



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE TECNOLOGIAS E GEOCIÊNCIAS
DEPARTAMENTO DE ENGENHARIA QUÍMICA
ENGENHARIA QUÍMICA

MARCOS GABRIEL LOPES DA SILVA

**APLICAÇÃO DE TÉCNICAS DE INTELIGÊNCIA ARTIFICIAL PARA CONTROLE
DE UM COLUNA DE DESTILAÇÃO: UMA ABORDAGEM COM ASPEN DINÂMICO
E PYTHON**

Recife

2024

MARCOS GABRIEL LOPES DA SILVA

**APLICAÇÃO DE TÉCNICAS DE INTELIGÊNCIA ARTIFICIAL PARA CONTROLE
DE UM COLUNA DE DESTILAÇÃO: UMA ABORDAGEM COM ASPEN DINÂMICO
E PYTHON**

Trabalho de Conclusão de Curso
apresentado ao Curso de Engenharia
Química da Universidade Federal de
Pernambuco, como requisito parcial para
obtenção do título de Engenheiro Químico.

Orientador (a): Dr. José Marcos Francisco da Silva

Recife

2024

Ficha de identificação da obra elaborada pelo autor,
através do programa de geração automática do SIB/UFPE

Silva, Marcos Gabriel Lopes da.

Aplicação de técnicas de inteligência artificial para controle de um coluna de destilação: uma abordagem com Aspen dinâmico e Python / Marcos Gabriel Lopes da Silva. - Recife, 2024.

86, tab.

Orientador(a): José Marcos Francisco da Silva

Trabalho de Conclusão de Curso (Graduação) - Universidade Federal de Pernambuco, Centro de Tecnologia e Geociências, Engenharia Química - Bacharelado, 2024.

Inclui referências, apêndices, anexos.

1. destilação. 2. simulação computacional. 3. dinâmica. 4. controle. 5. inteligência artificial. I. Silva, José Marcos Francisco da. (Orientação). II. Título.

660 CDD (22.ed.)

MARCOS GABRIEL LOPES DA SILVA

**APLICAÇÃO DE TÉCNICAS DE INTELIGÊNCIA ARTIFICIAL PARA CONTROLE
DE UM COLUNA DE DESTILAÇÃO: UMA ABORDAGEM COM ASPEN DINÂMICO
E PYTHON**

Trabalho de Conclusão de Curso
apresentado ao Curso de Engenharia
Química da Universidade Federal de
Pernambuco, como requisito parcial para
obtenção do título de Engenheiro Químico.

Aprovado em: 23/02/2024

BANCA EXAMINADORA

Profa. Dr. José Marcos Francisco da Silva (Orientador)
Universidade Federal de Pernambuco

Prof. Dr. Felipe Pedro da Costa Gomes (Examinador Interno)
Universidade Federal de Pernambuco

Prof. Dr. Allan de Almeida Albuquerque (Examinador Interno)
Universidade Federal de Pernambuco

Dedico este trabalho a meus avós Luiz Vicente, Estelita Nunes e Terezinha de Jesus.

AGRADECIMENTOS

Agradeço a Deus, Fonte de todas as bênçãos que tenho recebido ao longo da vida, pelo Seu constante cuidado, amor e pela graça e misericórdia que têm conduzido os meus passos.

Expresso minha profunda gratidão à minha família, em especial aos meus pais, Marcos Antônio e Ana Paula, e à minha querida irmã, Ana Rebeca, por todo o apoio e incentivo ao longo da minha jornada.

Ao meu orientador, professor José Marcos, manifesto minha sincera gratidão por sua orientação no desenvolvimento deste trabalho, sempre apontando os melhores caminhos para alcançar resultados claros e coerentes.

À professora Kátia, que foi minha orientadora de órgão por quatro anos, ao seu esposo Osmar e à Rebeca, amigas consolidadas ao longo dos seis anos em que estive no Centro de Educação Musical de Olinda. Até hoje, eles permanecem parte integrante da minha vida, oferecendo constante apoio e valiosos conselhos.

Agradeço também às amigas consolidadas durante o primeiro ano de graduação (ABI): Flávio Paiva, Kaynã Xavier, João Victor Albuquerque e João Guedes, que tiveram um papel essencial para minha permanência no curso de engenharia.

E não posso deixar de mencionar Aziz Teófilo, Carlos Belo, Erik Fellipe, Fernanda Azevedo, Hugo Tiburtino, Ícaro e José Alex. Essas amigas foram marcadas por frequentes reuniões de estudo e momentos de descontração envolvendo jogatinas de dominó, memes e piadas de qualidade duvidosa, que contribuíram significativamente para tornar o curso mais leve e agradável.

Agradeço aos professores Sérgio Lucena, Maria de Los Angeles e Sandra Sarmiento, com as quais tive o prazer de trabalhar, por todas as contribuições e incentivos em minha vida acadêmica.

Expresso minha gratidão a todos os integrantes do Laboratório de Bioprodutos e Bioprocessos, no qual realizei iniciação científica, em especial à minha orientadora professora Ester Ribeiro de Andrade e a Eliana, por todo apoio na realização das análises cromatográficas, conversas e conselhos.

Por fim, expresso meu profundo agradecimento a todos os meus irmãos em Cristo da Igreja Presbiteriana A Missão, com destaque para os integrantes do PG de jovens: Ana Cecília, Tiago, Vinicius, Letícia, Bruna e Arthur, com os quais tenho tido a oportunidade de compartilhar preciosos momentos de comunhão e edificação.

“O coração do homem considera o seu caminho, mas o SENHOR lhe dirige os passos.” (Provérbios 16:9)

RESUMO

O processo de separação de hidrocarbonetos desempenha um papel crucial na indústria petroquímica, dada sua ampla aplicação em diversos contextos industriais. A escolha apropriada do solvente empregado na separação é de suma importância para assegurar elevados níveis de pureza nos produtos desejados, destacando-se o sulfolano devido às suas propriedades físico-químicas favoráveis. A técnica de extração por solvente oferece vantagens operacionais e ambientais significativas, enquanto a destilação emerge como uma alternativa eficaz para a recuperação dos solventes utilizados. Todavia, o controle preciso das variáveis operacionais é essencial para garantir a eficácia do processo, representando um desafio considerável na indústria. Outrossim, a integração de técnicas de controle de processos com inteligência artificial revela-se promissora para otimizar a pureza dos produtos em sistemas de recuperação de solvente. Sendo assim, neste estudo, foi simulada, analisada, otimizada e controlada uma coluna recuperadora de solvente empregada no processo de separação de hidrocarbonetos. Ademais, métodos de inteligência artificial foram empregados para prever a pureza do produto de topo da coluna recuperadora de solvente. Como resultado deste trabalho, observou-se uma otimização no número de estágios empregados no processo de separação, uma significativa redução no consumo de vapor de água, bem como um controle eficiente da vazão de alimentação e da temperatura no refeedor por meio de métodos convencionais. Adicionalmente, constatou-se uma capacidade significativa dos modelos de floresta aleatória e regressão logística em prever as frações molares de benzeno e tolueno no produto de topo da coluna, utilizando variáveis facilmente mensuráveis no processo.

Palavras-chave: destilação; simulação computacional; dinâmica; controle; inteligência artificial.

ABSTRACT

The hydrocarbon separation process plays a crucial role in the petrochemical industry, given its wide application in various industrial contexts. The appropriate choice of solvent employed in the separation is of utmost importance to ensure high levels of purity in the desired products, with sulfolane standing out due to its favorable physicochemical properties. Solvent extraction technique offers significant operational and environmental advantages, while distillation emerges as an effective alternative for the recovery of the solvents used. However, precise control of operational variables is essential to ensure the effectiveness of the process, representing a considerable challenge in the industry. Furthermore, the integration of process control techniques with artificial intelligence proves promising for optimizing the purity of products in solvent recovery systems. Thus, in this study, a solvent recovery column employed in the hydrocarbon separation process was simulated, analyzed, optimized, and controlled. Additionally, artificial intelligence methods were employed to predict the purity of the top product of the solvent recovery column. As a result of this work, an optimization was observed in the number of stages employed in the separation process, a significant reduction in water vapor consumption, as well as efficient control of feed flow rate and temperature in the reboiler through conventional methods. Additionally, a significant capability of random forest and logistic regression models was observed in predicting the molar fractions of benzene and toluene in the top product of the column, using variables easily measurable in the process.

Keywords: distillation; computational simulation; dynamics; control; artificial intelligence.

LISTA DE ILUSTRAÇÕES

Figura 1 –	Representação simplificada de produtos obtidos a partir da destilação do petróleo bruto	20
Figura 2 –	Representação esquemática de uma coluna de destilação	22
Figura 3 –	Procedimento para simulação de sistemas multicomponentes	25
Figura 4–	Árvore de decisão para escolha de modelo termodinâmico para simulação de processo	26
Figura 5–	Coluna de destilação operando em estado estacionário e com refluxo máximo	27
Figura 6–	Coluna de destilação com alimentação, retirada de produtos, infinitos estágios de equilíbrio e um ponto de estrangulamento (Pinch-point zone)	30
Figura 7–	Representação de um estágio de Equilíbrio para modelagens rigorosas.	32
Figura 8–	Esquema do processo de separação de hidrocarbonetos	42
Figura 9–	Representação da coluna de destilação usada no Aspen Dinâmico.	45
Figura 10–	Estratégia convencional adotada para o controle da composição	46
Figura 11–	Razão de refluxo e Energia do referedor	50
Figura 12–	Mapa de curva residual para mistura Benzeno, Tolueno e sulfolane.	52
Figura 13–	Diagrama ternário para o sulfolane, heptano e benzeno.	52
Figura 14–	Diagrama ternário: água, tolueno e sulfolane.	53
Figura 15–	Diagrama ternário: água, benzeno e sulfolane	54

