4.21	Estudo da Linearidade e do <i>Bias</i> para XF1	51
4.22	Resultados do estudo Gage R&R para XF1	51
4.23	Resultados do estudo Gage R&R (Nested) para XF1	52
4.24	Gage Run Chart of Measurement by Part, Operator XF1	53
A.1	Resultado do teste de Wilcoxon	61
A.2	Resultado do teste de Levene	62
B.1	Coeficientes do modelo 3	65
B.2	QQ <i>plot</i> dos resíduos do modelo 3	65
B.3	Teste de normalidade para o modelo 3	66
B.4	Gráfico de dispersão do modelo 3	66
B.5	Resumo do modelo 4	66
B.6	Coeficientes do modelo 4	67
B.7	QQ <i>plot</i> dos resíduos do modelo 4	67
B.8	Teste de normalidade para o modelo 4	67
B.9	Gráfico de dispersão do modelo 4	68

Lista de Tabelas

3.1	Análise Descritiva (A + C)	16
3.2	Análise Descritiva (B + C)	16
3.3	Correlações A + C	20
3.4	Correlações B + C	20
4.1	Estudo da variabilidade de X dos tanques de armazenamento, F1 e G	29
4.2	Critérios de classificação para o teste de Durbin-Watson	31
4.3	Critérios de classificação para o valor de <i>Leverage</i>	31
4.4	Tabela ANOVA do modelo 2	34
4.5	Resultados Obtidos com o Modelo 3	39
4.6	Resultados Obtidos com o Modelo 4	40
4.7	Critérios de classificação para estudo Gage R&R	47
B.1	Resultados obtidos com o Modelo 2	63

Nomenclatura

Siglas e Acrónimos

APA	Agência portuguesa do ambiente
CFPP	Cold filter plug point
DCS	Distributed control system
ESGI	European study group with industry
LIE	Limite inferior de especificação
LSE	Limite superior de especificação
MSA	Measurement system analysis
ndc	Number of distinct categories
OAU	Óleo alimentar usado
OB	Óleo bruto
ON	Óleo neutro
PDCA	Plan-Do-Check-Act
TdB	Títulos de Biocombustíveis
tep	Tonelada equivalente de petróleo
TK	Tanque
TPN	Temperatura e pressão normais
VIF	Variation inflation fator

Símbolos

β	Coeficiente de regressão
C_p	Índice de capacidade de um processo
C_{pk}	Índice de capacidade de um processo
μ	Média
ε	Erro
σ	Desvio Padrão

Capítulo 1

Introdução

A preocupação com os impactos ambientais tem vindo a aumentar ao longo das últimas décadas e neste sentido, têm-se procurado encontrar fontes de energia renováveis alternativas, como é o caso da reutilização dos óleos alimentares usados (OAU). De acordo com a Agência Portuguesa do Ambiente (APA), em Portugal são consumidos cerca de 93 milhões de litros de óleo alimentar por ano [1]. O óleo alimentar usado tem impactos ambientais e económicos nefastos, como por exemplo, a poluição, o tratamento das águas e o entupimento de canalizações. Este óleo pode, no entanto, ser reciclado e utilizado como matéria prima para a produção de biodiesel, sabão e glicerina.

O biodiesel é um biocombustível líquido, considerado benéfico para o ambiente, que poderá representar uma resposta à problemática ambiental do uso de combustíveis fósseis, uma vez que é possível produzi-lo através de óleos alimentares usados. Contudo, a produção de biodiesel através destes resíduos envolve custos de produção muito elevados devido à sua complexidade e às propriedades dos óleos. Apesar disso, a Prio Energy, S.A aposta nesta área, possuindo uma fábrica de biodiesel, a Prio Bio, S.A.

A Prio Bio, S.A. ao longo dos últimos anos tem apostado na unidade de pré-tratamento de óleos alimentares usados (neutralização do óleo), tendo como objetivo a produção de biodiesel a partir de 100% desta matéria-prima.

A estratégia da Prio Bio, S.A. no recurso a óleos alimentares usados, promove a sustentabilidade de recursos e respetiva economia circular.

1.1 Enquadramento

O presente relatório de estágio enquadra-se na unidade curricular de Dissertação/ Estágio/ Projeto, pertencente ao plano curricular do Mestrado em Matemática e Aplicações do Departamento de Matemática da Universidade de Aveiro. O estágio decorrido na Prio Bio, S.A., teve a duração de sete meses.

Na Universidade de Aveiro decorreu a edição de 2017 do european study group with indus-

try (ESGI) onde foi proposto pela Prio Bio, S.A. um desafio, cujo tema era “To understand correlations between raw materials inputs (characteristic X) that define the recipe and highly influences the operational parameters to obtain the high quality neutral oil (outputs)”, tema referente ao processo de pré-tratamento de óleos alimentares usados (neutralização de óleo). Dada a curta durabilidade deste projeto e tendo por base os resultados finais com ele obtidos, surgiu a possibilidade de realizar um estágio curricular na Prio Bio, S.A., no sentido de analisar os dados do processo de forma exaustiva e tentar implementar melhorias no mesmo.

1.2 Objetivos

O objetivo deste estágio foi a aplicação da matemática na caracterização de um processo industrial com vista à sua melhoria. Para tal, o seu desenvolvimento e execução passou por 3 fases.

Na fase inicial, a autora deste trabalho conheceu a Prio Energy, S.A. através da formação de acolhimento. Teve também formação em procedimentos e regras de segurança nas instalações fabris. Depois foi-lhe apresentada a fábrica “Do Óleo da Cozinha ao Biodiesel”, focando o processo escolhido para o estágio: neutralização do óleo. Acompanhou os dados retirados do processo, assim como dos equipamentos de medição associados, e tomou conhecimento das análises realizadas diariamente no laboratório da fábrica relativas à neutralização do óleo.

Na segunda fase leu-se a memória descritiva do processo de neutralização do óleo. Analisou-se o relatório do caso de estudo apresentado na edição do ESGI 2017 na Universidade de Aveiro “To understand correlations between raw materials inputs (characteristic X) that define the recipe and highly influences the operational parameters to obtain the high quality neutral oil (outputs)”. Analisou-se, também, o histórico dos parâmetros de controlo fornecidos na edição do ESGI 2017 e posteriormente, fez-se a análise do histórico dos parâmetros de controlo do ano 2017.

Na terceira e última fase, procedeu-se à análise estatística dos vários dados recolhidos com vista à caracterização e melhoria do processo, incluindo:

- Análise exploratória (descritiva) inicial;
- Aplicação de alguns modelos estatísticos (tais como modelos de regressão, séries temporais, etc.) sugeridos a partir da análise exploratória inicial;
- Análise da adequação dos modelos;
- Conclusões.

1.3 Estrutura

Com o intuito de expor de forma clara o processo em estudo, este relatório encontra-se dividido em cinco capítulos.

No capítulo 2, apresenta-se a Prio Energy, S.A., fazendo-se uma pequena abordagem às 7 áreas de negócio em que esta empresa se encontra envolvida. Apresenta-se o biodiesel, explicando de que forma se pode obter, as suas vantagens, as suas aplicações e ainda os títulos de biocombustíveis (TdB) obtidos através da produção de biodiesel com OAU. Segue-se a descrição do processo de produção do biodiesel na Prio Bio, S.A. Por último, descreve-se detalhadamente o processo de neutralização do óleo que foi o foco do estágio curricular e do presente relatório, seguindo-se da apresentação e caracterização das variáveis em estudo.

No capítulo 3 realiza-se a análise exploratória de todos os dados do processo de neutralização do óleo referentes ao ano 2017 e início de 2018 (até 23 de fevereiro).

No capítulo 4 são apresentados e discutidos os resultados que foram sistematizados sob a forma do ciclo de melhoria contínua *Plan-Do-Check-Act (PDCA)*.

Os dados dos Capítulos 3 e 4 foram tratados com recurso aos *softwares* Excel 2016, SPSS versão 25, ACCEPT e Minitab 18.

O capítulo 5 consiste num resumo das conclusões tiradas no decorrer do estágio.

No final do relatório apresentam-se os anexos, dispostos da seguinte forma:

- Apêndice A: Análise do processo de neutralização;
- Apêndice B: Regressão Linear Múltipla.

Capítulo 2

Apresentação da Prio Energy, S.A.

Fundada em outubro de 2006, a Prio Energy é uma empresa de produção e distribuição de combustíveis. Com capital 100% português pertencente à Oxy Portugal, é a única gasolinheira ibérica com tripla certificação QSA (Qualidade, Segurança e Ambiente) [2]. A Prio Energy possui 7 áreas de negócio como se pode ver ilustrado na Figura 2.1.



Figura 2.1: Áreas de negócio da Prio Energy, S.A.

(Imagem cedida pela empresa Prio Energy, S.A.)

Desde 2006 que a Prio possui uma fábrica de biodiesel localizada no Porto de Aveiro conhecida como Prio Bio, S.A., iniciando a produção apenas no ano de 2007, cuja capacidade de produção anual é de 113880 toneladas, e conta ainda, com um laboratório de controlo que está em funcionamento 24h por dia para garantir a qualidade da produção. Esta preocupação com a qualidade dos produtos não está apenas relacionada com as especificações da Norma

Europeia de qualidade EN14214 que o biodiesel deve cumprir, mas também com a política da empresa, cujo foco é a inovação e a melhoria contínua dos seus produtos. Para tal, a Prio colabora com instituições de investigação e universidades no desenvolvimento de projetos que vão ao encontro destes objetivos [2].

2.1 Biodiesel

O biodiesel é a aposta da Prio Bio, S.A. por ser um combustível renovável e biodegradável, sendo por isso considerado “amigo do ambiente”. É produzido através de óleos virgens e óleos alimentares usados, sendo obtido por meio de uma reação química entre triglicerídeos e álcool, na presença de um catalisador, denominada por reação de transesterificação.



Figura 2.2: Reação de transesterificação

Desde o dia 1 de janeiro de 2010 que todo o gasóleo rodoviário (convencional), comercializado em Portugal, contém 7% de biodiesel (B7) [3]. Esta incorporação do biodiesel no gasóleo mostra-se vantajosa, uma vez que o biodiesel pode ser produzido através de óleos alimentares usados e por isso, consegue-se fazer o reaproveitamento dos mesmos, reduzindo assim a quantidade de resíduos depositados em esgotos e aterros sanitários, evitando custos de tratamento das águas, entupimento de canalizações e maus cheiros e diminuindo a utilização de combustíveis fósseis. De acordo com a APA, um litro de óleo polui até um milhão de litros de água [4]. No sentido de sensibilizar as pessoas para a reciclagem do óleo alimentar usado, foi criado em 2016 o projeto PRIO TOP LEVEL, que resulta de um consórcio entre a Prio Bio, S.A. e a empresa Hardlevel. No âmbito deste projeto foi implementada uma rede de oleões para permitir que as pessoas tenham os recursos necessários para reciclar o óleo, e são também realizadas várias ações de sensibilização, onde são dados a conhecer os impactos ambientais e económicos associados ao OAU, bem como as vantagens da recolha deste resíduo e o modo como pode ser feito.

O biodiesel face ao diesel convencional mostra-se profícuo, pois o índice de cetano é superior no biodiesel o que influencia de forma positiva a qualidade da combustão, sendo assim mais completa, reduzindo em cerca de 90% as emissões de CO_2 e em 98% de enxofre libertado para a atmosfera [5], que conseqüentemente reduz o efeito de estufa. Reduz também os ruídos do motor e melhora a sua eficiência e desempenho. Tem maior poder lubrificante reduzindo

o desgaste do motor, a probabilidade de avarias e diminuindo os custos de manutenção. É necessária uma fonte de calor acima dos 150°C para que ocorra a combustão do biodiesel, pelo que simplifica o seu transporte e armazenamento, uma vez que o risco de explosão é baixo. Contudo, o custo de produção de biodiesel é muito elevado.

Todas as entidades responsáveis pela comercialização de combustíveis fósseis para consumo final, estão obrigadas a contribuir para o cumprimento das metas de incorporação de biocombustíveis previstas por lei, designadamente 7,5% em teor energético [6]. A produção de biocombustíveis como é o caso do biodiesel gera TdB, que serve como prova para atingir esta meta. Assim, com o intuito de incentivar a produção de biodiesel através de resíduos como, por exemplo, óleos alimentares usados, por cada tonelada equivalente de petróleo (tep¹) de biocombustíveis sustentáveis incorporados no consumo, há lugar à emissão de 2 TdB [7]. Os TdB podem ser transacionados entre agentes económicos para comprovar a incorporação de biodiesel e, portanto, têm valor monetário. A produção de biodiesel com OAU é mais complexa do que com óleos virgens, no entanto, o uso de um resíduo na produção do biodiesel torna o produto economicamente mais viável e ambientalmente mais sustentável.

A aplicação do biodiesel em combustíveis não é única, pois o biodiesel pode também ser utilizado como componente na formulação de poliuretanos, sendo considerado um excelente solvente dada a sua viscosidade e solvência [2].

2.2 Produção de biodiesel na Prio Bio, S.A.

Ao longo deste capítulo, as variáveis serão apresentadas por letras devido à confidencialidade do processo.

Os óleos de A e B são usados em percentagem suficiente para conferirem determinadas características ao biodiesel final, de forma a cumprir a norma do biodiesel, EN 14214.

O óleo B é usado no verão e o óleo A, utilizado no inverno devido às suas propriedades que são favoráveis para o cumprimento das especificações de Cloud Point e Cold Filter Plug Point (CFPP) pertencentes à EN 14214.

Os óleos, sejam eles, óleos alimentares usados ou óleos virgens, são transportados por via marítima e via rodoviária desde o fornecedor (nacional ou internacional) até à unidade fabril. A Prio Bio, S.A. contém nas suas instalações cinco tanques de armazenamento destinados a esta matéria prima. A produção de biodiesel divide-se em vários processos como está representado na Figura 2.3.

¹A tep tem por objetivo a conversão de todas as formas de energia final para uma única forma de energia primária, permitindo calcular a energia primária total consumida anualmente por um país ou uma fábrica, relacionando esse consumo com a riqueza produzida.

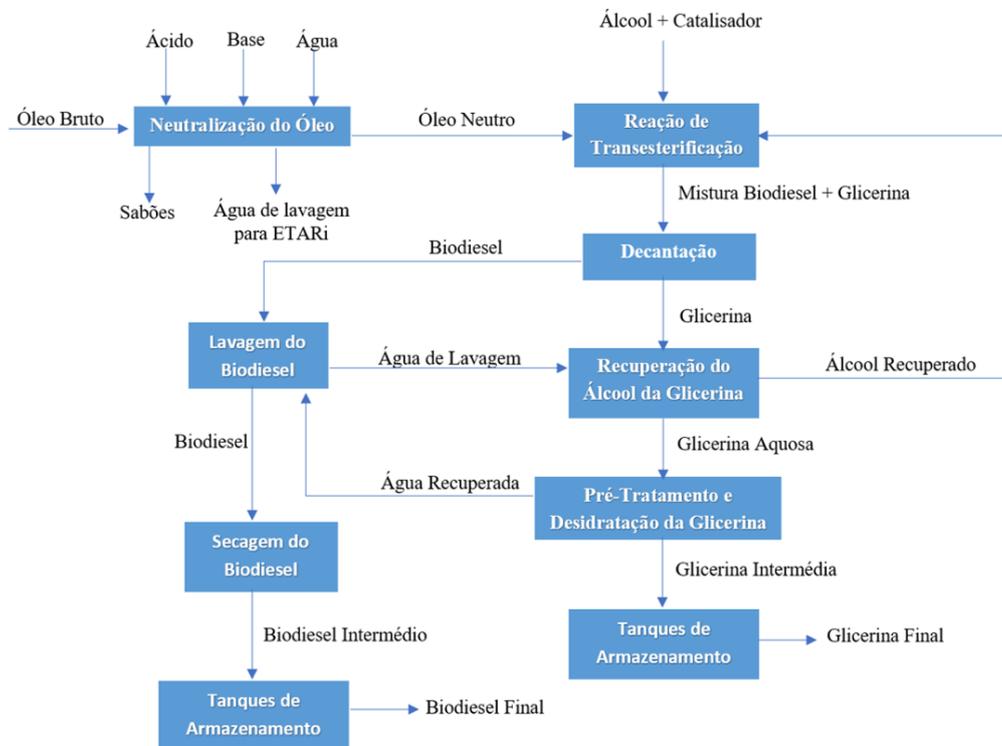


Figura 2.3: Processo de produção de biodiesel na Prio Bio, S.A.

A produção de biodiesel inicia-se com a neutralização destes óleos, seja para remoção de impurezas, seja para redução do valor X. Este processo será detalhado no subcapítulo 2.3. O óleo, agora designado óleo neutro é adicionado a um álcool (adicionado em excesso para melhorar o rendimento da reação) na presença de um catalisador para que ocorra a reação de transesterificação, em que os triglicerídeos reagindo com o álcool formam uma mistura de biodiesel e glicerina. Devido à diferença de densidades entre estes dois produtos é possível separá-los através de decantação (o álcool e catalisador que não reagiram têm afinidade com a fase de glicerina). Após este processo, o biodiesel e a glicerina seguem para diferentes processos. Ao biodiesel é adicionado ácido para neutralizar o excesso de catalisador e deste modo evitar a formação de espumas durante o processo que se segue, o processo de lavagem. O biodiesel é lavado em contracorrente com água, removendo algumas impurezas, Z, glicerol e álcool. Após a lavagem segue-se o processo de secagem do biodiesel, recorrendo a secadores em vácuo e com utilização de vapor. Esta produção segue para um tanque diário e após a análise laboratorial que valida alguns pontos da norma EN 14214 é transferido para o tanque de armazenamento. O biodiesel final vendido para os clientes é sempre certificado por laboratório externo que valida todos os pontos da norma EN 14214.

A glicerina resultante da produção do biodiesel é tratada para que seja possível comercia-

lizá-la. Após a decantação, a glicerina junto com a água de lavagem do biodiesel é acidificada e enviada para uma coluna de destilação para a recuperação do álcool que se encontrava em excesso na reação, sendo possível reutilizá-lo. De seguida a glicerina aquosa segue para um pré-tratamento e posteriormente para a desidratação. O pré-tratamento da glicerina neutraliza o pH. Por fim, tem-se o processo de desidratação da glicerina onde constam três evaporadores, permitindo no final ter um produto entre 80-85% de glicerol, produto muito valorizado no mercado da cosmética. A água resultante deste processo volta a ser usada para a lavagem de biodiesel.

É de notar que o álcool é reaproveitado, voltando a ser introduzido na reação, traduzindo-se numa poupança significativa de matéria prima fresca (álcool novo). Por sua vez, a água é utilizada na menor quantidade possível, sendo que o biodiesel é lavado com a água resultante da desidratação da glicerina.

2.3 Neutralização do óleo

Os óleos alimentares usados possuem um valor de X muito elevado devido às frituras, pois a água presente nos alimentos hidrolisa os triacilgliceróis [8]. O processo de neutralização do óleo visa reduzir a propriedade X, pelo que é um processo imprescindível na produção do biodiesel, dada a necessidade de evitar reações químicas secundárias indesejáveis durante a reação de transesterificação. A reação principal que ocorre no processo de neutralização é uma reação química entre a propriedade X do óleo bruto (OB) e a base, formando Z e reduzindo o valor de X do óleo neutro (ON). Esta reação pode traduzir-se de acordo com a Figura 2.4.

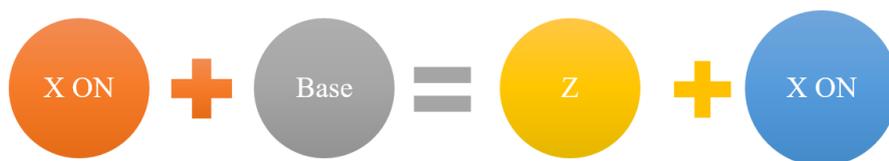


Figura 2.4: Reação química principal da neutralização do óleo

Como referido no subcapítulo 2.2, a Prio Bio, S.A. possui cinco tanques (TK) de armazenamento de óleos nas suas instalações (TK 1, TK 2, TK 3, TK 4 e TK 5, ver Figura 2.5) e ainda um tanque para o qual é bombeada a mistura dos óleos (tanque de receção F1, ver Figura 2.5). Os tanques contêm, no seu interior, os óleos depositados em camadas pois não existe agitador nos mesmos (representado na Figura 2.5 através de cruzes a vermelho e de traços a preto), o que implica a falta de homogeneização dos óleos e uma grande variabilidade do valor de X.

A percentagem bombeada de cada um dos tanques designa-se por formulação e esta é

definida consoante as necessidades da produção. Os tanques de armazenamento encontram-se armazenados no Tankfarm e é a partir daqui que o óleo é transferido, com o auxílio de bombas, para o tanque de receção (F1).

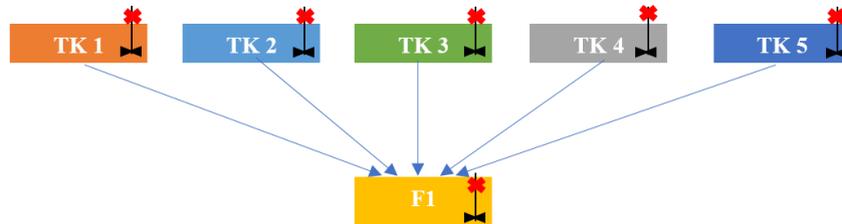


Figura 2.5: Tanques de armazenamento

Deste modo, o processo de neutralização é o primeiro processo na produção de biodiesel.

Na Figura 2.6 encontra-se ilustrado o processo de neutralização do óleo. Note-se que as setas azuis correspondem às variáveis analisadas ao longo do estágio, que serão posteriormente abordadas neste trabalho e encontram-se representadas nos respetivos pontos de amostragem. Ao nível dos inputs para o ajuste da receita o controlo é feito no distributed control system (DCS) por parte do chefe de turno. A variável I é ajustada de acordo com a liquidez dos sabões na primeira centrífuga, isto é, se os sabões estiverem muito líquidos deve-se aumentar a variável, caso contrário, o procedimento é inverso. Por sua vez, as variáveis G e E são ajustadas no início do processo, dado o facto de estas variáveis atuarem sobre a variável I e consequentemente sobre a variável L. Geralmente, a variável G era alterada às terças e sextas, pois eram os dias em que eram feitas as análises ao X dos cinco tanques de armazenamento. Assim, esta variável tomava valores consoante o cálculo estequiométrico de X destes tanques (TK 1, TK 2, TK 3, TK 4 e TK 5). A variável E atua sobre a variável I tal como a variável G sendo esta, no entanto, um ajuste mais fino. Esta variável era alterada de acordo com os resultados de XD3, pois face à base já doseada no processo devido à variável G, poderia ser necessário adicionar mais base se os resultados obtidos para XD3 assim o indicassem.

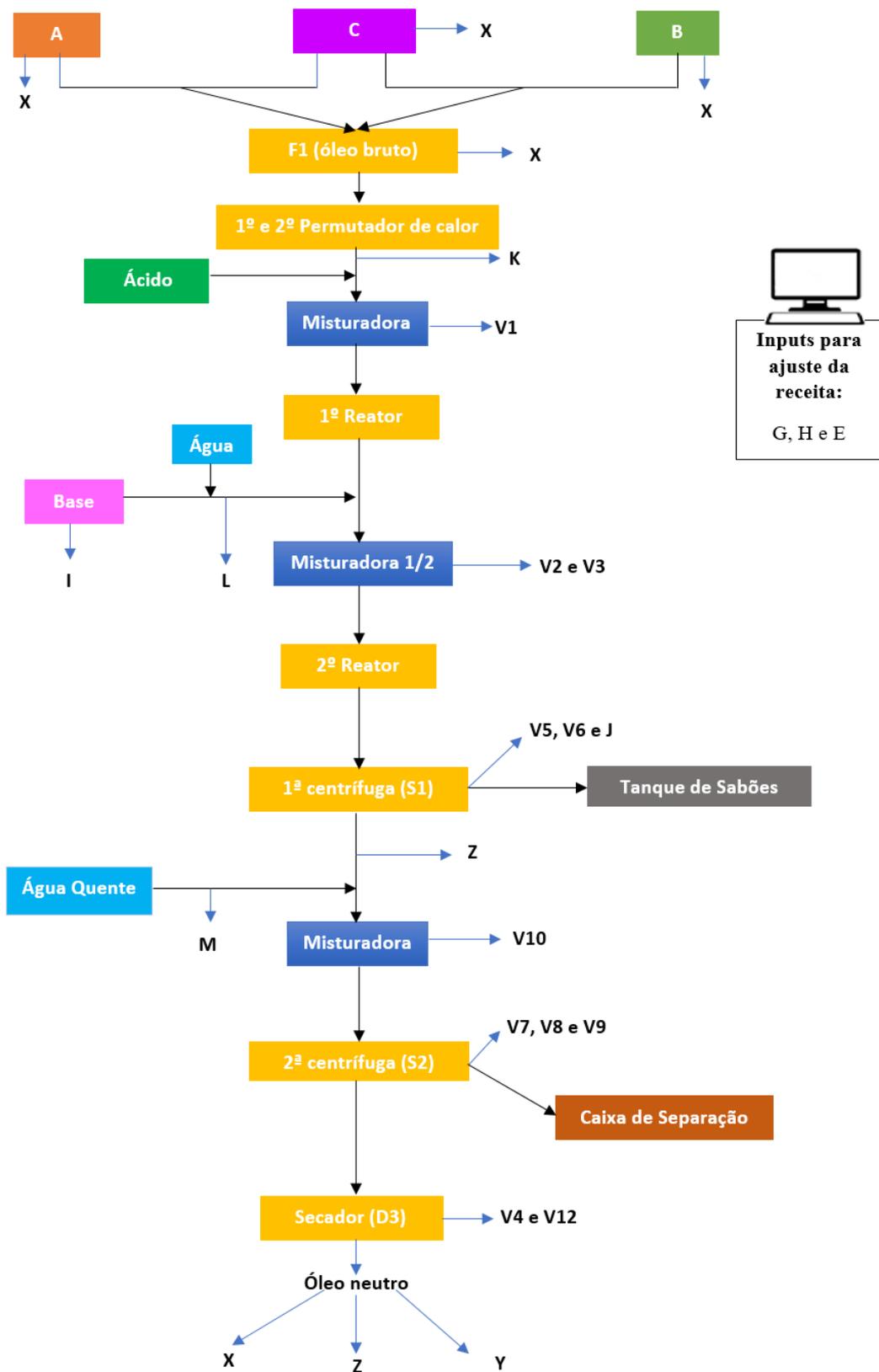


Figura 2.6: Processo de neutralização do óleo

Antes de ser aquecido, o óleo bruto é filtrado, por forma a remover as impurezas/resíduos provenientes, na sua maioria, dos óleos alimentares usados. O aquecimento é feito em duas partes, na primeira parte é pré-aquecido num permutador, sendo o calor proveniente da corrente de saída do processo, ou seja, do óleo neutro, desta forma conseguimos aproveitar o calor deste para aquecer o óleo que está a entrar e conseqüentemente arrefecer o óleo que está a sair. Como este pré-aquecimento não é suficiente, o óleo bruto é aquecido noutra permutador, onde é ejetado vapor, fazendo assim com que o óleo atinja os 90°C (temperatura pretendida). Concluído o aquecimento do óleo bruto, adiciona-se-lhe um ácido para que este quebre as ligações de fósforo presentes no óleo e conseqüentemente reduza o teor em fósforo do mesmo. A reação dá-se num tanque reator.

Dada a necessidade da redução de X, adiciona-se agora uma base ao óleo para que esta reaja com X e forme Z. Esta base é previamente diluída com água para que não seja tão concentrada. Mais uma vez, a reação entre o óleo e a base ocorre num tanque reator que possui um agitador para homogeneizar a mistura entre ambos e assim promover a reação. Esta reação entre X e a base constituem a reação principal do processo de neutralização do óleo.

Como consequência da reação anterior, existe agora a necessidade de remover o Z formado. Deste modo, o óleo irá para uma centrífuga, onde se dará a separação entre a fase pesada que é o Z e a fase leve que é o óleo. A centrífuga deve estar corretamente ajustada ao nível da variável J, pois este influenciará a quantidade de Z que serão removidos do óleo. O Z e as impurezas do óleo são a fase pesada. Esta circula junto à parede da centrífuga e, por isso, quanto maior o valor tomado pela variável J mais óleo, Z e impurezas são removidos, no entanto, há muitas perdas e a probabilidade de a centrífuga parar por entupimento é maior. No caso de estar demasiado fechado, não irá conseguir retirar as impurezas.

Ao óleo é agora adicionada água quente para proceder à lavagem do mesmo. A água misturada com o óleo será removida através de uma nova centrífuga, num processo semelhante ao processo da remoção de Z. A água de lavagem proveniente deste processo contém óleo pelo que vai para uma caixa de separação onde a água e o óleo se separam por decantação.

Para finalizar este processo e proceder à transesterificação, é necessário que o óleo passe pelo secador, pois ainda contém água. O secador tem como função secar o óleo evaporando a água através do efeito de vácuo, que faz com que o ponto de ebulição da água seja menor comparativamente ao seu ponto de ebulição em condições de temperatura e pressão normais (TPN). Ao entrar no secador, a água presente no óleo evapora-se. O óleo resultante desta última etapa é denominado por óleo neutro.

2.3.1 Dados Operacionais

O chefe de turno ajusta na sala de controlo (DCS) algumas variáveis de processo por meio de controladores. Essas variáveis são: E, G, H e J.

Além disso, as variáveis I, K, L e M encontram-se registadas numa base de dados com que estes controladores comunicam, designada por “ACCEPT” que permite o estudo das mesmas de 15 em 15 minutos.

2.3.2 *Check List em Gemba* (chão de fábrica)

A *check list em Gemba* é a *check list* feita em chão de fábrica, ou seja, *check list* feita no processo por parte de um operador com o objetivo de identificar problemas e procurar melhorias [9]. Esta *check list* é feita uma vez por turno (cada turno conta com 8h de trabalho) e os dados recolhidos pelo operador são enviados para a base de dados. As variáveis por ele registadas são V1, V2, V3, V4, V5, V6, V7, V8, V9, V10, V11 e V12.

2.3.3 Análise Laboratorial

Ao nível do laboratório são feitas análises de seis em seis horas, o que corresponde a uma vez por turno. As amostras são recolhidas por parte de um operador. Relativamente ao processo de neutralização do óleo, são alvo de análise laboratorial os seguintes aspetos:

- XF1 e XD3;
- ZS1 e ZD3;
- YD1.

Note-se que também são feitas análises laboratoriais aos óleos dos cinco tanques de armazenamento, contudo, contrariamente ao que acontece com o tanque F1 (tanque da mistura de óleos) onde são feitas múltiplas análises diárias, nestes tanques são apenas analisados os valores de X duas vezes por semana.

O registo das análises é enviado para a base de dados o que permite o cruzamento entre as variáveis e os resultados obtidos em laboratório.

Capítulo 3

Análise Exploratória

Os dados em estudo neste trabalho dizem respeito ao processo de neutralização do óleo e foram divididos segundo formulação, ou seja, divididos em A + C e B + C. Estes dados foram recolhidos entre o dia 1 de janeiro de 2017 e o dia 23 de fevereiro de 2018. À formulação A + C correspondem os períodos de 1 de janeiro de 2017 a 18 de março de 2017, e de 1 de outubro de 2017 a 23 de fevereiro de 2018, e à formulação B + C correspondem os dados de 18 de março de 2017 a 2 de outubro de 2017, contando assim com um total de 1123 instantes de observação. Os períodos de recolha são os seguintes: 00-05h, 05h-12h, 12-18h e 18-24h. Contudo, nem todos os dias contam com estes 4 períodos de recolha, sendo que na formulação A + C estão em falta 161 períodos e na formulação B + C estão em falta 122 períodos. Note-se que nestes valores não estão contabilizadas as paragens da fábrica para manutenção ou férias.

Os valores apresentados neste capítulo não correspondem aos valores reais por motivos de confidencialidade dos dados da empresa.

Numa primeira abordagem aos dados do processo de neutralização do óleo, começou-se por fazer a análise descritiva das variáveis em estudo e classificar as variáveis em diferentes grupos, sendo eles:

- Matéria Prima: percentagem de A, B e C e XA, XB, XC e XF1B;
- Qualidade do Produto Intermédio e Final: ZS1 e ZD3, YD3 e XF1;
- Ajuste da Receita: G, H e E;
- Dados Operacionais e Check List: K, I, L, M, V1, V2, V3, V4, V5, V6, V7, V8, V9, V10, V11 e V12.

Tabela 3.1: Análise Descritiva (A + C)

	%A	XA	% C	XC	XF1	ZS1	ZD3	YD3	XD3	G	H	E
Média	159,744	3,574	277,696	13,689	9,946	4041,286	880,339	2886,190	0,521	9,831	42,740	18,180
Erro-padrão	3,011	0,042	2,660	0,035	0,083	43,464	13,241	38,606	0,005	0,079	0,149	0,644
Mediana	129,980	3,410	303,287	13,635	10,329	4091,210	867,314	2816,236	0,516	10,182	42,460	17,331
Moda	129,980	2,409	303,287	13,865	10,606	2976,804	715,454	2924,119	0,490	11,265	38,994	0
Desvio-padrão	73,636	1,022	64,496	0,853	2,020	1056,641	322,169	938,526	0,115	1,923	3,638	15,724
Variância da amostra	1251,480	0,241	960,080	0,168	0,941	257690,930	23955,931	203299,770	0,003	0,853	3,054	57,067
Assimetria	9,203	4,252	-9,234	1,626	-7,747	-1,558	0,831	2,570	3,675	-8,970	1,635	4,079
Intervalo	389,940	5,767	346,614	4,779	11,980	6902,983	2003,947	6528,467	0,854	12,825	17,764	64,990
Mínimo	43,327	1,945	43,327	11,685	1,642	391,977	29,679	343,581	0,243	1,690	35,095	0
Máximo	433,267	7,712	389,940	16,464	13,622	7294,960	2033,625	6872,048	1,096	14,514	52,859	64,990

Tabela 3.2: Análise Descritiva (B + C)

	%B	XB	% C	XC	F1	ZS1	ZD3	YD3	XD3	G	H	E
Média	109,258	7,004	323,473	15,049	13,113	4375,695	905,029	3008,519	0,598	12,826	42,526	17,005
Erro-padrão	1,981	0,110	1,973	0,051	0,073	34,032	9,411	33,660	0,006	0,067	0,116	0,648
Mediana	108,317	6,213	324,950	15,021	13,238	4351,214	894,913	2855,879	0,598	12,998	42,027	12,998
Moda	86,653	10,138	346,614	14,168	12,803	3362,845	888,241	3340,489	0,611	12,998	41,160	0
Desvio-padrão	45,381	2,510	45,199	1,157	1,672	773,052	214,190	769,031	0,128	1,526	2,645	14,769
Variância da amostra	475,331	1,454	471,517	0,309	0,645	137930,856	10588,692	136499,837	0,004	0,537	1,615	50,341
Assimetria	10,122	2,418	-10,132	0,048	-3,705	-0,094	0,820	2,661	2,181	-3,809	5,961	3,471
Intervalo	281,624	8,553	281,624	7,747	10,342	5227,930	1467,042	4553,203	0,984	8,579	14,731	60,657
Mínimo	64,990	3,206	86,653	10,277	6,815	1448,585	232,188	996,947	0,091	7,019	38,127	0
Máximo	346,614	11,759	368,277	18,024	17,157	6676,514	1699,230	5550,150	1,075	15,598	52,859	60,657

Nas Tabelas 3.1 e 3.2 estão representadas as análises descritivas para ambas as formulações. Comparando a análise descritiva de A + C com a análise descritiva da B + C, pode-se verificar que durante o período de utilização do óleo B usou-se, em média, uma percentagem maior de óleo C e que o valor de X no tanque de receção F1 foram conseqüentemente mais elevados, o que leva também a um aumento da variável G. Além disto, todas as variáveis da qualidade de produto intermédio e final são também, em média, mais elevadas durante o período de laboração com a formulação B + C.

Posteriormente a esta análise, fez-se também a representação gráfica em caixas de bigodes e histogramas de algumas variáveis relacionadas com a propriedade X.

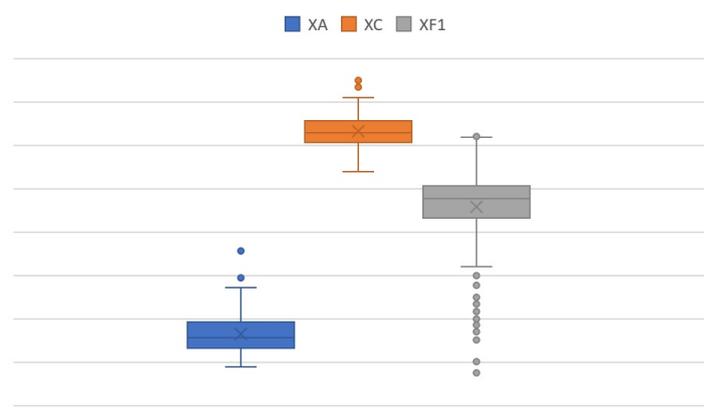


Figura 3.1: Caixa de bigodes das variáveis XA, XC e XF1, respetivamente

Da Figura 3.1, verifica-se que existem alguns *outliers* nas variáveis, sendo o maior número observado na variável XF1. A representação gráfica da variável XC aparenta ser simétrica, ao contrário das restantes, dado que a variável XA tem o primeiro quartil menor do que o terceiro e a variável XF1 tem o primeiro quartil maior do que o terceiro.

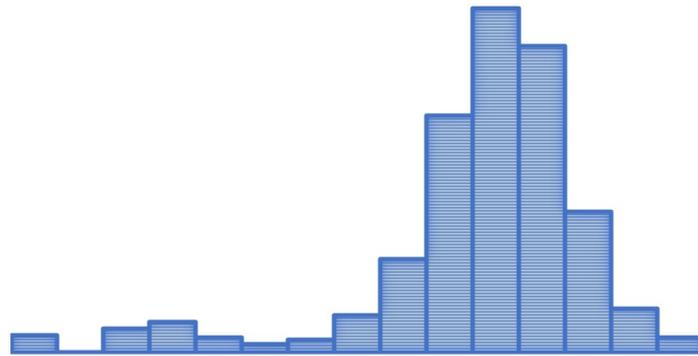


Figura 3.2: Histograma da variável XF1 para a formulação A + C

Na Figura 3.2 encontra-se ilustrado o histograma da variável XF1 para a formulação A + C. Deste histograma pode-se verificar que a propriedade X para a mistura de óleos encontraram-se, maioritariamente, compreendidos entre 9,874 e 10,697. Existem também muitos valores abaixo de 8,228, como se tinha verificado anteriormente na caixa de bigodes.

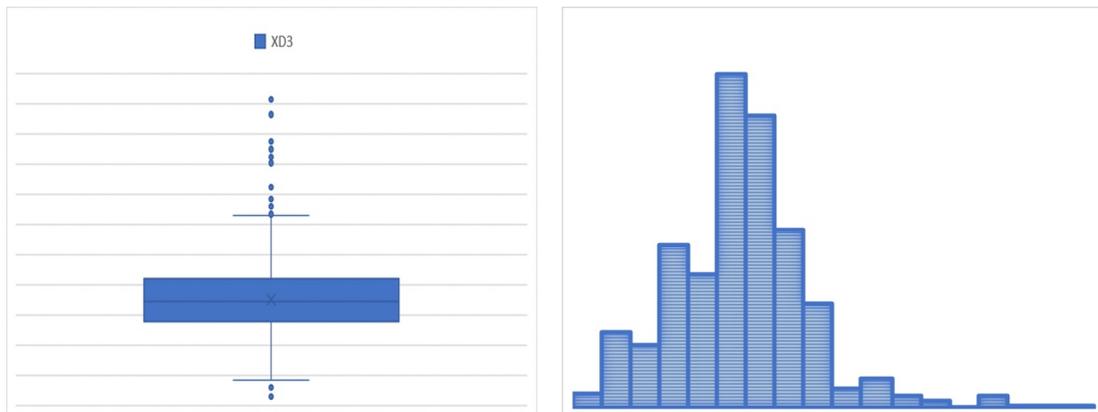


Figura 3.3: Caixa de bigodes e histograma para XD3, respectivamente (formulação A + C)

Ao contrário do que se verificou tanto na caixa de bigodes como no histograma da variável XF1, na caixa de bigodes da variável XD3 verifica-se a existência de muitos valores acima do limite superior, o que se reflete na representação do histograma.