Figura 16–	Temperatura ao longo da coluna.	55
Figura 17–	Distribuição molar das fases	55
Figura 18–	Fração molar dos componentes ao longo da coluna recuperadora	56
Figura 19–	Resposta do controlador FC_1 a perturbações na carga	60
Figura 20–	Resposta da temperatura no refeedor a perturbações na carga	60
Figura 21–	Matriz de confusão	62
Figura 22–	Matriz de confusão- método de Spearman.	62
Figura 23–	Matriz de confusão- método de Kendall	63
Figura 24–	Matriz de confusão- método de Pearson	63
Figura 25–	Dados do conjunto de treinamento e previsões realizadas pelo modelo Linear.	67
Figura 26–	Dados do conjunto de treinamento e previsões realizadas pelo modelo de árvore de decisão	67
Figura 27–	Dados do conjunto de testes e previsões realizadas pelo modelo de árvore de decisão	68
Figura 28–	Dados do conjunto de treinamento e previsões realizadas pelo modelo floresta aleatória	68
Figura 29–	Dados do conjunto de testes e previsões realizadas pelo modelo floresta aleatória	69
Figura 30–	Fração mássica de tolueno controle por IA.	69
Figura 30–	Fração mássica de benzeno controle por IA.	70

LISTA DE TABELAS

Tabela 1 –	Composição elementar do óleo cru de petróleo.	18
Tabela 2 –	Variáveis de interesse em uma coluna de destilação	36
Tabela 3 –	Composição da carga de alimentação da coluna extratora.	43
Tabela 4–	Composição da carga de alimentação da coluna recuperadora de solvente em estudo.	43
Tabela 5–	Funções dos controladores na coluna de destilação.	46
Tabela 6–	Sensores distribuídos ao longo da coluna de destilação.	47
Tabela 7–	Composição da carga de alimentação da coluna recuperadora de solvente	51
Tabela 8–	Valores gerados pela simulação na coluna de destilação sem degradação do solvente.	57
Tabela 9–	Valores da coluna recuperadora de solvente disponíveis na literatura	58
Tabela 10–	Parâmetros ótimos obtidos para os controladores	59
Tabela 11-	Métricas de desempenho para os modelos de aprendizagem de máquinas aplicados.	64

LISTA DE ABREVIATURAS E SIGLAS

ADM	Análise de dados multivariados
AED	Análise exploratória de dados
APC	Controle Avançado de Processos (<i>Advanced Process Control</i>)
DT	Árvore de Decisão (<i>Decision Tree</i>)
LDA	Análise discriminante linear
MAE	Erro médio absoluto
ML	Aprendizagem de máquinas (<i>Machine Learning</i>)
MPC	Controle Preditivo Baseado em Modelo (<i>Model Predictive Control</i>)
MSE	Erro quadrático médio
NRTL	Dois líquidos não aleatórios (<i>Non-Random Two-Liquid</i>)
PCA	Análise de componentes principais
PCR	Regressão de componentes principais
PLSR	Regressão de mínimos quadrados principais
QAV	Querosene de aviação
RF	Floresta aleatória (<i>Random Forest</i>)
SVM-C	Classificação por Máquinas de Vetores de Suporte

SUMÁRIO

1	INTRODUÇÃO	16
2	REVISÃO DE LITERATURA	18
2.1	COMPOSIÇÃO DA MATÉRIA-PRIMA E PRODUTOS DE UMA REFINARIA DE PETRÓLEO	18
2.2	DESTILAÇÃO	20
2.3	COLUNAS DE DESTILAÇÃO	22
2.4	DESTILAÇÃO EXTRATIVA	23
2.5	SIMULAÇÃO COMPUTACIONAL E MODELO TERMODINÂMICO	24
2.6	DETERMINAÇÃO APROXIMADA DO NÚMERO MÍNIMO DE ESTÁGIOS EM EQUILÍBRIO- MÉTODO DE FENSKE	26
2.7	DETERMINAÇÃO APROXIMADA DA RAZÃO DE REFLUXO MÍNIMO- MÉTODO DE UNDERWOOD	30
2.8	MODELAGEM RIGOROSA PARA ESTÁGIOS DE EQUILÍBRIO EM COLUNAS DE DESTILAÇÃO	32
2.9	CONTROLE DE PROCESSOS EM COLUNAS DE DESTILAÇÃO	34
2.10	ANÁLISE MULTIVARIADA E CONTROLE PREDITIVO DE PROCESSOS COM BASE EM MODELOS DE MACHINE LEARNING	37
3	METODOLOGIA	42
3.1	SIMULAÇÃO EM ESTADO ESTACIONÁRIO	42
3.2	SIMULAÇÃO DINÂMICA	44

3.3	IMPLEMENTAÇÃO DE SENSORES PARA GERAÇÃO DE DADOS	47
3.4	TRATAMENTO DE DADOS POR MODELOS DE MACHINE LEARNING	48
4	RESULTADOS E DISCUSSÃO	50
4.1	SIMULAÇÃO EM ESTADO ESTACIONÁRIO	50
4.2	DINÂMICA E CONTROLE CONVENCIONAL	58
4.3	APLICAÇÃO DE MÉTODOS DE INTELIGÊNCIA ARTIFICIAL NA COLUNA RECUPERADORA DE SOLVENTE	61
5	CONCLUSÃO	71
	REFERÊNCIAS	73

1 INTRODUÇÃO

O processo de separação de hidrocarbonetos é de extrema importância para a indústria petroquímica, dada a ampla utilização desses compostos em variados processos industriais. O tipo de solvente empregado para essa finalidade desempenha um papel crucial na obtenção de elevados níveis de pureza nos produtos de interesse. Dentro dessa perspectiva, destacam-se as propriedades físico-químicas do sulfolano, tornando-o um solvente altamente adequado para essa finalidade (Rappel et al., 2002).

A extração por solvente é um procedimento industrial que oferece diversas vantagens, tais como condições operacionais suaves, baixo consumo de energia, escalabilidade fácil e operação simplificada (Shi et al., 2024). Após a extração, a recuperação de solventes durante o processo é amplamente empregada para reduzir os custos operacionais e diminuir a geração de resíduos, aproveitando a disponibilidade generalizada de equipamentos de recuperação de solventes. Nesse contexto, a destilação se destaca como uma alternativa eficiente para separar e recuperar solventes de outros líquidos, com eficiência de remoção geralmente elevada, sendo aplicável tanto a misturas de solventes quanto a solventes individuais (Nemeth e Laszlo, 2020).

A destilação ainda permanece como um dos métodos de separação mais utilizados e versáteis nas indústrias petroquímica e química, representando cerca de 25-40% do consumo de energia nesses setores. Devido à sua eficiência energética relativamente baixa, este processo unitário é frequentemente um dos maiores consumidores de energia em processos industriais, portanto, requer a implementação de estratégias eficazes de controle para garantir a qualidade dos produtos, rendimentos, operação segura, limites ambientais e maximização do lucro (Al Kalbani e Zhang, 2023; Oravec et al., 2017).

Neste contexto, o Controle Preditivo Baseado em Modelo (Model Predictive Control - MPC) está na vanguarda das estratégias de controle, oferecendo a capacidade de otimizar o desempenho do controlador através da minimização de uma determinada função objetivo, levando em consideração futuros erros e restrições nas variáveis de processo. A adoção industrial do MPC começou na década de 1970, especialmente nos processos da indústria petrolífera. Inicialmente introduzido como Controle de Matriz Dinâmica (Dynamic Matrix Control - DMC), o MPC tem despertado

grande interesse entre pesquisadores nas últimas três décadas, resultando em avanços significativos nesta abordagem de controle (Oravec et al., 2017).

Por sua vez, o Controle Avançado de Processos (Advanced Process Control - APC) geralmente consiste em uma estratégia de controle ótimo implementada em conjunto com controladores básicos (PID). Esses controladores básicos garantem o rastreamento dos pontos de ajuste e a rejeição de perturbações. O APC supervisiona os controladores básicos, otimizando os pontos de ajuste e facilitando o desacoplamento e a minimização dos efeitos colaterais decorrentes das interações entre várias variáveis de processo (Oravec et al., 2017).

Em comparação com outras estratégias bem estabelecidas, o APC melhora os rendimentos do produto, reduz os custos operacionais e aprimora a segurança do processo. Além disso, o período típico de retorno do investimento para a implementação do APC é inferior a seis meses (Oravec et al., 2017).

Além disso, ao longo das décadas, a comunidade de engenharia química tem se beneficiado de avanços substanciais em modelagem, simulação e experimentação, resultando em uma vasta coleção de dados. Essa abundância de informações possibilita a aplicação eficaz de modelos de aprendizado de máquina (ML) para prever variáveis físicas e químicas, conforme destacado por Maarten et al. (2021).

No contexto da aplicação de técnicas de inteligência artificial em colunas de destilação, Kwon et al. (2021) e Zhao e Lu (2021) desenvolveram pesquisas voltadas para o controle da temperatura de operação das colunas de destilação, enquanto Uchoa et al. (2023) conduziram um estudo focado no controle de inundação da coluna.

Portanto, este trabalho tem como objetivo simular, analisar e controlar a vazão e a temperatura do refeedor de uma coluna de destilação usada na recuperação do sulfolano, além de prever a pureza dos compostos de topo por meio de métodos de inteligência artificial. Para isso, será empregado o software Aspen Plus® nos modos estacionário e dinâmico, juntamente com a linguagem de programação Python. A fim de viabilizar essa simulação, a carga será caracterizada como um sistema multicomponente, contendo dois compostos cíclicos, três alifáticos e o solvente sulfolano, com base em estudos e dados disponíveis na literatura.

2 REVISÃO DE LITERATURA

2.1 COMPOSIÇÃO DA MATÉRIA-PRIMA E PRODUTOS DE UMA REFINARIA DE PETRÓLEO

O refino do petróleo desempenha um papel crucial na sociedade contemporânea, dada a predominância de derivados como gasolina, diesel, querosene de aviação (QAV) e óleo combustível como fontes de energia para a maioria dos meios de transporte (Fahim et al., 2012).

Nas refinarias modernas, os processos de refino se dividem entre separação física e conversão química (Fahim et al., 2012).