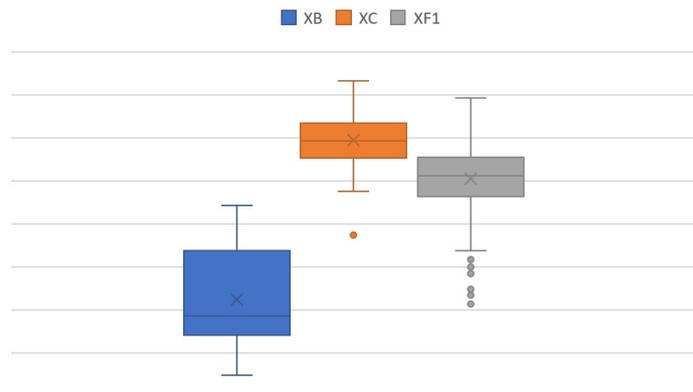


Figura 3.4: Caixa de bigodes das variáveis XB, XC e XF1, respetivamente

Na Figura 3.4 encontram-se representadas as caixas de bigodes das variáveis XB, XC e XF1 para a formulação B + C. Da análise das caixas, verifica-se para esta formulação os valores considerados *outliers* das variáveis são em menor quantidade comparativamente com a formulação A + C. A caixa de bigodes da variável XB é aquela que apresenta maior dispersão dos dados, pois o terceiro quartil é bastante superior face ao primeiro quartil, o que indica que a maioria dos valores se encontra acima da média. No caso da caixa de bigodes da variável XF1, esta aparenta ser simétrica para esta formulação.

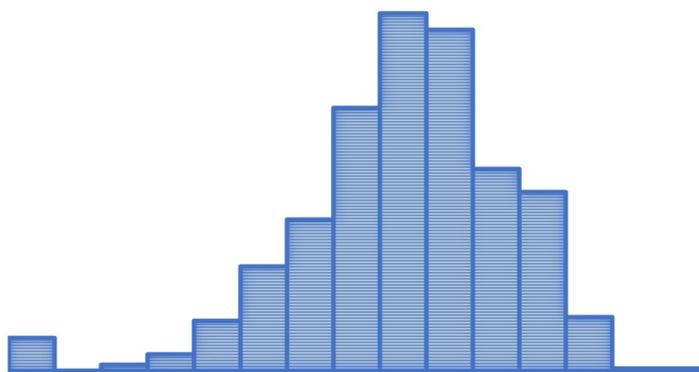


Figura 3.5: Histograma da variável XF1 para a formulação B + C

No histograma da variável XF1 (Figura 3.5) verifica-se que a propriedade X para a mistura de óleos encontraram-se, maioritariamente, compreendidos entre 12,968 e 13,444, o que representa um valor bastante superior ao obtido na formulação A + C. Com este gráfico, consegue-se verificar que o histograma não é simétrico, tendo na sua maioria, valores mais baixos do que a mediana.

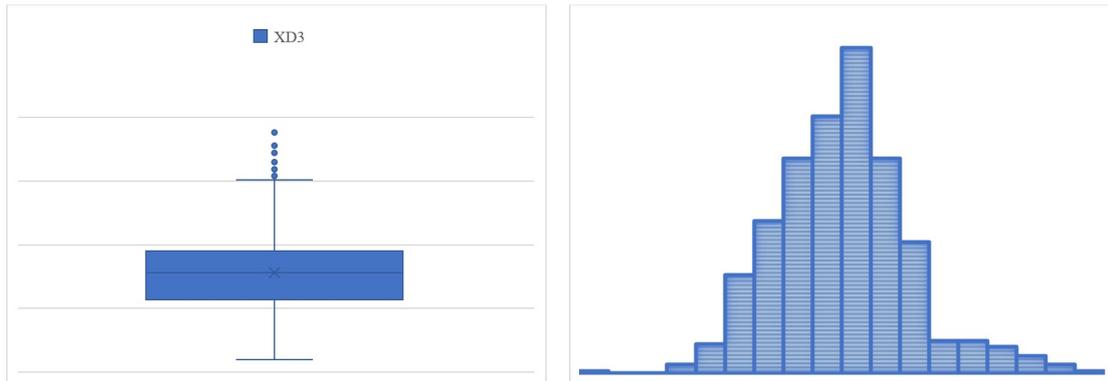


Figura 3.6: Caixa de bigodes e histograma para XD3, respetivamente (formulação B + C)

Analogamente ao que sucedeu na variável XD3 para a formulação A + C, também na formulação B + C se verifica a existência de muitos *outliers* (Figura 3.6), contudo, na caixa de bigodes verifica-se a existência de maior número de valores abaixo da mediana, pelo que é assimétrica, o que se pode também constatar no histograma da variável.

Como visto na análise descritiva das variáveis XF1 e G, para ambas as formulações, o seu comportamento ao longo do tempo é em geral semelhante (Figura 3.7 e Figura 3.8). No entanto, verificam-se algumas discrepâncias entre ambos, o que sucede principalmente devido ao facto de o valor da variável G ser alterada com pouca frequência e a variável XF1 ser medida quatro vezes ao dia, pelo que existe maior variabilidade nesta última variável comparativamente com a variável G. Note-se que na Figura 3.7 os intervalos entre as observações 76 a 199 e 205 a 268 não têm observações devido ao uso de óleo B.

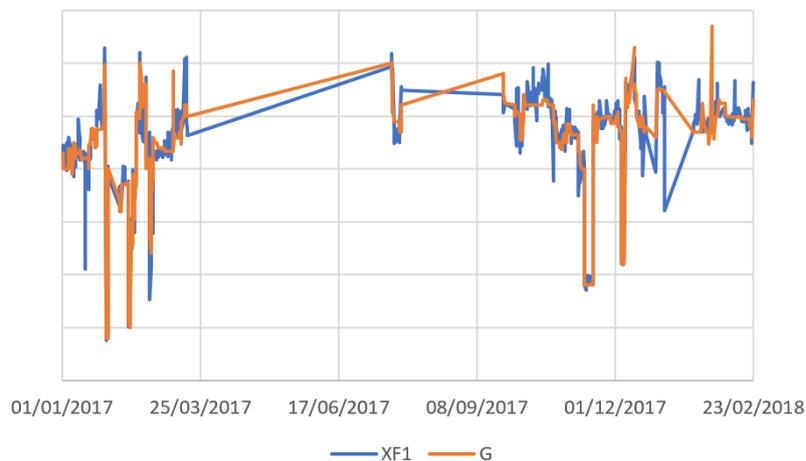


Figura 3.7: Série temporal para a formulação A + C

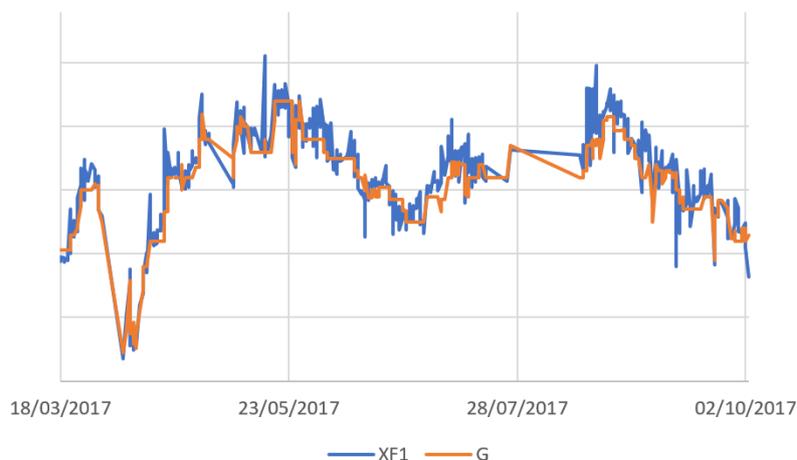


Figura 3.8: Série temporal para a formulação B + C

Concluída esta análise inicial, procurou-se estudar as correlações entre todas as variáveis do processo. Nas Tabelas 3.3 e 3.4 é possível ver algumas das correlações consideradas mais relevantes, isto é, iguais ou superiores a 0,5, para as formulações A + C e B + C, respetivamente.

Tabela 3.3: Correlações A + C

	%A	%C	XF1	ZS1	ZD3	YD3	G	H
%C	-4,322							
XF1	-3,922	3,833						
ZS1	-2,760	2,494	2,731					
ZD3	-2,420	2,187	2,213	3,321				
YD3					2,300			
G	-4,037	3,952	4,054	2,713				
I	-3,360	3,139	3,446	2,319			3,576	
L	-3,005	2,754	3,129				3,208	
V9								2,404
V11	-2,475	2,428	2,237		2,567	3,056	2,283	
V12					2,370	2,620		

Tabela 3.4: Correlações B + C

	%B	XB	%C	XF1	ZS1	ZD3	YD3	G
%C	-4,300							
XF1	-3,185		3,256					
ZD3					2,759			
YD3						2,458		
G	-3,451		3,505	4,048				
I	-3,187		3,236	3,721				3,800
L	-2,920		2,940	3,211				3,192
M		2,340						
V4		2,540						
V11							3,230	
V12							2,591	

Algumas das correlações existentes eram esperadas, no entanto, existiram correlações com valores mais baixos (menores que 0,5) do que era esperado e, por isso, já não foram consideradas para análise. Para ambas as formulações, observa-se que a %C está correlacionada com XF1, G, I e L, o que era esperado pois à medida que se aumenta a percentagem de óleo C é natural que o valor de X aumente e consequentemente aumente o doseamento de base no processo. Ao longo do estudo do processo junto dos profissionais, percebeu-se que o valor de X do óleo neutro é afetado, por exemplo, pela variável H, I e L adicionada ao óleo, no entanto, não se verificou uma correlação elevada (igual ou superior a 0,5) com as restantes variáveis em estudo.

Dado que a formulação em laboração na fábrica durante o estágio era A + C, todos os resultados apresentados a partir deste ponto são referentes apenas a esta formulação, pouco tendo a ver com a formulação B + C, o que impossibilita a generalização para a mesma. Desta forma, os dados utilizados correspondem às datas anteriormente mencionadas contando com um total de 525 observações.

Posteriormente à análise das correlações, recorreu-se à análise fatorial em componentes principais das variáveis que se considerou mais importantes, sendo estas: %C, XF1, ZS1, XD3, YD3, ZD3, G, H, E, I e J. Embora nem todas estas variáveis tenham apresentado uma correlação elevada, pensa-se que estas sejam as que têm uma maior influência na qualidade do produto final.

Análise Fatorial

A Análise em Componentes Principais é uma metodologia que visa estudar a estrutura da covariância dos dados.

No *Biplot* (Figura 3.9) encontram-se representados por um círculo azul os objetos e as setas pretas representam os *component loadings* (ajustados à escala de objetos). Num *Biplot* devem ser analisados 3 pontos: o comprimento das setas, o ângulo entre as setas e a projeção dos pontos [10]. Relativamente ao comprimento das setas nem todas aparentam ter o mesmo comprimento, sendo a seta da variável E a que parece ter menor comprimento. As setas de XD3, J e H aparentam ter comprimento semelhante. As setas das variáveis ZD3, ZS1, XF1 e %C também apresentam comprimentos semelhantes, tal como se verifica entre as variáveis, G e I. A seta da variável YD3 não se assemelha a nenhuma das restantes. Quanto maior o comprimento da seta, maior a percentagem de variância das variáveis que é explicada pela solução da análise fatorial, pelo que se conclui que a variável E é a que tem menor variância explicada e as variáveis G e XF1 são as variáveis que têm percentagem de variância explicada pela solução da análise fatorial.

Ao nível do ângulo entre as setas, sabe-se que quanto menor é o ângulo mais correlacionadas estão as variáveis. São exemplo disso, as variáveis G, XF1 e %C cujo ângulo formado

pelas suas setas é muito pequeno e como se viu anteriormente estas apresentavam uma elevada correlação entre si. Esta conclusão vai de encontro ao que era esperado, pois sempre que se aumenta a %C, o X no tanque de mistura (F1) aumenta e consequentemente a variável G também será aumentada.

No que diz respeito à projeção dos pontos, pode-se verificar que os objetos mais próximos entre si, dentro de um mesmo quadrante (entenda-se quadrante, um ângulo de 90 graus ou menos formado por 2 setas) apresentam resultados idênticos nas variáveis marcadas, no entanto, se estiverem num quadrante diferente a conclusão será precisamente a oposta. As observações que estão em direção contrária à direção de uma determinada seta, têm resultados abaixo da média para essa variável. As observações que se encontram posicionadas no seguimento de uma seta têm resultados acima da média para essa variável. Por exemplo, os objetos 27 e 28 encontram-se representados à frente da seta E, pelo que se conclui que estas observações tiveram valores superiores à média desta variável. No caso dos objetos mais afastados da origem do gráfico, ponto (0,0), também é possível tirar algumas conclusões relativamente aos mesmos, por exemplo, os objetos 391, 395 e 396 apresentam projeções na direção contrária às setas G, XF1, I e %C, o que significa que os resultados destas três observações estão abaixo da média para estas quatro variáveis. Todos os objetos que se encontram mais próximos da origem e mais próximos uns dos outros, têm resultados similares entre si.

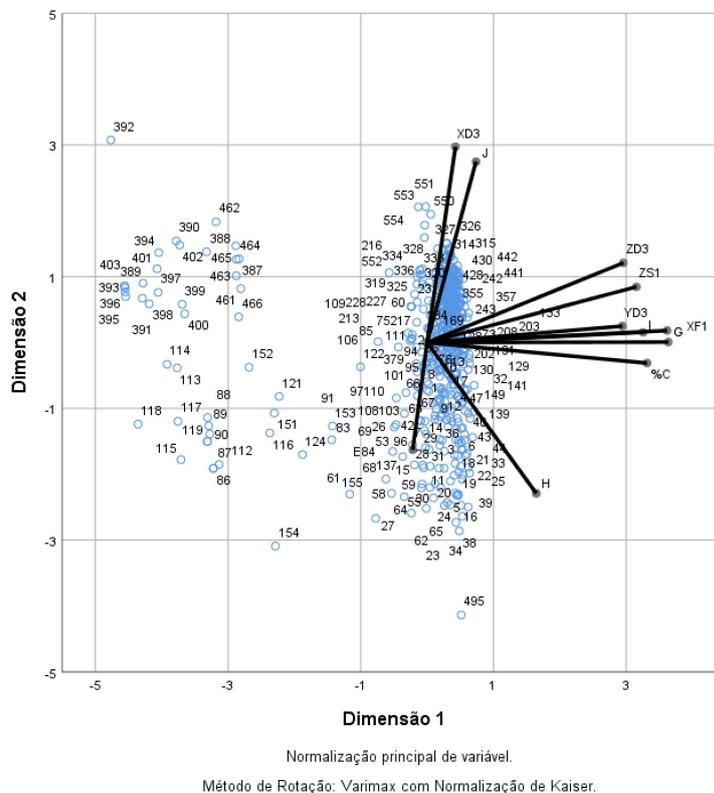


Figura 3.9: *Biplot*

Em suma, com esta análise percebeu-se que as variáveis G, XF1 e %C estão relacionadas entre si, o que era espectável, pois à medida que se aumenta a percentagem de óleo C na formulação, o X da mistura de óleos (F1) também irão aumentar, e uma vez, que a variável G custava tomar valores próximos da variável XF1, é natural que o seu aumento também aconteça. Da análise do *Biplot*, confirmou-se também que à medida que as variáveis H e E aumentam, o XD3 diminui.

A análise fatorial procura identificar um pequeno número de fatores que expliquem as relações existentes entre várias variáveis. Este é um método subjetivo, pois não existe apenas uma forma de interpretar os dados e considera que as variáveis observáveis podem ser agrupadas através das suas correlações, isto é, quanto maior a correlação entre as variáveis, mais provável será que exista um fator comum às duas que as explique. Este tipo de análise fatorial pressupõe vários passos [11].

Primeiramente, é necessário verificar se a técnica é adequada aos dados. Para isso, deve recorrer-se à análise da matriz de correlações, que tal como se verificou anteriormente no estudo das correlações entre as variáveis, não existem correlações nulas, pelo que as variáveis são correlacionadas entre si.

Segue-se a aplicação dos testes, Kaiser-Meyer-Olkin (KMO) e Bartlett. Estes testes avaliam a adequabilidade do método em relação aos dados e testam a hipótese de as variáveis não serem correlacionadas, o que implicaria que a matriz das correlações fosse igual à matriz identidade, pelo que o método não seria apropriado.

Medida Kaiser-Meyer-Olkin de adequação de amostragem.		,767
Teste de esfericidade de Bartlett	Aprox. Qui-quadrado	4068,759
	gl	55
	Sig.	,000

Figura 3.10: Teste KMO e Bartlett

No caso dos dados em análise neste trabalho, o valor do KMO é igual a 0.767 pelo que é aceitável, pois é superior a 0.5 e é ainda, considerado bom por alguns autores por se encontrar entre 0.7 e 0.8 [12]. No teste de Esfericidade de Bartlett o *p-value* foi inferior a 0.001 pelo que a matriz de correlações é estatisticamente diferente da matriz identidade, logo os dados são correlacionados entre si e não apenas consigo próprios.

Correlação anti-imagem	%C	,849 ^a	-.206	,008	-.044	-.316	,004	-.548	-.025	-.048	,140	-.279
XF1	-.206	,848 ^a	-.045	-.085	,089	-.250	-.538	,039	-.327	-.078	,071	
ZS1	,008	-.045	,815 ^a	-.631	,192	-.029	-.158	,069	,045	,125	-.026	
ZD3	-.044	-.085	-.631	,730 ^a	-.446	-.103	,159	-.186	,107	-.119	-.233	
YD3	-.316	,089	,192	-.446	,629 ^a	-.110	,152	-.107	,037	-.155	,313	
XD3	,004	-.250	-.029	-.103	-.110	,746 ^a	,181	,250	,218	,033	-.141	
G	-.548	-.538	-.158	,159	,152	,181	,736 ^a	-.149	,354	-.432	,207	
H	-.025	,039	,069	-.186	-.107	,250	-.149	,512 ^a	-.168	,327	,216	
E	-.048	-.327	,045	,107	,037	,218	,354	-.168	,524 ^a	-.180	,038	
I	,140	-.078	,125	-.119	-.155	,033	-.432	,327	-.180	,842 ^a	-.228	
J	-.279	,071	-.026	-.233	,313	-.141	,207	,216	,038	-.228	,662 ^a	

a. Medidas de adequação de amostragem (MSA)

Figura 3.11: Matriz anti-imagem

De acordo com a matriz anti-imagem e olhando apenas para os resultados na diagonal principal, verifica-se que são todos superiores a 0,5 pelo que todas as variáveis estariam em condições de serem incorporadas.

Verificada a adequacidade do método e das variáveis, estima-se o número de fatores a reter através do *scree plot* e do critério de Kaiser.

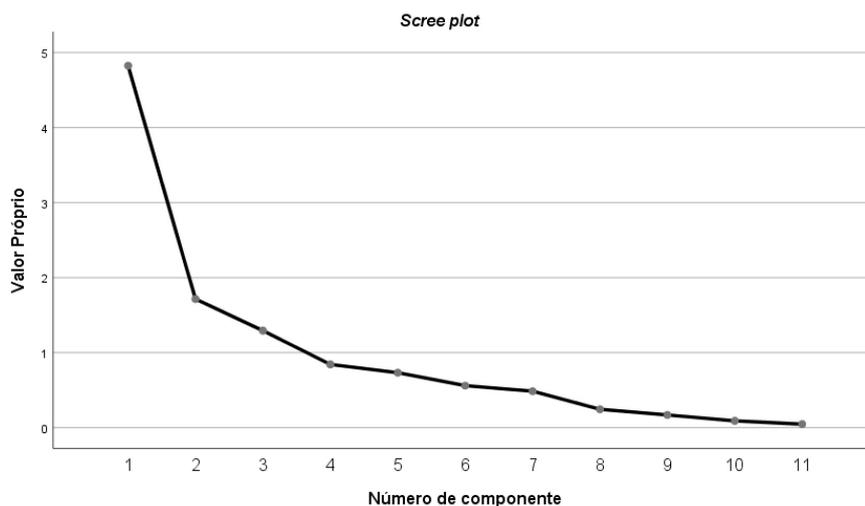


Figura 3.12: *Scree plot*

Com base na Figura 3.12 conclui-se que o número ideal de fatores a reter é igual a 3, dado que a partir daqui o declive da curva diminui bastante. O passo seguinte é escolher o método de extração que irá ser o método das componentes principais.