O óleo cru, matéria-prima fundamental para as refinarias de petróleo, exibe diferentes propriedades químicas e físicas, dependendo da região do mundo. Essa substância constitui-se como uma complexa mistura líquida, composta por uma variedade de hidrocarbonetos em proporções diversas. Além disso, é comum a presença de pequenas quantidades de impurezas, tais como enxofre, oxigênio, nitrogênio e metais (Fahim et al., 2012).

As impurezas encontradas no óleo cru são indesejáveis, uma vez que podem impactar a qualidade dos produtos resultantes e, adicionalmente, têm o potencial de contaminar e corroer catalisadores (Fahim et al., 2012). Portanto, a compreensão e gestão eficiente desses componentes indesejáveis tornam-se cruciais para assegurar a eficácia dos processos de refino e a qualidade dos derivados obtidos. A Tabela 1 apresenta a composição aproximada do óleo cru.

Tabela 1 – Composição elementar do óleo cru de petróleo.

Elemento	Composição (% em peso)
Carbono	83,0-87,0
Hidrogênio	10,0-14,0
Enxofre	0,05-6,0
Nitrogênio	0,1-0,2

Oxigênio	0,05-2,0
Ni	<120ppm
V	<1200ppm

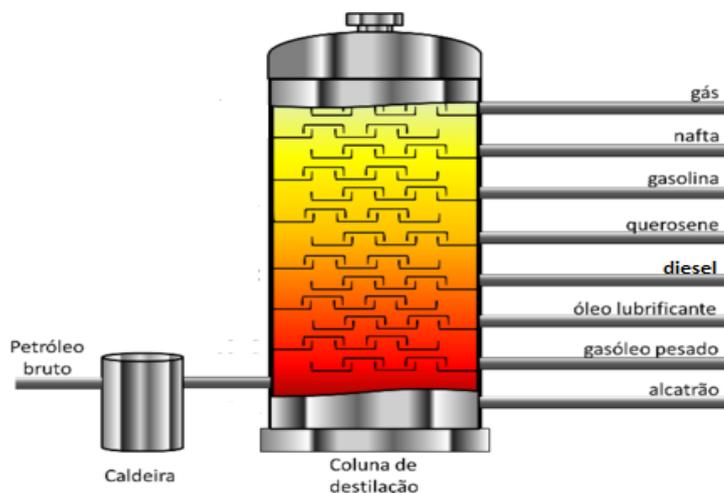
Fonte: adaptado de Fahim et al (2012).

No óleo cru, identificam-se três classes principais de hidrocarbonetos, conforme delineado por Roussel e Boulet (1995):

- Hidrocarbonetos saturados: compostos que apresentam exclusivamente ligações simples carbono-carbono e são conhecidos como parafinas (ou alcanos), quando de natureza acíclica, ou naftos (ou ciclanos), quando cíclicos.
- Hidrocarbonetos insaturados: categoria que compreende compostos que contêm ligações múltiplas carbono-carbono, tais como duplas, triplas ou ambas. São considerados insaturados, pois possuem uma menor proporção de átomos de hidrogênio por átomo de carbono em comparação com as parafinas. Os hidrocarbonetos insaturados são denominados olefinas.
- Hidrocarbonetos aromáticos: constituem uma classe distinta de compostos cíclicos insaturados, caracterizados pela presença de um ou mais anéis de benzeno. Essa categoria engloba compostos que exibem propriedades específicas associadas à estrutura do benzeno.

A Figura 1 apresenta de maneira simplificada os principais produtos obtidos a partir da destilação do petróleo bruto.

Figura 1 – Representação simplificada de produtos obtidos a partir da destilação do petróleo bruto.



Fonte: WikiCiências, 2023.

2.2 DESTILAÇÃO

A operação predominantemente utilizada na engenharia de processos é a destilação, destacada por sua capacidade de separar misturas com base em suas diferentes volatilidades, mediante a aplicação ou remoção controlada de calor (Mendez et al., 2022). Este processo envolve a interação entre uma fase vapor e uma fase líquida, resultando em uma transferência de massa e calor entre ambas as fases (FOUST et al., 1982). Embora líquido e vapor contenham, em grande parte, os mesmos componentes, suas proporções relativas variam (FOUST et al., 1982). Durante essa operação unitária, o líquido se encontra em seu ponto de bolha, enquanto o vapor está em equilíbrio no ponto de orvalho (FOUST et al., 1982). Ocorre uma transferência simultânea de massa do líquido para o vapor, através da vaporização de componentes leves, e do vapor para o líquido, por meio da condensação de componentes mais pesados (FOUST et al., 1982). Isso resulta em um aumento na concentração dos componentes mais leves na fase vapor e dos mais pesados na fase líquida.

O equilíbrio líquido-vapor de uma mistura é descrito pelo coeficiente de distribuição, conhecido como o K-fator. Para cada componente, o K é definido como

a razão entre as frações molares na fase vapor e líquida em equilíbrio (Seader et al. 2011). A expressão para o K é representada pela Equação 1.

$$K = \frac{y_i}{x_i} \quad (1)$$

Para cálculos de estágios de equilíbrio, como é o caso da destilação, os fatores de separação são definidos. No caso líquido-vapor, a volatilidade relativa (α) entre os componentes i e j é dada pela equação 2 (Seader et al. 2011).

$$\alpha = \frac{K_i}{K_j} \quad (2)$$

Ademais, a fugacidade parcial do componente i no vapor (\bar{f}_{iV}) é igual a fugacidade parcial do componente i na fase líquida (\bar{f}_{iL}), conforme mostra a Equação 3 (Seader et al. 2011).

$$\bar{f}_{iV} = \bar{f}_{iL} \quad (3)$$

Onde f_{iV} e f_{iL} são obtidas pelas equações 4 e 5 respectivamente.

$$\bar{f}_{iL} = \bar{\phi}_{iL} x_i P \quad (4)$$

$$\bar{f}_{iV} = \bar{\phi}_{iV} y_i P \quad (5)$$

Portanto, para uma modelagem rigorosa o K- valor pode ser reescrito conforme as equações 6 e 7

$$K_i = \frac{\bar{\phi}_{iL}}{\bar{\phi}_{iV}} \quad (6)$$

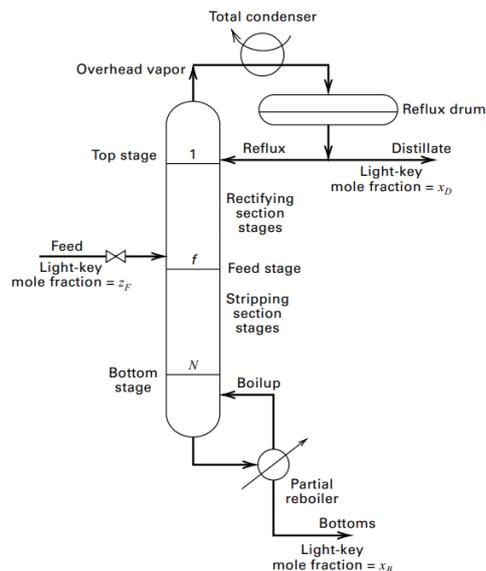
$$K_i = \frac{\gamma_{iL} f_{iL}^o}{\phi_{iV} P} = \frac{\gamma_{iL} \phi_{iL}}{\phi_{iV}} \quad (7)$$

onde Phi é o coeficiente de fugacidade da espécie pura.

2.3 COLUNAS DE DESTILAÇÃO

A destilação emerge como a operação fundamental nos domínios petroquímico e farmacêutico, essencial para a separação de misturas complexas. Esta operação é efetuada em uma coluna vertical, cujo processo visa a segregação de componentes presentes na mistura. A estrutura da coluna destilatória compreende a pratos (ou recheios) destinados a otimização da eficiência da separação, um refeedor responsável por fornecer calor para a vaporização, localizado na base da coluna, um condensador projetado para resfriar e condensar o vapor proveniente do ápice da coluna e, por fim, um tambor de refluxo, incumbido de reter o vapor condensado para permitir o reciclo do refluxo líquido proveniente do topo da coluna (Mendez et al., 2022). A Figura 2 mostra o esquema de uma coluna de destilação em operação.

Figura 2 – Representação esquemática de uma coluna de destilação.



Fonte: Seader et al (2011).

A substância líquida passível de processamento é designada como 'alimentação' e pode ser introduzida nas proximidades do ponto médio da coluna, em

um estágio denominado 'prato de alimentação'. Dentro desse contexto, a coluna é subdividida em duas seções distintas: a seção de enriquecimento ou retificação, localizada acima, e a seção de extração ou esgotamento, posicionada abaixo do prato de alimentação (Mendez et al., 2022).

2.4 DESTILAÇÃO EXTRATIVA

Os azeótropos são misturas de composição específica cujas temperaturas de ebulição resultam em proporções idênticas de líquido e vapor, mantidas sob uma pressão constante. A temperatura de ebulição dessas misturas pode ser inferior ou superior à temperatura de ebulição dos compostos individuais presentes na mistura original, impedindo a recuperação eficaz de cada componente por meio de destilação simples (Gerbaud et al., 2019).

Como resposta a essa limitação, diversos métodos de destilação foram desenvolvidos, muitos dos quais incorporam a adição de um terceiro componente, denominado solvente ou extrator. A destilação azeotrópica e a destilação extrativa são exemplos dessas abordagens (Gerbaud et al., 2019).

O solvente, ao ser introduzido como terceiro componente na mistura a ser separada, desempenha o papel crucial de interromper o fenômeno azeotrópico, modificando a volatilidade relativa dos constituintes da mistura (Ma et al., 2019). A seleção adequada do solvente requer considerações meticulosas em relação a propriedades como estabilidade, não corrosividade, custo acessível, seletividade, baixo ponto de ebulição, baixa toxicidade, entre outras (Ma et al., 2019).

A destilação extrativa, extensivamente estudada e implementada na indústria, demanda uma modificação substancial na volatilidade relativa dos componentes a serem separados, quando o solvente é introduzido (Ma et al., 2019).