Componente	Variância total explicada								
	Autovalores iniciais			Somadas de extração de carregamentos ao quadrado			Somadas de rotação de carregamentos ao quadrado		
	Total	% de variância	% cumulativa	Total	% de variância	% cumulativa	Total	% de variância	% cumulativa
1	4,826	43,868	43,868	4,826	43,868	43,868	3,788	34,439	34,439
2	1,714	15,581	59,450	1,714	15,581	59,450	2,171	19,737	54,176
3	1,293	11,751	71,200	1,293	11,751	71,200	1,873	17,024	71,200
4	,843	7,662	78,862						
5	,732	6,652	85,515						
6	,560	5,091	90,606						
7	,485	4,410	95,016						
8	,244	2,219	97,235						
9	,168	1,529	98,765						
10	,090	,816	99,581						
11	,046	,419	100,000						

Método de Extração: análise de Componente Principal.

Figura 3.13: Variância total explicada

Assim como na análise do *scree plot*, também através do critério de Kaiser se concluiu que o número de fatores a reter é 3, pois o valor próprio dos fatores é igual ou superior a 1 até ao terceiro fator e inferior a 1 depois deste. Deste modo, tem-se que o 1º fator explica 43,87% da variância dos dados, o 2º fator explica 15,58% e o 3º fator explica 11,75% da variância dos dados, perfazendo assim um total de 71,2% da variância dos dados explicada por estes 3 fatores retidos.

Comunalidades

	Inicial	Extração
%C	1,000	,895
XF1	1,000	,912
ZS1	1,000	,659
ZD3	1,000	,801
YD3	1,000	,636
XD3	1,000	,577
G	1,000	,934
H	1,000	,716
E	1,000	,333
I	1,000	,799
J	1,000	,570

Método de Extração: análise de Componente Principal.

Figura 3.14: Comunalidades

A coluna das comunalidades (Figura 3.14) contém a informação relativa à variabilidade explicada por cada uma das variáveis em análise, pelo que os seus valores devem ser superiores a 0.5. Pela análise desta tabela, verifica-se que todas as variáveis têm uma forte relação com os fatores retidos, exceto a variável E cujos fatores retidos apenas explicam 33% da variância

desta variável, aproximadamente. Note-se que estes resultados estão de acordo com os obtidos na análise do *Biplot*, tal como era espectável que acontecesse.

Matriz de componente ^a				Matriz de componente rotativa ^a			
	Componente				Componente		
	1	2	3		1	2	3
%C	,901	,252	-,140	G	,951	,173	,009
XF1	,880	,266	-,257	XF1	,936	,181	,050
G	,878	,310	-,258	%C	,899	,293	,033
I	,816	,092	-,353	I	,861	,090	,222
ZS1	,771	-,104	,232	ZD3	,350	,805	,175
ZD3	,734	-,146	,491	YD3	,194	,743	-,216
H	-,105	,707	,454	ZS1	,505	,596	,221
XD3	,401	-,620	,180	E	-,001	-,437	-,377
J	,417	-,605	-,177	H	-,026	,240	-,811
E	-,312	,407	-,265	J	,204	,138	,714
YD3	,476	,167	,618	XD3	,034	,438	,620

Método de Extração: análise de Componente Principal.
a. 3 componentes extraídos.

Método de Extração: análise de Componente Principal.
Método de Rotação: Varimax com Normalização de Kaiser.^a
a. Rotação convergida em 5 iterações.

Figura 3.15: Matriz de componentes sem rotação e com rotação *Varimax*, respetivamente

Nas matrizes de componentes (Figura 3.15) são apresentados os factor *loadings*, que são as correlações entre as variáveis e os fatores [13]. Na matriz de componentes, verifica-se que as variáveis %C, XF1, G, I, ZS1 e ZD3 são explicadas pelo fator 1. O fator 2 explica a variável H, e explica negativamente as variáveis XD3, J e E. Por último, o fator 3 explica a variável YD3. Contudo não é possível associar estes fatores a nenhuma estratégia traçada pelos chefes de turno.

O próximo passo é fazer a rotação. Existem vários tipos de rotação e não há nenhuma que seja a mais indicada a aplicar, por isso, deve-se escolher a que facilita a leitura dos dados e facilita a associação dos fatores. Neste caso, optou-se por usar a rotação *Varimax*. Este método maximiza a soma das variâncias dos *loadings* necessários na matriz de componentes [13]. Com esta rotação, tendem a existir alguns *loadings* altos e outros próximos de zero.

Da análise da matriz de componente rotativa com rotação *Varimax* (Figura 3.15), verifica-se que as variáveis G, XF1, %C, I são explicadas pelo fator 1, enquanto que as variáveis ZD3, YD3, ZS1 e E são explicadas pelo fator 2. Por último, o fator 3 explica a variável J, XD3 e explica negativamente a variável H.

Capítulo 4

Análise Estatística da variável X

Após a análise exploratória da unidade de pré-tratamento e avaliando todas as variáveis, concluiu-se que a propriedade X do óleo bruto e do óleo neutro condiciona todos os processos, tanto na determinação da receita na unidade de neutralização, como na forma como esta propriedade do óleo neutro influencia a reação de transesterificação. Dada a criticidade desta variável, este capítulo destina-se a avaliar o comportamento de X e aplicar o ciclo *PDCA* com vista à melhoria dos resultados do XD3 e XF1. O ciclo *PDCA* é uma ferramenta de gestão de problemas com vista à melhoria contínua dos processos.



Figura 4.1: Ciclo *PDCA*

Todos os valores apresentados neste capítulo, não correspondem aos valores reais por motivos de confidencialidade.

4.1 1ª Fase do ciclo PDCA- *Plan* (Planeamento)

Esta fase passa por identificar os problemas relacionados com a variável crítica X. O óleo neutro apresenta oscilações no seu valor de X. Este X resulta do acerto de uma receita com base no X do óleo bruto. Este acerto é feito pelo chefe de turno após as análises laboratoriais.

Junto com estes profissionais tentou-se identificar as dificuldades no acerto desta receita.

O primeiro problema identificado pela autora desta dissertação ao entrevistar todos os 5 chefes de turno foi a não coerência no ajuste da receita, uma vez que existem duas variáveis G e E, ambas a atuarem sobre a variável I. Verificou-se que a existência destas variáveis para o mesmo efeito levava a alguma confusão por parte dos chefes de turno, ou seja, não se encontra definido qual a variável a alterar e como proceder à sua alteração, uma vez que a variável G é mais representativa do valor de X no tanque da mistura de óleos (F1) do que a variável E, contudo, a variável E resulta num ajuste mais fino, e não existe nenhuma recomendação por parte do fabricante da tecnologia quanto à forma como este ajuste deve ser feito.

O segundo problema identificado pela autora é que a indicação que estes profissionais tinham relativamente ao *target* dos X do óleo neutro não era um valor, mas sim uma gama [P-Q]. Significa isto que, os profissionais não tomavam nenhuma ação se os valores de X do óleo neutro estivessem dentro deste intervalo, o que acabou por se traduzir numa maior oscilação desta variável (Figura 4.2).

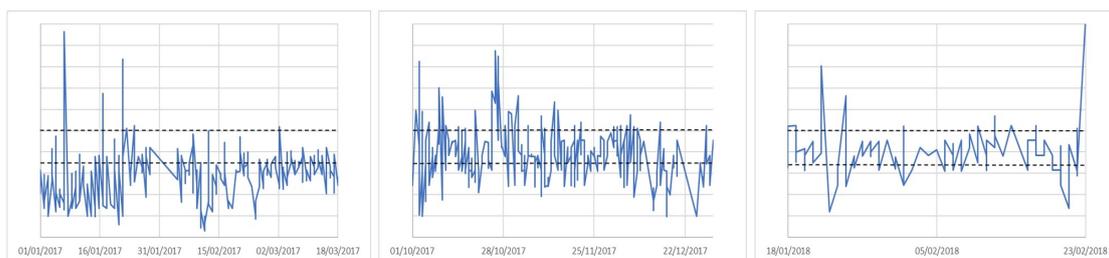


Figura 4.2: Gráficos de dispersão de XD3

Note-se que na Figura 4.2, os gráficos de dispersão foram apresentados de modo dividido, pois os períodos de tempo em falta correspondem ao período de laboração com óleo B e a uma paragem da fábrica, respetivamente.

Por último, o terceiro problema identificado por estes chefes de turno foi a diferença entre os valores de X nos tanques de armazenamento, sem agitação, e no tanque da mistura de óleos (F1) que os conduzem a perceções erradas, no que diz respeito ao ajuste da variável G. Identificou-se a necessidade de realizar um estudo para verificar a discrepância existente.

4.2 2ª Fase do ciclo PDCA – *Do*(Fazer)

Relativamente ao primeiro problema, sugeriu-se ao responsável de produção eliminar a variável E colocando-a a zero, e fazer os ajustes à receita recorrendo apenas à variável G.

No segundo problema a sugestão passou por eliminar o intervalo [P-Q] e colocar como valor de referência R. Desta forma, os profissionais passaram a efetuar sempre um ajuste obtendo valores menos variáveis para a variável X do óleo neutro.

No terceiro problema elaborou-se um estudo aos tanques de armazenamento comparando este com o tanque mistura de óleos F1. A Tabela 4.1 contém os resultados de X correspondentes ao estudo feito aos quatro tanques de armazenamento (tanques que estavam em produção à data do teste) e ao tanque F1, os valores introduzidos na variável G, bem como a diferença entre esta variável e XF1. Através da Tabela 4.1 pode-se concluir que, de hora a hora, o valor de X está sempre a modificar, o que se deve à falta de agitadores e consequentemente à falta de homogeneização dos óleos como já se tinha referido no subcapítulo 2.2.

Tabela 4.1: Estudo da variabilidade de X dos tanques de armazenamento, F1 e G

17/01/2018	8h	9h30	10h20	11h20	14h25	15h05	16h30	Média	Desvio- Padrão
Tanque 1	3,501	3,509	3,605	3,644	3,579	3,418	3,457	3,531	0,082
Tanque 2	13,739	13,652	13,865	13,943	13,592	14,662	13,561	13,860	0,381
Tanque 3	13,670	13,167	13,639	13,648	13,449	13,518	13,345	13,492	0,186
Tanque 4	14,753	14,618	14,826	14,688	14,623	14,813	14,259	14,653	0,195
XF1	11,612	10,949	10,689	10,537	10,767	11,061	10,663	10,897	0,360
G	10,615	10,615	10,615	10,615	10,745	10,745	10,745	10,680	0,069
XF1-G	-0,997	-0,334	-0,074	0,078	-0,022	-0,316	0,082		

Ainda relativamente ao estudo da variabilidade de X nos tanques de armazenamento e F1, é possível comprovar, através da Tabela 4.1, que existe alguma diferença entre o valor de X da mistura dos óleos (F1) e da variável G ajustada por parte do chefe de turno. De modo a verificar este resultado, realizou-se o teste de Wilcoxon (*signed ranks*) que testa se a mediana das diferenças é ou não igual a zero. O resultado obtido para o teste foi $p\text{-value}=0,237$ (ver Apêndice A), pelo que a hipótese nula não pode ser rejeitada, ou seja, a mediana das diferenças não pode ser considerada diferente de zero. Dada a pequena dimensão da amostra e consequente baixa potência do teste, não se pode saber se existem ou não diferenças entre as medianas das populações de onde foram retiradas as amostras.

Além do teste de Wilcoxon, realizou-se também um teste de variâncias, o teste de Levene. Este teste é utilizado para avaliar se existe homogeneidade entre as variâncias das amostras. Este teste foi realizado para as variáveis G e XF1, pois esta última varia muito comparativamente à anterior. Como o $p\text{-value}$ do teste de Levene foi igual a 0,337 (ver Apêndice A), então não se rejeita a hipótese nula, pelo que se pode concluir que as variâncias de ambas as variáveis são iguais.

Numa tentativa de análise deste problema, ou seja, de encontrar uma forma de ajustar devidamente a variável G tentou-se uma nova abordagem que passou por recorrer à Regressão Linear Múltipla, para tentar encontrar um modelo que permitisse conhecer melhor o processo

e assim ajustar da melhor forma a receita, diminuindo a variabilidade do processo.

4.3 3ª Fase do ciclo PDCA- *Check* (Verificar)

4.3.1 Regressão Linear Múltipla

A Regressão Linear é uma técnica utilizada para modelar relações entre variáveis e prever o valor da variável dependente, Y, a partir de um conjunto de variáveis independentes. O conjunto de variáveis independentes pode ser composto por apenas uma variável e, neste caso tem-se uma regressão linear simples, ou por duas ou mais variáveis, tendo assim uma regressão linear múltipla.

Neste trabalho irá utilizar-se apenas a regressão linear univariada múltipla, ou seja, regressão linear múltipla com apenas uma variável dependente. As expressões apresentadas a seguir são, unicamente, respeitantes a este tipo de regressão.

A relação entre as variáveis independentes e a variável dependente pode ser traduzida numa equação linear da seguinte forma:

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_nx_n + \varepsilon \quad (4.1)$$

onde β_i são os coeficientes de regressão (declives parciais), ε representa os erros de média nula, variância constante e independentes (não correlacionados) e β_0 é a ordenada na origem.

Com vista à melhoria dos resultados do valor de X do óleo neutro, recorreu-se à regressão linear múltipla para tentar construir um modelo que permita estabilizar e melhorar estes valores, tornando possível eliminar ou atenuar as oscilações na reação principal de transesterificação, que sucedem devido a reações secundárias.

Para que os modelos de regressão possam ser aplicados é necessário que cumpram determinados pressupostos. Em primeiro lugar deve-se verificar a qualidade do ajuste do modelo através do valor do R^2 e R_a^2 . O R^2 é o coeficiente de determinação do efeito das variáveis independentes sobre a variável dependente, e mede a variabilidade total explicada pela regressão. Por sua vez, o R_a^2 mede qual seria a variabilidade explicada pelo modelo de regressão se a população derivasse da amostra tirada. Considera-se um bom ajuste se R^2 for igual ou superior a 0,5.

O próximo passo é verificar se o modelo é significativo, recorrendo para isso à Tabela ANOVA. Este teste determina se existe uma relação linear entre a variável independente e alguma das variáveis regressoras. Desta forma, se o *p-value* é menor que $\alpha=0,05$ significa que, pelo menos, uma das variáveis independentes tem efeito significativo sobre a variação da variável dependente [14].

Uma vez determinada que existe, pelo menos, uma das variáveis regressoras é importante, deve-se determinar qual ou quais são. Para isso, analisa-se o teste individual feito para cada um dos coeficientes do modelo. Como o método utilizado na regressão foi o método *stepwise*,

apenas se irá verificar o coeficiente da constante, pois o método stepwise apenas insere no modelo as variáveis que consideradas significativas.

Deve-se ainda, verificar o cumprimento do teste de Durbin-Watson. Este teste é usado para verificar a autocorrelação dos erros, isto é, se os erros de observações adjacentes são correlacionados (não independentes). O valor deste teste encontra-se entre 0 e 4 [12]. Os critérios de classificação encontram-se na Tabela 4.2.

Tabela 4.2: Critérios de classificação para o teste de Durbin-Watson

Durbin-Watson (d)	Resultado
d<2	Autocorrelação positiva
d ≈ 2	Não existe autocorrelação
d>2	Autocorrelação negativa

Outro ponto a avaliar é o valor centrado de influência, *Leverage*. O valor de *Leverage* permite estimar a possibilidade de existência de um *outlier* multivariado. A análise deste resultado é feita de acordo com a Tabela 4.3 [15].

Tabela 4.3: Critérios de classificação para o valor de *Leverage*

Leverage	Resultado
<0.2	Aceitável
0.2 - 0.5	Arriscado
>0.5	Existe um <i>outlier</i> multivariado

Além dos pressupostos mencionados acima, é ainda necessário verificar a normalidade dos resíduos e se estes são aleatoriamente distribuídos (não correlacionados). A normalidade dos resíduos pode ser vista através de um gráfico de probabilidade normal, e pode-se afirmar que os erros são normais, se estes se encontram dispostos mais ou menos na diagonal principal. É também possível recorrer ao teste de Kolmogorov-Smirnov para testar a normalidade dos resíduos. A análise da aleatoriedade dos resíduos é verificada através de um gráfico de dispersão.

Por último, averigua-se a existência de multicolinearidade entre as variáveis, isto é, se as variáveis independentes são fortemente correlacionadas entre si. Para tal, analisa-se o fator de inflação de variância (VIF do inglês variation inflation factor) que se for superior a 10 indica a existência de problemas de multicolinearidade [14].

Os modelos apresentados neste trabalho foram construídos com base nos dados correspondentes à formulação A + C recolhidos nos períodos de 1 de janeiro de 2017 a 18 de março de 2017 e 1 de outubro de 2017 a 23 de fevereiro de 2018.

Estes modelos foram testados para esta mesma formulação, pois era a que se encontrava em laboração na fábrica durante o período de estágio. É ainda de realçar que os modelos

não foram testados no mesmo conjunto de dados, pois foram testados individualmente em tempo real, ou seja, apenas era aplicado um novo modelo quando se verificava que o anterior não se ajustava da forma pretendida. Foi utilizado o *software* estatístico SPSS versão 25 e os modelos de regressão foram obtidos através do método *stepwise*.

Ao longo deste capítulo considere-se $\alpha=0,05$.

Uma vez que se procurava estimar os valores de X do óleo neutro, o primeiro passo foi utilizar a variável XD3 como variável dependente. Embora as correlações com esta variável tenham sido baixas (abaixo de 0,5) como se viu no Capítulo 3, utilizou-se como variáveis independentes para o modelo de regressão as seguintes variáveis: % C, XF1, ZS1, YD3, ZD3, G, H e E, pois acredita-se que estas variáveis tenham influência no resultado de X do óleo neutro.

Resumo do modelo^e

Modelo	R	R quadrado	R quadrado ajustado	Erro padrão da estimativa	Mudança de R quadrado	Estatísticas de mudança			Sig. Mudança F	Durbin-Watson
						Mudança F	df1	df2		
1	,354 ^a	,125	,124	,0241525946	,125	80,505	1	562	,000	
2	,493 ^b	,243	,240	,0224875973	,118	87,303	1	561	,000	
3	,511 ^c	,261	,257	,0222375831	,018	13,685	1	560	,000	
4	,516 ^d	,267	,261	,0221729366	,006	4,270	1	559	,039	1,679

a. Preditores: (Constante), ZD3
b. Preditores: (Constante), ZD3, H
c. Preditores: (Constante), ZD3, H, E
d. Preditores: (Constante), ZD3, H, E, YD3
e. Variável Dependente: XD3

Figura 4.3: Resumo do modelo 1 para XD3

Contrariamente ao que era esperado, quando se utiliza a variável XD3 como variável dependente o ajuste do modelo não é bom, como se pode avaliar pelo valor obtido para R^2 (igual a 0,267).

Assim, como a redução de X se dá através da reação química entre o X do óleo bruto com a base (Figura 2.4) achou-se profícuo tentar utilizar a variável G como variável dependente, devido ao facto de esta variável controlar a variável I e de apresentar elevadas correlações com algumas variáveis que se consideram importantes no processo, tais como: %C, XF1, ZS1 e I. Contudo, introduziu-se como conjunto de variáveis independentes não só as que apresentavam elevada correlação com a variável dependente, bem como as que se acredita que afetam o valor de X do óleo neutro, como se ilustra na Figura 4.4.

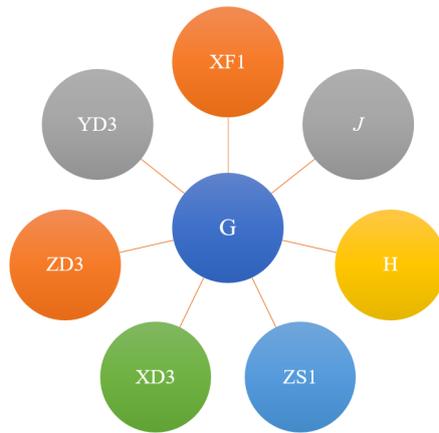


Figura 4.4: Esquema de variáveis relacionadas com G

Com base nas variáveis descritas anteriormente para o conjunto de variáveis independentes e na variável dependente G, obteve-se os outputs apresentados em seguida.