Essa pode ser categorizada em duas modalidades operacionais principais: destilação extrativa contínua e destilação extrativa descontínua (Lei et al., 2003).

No caso da destilação extrativa contínua, duas colunas principais são empregadas: a coluna extrativa e a coluna de recuperação de solvente. A mistura a ser separada é introduzida na seção inferior da coluna extrativa, enquanto o solvente é inserido na seção superior da coluna de recuperação. A obtenção de dois produtos finais ocorre no topo de ambas as colunas, e o solvente é reciclado da parte inferior da coluna de recuperação de solvente para a coluna extrativa (Ma et al., 2019).

Já na destilação extrativa em batelada, tanto a mistura a ser separada quanto o solvente são carregados simultaneamente no refeedor, com os produtos sendo obtidos no topo da coluna. A determinação precisa da proporção adequada entre a alimentação e o solvente é crucial para alcançar a máxima eficiência (Ma et al., 2019).

2.5 SIMULAÇÃO COMPUTACIONAL E MODELO TERMODINÂMICO

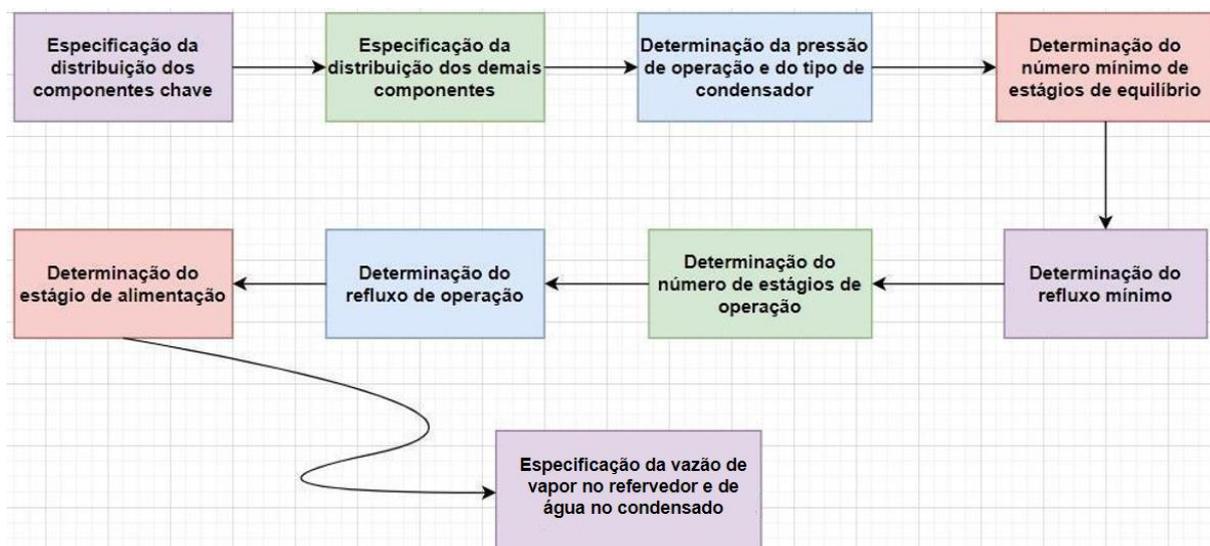
A simulação computacional é uma ferramenta essencial para engenheiros químicos envolvidos na concepção e otimização de processos químicos e industriais complexos. Essa abordagem permite a exploração de diversos cenários, a manipulação de variáveis de processo e a configuração de sistemas, sem a priori, necessidade de construção de protótipos físicos. Essa metodologia proporciona economia notável de tempo e recursos, pois a detecção e resolução de problemas durante a fase de simulação mostram-se mais econômicas em comparação com intervenções em uma planta real.

A simulação não apenas aumenta a eficiência do processo, mas também permite a otimização do uso de matérias-primas, energia e recursos, contribuindo para a sustentabilidade ambiental. No contexto do desenvolvimento e teste de estratégias de controle de processos, a simulação desempenha um papel crucial, assegurando a eficácia e estabilidade do sistema em condições operacionais reais.

Especificamente na simulação do processo de extração, a aplicação de métodos empíricos para caracterização de todos os componentes da mistura de hidrocarbonetos requer uma extensa base de informações experimentais (Fouladvand et al., 2021).

Quanto à simulação de sistemas de separação multicomponente e de múltiplos estágios, Seader et al. (2011) propõem um conjunto de etapas que estão representadas no fluxograma da Figura 3.

Figura 3 – Procedimento para simulação de sistemas multicomponentes.



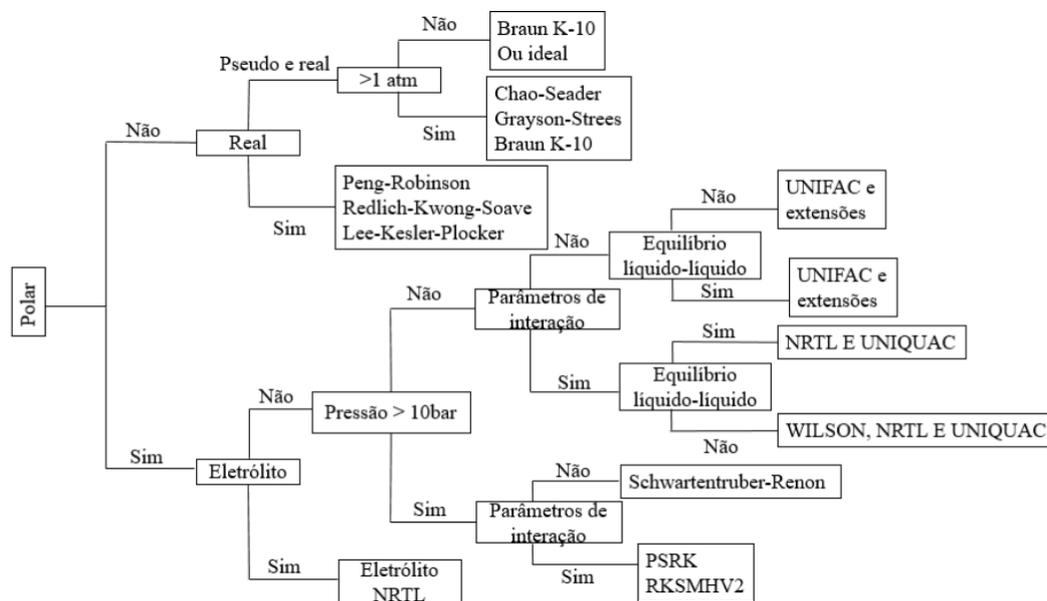
Fonte: Adaptado Seader et al (2011).

Nesse contexto, a termodinâmica desempenha um papel fundamental na análise e descrição de operações e processos, fornecendo um arcabouço teórico para compreender e quantificar fenômenos relacionados à energia e equilíbrio em sistemas. A compreensão do comportamento microscópico da matéria é crucial para a determinação de propriedades termodinâmicas, as quais desempenham um papel essencial em toda engenharia de processos (SMITH et al., 2007). A termodinâmica também é um fator essencial para as operações de separação, abrangendo considerações relacionadas à transferência de calor, equilíbrio de fases, atividade biológica e dimensionamento de equipamentos (SEADER et al., 2010).

Em projetos e simulações de processos de extração líquido-líquido, é imperativo estabelecer um modelo termodinâmico consistente para representar o equilíbrio vapor-líquido-líquido (VLE). A literatura científica apresenta uma variedade de abordagens que empregam modelos termodinâmicos como NRTL, UNIQUAC e UNIFAC para essa finalidade (Fouladvand et al., 2021).

A Figura 4 apresenta uma árvore de decisão que orienta a escolha do modelo termodinâmico apropriado em diferentes contextos. Em situações em que a pressão adotada na simulação não excede 10 bar, os parâmetros binários de interação binárias são conhecidos e o equilíbrio vapor-líquido-líquido é observado, percebe-se que o modelo NRTL é adequado para realização de simulações (Carlson, 1996).

Figura 4 – Árvore de decisão para escolha de modelo termodinâmico para simulação de processo.



Fonte: Adaptado de Carlson (1996) e AL-MALAH (2017).

O modelo NRTL, proposto por Renon e Prausnitz em 1968, destaca-se como uma ferramenta aplicável em sistemas parcialmente miscíveis, conforme discutido por Martins et al. (2010). Neste contexto, para a condução deste estudo, foi deliberadamente selecionado o modelo NRTL, alinhado aos dados disponibilizados por Mesquita et al. (2015) e Rapel et al. (2002). Esses dados abrangem o equilíbrio de fases entre componentes alifáticos, aromáticos e sulfolane, proporcionando uma base consistente para as análises termodinâmicas envolvidas no escopo do presente trabalho.

2.6 DETERMINAÇÃO APROXIMADA DO NÚMERO MÍNIMO DE ESTÁGIOS EM EQUILÍBRIO- MÉTODO DE FENSKE

Um estágio de equilíbrio é considerado uma unidade de equipamento onde ocorre o íntimo contato de duas fases distintas, seguido pela sua subsequente separação. Nesse processo, a difusão desencadeia a redistribuição dos componentes da mistura entre as fases (FOUST et al., 1982).

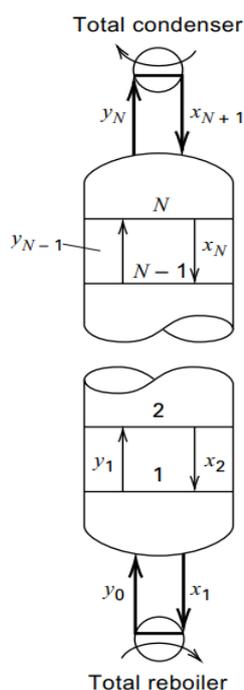
Apesar da existência de métodos rigorosos para resolver problemas de separação em sistemas com múltiplos componentes, métodos aproximados ainda são

empregados nas fases iniciais de projetos. Eles são utilizados em estudos paramétricos para determinar condições ideais de projeto, na síntese de processos para estabelecer sequências de separação ótimas e no fornecimento de aproximações iniciais para métodos iterativos mais rigorosos (SEADER et al., 2011).