Resumo do modelo^h

Modelo	R	R quadrado	R quadrado ajustado	Erro padrão da estimativa	Mudança de R quadrado	Estatísticas de mudança			Sig. Mudança F	Durbin-Watson
						Mudança F	df1	df2		
1	,941 ^a	,885	,885	,1526792767	,885	4001,441	1	520	,000	
2	,948 ^b	,899	,899	,1432764711	,014	71,492	1	519	,000	
3	,951 ^c	,904	,903	,1400512828	,005	25,179	1	518	,000	
4	,952 ^d	,906	,905	,1387273337	,002	10,934	1	517	,001	
5	,953 ^e	,908	,907	,1374225259	,002	10,864	1	516	,001	
6	,953 ^f	,909	,908	,1365861203	,001	7,339	1	515	,007	
7	,954 ^g	,910	,908	,1361368792	,001	4,405	1	514	,036	1,407

a. Preditores: (Constante), XF1
b. Preditores: (Constante), XF1, I
c. Preditores: (Constante), XF1, I, H
d. Preditores: (Constante), XF1, I, H, XD3
e. Preditores: (Constante), XF1, I, H, XD3, ZS1
f. Preditores: (Constante), XF1, I, H, XD3, ZS1, ZD3
g. Preditores: (Constante), XF1, I, H, XD3, ZS1, ZD3, J
h. Variável Dependente: G

Figura 4.5: Resumo do modelo 2

Na Figura 4.5 é apresentado o quadro sumário do modelo. Deste quadro resumo do modelo 2, verifica-se que o valor de R^2 é igual a 0,906, pelo que se pode dizer que o modelo se ajusta bem aos dados, e que 90,6% da variabilidade total é explicada pelo modelo de regressão. Pode-se também avaliar o valor do teste de Durbin-Watson, que é igual a 1,407 pelo que existe autocorrelação positiva, todavia, o que era desejável era que não existisse autocorrelação [12].

Tabela 4.4: Tabela ANOVA do modelo 2

ANOVA ^a						
Modelo		Soma dos Quadrados	df	Quadrado Médio	Z	Sig
1	Regressão	93,277	1	93,277	4001,441	,000 ^b
	Resíduo	12,122	520	,023		
	Total	105,399	521			
2	Regressão	94,745	2	47,373	2307,686	,000 ^c
	Resíduo	10,654	519	,021		
	Total	105,399	521			
3	Regressão	95,239	3	31,746	1618,523	,000 ^d
	Resíduo	10,160	518	,020		
	Total	105,399	521			
4	Regressão	95,449	4	23,862	1239,906	,000 ^e
	Resíduo	9,950	517	,019		
	Total	105,399	521			
5	Regressão	95,655	5	19,131	1013,024	,000 ^f
	Resíduo	9,745	516	,019		
	Total	105,399	521			
6	Regressão	95,791	6	15,965	855,780	,000 ^g
	Resíduo	9,608	515	,019		
	Total	105,399	521			
7	Regressão	95,873	7	13,696	739,004	,000 ^h
	Resíduo	9,526	514	,019		
	Total	105,399	521			

- a. Variável Dependente: G
- b. Preditores: (Constante), XF1
- c. Preditores: (Constante), XF1, I
- d. Preditores: (Constante), XF1, I, H
- e. Preditores: (Constante), XF1, I, H, XD3
- f. Preditores: (Constante), XF1, I, H, XD3, ZS1
- g. Preditores: (Constante), XF1, I, H, XD3, ZS1, ZD3
- h. Preditores: (Constante), XF1, I, H, XD3, ZS1, ZD3, J

Na tabela ANOVA (Tabela 4.4) avalia-se a significância do modelo. Como é possível averiguar, o modelo é significativo, pelo que existe pelo menos uma variável independente que tem efeito significativo sobre a variação da variável dependente, pois o seu $p\text{-value} < 0,001 < \alpha$. Contudo, a significância do modelo já era esperada, dada a utilização do método *stepwise*.

Coeficientes ^a										
Modelo		Coeficientes não padronizados		Coeficientes padronizados		95,0% Intervalo de Confiança para B		Estatísticas de colinearidade		
		B	Erro Erro	Beta	t	Sig.	Limite inferior	Limite superior	Tolerância	VIF
1	(Constante)	,203	,033		6,106	,000	,138	,269		
	XF1	,901	,014	,941	63,257	,000	,873	,929	1,000	1,000
2	(Constante)	,130	,032		4,012	,000	,066	,194		
	XF1	,747	,023	,780	33,043	,000	,702	,791	,350	2,860
	I	,006	,001	,200	8,455	,000	,005	,008	,350	2,860
3	(Constante)	-,256	,083		-3,075	,002	-,419	-,092		
	XF1	,719	,023	,751	31,610	,000	,675	,764	,330	3,035
	I	,007	,001	,233	9,703	,000	,006	,009	,323	3,098
	H	,038	,008	,071	5,018	,000	,023	,053	,923	1,084
4	(Constante)	-,090	,096		-,936	,350	-,279	,099		
	XF1	,736	,023	,768	31,885	,000	,690	,781	,315	3,178
	I	,007	,001	,227	9,512	,000	,006	,008	,321	3,117
	H	,029	,008	,054	3,635	,000	,013	,045	,815	1,227
	XD3	-,834	,252	-,049	-3,307	,001	-1,329	-,338	,833	1,200

Figura 4.6: Coeficientes do modelo 2

Ao nível dos coeficientes (Figura 4.6), dada a utilização do método *stepwise*, as variáveis independentes inseridas no modelo, já são as que o método considera que são mais significativas e que melhor explicam os dados, pelo que os seus coeficientes são menores que 0,05.

Contudo, a constante não é significativa pois o seu $p\text{-value}=0,350$, logo o modelo pode ser descrito da seguinte forma:

Modelo 2:

$$G = 0,736 (XF1) + 0,007 (I) + 0,029 (H) - 0,834 (XD3)$$

Ainda na Figura 4.6, deve-se avaliar o VIF. Este parâmetro deve tomar valor inferiores a cinco, o que de facto se verifica para o modelo escolhido, assim, pode-se considerar que não existem problemas de multicolinearidade entre as variáveis.

Estatísticas de resíduos^a

	Mínimo	Máximo	Média	Erro Desvio	N
Valor previsto	,4966396689	3,011677504	2,265480843	,4289724586	522
Erro Valor previsto	-4,123	1,739	,000	1,000	522
Erro padrão do valor previsto	,008	,041	,016	,005	522
Valor previsto ajustado	,5013681650	3,011844397	2,265410172	,4287675830	522
Resíduo	-,626173675	,9806593060	,0000000000	,1352192393	522
Erro Resíduo	-4,600	7,203	,000	,993	522
Resíduos Resíduo	-4,655	7,241	,000	1,004	522
de Estud.	-,641214848	,9908015132	,0000706709	,1381590172	522
Resíduos de Estud.	-4,751	7,633	,001	1,015	522
Mahal. Distância	,727	46,239	6,987	6,011	522
Distância de Cook	,000	,177	,003	,011	522
Valor de ponto alavanca centralizado	,001	,089	,013	,012	522

a. Variável Dependente: G

Figura 4.7: Estatísticas dos resíduos do modelo 2

O valor máximo centralizado de alavacagem (Figura 4.7) é igual a 0,089, ou seja, é menor que 0,2, então pelos critérios mencionados na Tabela 4.3, não indica a presença de um *outlier* multivariado.

O gráfico de probabilidade normal dos resíduos padronizados sugere que estes sejam normais, pois encontram-se maioritariamente dispostos sobre a diagonal. O teste de Kolmogorov-Smirnov destina-se a averiguar se uma amostra provém de uma determinada população, que neste caso seria uma distribuição normal [16].

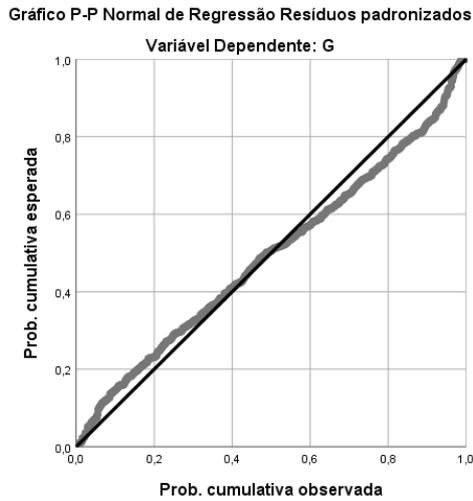


Figura 4.8: QQ plot dos resíduos do modelo 2

Assim, para confirmar esta probabilidade dos resíduos, aplicou-se este teste de normalidade à variável “res-1” (coluna dos resíduos) com o qual obtemos um *p-value* igual a 0,000 pelo que se rejeita a hipótese nula, ou seja, rejeita-se a normalidade dos resíduos.

Testes de Normalidade

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Estatística	df	Sig.	Estatística	df	Sig.
Unstandardized Residual	,075	522	,000	,912	522	,000

a. Correlação de Significância de Lilliefors

Figura 4.9: Teste de normalidade para o modelo 2

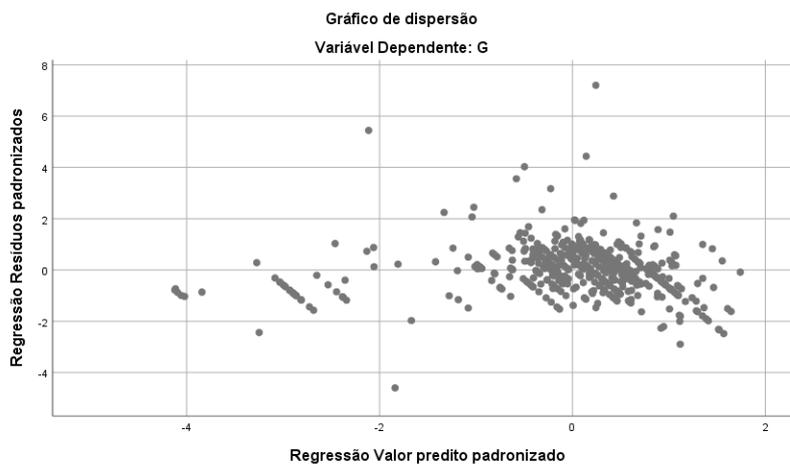


Figura 4.10: Gráfico de dispersão do modelo 2

O gráfico de dispersão (Figura 4.10) apresenta os pontos dispersos, logo os resíduos não são correlacionados. No gráfico é possível observar pontos dispostos sobre retas de declive negativo, todas elas paralelas, o que acontece devido ao facto de esta variável sofrer poucas alterações, pois a variável G apenas era alterada quando haviam novas análises dos tanques, e nem sempre isso acontecia, pois existia também a variável E que também era utilizada no ajuste da base doseada do processo.

Como foi explicado no subcapítulo 4.2., o valor de referência para a variável XD3 era R, pelo que no modelo 2 se substituiu a variável XD3 por este valor, pois pensou-se que assim seria possível estimar o valor correto para a variável G de modo a que os valores de X do óleo neutro estivessem próximos do valor de referência. Os resultados obtidos são apresentados na Tabela B.1.

Os resultados não foram ao encontro do desejado, isto é, o modelo não se adaptou devidamente às alterações que ocorrem no processo. Quando os valores do XF1 aumentam, ambos os modelos deveriam aumentar o valor do G, pois seria necessário dosear mais base para reagir com X, e o contrário também deveria acontecer quando estes valores diminuem. Outra situação que se verificou foi quando se diminuía a variável H e o XF1 aumentava, era esperado que o valor de G (valor este que era calculado através dos modelos de regressão) tivesse um aumento mais acentuado, pois se a base é menos concentrada, então nesta situação precisava-se de a dosear em maior quantidade.

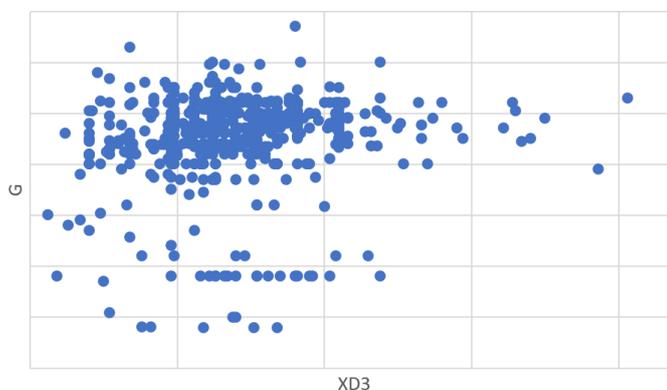


Figura 4.11: Gráfico de dispersão

Na Figura 4.11, encontra-se representado o gráfico de dispersão entre a variável G e a variável XD3. Através deste gráfico é possível perceber porque não foi possível utilizar um modelo de regressão linear com a variável dependente G para estimar os valores de X do óleo neutro, dado não existir uma relação linear evidente entre ambas as variáveis.

Assim, percebeu-se que se estava a utilizar a variável G como uma variável de saída, porque se estava a dizer que esta variável dependia de todas as restantes, quando isso não era totalmente verdade.

Após esta análise, constatou-se que a variável G era afetada pelo valor de X dos tanques do óleo A e C, pelo que seria pertinente introduzir estas variáveis num novo modelo. Além disso, como já tinha sido mencionado no subcapítulo 4.1., a variável G é também representativa do valor de X no tanque da mistura de óleos (F1), pelo que a variável XF1 também foi considerada para o novo modelo.

O output obtido para o novo modelo é apresentado na Figura 4.12. Este modelo explica 91,4% da variabilidade total dos dados, pelo que se pode dizer que este modelo teve um bom ajustamento aos dados. O teste de Durbin-Watson foi igual a 1,233 podendo-se afirmar que existe autocorrelação positiva, contudo, o que era pretendido era que não existisse autocorrelação. Note-se que as variáveis C_Cal e A_Cal foram obtidas através do cálculo %C*XC e %A*XA, respetivamente.

Resumo do modelo^d

Modelo	R	R quadrado	R quadrado ajustado	Erro padrão da estimativa	Mudança de R quadrado	Estatísticas de mudança			Sig. Mudança F	Durbin-Watson
						Mudança F	df1	df2		
1	,936 ^a	,875	,875	,1570657342	,875	4163,073	1	593	,000	
2	,956 ^b	,914	,914	,1302683290	,039	270,065	1	592	,000	
3	,959 ^c	,919	,919	,1265216947	,005	36,580	1	591	,000	1,233

a. Preditores: (Constante), XF1
b. Preditores: (Constante), XF1, C_Cal
c. Preditores: (Constante), XF1, C_Cal, A_Cal
d. Variável Dependente: G

Figura 4.12: Resumo do modelo 3

O teste feito à significância do modelo também foi verificado, pelo que se obteve o seguinte modelo, onde a variável A_Cal não foi considerada.

Modelo 3:

$$G = 0,412 + 0,494 (XF1) + 0,363 (\%C*XC)$$

O valor dos VIF para este modelo também foi menor do que dez, pelo que não há indicações de problemas de multicolinearidade. O ponto máximo de alavancagem foi de 0,055 não havendo indicação de existência de *outliers*. No Q-Q *plot* os pontos encontram-se ligeiramente afastados da diagonal principal. Através do teste de Kolmogorov-Smirnov concluiu-se que os resíduos não são normais, pois rejeita-se a hipótese nula ($p\text{-value}=0,000$).

Os resultados obtidos com este modelo foram testados num novo conjunto de dados e encontram-se na Tabela 4.5.

Tabela 4.5: Resultados Obtidos com o Modelo 3

Data	Hora	X-TK 2	X-TK 3	X-TK 5	XC	XF1	G	XD3
Formulação do OAU: 25% TK 2+ 26% TK 3 + 10% TK 5								
20/03/2018	9h	14,458	13,483	16,672	10,134	11,135	10,962	0,572
20/03/2018	15h	14,878	13,795	16,217	10,307	11,191	11,048	0,490
Formulação do OAU: 45% TK 2 + 16% TK 3 + 10% TK 5								
20/03/2018	20h	15,008	13,336	16,230	10,511	11,499	11,265	0,607
21/03/2018	9h	14,575	12,469	15,944	10,147	11,148	10,962	0,581
21/03/2018	15h	14,887	13,015	16,113	10,394	11,564	11,265	0,550
21/03/2018	21h	14,839	12,300	16,074	10,251	11,178	11,048	0,507
Formulação do OAU: 35% TK 2 + 26% TK 3 + 10% TK 5								
22/03/2018	9h	14,545	12,361	15,806	9,887	10,832	10,745	0,594
22/03/2018	15h	14,922	12,110	15,862	9,956	10,463	10,572	0,589
22/03/2018	17h30	14,510	12,040	16,174	9,826	11,027	10,788	0,542
22/03/2018	21h	13,778	12,114	15,416	9,515	9,974	10,182	0,529
23/03/2018	1h	13,765	11,807	15,151	9,402	9,961	10,138	0,485
23/03/2018	5h	13,960	12,153	16,624	9,710	10,346	10,442	0,633
23/03/2018	9h	14,441	12,305	15,823	9,835	10,836	10,702	0,654
23/03/2018	15h	13,778	12,396	16,126	9,658	10,602	10,528	0,520

Analogamente ao que sucedeu com o modelo 2, também a aplicação do modelo 3 não se mostrou vantajosa nem apropriada, pois em muitos casos não se ajustava da forma pretendida.

A título de curiosidade, experimentou-se obter um modelo onde a variável dependente era o valor de X do óleo bruto com o objetivo de comparar o valor estimado com o valor real medido em laboratório. Todos os pressupostos deste modelo foram satisfeitos (consultar Apêndice B), pelo que se pode escrever o modelo na seguinte forma:

Modelo 4:

$$XF1 = 0,396 + 0,861 (\%C * XC) + 0,619 (\%A * XA)$$

Na Tabela 4.6 são apresentados os resultados obtidos com o modelo 4. Atente-se que a coluna XF1 (real) corresponde ao valor medido em laboratório e a coluna XF1 (teórico) corresponde ao valor calculado através do modelo 4.

Tabela 4.6: Resultados Obtidos com o Modelo 4

Data	Hora	X- TK 1	X-TK 2	X-TK 3	X-TK 5	XC	XF1 (real)	XF1 (teórico)	XD3
Formulação do OAU: 25% TK 2 + 26% TK 3 + 10% TK 5									
20/03/2018	9h	2,964	14,458	13,483	16,672	10,134	11,135	10,975	0,572
20/03/2018	15h	3,085	14,878	13,795	16,217	10,307	11,191	11,144	0,490
Formulação do OAU: 45% TK 2 + 16% TK 3 + 10% TK 5									
20/03/2018	20h	3,141	15,008	13,336	16,230	10,511	11,499	11,330	0,607
21/03/2018	9h	2,968	14,575	12,469	15,944	10,147	11,148	10,988	0,581
Formulação do OAU: 35% TK 2 + 26% TK 3 + 10% TK 5									
22/03/2018	9h	3,055	14,545	12,361	15,806	9,887	10,832	10,775	0,594
22/03/2018	17h30	2,864	14,510	12,040	16,174	9,826	11,027	10,689	0,542
22/03/2018	21h	2,964	13,778	12,114	15,416	9,515	9,974	10,437	0,529
23/03/2018	1h	2,678	13,765	11,807	15,151	9,402	9,961	10,290	0,485
23/03/2018	5h	2,647	13,960	12,153	16,624	9,710	10,346	10,550	0,633
23/03/2018	9h	3,219	14,441	12,305	15,823	9,835	10,836	10,762	0,654
23/03/2018	15h	2,747	13,778	12,396	16,126	9,658	10,602	10,524	0,520

Com o modelo 4 conseguiu-se perceber que o valor esperado de X do óleo bruto difere muito do valor real (medido em laboratório). Isto deve-se, principalmente, à variabilidade de X do óleo alimentar usado uma vez que não há homogeneização dos tanques, o que impossibilita a previsão do mesmo.

Face à dificuldade sentida na implementação de um modelo de regressão para controlo estatístico do processo, e com o intuito de perceber de que forma a remoção da variável E e a introdução de um valor de referência foram ou não benéficos para a melhoria do processo, aplicou-se uma ferramenta da metodologia 6-sigma, o estudo de capacidade do processo. A metodologia 6-sigma foi criada para reduzir as perdas e aumentar a eficiência dos processos [17]. O estudo de capacidade foi realizado para o valor de X do óleo neutro, recorrendo para isso ao *software* ACCEPT utilizado pela Prio Bio, S.A.

4.3.2 Estudo de Capacidade do Processo

O estudo da capacidade de um processo permite avaliar a potencialidade do processo, isto é, se ele é capaz de produzir dentro das especificações estabelecidas. O controlo estatístico de um processo permite monitorizar o comportamento do processo através de cartas de controlo e reduzir a sua variabilidade [18]. É, por isso, uma ferramenta importante na metodologia 6-Sigma, dada à sua capacidade de identificar problemas.