O número mínimo de estágios de equilíbrio corresponde ao refluxo total, uma condição alcançada na prática quando uma coluna recebe uma corrente de alimentação e atinge o estado estacionário, caracterizado pela ausência mudanças na alimentação e retirada de produtos (SEADER et al., 2011).

A Figura 5 apresenta uma coluna de destilação com refluxo total e em estado estacionário, conforme as considerações adotadas por Fenske.

Figura 5 – Coluna de destilação operando em estado estacionário e com refluxo máximo.



Fonte: Seader et al (2011)

Com a abordagem utilizada na Figura 5, os estágios são numerados de baixo para cima. Todo vapor proveniente do topo do estágio N é condensado e retorna ao estágio N como refluxo. Simultaneamente, todo líquido que sai do estágio 1 é vaporizado no reboiler e retorna ao estágio 1 como vapor. Para manter a operação em estado estacionário dentro da coluna, a entrada de calor através do reboiler e

a saída de calor no condensador devem ser iguais, assumindo idealidade, ou seja, sem perdas de calor (Seader et al. 2011).

A descrição matemática de uma equação para o número mínimo de estágios de equilíbrio envolve a definição do valor de K e a fração molar do líquido e vapor (x_i, y_i) entre os estágios, considerando o componente i no estágio 1 (Seader et al. 2011).

Sabe-se que:

$$y_{i,1} = K_{i,1} \cdot x_{i,1} \quad (9)$$

$$y_{i,1} = x_{i,2} \quad (10)$$

Combinando as Equações 9 e 10 obtêm-se a equação 11.

$$x_{i,2} = K_{i,1} x_{i,1} \quad (11)$$

onde, por analogia resulta na Equação 12.

$$y_{i,2} = K_{i,2} x_{i,2} \quad (12)$$

Ao substituir a Equação 11 na Equação 12, obtêm-se a Equação 13.

$$y_{i,2} = K_{i,2} K_{i,1} x_{i,1} \quad (13)$$

Portanto, por analogia a Equação 13 para o componente i no estágio N obtêm-se a Equação 14.

$$y_{i,N} = K_{i,N} K_{i,N-1} \dots K_{i,2} K_{i,1} x_{i,1} \quad (14)$$

Analogamente para o componente j, obtêm-se a Equação 15.

$$y_{j,N} = K_{j,N}K_{j,N-1} \dots K_{j,2}K_{j,1}x_{j,1} \quad (15)$$

Dividindo a Equação 14 pela 15 é possível obter a Equação 16, onde α é a volatilidade relativa.

$$\frac{y_{i,N}}{y_{j,N}} = \alpha_N \alpha_{N-1} \dots \alpha_1 \alpha_2 \frac{x_{i,1}}{x_{j,1}} \quad (16)$$

$$\left(\frac{x_{i,N+1}}{x_{i,1}} \right) \left(\frac{x_{j,1}}{x_{j,N+1}} \right) = \alpha^N \quad (17)$$

Finalmente, obtêm-se a equação de Fenske para determinação do número mínimo de estágios, representada na equação 18.

$$N_{min} = \frac{\log \left\{ \left[\left(\frac{x_{i,N+1}}{x_{i,1}} \right) \right] \left[\left(\frac{x_{j,1}}{x_{j,N+1}} \right) \right] \right\}}{\log \alpha_{i,j}} \quad (18)$$

Uma formulação mais conveniente para aplicar a equação de Fenske envolve a substituição do produto da razão das frações molares pela expressão que representa a razão do fluxo molar dos produtos em relação ao destilado (d_i, d_j) e ao produto de base (b_i, b_j). Dessa maneira, a equação de Fenske pode ser reformulada, conforme a Equação 19 (Seader et al., 2011).

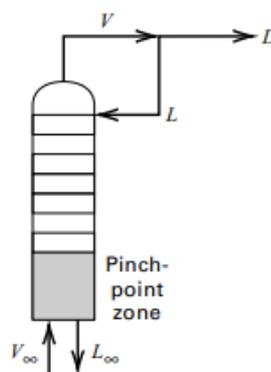
$$N_{min} = \frac{\log \left\{ \left[\left(\frac{d_i}{d_j} \right) \right] \left[\left(\frac{b_i}{b_j} \right) \right] \right\}}{\log \alpha_m} \quad (19)$$

$$\alpha_m = [(\alpha_{i,j})_N (\alpha_{i,j})_1]^{1/2} \quad (20)$$

2.7 DETERMINAÇÃO APROXIMADA DA RAZÃO DE REFLUXO MÍNIMO- MÉTODO DE UNDERWOOD

O conceito de refluxo mínimo é uma condição limitante valiosa em processos de destilação. Ao contrário das considerações feitas para a determinação do número mínimo de estágios, essa condição considera a presença de uma corrente de alimentação e a retirada de produtos, mas a coluna é tratada como fictícia, exigindo um número infinito de estágios de equilíbrio. A Figura 6 apresenta a configuração de uma coluna de destilação com uma zona de pinch-point (Seader et al., 2011).

Figura 6 – Coluna de destilação com alimentação, retirada de produtos, infinitos estágios de equilíbrio e um ponto de estrangulamento (Pinch-point zone)



Fonte: Adaptado de Seader et al (2011).

A classificação proposta por Shiras, Hanson e Gibson (1950) diferencia sistemas multicomponentes em duas categorias: Classe 1, caracterizada por um pinch point (ponto de estrangulamento), e Classe 2, que apresenta dois pinch points (Seader et al. 2011). Os pinch points referem-se a regiões específicas na coluna de destilação onde as composições do líquido e do vapor estão extremamente próximas, dificultando a separação eficiente dos componentes (SEADER et al., 2011).

Na Classe 1, presume-se que todos os produtos provenientes da alimentação são distribuídos entre o destilado e o produto de base. A separação ocorre em intervalos estreitos de pontos de ebulição ou quando não há uma distinção clara no grau de separação entre os principais componentes (Seader et al. 2011).

Na Classe 2, um ou mais componentes aparecem exclusivamente no destilado ou no produto de base. Caso nem o destilado nem o produto de base contenham todos os componentes da alimentação, surgem dois pinch points, um na seção de

retificação e outro na seção de esgotamento, distantes do estágio de alimentação. Os produtos mais pesados, ausentes no destilado, são retirados pelos estágios entre o estágio de alimentação e o pinch point na seção de retificação. Enquanto isso, os produtos mais leves, que não estão presentes no produto de base, são removidos pelos estágios entre a alimentação e o pinch point na seção de esgotamento (Seader et al., 2011).

Considerando um pinch point na seção de retificação, seja ou não no estágio de alimentação, o balanço de massa para os componentes em todos os estágios da coluna é expresso de acordo com as formulações apresentadas por Seader et al. (2011).

Considerando a Figura 6, o balanço global e o balanço por componente são representados respectivamente pelas equações 21 e 22, onde L e D são as vazões de refluxo e do destilado respectivamente.

$$y_{i,\infty} = x_{i,\infty}L_{\infty} + x_{i,D}D \quad (21)$$

$$V_{\infty} = L_{\infty} + D \quad (22)$$

A composição de equilíbrio não se altera na região de pinch point, e a relação de equilíbrio é determinada pela Equação 23.

$$y_{i,\infty} = k_{i,\infty}x_{i,\infty} \quad (23)$$

Combinando as equações 22 e 23, obtêm-se a equação de Underwood para o refluxo mínimo (Equação 24).

$$\frac{L_{\infty}}{D} = \frac{\left[\left(\frac{x_{i,D}}{x_{i,\infty}} \right) - (\alpha_{i,j})_{\infty} \left(\frac{x_{j,D}}{x_{j,\infty}} \right) \right]}{(\alpha_{i,j})_{\infty} - 1} \quad (24)$$

2.8 MODELAGEM RIGOROSA PARA ESTÁGIOS DE EQUILÍBRIO EM COLUNAS DE DESTILAÇÃO

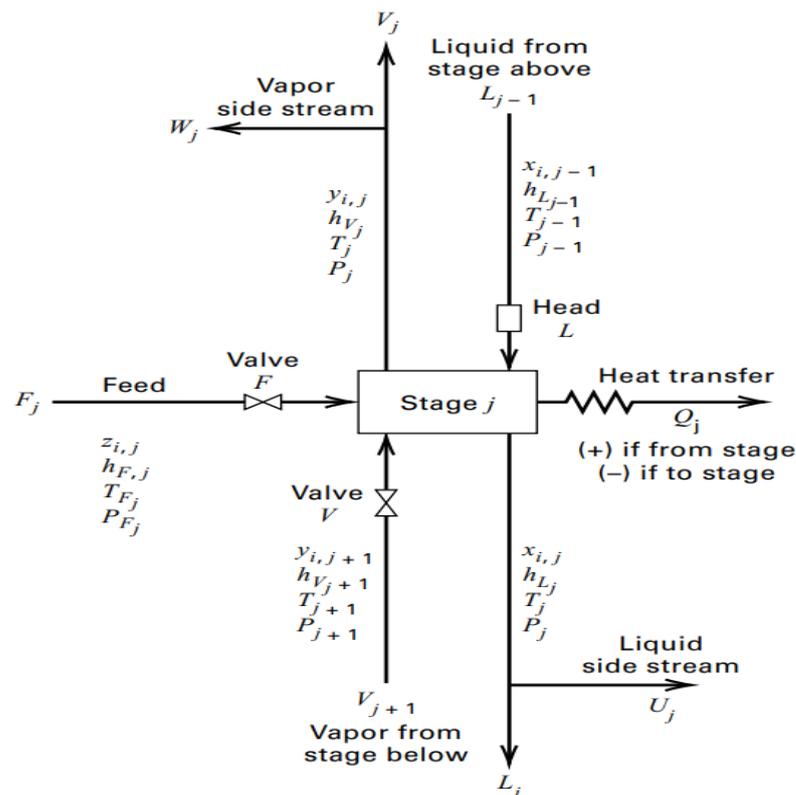
O projeto final de equipamentos para separação em múltiplos estágios e multicomponentes requer a determinação de temperaturas, pressões, taxas de fluxo, composições e taxas de transferência de calor em cada estágio. Essas determinações são obtidas por meio da resolução de sistemas algébricos não lineares, fundamentados nos balanços de massa e energia em cada estágio (Seader et al., 2011).

Para a modelagem rigorosa de uma coluna de destilação, são consideradas as seguintes hipóteses (Seader et al., 2011):

1. O equilíbrio de fases é alcançado em cada estágio.
2. Não ocorrem reações químicas em nenhum dos estágios da coluna.
3. O arraste de gotas líquidas no vapor e a oclusão de bolhas de vapor no líquido são insignificantes.

A Figura 7 mostra a representação esquemática de um estágio de equilíbrio.

Figura 6 – Representação de um estágio de equilíbrio para modelagens rigorosas.



Fonte: Seader et al (2011).

A entrada no estágio j consiste em uma alimentação monofásica ou bifásica, com vazão molar F_j , e composição global expressa em frações molares $z_{i,j}$ do componente i . Além disso, são especificadas a temperatura T_{Fj} , a pressão P_{Fj} e a entalpia molar global correspondente h_{Fj} . A pressão de alimentação é igual ou superior à pressão do estágio P_j (Seader et al., 2011).

O estágio j também recebe líquido proveniente do estágio $j-1$, com vazão molar L_{j-1} , frações molares $x_{i,j-1}$, entalpia $h_{L,j-1}$, temperatura T_{j-1} e pressão P_{j-1} , sendo esta última igual ou menor que a pressão do estágio j . A pressão do líquido proveniente do estágio $j-1$ aumenta adiabaticamente (Seader et al., 2011). Esse mesmo processo ocorre com o vapor do estágio $j+1$ abaixo do estágio j .

Dessa maneira, é factível estabelecer expressões para o balanço de massa, o balanço de energia e o equilíbrio de fases em cada estágio. As respectivas formulações são apresentadas nas Equações 25 à 29.

$$M_{i,j} = L_{j-1}x_{i,j-1} + V_{j+1}y_{i,j-1} + F_j z_{i,j} - (L_j + U_j)x_{i,j} - (V_j + W_j)y_{i,j} - Q_j = 0 \quad (25)$$

$$E_{i,j} = y_{i,j} - K_{i,j}x_{i,j} = 0 \quad (26)$$

$$(S_y)_j = \sum_{i=1}^C y_{i,j} - 1.0 = 0 \quad (27)$$

$$(S_x)_j = \sum_{i=1}^C x_{i,j} - 1.0 = 0 \quad (28)$$

$$H_j = L_{j-1}H_{j-1} + V_{j+1}h_{V,j-1} + F_j h_{Fj} - (L_j + U_j)h_{L,j} - (V_j + W_j)h_{V,j} - Q_j = 0 \quad (29)$$

onde $M_{i,j}$ é o balanço de massa para cada componente, $E_{i,j}$ a relação de equilíbrio de fases para cada componente, S é o somatório das frações molares e H é o balanço de energia.

2.9 CONTROLE DE PROCESSOS EM COLUNAS DE DESTILAÇÃO

As modernas instalações de processamento nas indústrias química, petrolífera, alimentícia e farmacêutica encontram-se altamente automatizadas para atender aos padrões contemporâneos de segurança, produtividade, qualidade e preservação ambiental. Atualmente, os operadores dessas instalações, principalmente na produção química, desempenham um papel mais orientado à intervenção por exceção, agindo apenas quando o sistema automatizado encontra situações não previamente programadas. O avanço e a acessibilidade aprimorada de computadores, sensores e sistemas de automação nas últimas décadas levaram a um significativo investimento das indústrias de processos na automação (Meadowcrof, 2022).

O controle de um processo envolve a regulação de variáveis para mantê-las em um valor desejado (*setpoint*) ou próximo a ele. Essa técnica é realizada por meio de um controlador de processo, que age em resposta ao erro entre o valor desejado (*setpoint*) e o valor medido pelo sensor (medidor) (Simonelli et al, 2017).

A destilação azeotrópica é altamente sensível a perturbações, podendo resultar na degradação de produtos ou dificultar a retomada da operação normal. A implementação de técnicas avançadas de controle é frequentemente adotada para minimizar as operações transitórias. O monitoramento e diagnóstico de processos tornaram-se cruciais na indústria química devido à crescente complexidade das operações e do *controle multiloop* (Dantas e Brito, 2009).

A análise dinâmica de um processo implica no estudo das mudanças que ocorrem ao longo do tempo nas variáveis de interesse, seja em processos químicos ou outros. Essa análise concentra-se nos aspectos temporais, como taxas de variação, transientes e resposta do sistema a perturbações

Alguns dos principais objetivos da análise dinâmica incluem:

1. Avaliar a estabilidade do sistema para determinar sua capacidade de retornar a um comportamento desejado após perturbações.
2. Analisar as mudanças nas variáveis do processo em resposta a alterações nas condições operacionais.
3. Projetar estratégias de controle com base na análise dinâmica para manter o processo dentro dos limites desejados.

O design de um sistema de controle de processos visa minimizar as perdas econômicas decorrentes de perturbações ao processo. O sistema de controle tem

como objetivo eliminar desvios entre a variável controlada e o valor desejado (setpoint – SP), mantendo-se estável, mesmo diante das não linearidades do processo (Campos; Teixeira, 2006).

De acordo com Campos e Teixeira (2006), o projeto de um sistema de controle pode ser dividido em três etapas:

1. Projeto básico: Análise do processo e definição de uma estratégia de controle, incluindo a instrumentação necessária.

2. Detalhamento e implantação: Cuidados com a instalação da instrumentação e configuração do controle no sistema digital.

3. Fase de operação e manutenção: Ajuste dos controladores e possíveis alterações em função de mudanças operacionais

Os controladores mais amplamente utilizados industrialmente são os controladores de feedback, especialmente os Proporcionais-Integrativos-Derivativos (PID), que calculam o erro entre o valor de referência (setpoint) e a variável controlada, emitindo um sinal para corrigir a diferença calculada. Vale ressaltar que os módulos do PID podem ser utilizados separadamente, como controladores proporcionais (P) e controladores proporcionais e integrais (PI) (Seaborg, 2010).

Nos controladores proporcionais, a saída do sinal é proporcional ao erro calculado, conforme representado pela Equação 27 (Seaborg, 2010).

$$p(t) = p + K_c e(t) \quad (27)$$

onde $p(t)$ é a saída do controlador, p é o valor de polarização em estado estacionário, K o ganho do controlador, $e(t)$ o erro associado e τ é um parâmetro chamado de tempo integral.

Nos controladores proporcionais integrais, o sinal de saída é dependente da integral do erro do sinal em relação ao tempo, conforme a Equação 28 (Seaborg, 2010).

$$p(t) = p + \frac{1}{\tau} \int e(t^*) dt^* \quad (28)$$

Os controladores derivativos atuam antecipando futuros erros que possam ocorrer no sistema através da taxa de mudanças, como observado na equação 29 (Seaborg, 2010).

$$p(t) = p + \tau \frac{de(t)}{dt} \quad (29)$$

Numa coluna de destilação é comum ter variáveis controladas e variáveis manipuláveis, a Tabela 2 apresenta algumas destas variáveis.

Tabela 2 – Variáveis de interesse em uma coluna de destilação

Variáveis Controladas	Variáveis Manipuladas
Temperatura do prato sensível	Carga térmica do refeedor
Composição do destilado	Vazão do produto de topo
Pressão da coluna ou prato	Vazão do fluido refrigerante
Composição do produto de fundo	Vazão do produto de fundo
Nível do vaso de refluxo	Razão de refluxo
Nível do refeedor	Vazão de vapor de aquecimento

Fonte: Adaptado de Santos, 2019.

Pensando nas estratégias de controle que utilizam a temperatura do prato sensível para inferir a composição do produto, é necessário selecionar corretamente o prato ou o controle não será efetivo. Deve-se lembrar que, devem ser identificados 2 pratos sensíveis: um na seção de retificação e outro na seção de esgotamento. Diversos estudos propõem diferentes métodos para determinar o prato sensível (Santos, 2019). O mais simples deles, é o Método da Inclinação dos Pratos, proposto por Luyben (2013). Neste método é calculada a diferença entre dois pratos seguintes por toda coluna. Os pratos que apresentarem maiores diferenças, são os pratos sensíveis.

2.10 ANÁLISE MULTIVARIADA E CONTROLE PREDITIVO DE PROCESSOS COM BASE EM MODELOS DE MACHINE LEARNING

A implementação de sistemas de controle eficazes é essencial para garantir a conformidade rigorosa das especificações de composição em colunas de destilação. Convencionalmente, controladores PID são empregados nesses contextos para supervisionar a pureza dos produtos de base e topo. Contudo, dada a complexidade intrínseca, a natureza não linear e multivariável dos sistemas de destilação pode demandar estratégias mais avançadas para garantir um bom desempenho (Júnior et al., 2023).

Avanços significativos em modelagem, simulação e experimentação ao longo das décadas proporcionaram à comunidade de engenharia química uma vasta quantidade de dados. Essa riqueza de dados viabiliza a utilização de modelos de *machine learning* (ML) para prever variáveis químicas com eficácia (MAARTEN, et al., 2021). Os modelos de ML, sendo modelos matemáticos estatísticos, têm a capacidade de identificar padrões nos dados sem a necessidade de programação explícita baseada em regras (MAARTEN, et al., 2021).

Análise de dados multivariados (ADM) é a investigação de muitas variáveis simultaneamente, a fim de compreender as relações que podem existir entre elas. Métodos de análise de dados multivariados existem há décadas, porém, até recentemente, foram utilizados principalmente em laboratórios e grupos técnicos especializados, raramente sendo aplicados a processos de produção (Bray, 2013).