As cartas de controlo permitem monitorizar o processo de forma a indicar pontos de melhoria. Estas são definidas por um gráfico que mostra a evolução ao longo do tempo de uma estatística referente a uma determinada característica de qualidade. Como este estudo incidiu sobre um produto que é resultado de um processo químico, no qual não é possível a recolha de amostras com mais de uma observação, as cartas de controlo utilizadas foram: a carta X e a carta MR. A carta X é a carta de observações individuais e a carta MR é a carta de amplitudes móveis (valor absoluto da diferença entre dois valores consecutivos à medida que eles vêm do processo [19]). Através destas cartas pode-se identificar se o processo está ou não sob controlo estatístico.

Para que se possa afirmar que um processo está sob controlo estatístico, os pontos devem encontrar-se dentro dos limites da carta, dispostos de forma aleatória e não deve ser possível identificar comportamentos não aleatórios. Um ponto fora dos limites de controlo ou um comportamento não aleatório, evidenciam a existência de causas especiais [18].

Além das cartas de controlo deve-se também analisar o histograma e o *QQ plot*. O histograma permite visualizar o número de ocorrências dos valores obtidos. O *QQ plot* permite avaliar se a distribuição de uma amostra se aproxima de uma dada distribuição. Se os pontos se encontrarem dispostos muito perto da reta, pode-se dizer que as duas distribuições que estão a ser comparadas, são semelhantes.

De forma sumária, a análise do histograma, do *QQ plot* e das cartas de controlo traduzem o desempenho do processo. Contudo, deve-se também calcular os valores dos índices de capacidade C_p e C_{pk} , para averiguar se o processo tem capacidade de produzir de acordo com as especificações.

O índice de capacidade C_p é um indicador matemático que compara a diferença entre o limite superior de especificação (LSE) e o limite inferior de especificação (LIE) com a variabilidade do processo [20], e que se pode traduzir na Equação (4.2).

$$C_p = \frac{LSE - LIE}{6\sigma} \quad (4.2)$$

Em geral, considera-se que $C_p=1,33$ é um valor aceitável para processos ou máquinas [18]. No entanto, a metodologia 6-Sigma exige que a capacidade seja pelo menos igual a dois. Por sua vez, o índice de capacidade C_{pk} considera a localização média do processo, ao contrário do que acontecia com o C_p . Há um índice de capacidade do processo superior e outro inferior, que se podem escrever de acordo com as Equações (4.3) e (4.4), respetivamente.

$$(C_{pk})_S = \frac{LSE - \mu}{3\sigma} \quad (4.3)$$

$$(C_{pk})_I = \frac{\mu - LIE}{3\sigma} \quad (4.4)$$

Assim, o C_{pk} índice é dado por:

$$C_{pk} = \min((C_{pk})_I, (C_{pk})_S) \quad (4.5)$$

É desejável que os índices C_p e C_{pk} tomem valores superior a um, pois desta forma temos a indicação de que o processo é capaz de produzir perto de 100% de produtos aceitáveis, isto é, dentro das especificações.

Neste trabalho, o estudo da capacidade do processo foi aplicado ao X do óleo neutro (D3), com o intuito de comparar a variabilidade própria do produto com as especificações para o mesmo. Neste estudo de capacidade é possível ver a carta X, a carta MR, o histograma, o QQ *plot* da distribuição normal e os dados de processo (Cartas de Controle, Descrição, Estatística e Capacidade) desta variável. É de realçar que o Sigma apresentado, neste estudo de capacidade, diz respeito ao desvio-padrão *within*, enquanto que o St. Dev(ind) corresponde ao desvio-padrão *overall*. Para o cálculo dos índices C_p e C_{pk} é utilizado o Sigma, pois estes índices correspondem a medidas de curto prazo.

Os resultados a seguir apresentados foram obtidos com recurso ao *software* ACCEPT, disponível nas instalações da unidade fabril onde decorreu o estágio.

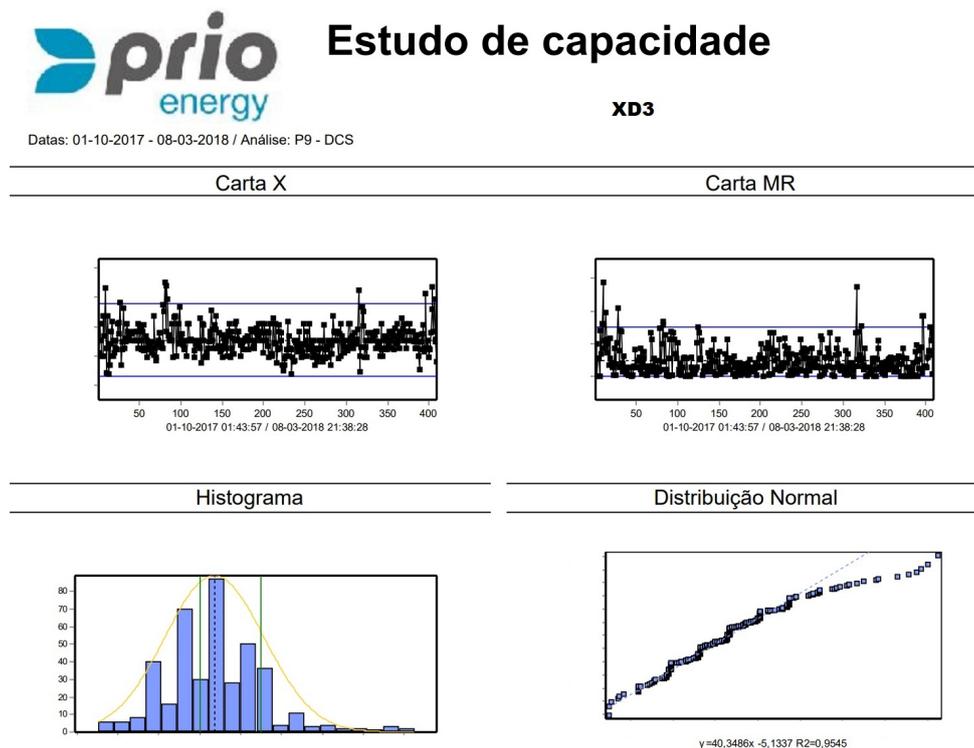


Figura 4.13: Estudo de capacidade para XD3 (01-10-2017 a 08-03-2018)

Na Figura 4.13, encontra-se ilustrado o estudo de capacidade para o período de 1 de outubro de 2017 a 8 de março de 2018. Este período corresponde ao período em que iniciaram a produção com o óleo A até à data em que foram implementadas as sugestões de melhoria.

Da análise das cartas de controlo X e MR conclui-se que o processo não está sob controlo estatístico, tendo em conta que existem, na carta X, vários pontos fora dos limites de controlo. Pelo histograma da Figura 4.13, observa-se que a classe com maior frequência inclui o valor 0,56 (aproximadamente). O *QQ plot*, evidencia a aparente assimetria da distribuição, pois a cauda direita é mais pesada que o esperado para uma distribuição normal e este “peso” traduz diretamente os resultados excessivamente elevados desta variável. No entanto, em 50% dos casos o processo conseguiu produzir dentro dos limites especificados. Também o índice de capacidade sugere que o processo se encontra fora de controlo.

Analisado este período, a análise que se segue é respeitante ao período subsequente à implementação das sugestões de melhoria, nomeadamente, a remoção da variável E e introdução do valor alvo R.

O período agora analisado está compreendido entre 8 de março de 2018 e 21 de maio de 2018.

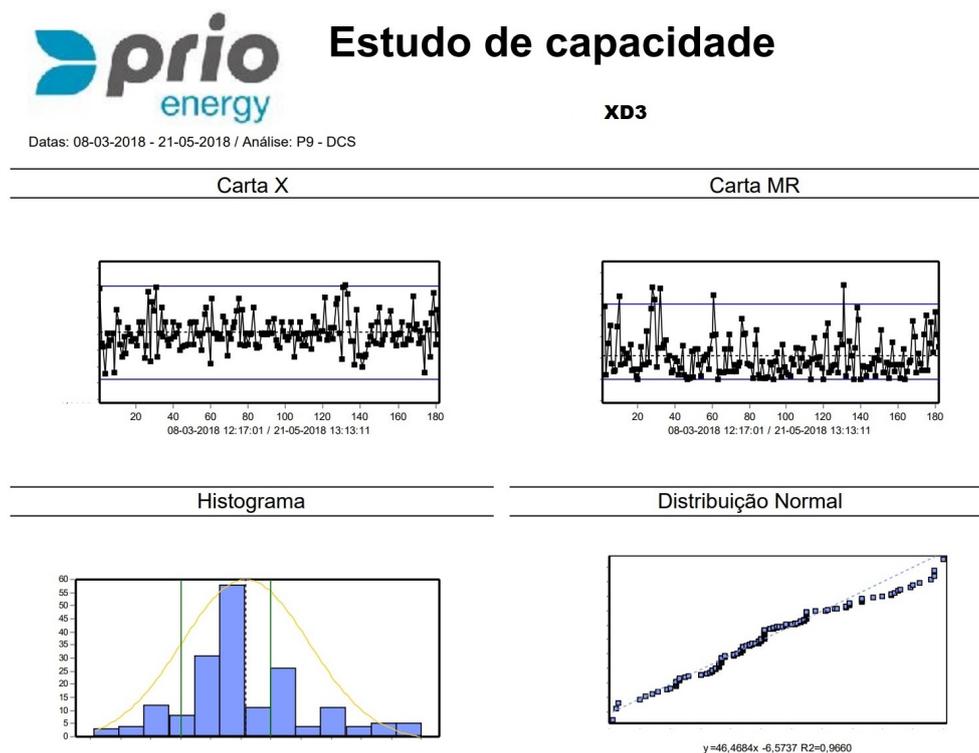


Figura 4.14: Estudo de capacidade de XD3 (08-03-2018 a 21-05-2018)

Embora a sugestão para o segundo problema, mencionado em 4.2, indicasse a remoção do intervalo [P-Q], este intervalo continua a existir tendo sido apenas alterada a metodologia da ação, isto é, a forma como os profissionais deveriam agir face aos resultados do XD3 apontando sempre para um valor alvo. É assim de notar que, o estudo de capacidade utiliza este mesmo intervalo para os seus cálculos no período posterior às alterações.

Na Figura 4.14 encontram-se ilustrados os resultados das cartas de controlo X e MR, das quais se conclui que o processo não está sob controlo estatístico, dada a existência de um ponto fora dos limites de controlo (observação nº 132). No histograma, verifica-se que a classe com maior frequência inclui o valor R que era o valor alvo. Mais uma vez, o *QQ plot*, evidencia a aparente assimetria da distribuição (cauda direita é mais pesada) que não era esperado para uma distribuição normal e sendo este “peso” traduzido diretamente pelos resultados elevados da variável XD3. Todavia, em 54% dos casos (aproximadamente) o processo foi capaz de produzir de acordo com as especificações.

O índice de capacidade obtido sugere, da mesma forma que na Figura 4.13 que o processo se encontre fora de controlo. Ao nível do desvio padrão, este é agora mais baixo do que no período anterior às alterações, o que significa que a variação do processo reduziu. Note-se que para este estudo em específico a análise do $(C_{pk})_I$ não é muito relevante, uma vez que apenas interessa analisar quando é que o valor de X do óleo neutro esteve acima do limite superior.

Com este estudo e tendo em conta a média dos resultados para o período de 8 de março de 2018 a 21 de maio de 2018, pode-se concluir que a implementação do valor objetivo R e a remoção da variável E se mostraram positivas no que diz respeito à melhoria do processo, dado que as oscilações do mesmo são agora menores (valores de $(C_{pk})_S$ e do desvio padrão (Sigma) mais baixos).

Findado o estudo de capacidade do processo para a variável XD3, com o qual se verificou que o processo não estava sob controlo, e relembrando que os resultados obtidos com a Regressão Linear Múltipla foram muito insatisfatórios, e não permitiram dar resposta aos problemas colocados, decidiu-se proceder à análise do sistema de medição (MSA do inglês measurement system analysis) utilizado em laboratório, para tentar perceber o porquê da variabilidade de X do óleo bruto e neutro.

4.4 4ª Fase do ciclo PDCA – *Act* (Atuar)

4.4.1 Análise do Sistema de Medição

Após a análise dos valores de X dos óleos bruto e neutro obtidos durante as experiências com os modelos de regressão, e tendo também por base as conclusões conseguidas com o estudo de capacidade do processo feito para o X do óleo neutro, decidiu-se fazer um MSA para estas duas variáveis, para assim avaliar se os valores obtidos nas análises laboratoriais poderiam ser considerados aceitáveis.

O MSA é um método utilizado para avaliar o grau de exatidão, variabilidade e estabilidade de um sistema de medição. Este é um método essencial para um processo de melhoria da qualidade, pois permite avaliar se os valores obtidos são fiáveis.

Um sistema de medição deve ser caracterizado quando à sua estabilidade em função do valor alvo (exatidão) e quanto à sua variabilidade. Relativamente aos problemas que afetam a exatidão dos resultados devem ser avaliados 3 aspetos: a estabilidade, o *viés* (*Bias* em inglês) e a linearidade. A estabilidade é a capacidade de um sistema de medição reproduzir os mesmos valores ao longo do tempo em que é medida a amostra. O *viés* é a diferença entre a média observada das medições e o valor convencionalmente verdadeiro [21]. Por último, a linearidade é uma medida de consistência da exatidão ao longo da gama de medição.

Por sua vez, no que diz respeito aos problemas que afetam a variabilidade dos resultados, deve ser analisada a repetibilidade e a reprodutibilidade (R&R). A repetibilidade está relacionada com a variância dos resultados obtidos pelo mesmo operador nas mesmas condições de medição, ou seja, a capacidade de um mesmo operador ler o resultado da mesma forma várias vezes. A reprodutibilidade está relacionada com a variância dos resultados obtidos quando a análise é feita por diferentes operadores nas mesmas condições de medição.

A análise da reprodutibilidade e repetibilidade têm extrema importância no contexto de melhoria da qualidade dos processos produtivos, pois se a variação do sistema de medição for elevada e os intervenientes no processo não tiverem esse conhecimento, pode-se incorrer numa conclusão errada do estudo de capacidade.

A análise de um sistema de medição pressupõe vários procedimentos, como se encontra ilustrado na Figura 4.15.



Figura 4.15: Procedimentos MSA

De acordo com os procedimentos mencionados na Figura 4.15, o MSA realizado no laboratório da Prio Bio, S.A., iniciou-se através da preparação das amostras e da medição, ou seja, identificaram-se de forma correta as amostras colocando a parte do processo a que pertenciam, a hora e dia em que foram recolhidas. Definiu-se que esta análise seria feita por duas analistas que conheçam bem os equipamentos e procedimentos para a análise e que estariam responsáveis por registar devidamente os resultados. As duas analistas realizariam três ensaios para cada uma das amostras recolhidas, cujo total seriam dez amostras.

O segundo passo corresponde à estabilidade. Neste passo, devem-se traçar os pontos (valores das amostras) numa carta de controlo de valores individuais e amplitudes móveis e avaliar se os dados estão “sob controlo estatístico”, isto é, se não existem tendências não aleatórias e se os parâmetros de distribuição normal se mantêm constantes.

No terceiro passo, avalia-se o *viés* e a linearidade. Inicialmente, deve-se calcular o *viés* para os valores em estudo, traçando posteriormente um gráfico “*Bias*” versus “valores de referência” e efetuando uma regressão linear aos dados. Os valores de referência para esta análise foram os valores registados na base de dados da fábrica, sendo que resultaram de amostras diferentes das utilizadas para o estudo, contudo recolhidas nos mesmos momentos (dia e horas). No final faz-se uma regressão linear aos dados, seguida de um estudo do valor do *viés* e um teste de hipótese ao declive da reta de regressão, do qual retiraremos informações sobre a linearidade. No estudo do valor do *viés* tem-se que:

$$\begin{cases} H_0 : Bias_i = 0 & \text{(em cada ponto } i \text{ da gama de medição)} \\ H_1 : Bias_i \neq 0 & \text{(em cada ponto } i \text{ da gama de medição)} \end{cases}$$

Se o *p-value* for menor que α , então rejeita-se a hipótese nula e conclui-se que existe um problema de *viés* naquele ponto da gama [21].

Por sua vez, para o teste de hipótese ao declive da reta de regressão tem-se:

$$\begin{cases} H_0 : \text{O declive é zero} \\ H_1 : \text{O declive é diferente de zero} \end{cases}$$

Se o *p-value* for menor que α , então o declive é diferente de zero e a linearidade é elevada pelo que é necessário rever o sistema de medição [21].

O último passo é o estudo Gage R&R, ou seja, o estudo de repetibilidade e reprodutibilidade que é utilizado para estudar a variabilidade do sistema de medição. Este tem como objetivo comparar os instrumentos de medição e/ou analistas e assegurar a confiança nas medições. Para avaliar a variação total de um processo é necessário ter em conta as fontes de variação quer do processo de fabrico quer do sistema de medição.



Figura 4.16: Variação total de um processo [17]

É na análise do sistema de medição que se avalia a repetibilidade e a reprodutibilidade, sendo esta análise feita através dos seguintes gráficos e cartas: Components of Variation, Measurement by Part, Measurement by Operator, Xbar chart by Operator, R chart by Operator, Gage Run Chart of Measurement by Part, Operator.

Ao contrário do que normalmente acontece com as cartas de controlo, no MSA, o tempo não é a escala do eixo horizontal. Nestas cartas, o eixo x é segmentado em regiões para as diferentes analistas e para as suas medições. A carta Xbar chart by Operator é a carta das médias que traduz a diferença entre as partes, ou seja, compara a variação entre amostras. Pretende-se neste caso que haja o maior número de pontos fora de controlo, pois isso indica que a variabilidade relativa ao Part é maior comparativamente com a repetibilidade e reprodutibilidade [18]. A carta R chart by Operator é a carta de amplitudes que traduz a repetibilidade das medições. Ao contrário do que acontece com a carta Xbar, aqui interessa que todos os pontos estejam sob controlo estatístico, o que indica consistência nas medições efetuadas nas mesmas condições de medição.

O gráfico Components of Variation permite identificar se as principais fontes de variação se devem à repetibilidade, reprodutibilidade ou ao processo de fabrico, através de um gráfico de barras.

O gráfico Measurement by Part permite avaliar a variabilidade entre as peças e a dispersão destes dados.

O gráfico Measurement by Operator permite avaliar as diferenças entre operadores.

Por último, o gráfico Gage Run Chart of Measurement by Part, Operator permite avaliar a repetibilidade e reprodutibilidade.

Deve-se também analisar o valor obtido para o número de categorias distintas (ndc do inglês number of distinct categories), pois este parâmetro avalia se a resolução do equipamento é adequada. Deve ser preferencialmente superior a cinco para que se possa considerar que o sistema de medição efetua uma análise aceitável do processo [21].

A Tabela 4.7 indica de que forma se deve classificar o sistema de medição de acordo com os resultados obtidos [22].

Tabela 4.7: Critérios de classificação para estudo Gage R&R

Sistema de Medição	ndc	%Study Var
Rejeitado	1	$\leq 10\%$
Pobre/Marginal	2 - 4	10% - 30%
Aceite	5 ou mais	$> 30\%$

Como referido no estudo de capacidade, a recolha das amostras apenas permite uma observação, logo este tipo de amostras são destrutivas, pelo que se utilizou o método Nested-ANOVA para o estudo Gage R&R. Os dados foram tratados no *software* Minitab 18.

Resultados do MSA para XD3

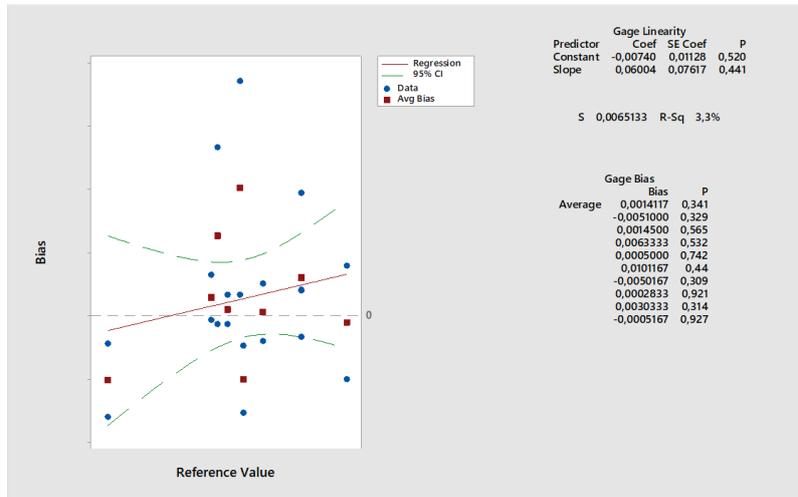


Figura 4.17: Estudo da Linearidade e do *Bias* para XD3

Pela análise da Figura 4.17, pode-se concluir que a linearidade é baixa pois $p\text{-value} > \alpha$ e por isso a hipótese nula não é rejeitada. Através da coluna “P” em “Gage Bias” verifica-se que todos os valores são superiores a α pelo que não se rejeita a hipótese nula, ou seja, o *viés* pode ser considerado nulo em todos os pontos da gama de medição, o que significa que não existem diferenças significativas entre a média observada das medições e o valor de referência.