A maioria dos processos de fabricação química é altamente multivariada devido às complexas reações envolvidas, ou seja, há um grande número de variáveis que são tipicamente muito interativas. Sistemas complexos exigem múltiplas medições para serem totalmente compreendidos (Bray, 2013).

No entanto, as ferramentas de Controle Estatístico de Processo (CEP) utilizadas em engenharia química ainda dependem em grande parte de métodos univariados (ou seja, uma variável por vez), que não apresentam o quadro completo de processos químicos complexos, apesar da coleta massiva de dados por meio de instrumentos e sistemas de controle. Essas ferramentas de CEP utilizam abordagens estatísticas tradicionais, como média, desvio padrão e teste t de Student, que analisam apenas variáveis individuais isoladamente (Bray, 2013).

Enquanto estatísticas univariadas podem ser úteis para investigar sistemas simples, elas tendem a falhar quando sistemas mais complexos são analisados. Isso ocorre porque elas não conseguem detectar relações que podem existir entre as variáveis em estudo, tratando todas como independentes umas das outras (Bray, 2013).

As principais técnicas multivariadas incluem Análise Exploratória de Dados, métodos de Regressão/Predição e métodos de Classificação (Bray, 2013).

A Análise Exploratória de Dados (AED) busca encontrar a estrutura oculta ou padrões subjacentes em conjuntos de dados grandes e complexos. Isso proporciona uma compreensão melhor do processo e pode levar a *insights* que não seriam observados de outra forma. Métodos de AED incluem análise de agrupamento (*Cluster analysis*) e Análise de Componentes Principais (PCA). Um exemplo de aplicação da análise exploratória de dados é verificar a presença de contaminantes em um processo ou matéria-prima, ou identificar subprodutos causados por configurações de processo incorretas (Bray, 2013).

A análise de regressão envolve o desenvolvimento de um modelo a partir de dados disponíveis para prever uma resposta desejada ou respostas para medições futuras. A regressão multivariada é uma extensão do caso simples do modelo linear univariado, em que existem muitas variáveis independentes e pelo menos uma variável dependente. Métodos de regressão incluem Regressão Linear Múltipla (MLR), Regressão de Componentes Principais (PCR) e Regressão de Mínimos Quadrados Parciais (PLSR). Aplicações comuns incluem prever pureza, rendimento ou qualidade do produto final a partir da qualidade dos materiais brutos de entrada (Bray, 2013).

A classificação é a separação (ou ordenação) de um grupo de objetos em uma ou mais classes com base em características distintivas nos objetos. Métodos de classificação incluem Análise Discriminante Linear (LDA) e Classificação por Máquinas de Vetores de Suporte (SVM-C). Exemplos de aplicações incluem agrupar produtos de acordo com características semelhantes ou categorias de qualidade (Bray, 2013).

A análise de dados multivariados é altamente visual por natureza. Em vez de simplesmente exibir estatísticas, ela mostra dados plotados de várias formas, tornando os padrões mais fáceis de serem identificados e auxiliando na interpretação (Bray, 2013).

Por outro lado, os gráficos de controle univariados tradicionais mostram muitas variáveis diferentes simultaneamente, tornando extremamente difícil obter uma imagem clara e completa. Os gráficos de controle multivariados condensam todos esses dados em um ou dois gráficos, levando em consideração as interações complexas entre as variáveis. Se o processo começar a se desviar, é possível "aprofundar" em amostras ou valores atípicos específicos para identificar rapidamente a causa raiz do problema, com base em uma combinação de diagnósticos multivariados e univariados (Bray, 2013).

A abordagem ML é sustentada por três pilares essenciais: os dados, a representação desses dados no modelo e a estratégia de modelagem (Maarten, et al., 2021). Praticamente qualquer conjunto de dados contendo resultados de experimentos, cálculos de primeiros princípios ou modelos de simulação complexos pode ser empregado para treinar um modelo (Maarten, et al., 2021).

Os dados de entrada consistem em parâmetros fundamentais do processo de interesse, como pressão, temperatura, tempo de residência e cargas de alimentação, composição da carga (Maarten, et al., 2021).

Além disso, a escolha adequada de modelos de aprendizagem de máquinas é crucial para a eficácia do sistema. Essa seleção pode ser realizada por meio de medidas de desempenho específicas para o modelo em questão.

Os métodos de regressão são empregados quando se busca descrever a relação entre uma variável de resposta e uma ou mais variáveis explicativas em uma análise de dados (Silveira, 2021). Através de uma análise de regressão, é possível explorar tanto a direção (positiva ou negativa) quanto a magnitude (fraca ou forte) da associação entre a variável dependente (Y) e a variável independente (X). Também permite a previsão dos valores da variável dependente com base na variável independente (Figueira, 2006).

A regressão logística difere dos modelos de regressão linear pelo fato de a variável dependente ser qualitativa e binária. Na regressão logística, a variável resposta assume apenas valores 0 e 1, sendo geralmente '1' para a ocorrência do evento de interesse e '0' para sua ausência, ou seja, '1' corresponde ao sucesso e '0' ao fracasso. Portanto, o valor previsto de Y sempre estará no intervalo $0 \leq Y \leq 1$. Neste modelo, a curva logística é ajustada aos dados, permitindo o cálculo da probabilidade de ocorrência de um evento de interesse (Silveira, 2021).

Nos modelos de regressão linear o método mais utilizado para a estimação dos parâmetros é o de mínimos quadrados. Neste método, os coeficientes são estimados pelos valores que minimizam a soma das diferenças quadradas entre os valores observados e os valores preditos. Entretanto, devido à ausência de relação linear em regressão logística, o método de mínimos quadrados não é apropriado para estimar os coeficientes, de modo que, utiliza-se o método da máxima verossimilhança (Silveira, 2021).

A Árvore de Decisão (*Decision Tree-DT*) é uma das ferramentas utilizadas para fundamentar a tomada de decisões em diversos setores industriais (Sishi et al, 2021). A DT utiliza uma abordagem de modelagem em formato de árvore para percorrer os possíveis resultados de uma sequência de escolhas relacionadas (Yan et al., 2016). As DTs também são empregadas como ferramentas de Mineração de Dados (DM) para classificar e prever variáveis de classe (Sishi et al, 2021). Elas constroem modelos de classificação ou regressão utilizando uma estrutura de árvore. Uma árvore de classificação e uma árvore de regressão são utilizadas para variáveis-alvo discretas e contínuas, respectivamente (De Oña et al., 2014; Yan et al., 2016).

Um conjunto de dados é particionado em subgrupos menores que gradualmente se desenvolvem em uma árvore associada com nós de decisão e nós folha (Sishi et al, 2021). Um nó de decisão abrange duas ou mais subdivisões, cada uma representando valores para o elemento testado. Por outro lado, um nó folha denota uma decisão no final aritmético (Sishi et al, 2021). O nó localizado no topo da árvore, considerado o melhor preditor, é chamado de nó raiz (Zhao et al., 2014). As DTs podem acomodar dados estatísticos e categóricos (Candanedo and Feldheim, 2016).

As DTs são elogiadas em relação a outras técnicas de modelagem, principalmente, devido à vantagem de poderem ser apresentadas por meio de imagens bifurcadas visuais, que claramente representam regras *If-Then* eficazes (Yan et al., 2016). Cada folha em uma árvore de decisão representa uma regra de decisão que abrange metadados (Sishi et al, 2021).

Já um classificador/regressor de floresta aleatória (*Random Forest-RF*) é um classificador de conjunto que produz múltiplas árvores de decisão, usando um subconjunto selecionado aleatoriamente de amostras e variáveis de treinamento. As árvores são criadas desenhando um subconjunto de amostras de treinamento por meio de substituição (uma abordagem de *bagging*). Isto significa que a mesma

amostra pode ser selecionada diversas vezes, enquanto outras podem não ser selecionadas. Cerca de dois terços das amostras são usadas para treinar as árvores e o terço restante é usado em um cruzamento interno-técnica de validação para estimar o desempenho do modelo de RF resultante (Belgiu, 2016).

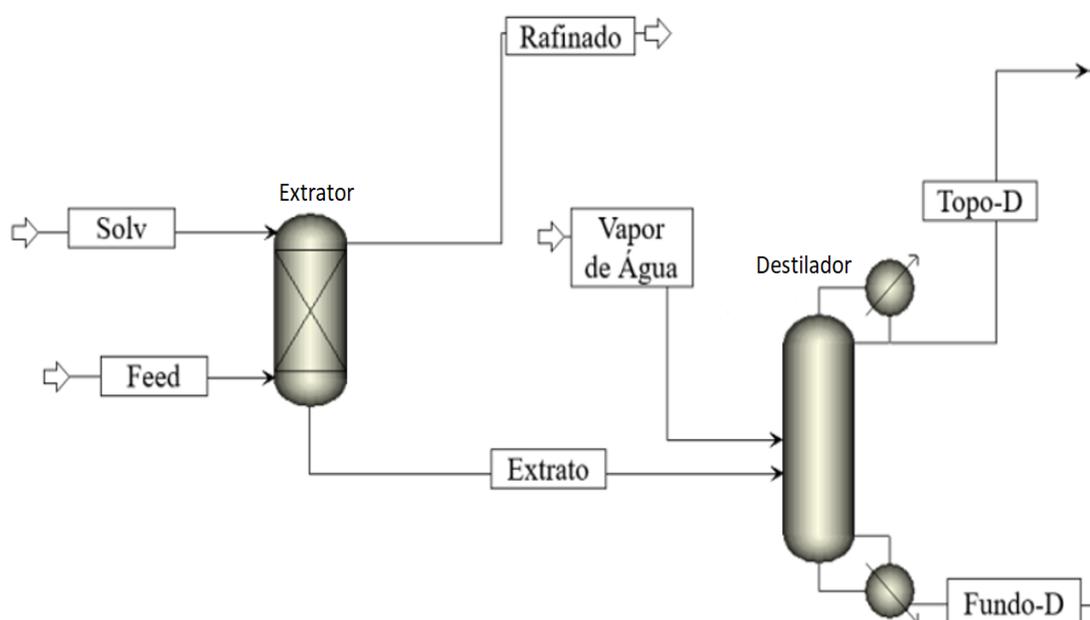
Uma métrica comum de desempenho para problemas de regressão é a Raiz do Erro Quadrático Médio (MSE, em inglês), fornecendo uma avaliação da quantidade de erros nas previsões do sistema, com ênfase nos erros mais significativos (GÉRON, 2019). O Erro Médio Absoluto (MAE, em inglês) é outra métrica útil que calcula o erro absoluto médio entre os valores observados e as previsões, oferecendo uma medida robusta do desempenho do modelo (GÉRON, 2019).