Gage R&R Study - Nested ANOVA

Gage R&R (Nested) for Measurement

Source	DF	SS	MS	F	P
Operator	1	0,0002428	0,0002428	0,1637	0,691
Part (Operator)	18	0,0266946	0,0014830	61,9920	0,000
Repeatability	40	0,0009569	0,0000239		
Total	59	0,0278943			

Variance Components

Source	VarComp	%Contribution (of VarComp)
Total Gage R&R	0,0000239	4,69
Repeatability	0,0000239	4,69
Reproducibility	0,0000000	0,00
Part-To-Part	0,0004864	95,31
Total Variation	0,0005103	100,00

Gage Evaluation

Source	StdDev (SD)	Study Var (6 × SD)	%Study Var (%SV)
Total Gage R&R	0,0048911	0,029347	21,65
Repeatability	0,0048911	0,029347	21,65
Reproducibility	0,0000000	0,000000	0,00
Part-To-Part	0,0220538	0,132323	97,63
Total Variation	0,0225897	0,135538	100,00

Number of Distinct Categories = 6

Figura 4.18: Resultados do estudo Gage R&R para XD3

Na Figura 4.18 apresentam-se os resultados do estudo Gage R&R para a variável XD3 (óleo neutro).

O p -value correspondente ao operador em “Gage R&R (Nested) for Measurement” é igual a 0,691 o que indica que não há diferenças significativas entre os operadores [23]. Por sua vez, para a variável Part(Operator) pode-se concluir que pelo menos uma média é estatisticamente diferente.

De acordo com a estatística F, tem-se que o fator Part(Operator) tem um contributo maior para a variabilidade da medição.

Apenas 4,69% da variação dos dados se deve ao sistema de medição, sendo que os restantes 95,31% se devem às características do processo de fabrico. O valor da “%Study Var” é de 21,65% o que segundo a Tabela 4.7 indica que o sistema de medição deve ser considerado marginal e apenas deve ser aceite sob condição, contudo, esta é uma variável crítica para a qualidade do produto final, pelo que neste caso, não deverá ser aceite sob nenhuma condição. O valor de ndc obtido é satisfatório, pelo que o método utilizado para fazer esta análise pode ser considerado adequado.

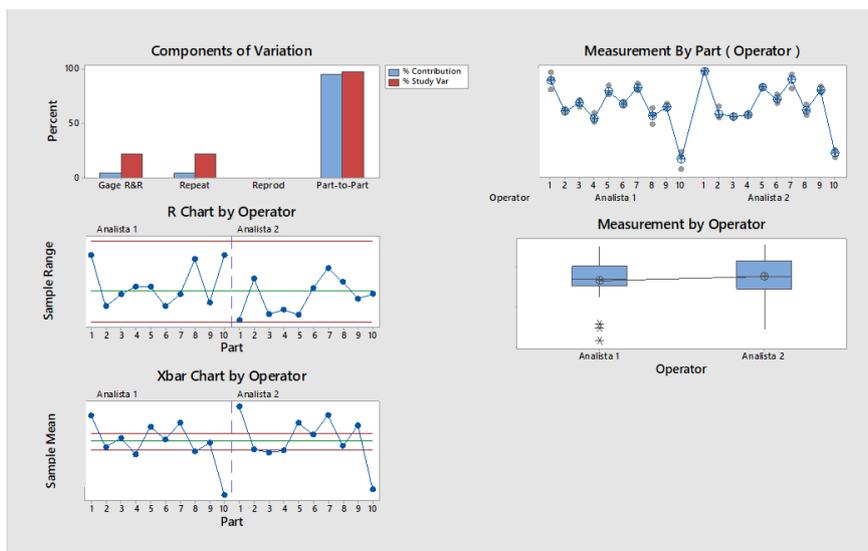


Figura 4.19: Resultados do estudo Gage R&R (Nested) para XD3

Do gráfico “Components of variation” verifica-se que a variação se deve essencialmente à diferença entre as peças (Part-to-Part).

Através da carta “R Chart by Operator” pode-se verificar que todos os pontos se encontram dentro dos limites de controlo da carta, logo pode-se dizer que a repetibilidade de cada analista está sob controlo. Através das conclusões obtidas com esta carta pode-se validar os limites da carta das médias “Xbar Chart by Operator”. Na carta das médias, pode observar-se que para ambas as analistas, pelo menos 50% dos dados estão fora dos limites de controlo, o que indica que o sistema de medição tem baixa variação [23].

No gráfico “Measurement by Part” encontra-se ilustrados os ensaios feitos para cada medição por cada analista. Deste gráfico verifica-se que os resultados obtidos para os diferentes ensaios foram semelhantes, contudo, a análise da amostra 3 apresenta diferenças entre as analistas. Verifica-se também que as amostras variam muito entre si.

O gráfico “Measurement by Operator” mostra que o valor médio para cada analista é ligeiramente diferente, sendo mais baixo para a analista 1. Pode-se ainda dizer que no caso da analista 1, poderão ter existido alguns erros de registo.

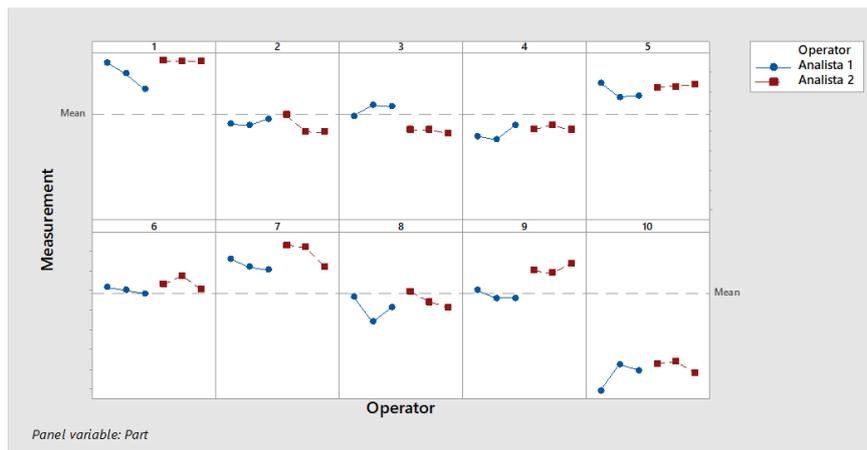


Figura 4.20: Gage Run Chart of Measurement by Part, Operator para XD3

O gráfico ilustrado na Figura 4.20 não evidencia problemas significativos nem de reprodutibilidade nem de repetibilidade para a variável XD3. Contudo, nas amostras 1,8 e 10 é possível verificar que a analista 1 teve dificuldades em obter valores semelhantes para os diferentes ensaios, ou seja, ocorreram aqui problemas de repetibilidade. Para estas mesmas amostras, verificam-se também problemas de reprodutibilidade, isto é, diferenças entre as analistas.

Resultados do MSA para XF1

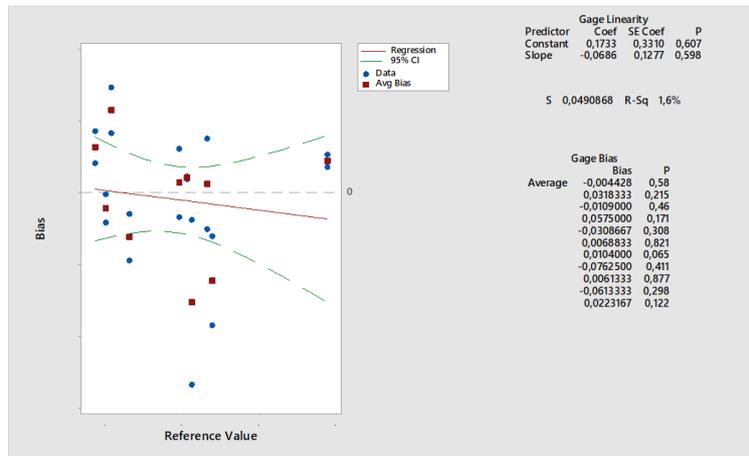


Figura 4.21: Estudo da Linearidade e do *Bias* para XF1

Analisando a Figura 4.21 e de forma análoga ao que sucedeu no estudo da linearidade e do *Bias* para o XD3, tem-se que o $p\text{-value} > \alpha$ pelo que não se rejeita a hipótese nula, logo o declive da reta poderá não ser nula, não existindo assim um problema de linearidade. Na coluna “P” em “Gage *Bias*” verifica-se que todos os valores são superiores a α pelo que o *viés* pode ser nulo em todos os pontos da gama de medição.

Gage R&R Study - Nested ANOVA

Gage R&R (Nested) for Measurement

Source	DF	SS	MS	F	P
Operator	1	0,017282	0,0172822	0,6258	0,439
Part (Operator)	18	0,497086	0,0276159	10,1432	0,000
Repeatability	40	0,108904	0,0027226		
Total	59	0,623272			

Variance Components

Source	VarComp	%Contribution (of VarComp)
Total Gage R&R	0,0027226	24,71
Repeatability	0,0027226	24,71
Reproducibility	0,0000000	0,00
Part-To-Part	0,0082978	75,29
Total Variation	0,0110204	100,00

Gage Evaluation

Source	StdDev (SD)	Study Var (6 × SD)	%Study Var (%SV)
Total Gage R&R	0,052178	0,313071	49,70
Repeatability	0,052178	0,313071	49,70
Reproducibility	0,000000	0,000000	0,00
Part-To-Part	0,091092	0,546553	86,77
Total Variation	0,104978	0,629868	100,00

Number of Distinct Categories = 2

Figura 4.22: Resultados do estudo Gage R&R para XF1

Na Figura 4.22 apresentam-se os resultados do estudo Gage R&R para a variável XF1. O *p-value* respeitante ao operador em “Gage R&R (Nested) for Measurement” é igual a 0,439 o que indica que não há diferenças entre os operadores. Para a variável Part(Operator) pode-se concluir que, pelo menos, uma média é estatisticamente diferente.

Olhando para a estatística F, novamente se verifica que é mais provável que o fator Part(Operator) tenha maior contributo para a variabilidade da medição.

Um total de 24,71% da variação nos dados deve-se à repetibilidade e reprodutibilidade, sendo que os restantes 75,29% se devem às características do processo de fabrico. O valor da “%Study Var” é igual a 49,70% o que indica que o sistema de medição deve ser rejeitado. O valor de *ndc* é igual a dois, pelo que os dados podem ser divididos em grupos altos e baixos, o que é equivalente a dados de atributos [21]. Neste caso, o *ndc* indica que o sistema é pobre e que pode dar origem a cartas de controlo pouco sensíveis.

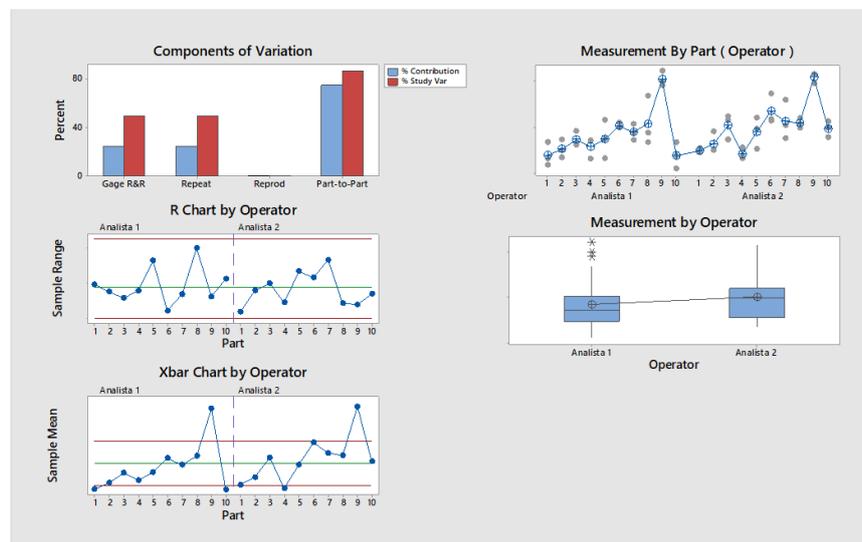


Figura 4.23: Resultados do estudo Gage R&R (Nested) para XF1

Na Figura 4.23 apresentam-se os gráficos: Components of variation, Measurement by Part (Operator), Measurement by Operator e as cartas R Chart by Operator e Xbar Chart by Operator obtidos para a variável XF1.

Analogamente aos resultados obtidos para o estudo Gage R&R para o XD3, também no gráfico “Components of Variation” se pode verificar que a variação se deve principalmente à diferença entre as peças (Part-to-Part).

Na carta “R Chart by Operator” verifica-se que todos os pontos se encontram dentro dos limites de controlo da carta, o que indica que a repetibilidade de cada analista está sob controlo. Com esta conclusão pode-se validar os limites da carta das médias “Xbar Chart by Operator”. Nesta carta observa-se que para ambas as analistas, apenas dois pontos se encontram fora dos limites da carta, o que indica variação no sistema de medição. O gráfico

“Measurement by Part” indica variabilidade em todas as medições e entre as amostras.

O gráfico “Measurement by Operator” mostra que o valor médio obtido pela analista 1 é inferior ao valor médio obtido pela analista 2. Pode-se também para esta variável, dizer que no caso da analista 1, poderão ter existido alguns erros de registo.

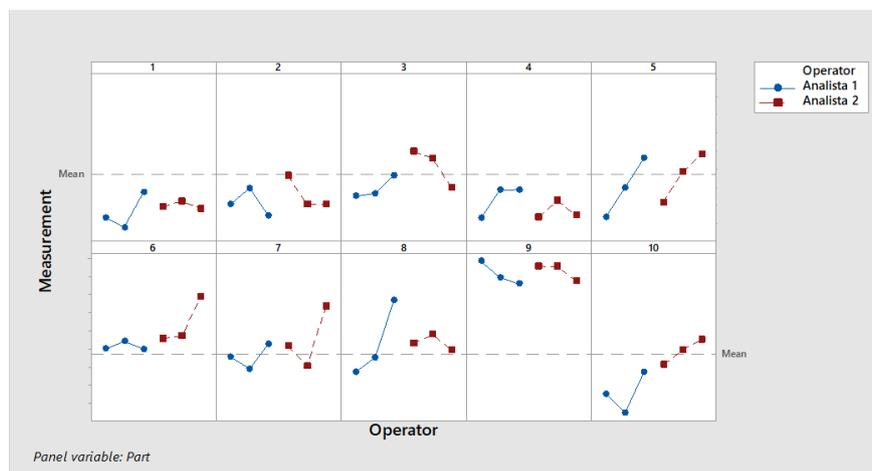


Figura 4.24: Gage Run Chart of Measurement by Part, Operator XF1

No gráfico representado na Figura 4.24 é visível que existem alguns problemas de reprodutibilidade, ou seja, diferenças entre os resultados obtidos pelas diferentes analistas. Esta leitura pode ser verificada principalmente nas amostras 1, 8 e 10, pois a representação gráfica dos resultados não é semelhante para ambas as analistas. Ao nível da repetibilidade, as diferenças mais notórias da analista 1 ocorreram nas amostras 5 e 8, já no caso da analista 2 ocorreram nas amostras 2, 5, 6 e 7.

Desta forma pode-se afirmar que para a variável XF1, os valores medidos também são afetados pela analista que efetua a análise laboratorial dos mesmos.

Capítulo 5

Conclusões e Oportunidades Futuras

Ao longo deste relatório de estágio, estudou-se o processo de neutralização do óleo utilizado na produção de biodiesel na Prio Bio, S.A., que consiste em reduzir o valor de X do óleo bruto por forma a evitar reações secundárias indesejáveis na reação de transesterificação.

Numa fase inicial, fez-se uma análise exploratória de todas as variáveis do processo, sendo que estas se encontravam divididas em duas formulações: A + C e B + C. Desta análise, verificou-se que a propriedade X do óleo B é, em média, igual a 7,02 sendo mais elevada que no óleo A, cuja média é igual a 3,6. Em média, o XC foi também mais elevado no período de laboração com o óleo B, sendo o seu valor igual a 15,03, enquanto que durante o período de utilização do óleo A foi em média, 13,69. O tanque F1 reflete estas misturas distintas em cada época do ano pelo que para a formulação B + C, a sua média é igual a 13,13 e para a formulação A + C a sua média foi de 9,96, Estes valores refletem-se também com o aumento da incorporação de óleo C que no verão foi de 323,48 e no inverno foi igual a 277,68. Em média, a qualidade do produto intermédio e final foi melhor (resultados mais baixos) durante o inverno, ou seja, para a formulação A + C. Desta análise rapidamente se percebeu que a percentagem de óleo C e os valores de XF1 e G estão correlacionados entre si, o que se traduziu no estudo de correlações feito, uma vez que estas correlações têm valores acima de 0,8. Do estudo das correlações, verificou-se ainda que o X do óleo neutro, apresentava correlações muito baixas com todas as variáveis do processo. É exemplo disso, a correlação entre a variável I e XD3, que pela reação química principal do processo de neutralização (Figura 2.4), se verifica que esta variável influencia diretamente no resultado de X do óleo neutro, contudo, a correlação entre ambas foi de 0,42 para a formulação B + C e de 0,83 para a formulação A + C, o que não era espectável que acontecesse. Todavia, é preciso ter sempre em mente que este é um processo industrial com muita variabilidade, pelo que os resultados são complexos e não refletem o esperado, em muitos dos casos.

Dada a importância de X no processo de neutralização, fez-se um estudo estatístico para

a variável X. Este estudo permitiu identificar 3 problemas:

1. existência de duas variáveis a atuar sobre a variável I para neutralizar X: variáveis G e E;
2. existência de uma gama, ao invés de um valor alvo na qual devia estar XD3;
3. variabilidade de X nos tanques de armazenamento e F1, que ao longo de 8h30 tiveram um desvio-padrão considerável: 0,360, 0,082, 0,381, 0,186 e 0,195, para os tanques F1, TK 1, TK 2, TK 3 e TK 4, respetivamente.

Deste modo, a autora sugeriu três possíveis soluções para estes problemas. A variável E foi eliminada, introduziu-se como valor alvo R e por último, a sugestão passou por utilizar a regressão linear múltipla para tentar encontrar um modelo que auxiliasse os chefes de turno a proceder a um ajuste de receita mais correto. Contudo, após serem testados vários modelos de regressão concluiu-se que estes não eram capazes de se adaptar da forma esperada às oscilações do processo. Desta forma, e para tentar perceber o porquê da implementação de um modelo não ter funcionado, recorreu-se a um estudo de capacidade do processo para a variável XD3. Realizou-se um estudo de capacidade para o período anterior à implementação das sugestões de forma a comparar o antes e o depois. Deste estudo, verificou-se que o processo não estava sob controlo estatístico, contudo, as oscilações do processo eram agora menores pois o seu valor de sigma é igual a 0,082 e o desvio-padrão é igual a 0,091, sendo que anteriormente estes tinham valores de 0,091 e 0,104, respetivamente. Aproveitou-se também este estudo de capacidade, para perceber se a remoção da variável e introdução de um valor alvo se tinham mostrado profícuos, o que se verificou ser verdade.

Uma vez que o processo estava fora de controlo estatístico e dada a variabilidade que existe entre os resultados, foi-se validar o sistema de medição usado em laboratório através de um MSA às variáveis XF1 e XD3. Para a variável XD3 percebeu-se que a maior percentagem da variação se devia essencialmente ao próprio processo de fabrico, pois este contava com 95,31%, enquanto que a repetibilidade teve um valor de 4,69% e a reprodutibilidade 0%. O valor de ndc foi igual a 6 o que indica que o sistema de medição consegue distinguir diferentes grupos, contudo para a %Study Var, a repetibilidade obteve um valor de 21,65% que segundo os critérios de classificação indica que o sistema de medição é marginal e apenas deve ser aceite sob condição, pois está entre 10% e 30%. Também para a variável XF1 se concluiu que a maior percentagem de variação se devia ao próprio processo, tendo este um valor de 75,29%, contudo para esta variável, a repetibilidade tem maior destaque comparativamente à anterior, pois contribui em 24,7% para a variação total e a reprodutibilidade não contribui para a variação, visto que o seu valor foi de 0%. Ao nível do ndc, este foi igual a 2, pelo que o sistema de medição deve ser considerado pobre, pois o seu ndc é inferior a cinco. Para a %Study Var, o valor obtido para a repetibilidade foi igual a 49,70%, assim, uma vez que este valor é superior a 30% e pela Tabela 4.7, o sistema de medição deve ser rejeitado.

Através deste estudo percebeu-se a dificuldade na implementação de um modelo de regressão no processo.

Como propostas de melhoria para a empresa, sugere-se a aquisição de agitadores para os tanques de armazenamento e para o tanque de mistura de óleos F1. Sugere-se também a implementação de um sensor online para determinação de X (valores recebidos ao minuto) com ajuste automático da receita. Na impossibilidade de avançar de imediato com estas soluções, dado o investimento económico e disponibilidade da fábrica para serem implementadas, sugere-se um aumento do número de análises a realizar durante o turno, pois esta medida irá auxiliar os chefes de turno a proceder a um ajuste mais correto da receita. Por último, e tendo por base os resultados obtidos no MSA sugere-se a revisão do sistema de análise da propriedade X, para que seja possível obter resultados mais fidedignos que os atuais.

Bibliografia

- [1] Ambiente, A. P. do. (2016). Óleos alimentares usados – Rede de Recolha Seletiva Municipal.
- [2] Prio Energy. Obtido a 18 de Abril de 2018, de <https://www.prioenergy.com>
- [3] ENMC. Incorporação: Metas e Obrigações. Obtido a 18 de Abril de 2018, de <http://www.enmc.pt/pt-PT/atividades/biocombustiveis/explicador/incorporacao-metas-e-obrigacoes-2/>
- [4] Ambiente, A. P. do. (2011). Recolha de Óleos Alimentares Usados. Obtido a 21 de Novembro de 2017, de https://www.apambiente.pt/_zdata/Políticas/Residuos/FluxosEspecificosResiduos/OAU
- [5] dos Vultos, D. A. A. N. (2012). Desenvolvimento Sustentável : A utilização do Biodiesel no Concelho do Seixal. Faculdade de Ciências Sociais e Humanas: Universidade Nova de Lisboa.
- [6] ENMC. Metas e Obrigações de Incorporação de Biocombustível. Obtido a 4 de Junho de 2018, de <http://www.enmc.pt/pt-PT/atividades/biocombustiveis/indicadores/metas-e-obrigacoes-de-incorporacao-de-biocombustivel/>
- [7] ENMC. (2017). A redução da utilização do óleo de palma nos biocombustíveis em Portugal. Obtido a 4 de Junho de 2018, de <http://www.enmc.pt/en-GB/news/news-archive/a-reducao-da-utilizacao-do-oleo-de-palma-nos-biocombustiveis-em-portugal/>
- [8] Mendes, S. M. da C. (2006). Controlo de qualidade de óleos de fritura em restauração colectiva Validação do plano HACCP. Faculdade de Farmácia da Universidade do Porto. Obtido de <https://repositorio-aberto.up.pt/handle/10216/63991>
- [9] Tervene. (2017). A simple guide to Gemba Walk, 2. Obtido de <http://tervene.com/wp-content/uploads/2017/06/A-simple-guide-to-Gemba-Walk-ebook.pdf>
- [10] Freitas, A. (2016). Apontamentos de Análise Multivariada. Universidade de Aveiro.

- [11] Pestana, M. H., & Gageiro, J. N. (2003). *Análise de Dados para Ciências Sociais- A complementaridade do SPSS* (3.a ed.). Lisboa: Edições Sílabo, Lda.
- [12] Field, A. (2011). *Discovering Statistics using SPSS* (3.a ed.). SAGE Publications.
- [13] Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate Data Analysis* (5.a ed.). Upper Saddle River (NJ): Prentice-Hall.
- [14] Montgomery, D. C., Peck, E. A., & Vining, G. G. (2006). *Introduction to Linear Regression Analysis* (4.a ed.). New York (NY): John Wiley & Sons, Inc.
- [15] Marôco, J. (2010). *Análise Estatística com o SPSS Statistics* (6.a ed.). Pêro Pinheiro : Report Number.
- [16] Testes não paramétricos. (2003). Obtido de <http://www.mat.uc.pt/cmtm/ECwww/TestesNP.pdf>
- [17] Henriques, S. (2016). *Estabilização do Valor Ácido do Biodiesel*.
- [18] Pereira, Z. L., & Requeijo, J. G. (2008). *Qualidade: Planeamento e Controlo Estatístico de Processos*. Lisboa: FCT-UNL e Prefácio.
- [19] Sigma, S. I-MR Charts. Obtido a 4 de Junho de 2018, de <http://www.six-sigma-material.com/I-MR.html>
- [20] Pereira, N. J. R. (1999). *Estudo da Capacidade do Processo*. Faculdade de Engenharia da Universidade do Porto. Obtido de <https://repositorio-aberto.up.pt/handle/10216/59364>
- [21] III, F. W. B. (2003). *Implementing Six Sigma: smarter solutions using statistical methods* (2.a ed.). Hoboken (NJ): John Wiley & Sons.
- [22] Dias, P. A. B. (2009). *Utilização da Metodologia 6 Sigma num Processo Industrial*. Universidade de Aveiro. Obtido de <https://ria.ua.pt/handle/10773/1686>
- [23] Minitab. Suport Minitab. Obtido a 21 de Maio de 2018, de <https://support.minitab.com/en-us/minitab/18/help-and-how-to/quality-and-process-improvement/measurement-system-analysis/how-to/gage-study/nested-gage-r-r-study/interpret-the-results/all-statistics-and-graphs/#xbar-chart>
- [24] Martifer. (2008). *Memória descritiva do processo de neutralização do óleo*.
- [25] Antunes, A. (2015). *Optimização do processo de neutralização para processamento de óleos vegetais usados*.
- [26] Vieira, J. C. (2001). *Controlo Estatístico do Processo*. Lisboa: IPQ.

Apêndice A

Análise do processo de neutralização

Teste de Classificações Assinadas por Wilcoxon

		Postos		
		N	Posto Médio	Soma de Classificações
F01 - FFA	Classificações Negativas	2 ^a	3,50	7,00
	Classificações Positivas	5 ^b	4,20	21,00
	Empates	0 ^c		
	Total	7		

a. F01 < FFA

b. F01 > FFA

c. F01 = FFA

Estadísticas de teste^a

	F01 - FFA
Z	-1,183 ^b
Significância Sig. (bilateral)	,237

a. Teste de Classificações Assinadas por Wilcoxon

b. Com base em postos negativos.

Figura A.1: Resultado do teste de Wilcoxon

Teste-T

Estatísticas de grupo

	FFADCS	N	Média	Erro Desvio	Erro padrão da média
@09F01FFA	2,45	4	2,52650	,109595	,054797
	2,48	3	2,49967	,047721	,027552

Teste de amostras independentes

		Teste de Levene para igualdade de variâncias		teste-t para igualdade de Médias					95% Intervalo de Confiança da Diferença	
		Z	Sig.	t	df	Sig. (2 extremidades)	Diferença média	Erro padrão de diferença	Inferior	Superior
@09F01FFA	Variâncias iguais assumidas	1,126	,337	,390	5	,713	,026833	,068813	-,150056	,203723
	Variâncias iguais não assumidas			,437	4,297	,693	,026833	,061334	-,138919	,192586

Figura A.2: Resultado do teste de Levene

Apêndice B

Regressão Linear Múltipla

Tabela B.1: Resultados obtidos com o modelo 2

Data	Hora	XF1	I	H	G	XD3
09/03/2018	15h	11,308	356,925	47,659	11,698	0,581
09/03/2018	17h	11,187	357,619	47,659	11,612	0,550
09/03/2018	21h	10,702	355,149	46,360	11,222	0,611
10/03/2018	1h	11,226	354,022	46,360	11,568	0,611
10/03/2018	5h	10,944	354,022	46,360	11,352	0,611
10/03/2018	7h	11,525	348,477	46,360	11,742	0,611
10/03/2018	11h45	10,832	358,442	46,360	11,308	0,490
10/03/2018	15h	11,196	346,830	46,360	11,525	0,702
10/03/2018	17h30	13,345	350,600	45,060	13,085	0,611
10/03/2018	21h	10,940	354,586	43,760	11,308	0,550
11/03/2018	1h	10,749	347,523	43,760	11,092	0,732
11/03/2018	5h	10,905	347,523	43,760	11,222	0,672
11/03/2018	7h	11,373	347,610	43,760	11,568	0,550
11/03/2018	12h	11,473	354,022	43,760	11,698	1,209
11/03/2018	15h	11,915	355,409	43,760	12,001	1,126
11/03/2018	17h30	11,915	354,802	42,460	11,958	1,252
11/03/2018	21h	12,157	360,911	42,460	12,218	1,252
12/03/2018	1h	11,798	388,727	42,460	12,131	0,485
12/03/2018	5h	11,456	368,147	43,327	11,742	0,667
12/03/2018	9h	11,729	368,060	43,327	11,958	0,589
12/03/2018	14h40h	12,157	360,911	43,327	12,218	1,005
12/03/2018	17h30	11,763	391,673	44,627	12,175	0,581
12/03/2018	21h	11,560	388,337	44,627	12,001	0,451
13/03/2018	1h	11,560	373,563	44,627	11,915	0,490
13/03/2018	5h	10,996	362,774	43,327	11,395	0,546

Continued on next page

Tabela B.1 – continued from previous page

Data	Hora	XF1	I	H	G	XD3
13/03/2018	9h	12,222	347,827	43,327	12,175	0,728
14/03/2018	9h	11,425	378,762	45,493	11,872	0,360
14/03/2018	15h	11,443	376,422	45,493	11,872	0,550
14/03/2018	17h30	11,716	359,482	44,193	11,915	0,572
14/03/2018	21h	11,473	359,698	44,193	11,742	0,641
15/03/2018	1h	11,178	358,398	43,327	11,482	0,607
15/03/2018	9h	11,508	355,712	42,027	11,655	0,724
15/03/2018	12h	11,655	358,528	38,127	11,698	0,810
15/03/2018	15h	11,551	364,334	40,294	11,698	0,611
15/03/2018	16h30	11,486	364,464	41,160	11,698	0,524
15/03/2018	21h	10,845	361,865	39,861	11,178	0,568
16/03/2018	1h	11,373	358,485	47,659	11,742	0,581
16/03/2018	3h	10,840	358,442	45,493	11,308	0,581
16/03/2018	9h	10,827	359,872	38,994	11,092	0,693
16/03/2018	18h	11,365	360,998	38,994	11,525	0,706
16/03/2018	21h	11,131	369,837	38,994	11,395	0,542
17/03/2018	1h	10,754	352,289	38,994	11,005	0,672
17/03/2018	5h	10,940	353,416	38,994	11,135	0,611
17/03/2018	7h	11,525	353,373	38,994	11,568	0,672
17/03/2018	15h	9,939	342,541	38,994	10,355	0,477
17/03/2018	17h30	9,939	337,038	38,994	10,312	0,546
17/03/2018	20h	10,351	326,120	38,994	10,528	0,832
18/03/2018	1h	10,741	347,957	38,994	10,962	0,459
18/03/2018	5h	10,242	341,414	38,994	10,572	0,672
18/03/2018	7h	10,758	342,498	38,994	10,962	0,780
18/03/2018	15h	10,862	354,846	34,661	10,962	0,728
18/03/2018	21h	10,953	366,241	33,795	11,092	0,854
19/03/2018	2h	11,464	375,296	33,795	11,525	0,485
19/03/2018	5h	13,427	369,880	33,795	12,955	0,611
19/03/2018	9h	12,448	369,577	33,795	12,218	0,750

Coefficientes^a

Modelo		Coeficientes não padronizados		Coeficientes padronizados		t	Sig.	95,0% Intervalo de Confiança para B		Estatísticas de colinearidade	
		B	Erro Erro	Beta				Limite inferior	Limite superior	Tolerância	VIF
1	(Constante)	,222	,032			6,844	,000	,158	,285		
	XF1	,892	,014	,936		64,522	,000	,865	,919	1,000	1,000
2	(Constante)	,412	,029			14,080	,000	,354	,469		
	XF1	,494	,027	,518		18,433	,000	,441	,547	,183	5,462
	C_Cal	,363	,022	,462		16,434	,000	,320	,407	,183	5,462
3	(Constante)	,212	,044			4,863	,000	,126	,297		
	XF1	,447	,027	,469		16,452	,000	,394	,500	,168	5,950
	C_Cal	,464	,027	,590		17,066	,000	,411	,518	,114	8,775
	A_Cal	,355	,059	,112		6,048	,000	,240	,470	,400	2,502

a. Variável Dependente: G

Figura B.1: Coeficientes do modelo 3

Na Figura B.1 é apresentada a tabela dos coeficientes obtidos para o modelo 3. Como se pode verificar, os coeficientes são significativos, incluindo a constante.

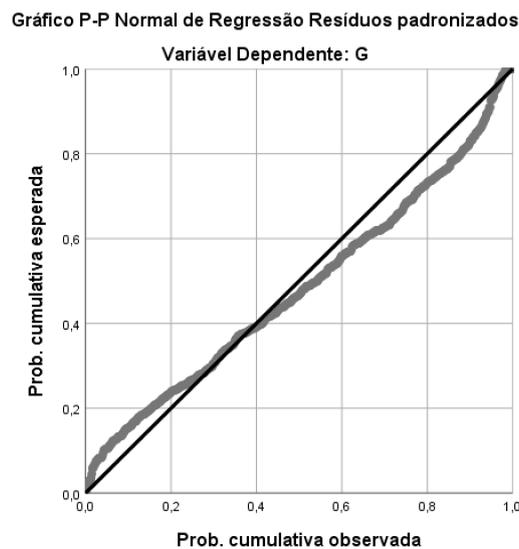


Figura B.2: QQ plot dos resíduos do modelo 3

O gráfico de probabilidade normal do modelo 3, indica que os resíduos podem não ser normais, uma vez que não se encontram muitos próximos da diagonal principal. O teste de Kolmogorov-Smirnov indicou que os resíduos não eram normais.

Testes de Normalidade

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Estatística	df	Sig.	Estatística	df	Sig.
Unstandardized Residual	,084	595	,000	,865	595	,000

a. Correlação de Significância de Lilliefors

Figura B.3: Teste de normalidade para o modelo 3

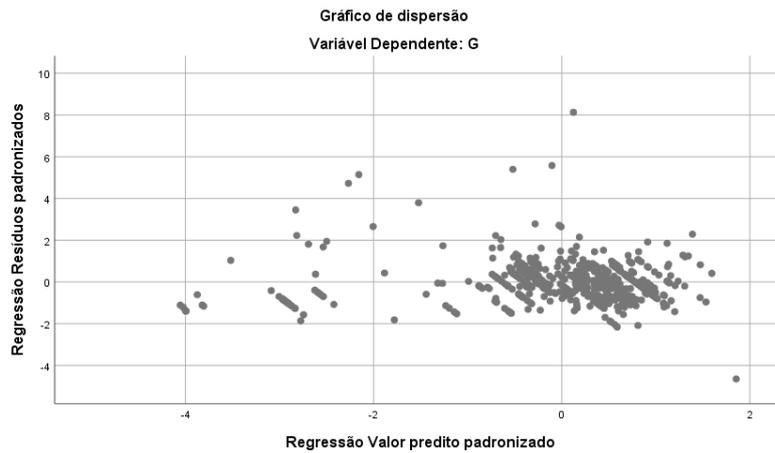


Figura B.4: Gráfico de dispersão do modelo 3

O gráfico de dispersão do modelo 3 (Figura B.4) indica que os resíduos são não correlacionados.

Resumo do modelo^c

Modelo	R	R quadrado	R quadrado ajustado	Erro padrão da estimativa	Mudança de R quadrado	Estatísticas de mudança			Sig. Mudança F	Durbin-Watson
						Mudança F	df1	df2		
1	,904 ^a	,817	,817	,1994549492	,817	2655,925	1	594	,000	
2	,912 ^b	,832	,832	,1912536197	,015	53,036	1	593	,000	1,056

a. Preditores: (Constante), C_Cal

b. Preditores: (Constante), C_Cal, A_Cal

c. Variável Dependente: XF1

Figura B.5: Resumo do modelo 4

Através do quadro resumo obtido para o modelo 4, verifica-se que 83,2% da variabilidade total é explicada pelo modelo de regressão. O valor obtido para o teste de Durbin-Watson indica a existência de autocorrelação positiva nos erros do modelo.

Coeficientes ^a										
Modelo		Coeficientes não padronizados		Coeficientes padronizados			95,0% Intervalo de Confiança para B		Estatísticas de colinearidade	
		B	Erro Erro	Beta	t	Sig.	Limite inferior	Limite superior	Tolerância	VIF
1	(Constante)	,810	,030		27,060	,000	,752	,869		
	C_Cal	,745	,014	,904	51,536	,000	,717	,774	1,000	1,000
2	(Constante)	,396	,064		6,200	,000	,270	,521		
	C_Cal	,861	,021	1,044	40,893	,000	,819	,902	,434	2,303
	A_Cal	,619	,085	,186	7,283	,000	,452	,785	,434	2,303

a. Variável Dependente: XF1

Figura B.6: Coeficientes do modelo 4

Também o teste de significância do modelo 4 verificou-se o critério pretendido, ou seja, os coeficientes são todos estatisticamente significativos.

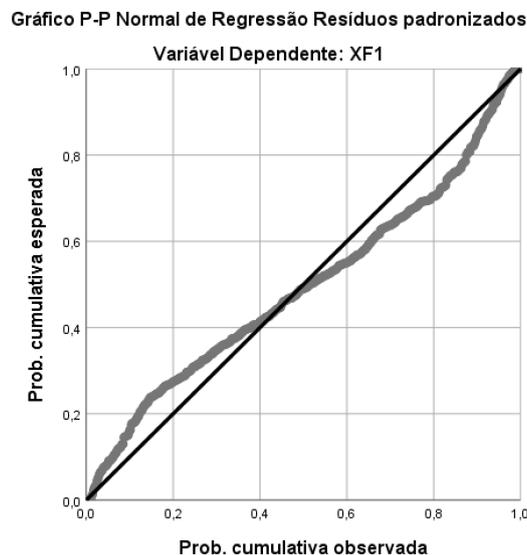


Figura B.7: QQ plot dos resíduos do modelo 4

Assim como aconteceu com o gráfico de probabilidade normal do modelo 3, também para o modelo 4 (Figura B.7) se verifica que os resíduos se encontram afastados da diagonal principal. O teste de Kolmogorov-Smirnov teve um *p-value* igual a 0,000 pelo que é rejeitada a normalidade dos resíduos.

	Testes de Normalidade					
	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Estatística	df	Sig.	Estatística	df	Sig.
Unstandardized Residual	,101	596	,000	,905	596	,000

a. Correção de Significância de Lilliefors

Figura B.8: Teste de normalidade para o modelo 4

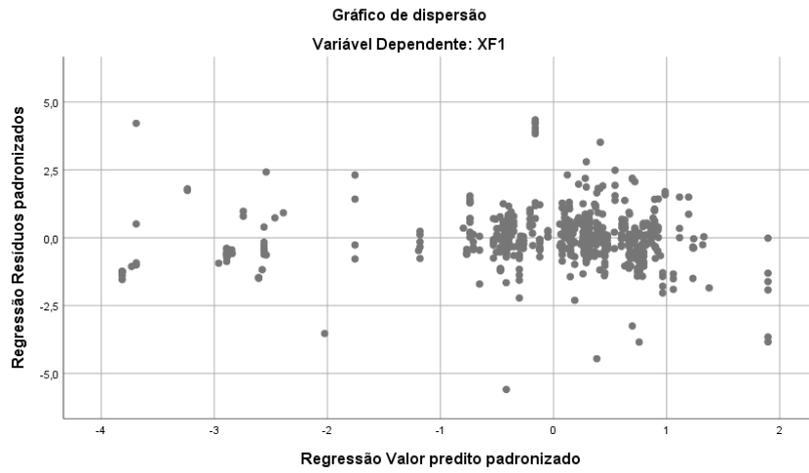


Figura B.9: Gráfico de dispersão do modelo 4

O gráfico de dispersão do modelo 4 confirma a independência dos resíduos deste modelo, pois não são correlacionados.