3 METODOLOGIA

3.1 SIMULAÇÃO EM ESTADO ESTACIONÁRIO

A metodologia empregada neste estudo consistiu na simulação de uma planta de separação de hidrocarbonetos e recuperação de solvente, utilizando o software Aspen Plus® e o modelo termodinâmico NRTL, com o ajuste dos parâmetros binários. Esse ajuste foi realizado fixando o terceiro parâmetro, α_{ij} , em -1, conforme descrito por Fernandez (2017), utilizando os dados disponibilizados por Mesquita et al. (2015) e Rappel (2002). A Figura 8 ilustra a configuração da planta simulada, composta por uma coluna de extração contendo 28 estágios. Nessa coluna, o sulfolano foi escolhido como solvente e adicionado no estágio inicial, enquanto a carga de alimentação foi introduzida no estágio final, conforme proposto por Fernandez (2017).

Figura 8 –Esquema do processo de separação de hidrocarbonetos.



Fonte: Adaptado de Silva (2016), Fernandez (2017) e Lima (2019).

A composição detalhada da carga utilizada na alimentação da coluna extratora é descrita na Tabela 3. Esta composição representa um sistema complexo que inclui hidrocarbonetos e sulfolano, conforme documentado por Rappel et al. (2002), Mesquita (2015), Fernandez (2017) e Lima (2019).

Tabela 3 – Composição da carga de alimentação da coluna extratora.

Composto	Vazão mássica (kg/h)	%, em massa	Ponto de ebulição (°C)
Sulfolane	80000	-	285,0
N-hexano	4278,08	9	68,0
N-heptano	1543,86	3	98,4
N-octano	674,64	1	125,6
Benzeno	28999,5	58	110,6
Tolueno	14628,07	29	80,1

Fonte: Adaptado de Fernández (2017).

Dado o propósito deste estudo na análise da coluna de destilação, especificamente voltada para a recuperação do solvente, optou-se por utilizar o extrato proveniente da torre de extração como matéria-prima para alimentar a coluna recuperadora, após a simulação individual desta coluna. A composição da alimentação da coluna recuperadora de solvente está representada na Tabela 4.

Tabela 4 – Composição da carga de alimentação da coluna recuperadora de solvente em estudo.

Composto	Vazão mássica (kg/h)	%, em massa	Ponto de ebulição (°C)
Sulfolane	80000	64,5	285,0
N-hexano	0	-	68,0
N-heptano	780,641	0,63	98,4
N-octano	0	-	125,6

Benzeno	28999,5	23,4	110,6
Tolueno	14174,7	11,4	80,1
Água	12000	-	100

Fonte: O autor (2023).

A metodologia empregada neste estudo envolveu uma abordagem diferenciada para operar a coluna recuperadora de solvente, em contraste a metodologia descrita por Fernandez (2017) e Lima (2019). Em particular, foram consideradas condições de pressão distintas de 1 bar, conforme sugerido por Wang et al. (2019), resultando em pressões operacionais de 0,75 bar no refeedor e 0,5 bar no condensador. Vale ressaltar que com essa abordagem, o número de pratos da coluna caiu de 24 para 12.

Inicialmente, para determinar a razão de refluxo e o calor consumido pelo refeedor, utilizou-se uma coluna do bloco DSTWU (método de FUG) no software Aspen Plus® devido à sua menor complexidade. A carga de alimentação, as pressões operacionais e as características dos componentes presentes, como o benzeno (componente de menor polaridade) e o sulfolane (componente de alta polaridade), foram consideradas para estimar a razão de refluxo mínima e o número mínimo de estágios necessários para o processo.

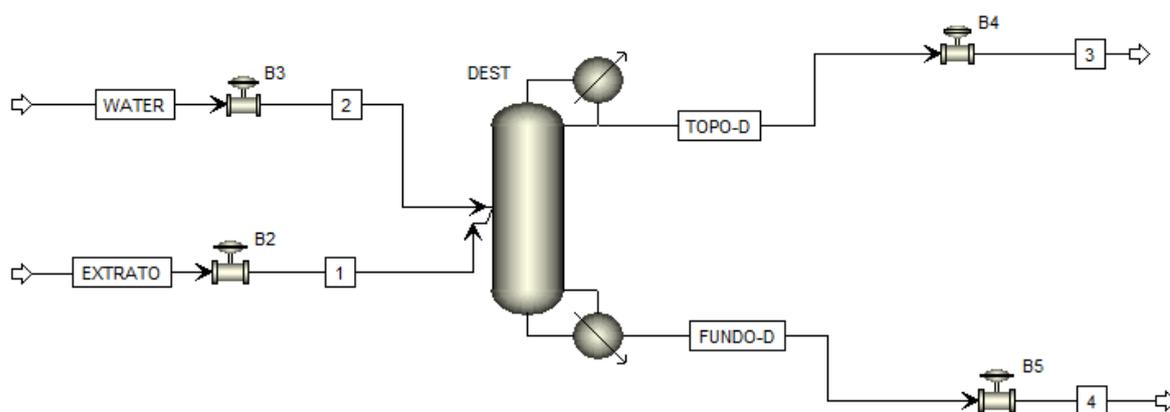
Através da análise dos resultados da coluna simplificada e com o objetivo de obter uma simulação mais precisa, a coluna RadFrac foi integrada ao software Aspen Plus®. Com esta ferramenta, tornou-se viável realizar análises detalhadas e otimizações para a coluna de recuperação de solvente.

3.2 SIMULAÇÃO DINÂMICA

Após a conclusão da simulação em estado estacionário (E.E.), foi habilitado o modo dinâmico no software Aspen Plus® para a continuar da análise. Esta etapa permitiu o dimensionamento do vaso de refluxo, da base da coluna (sump) e a análise da hidráulica correspondente. O dimensionamento dos vasos envolveu a consideração de parâmetros como comprimento e diâmetro, fundamentados na vazão de líquido no primeiro estágio da coluna de destilação, fornecida pelo Aspen, utilizando um tempo de residência de 10 minutos e uma razão L/D igual a 2.

No contexto da simulação dinâmica, foram incorporadas válvulas para simular a perda de carga nas correntes do sistema e para configurar os controladores. Essas válvulas desempenham um papel crucial na dinâmica do sistema, recebendo dados das variáveis de resposta e atuando nas variáveis manipuláveis. A representação gráfica da coluna de destilação no módulo dinâmico consta na Figura 9.

Figura 9– Representação da coluna de destilação usada no Aspen Dinâmico.



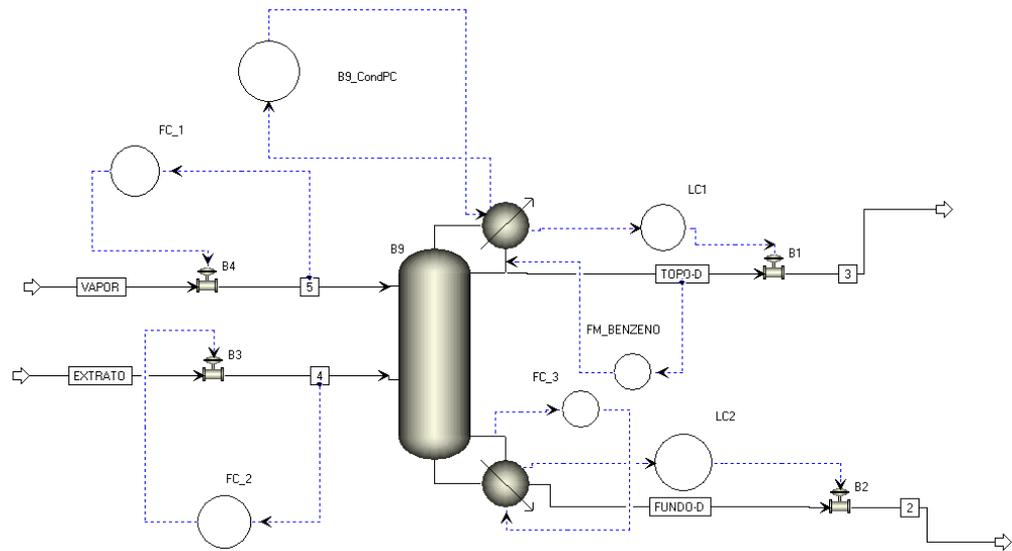
Fonte: O autor (2023).

A pesquisa propôs avaliar a eficácia da estratégia de controle da pureza do produto de topo em um sistema dinâmico complexo de destilação. Para isso, foram empregados controladores do tipo PID para regular as variáveis de fundo, topo, pressão e vazão de alimentação, com foco na composição do benzeno no produto de topo como variável de interesse principal.

É importante ressaltar que as válvulas de alimentação desempenham uma função adicional, mantendo a pressão das correntes no mesmo nível do prato. Isso se deve ao fato de que a dinâmica da pressão é simulada com maior precisão no Aspen Dinâmico, o que torna essas válvulas essenciais para a estabilidade do processo.

A estratégia de controle associada à composição dos produtos é ilustrada na Figura 10, enquanto a Tabela 5 fornece uma descrição detalhada das funções específicas de cada controlador. Essas abordagens foram cruciais para investigar a dinâmica do processo e avaliar a efetividade do controle nas condições dinâmicas simuladas.

Figura 10 – Estratégia convencional adotada para o controle da composição.



Fonte: O autor (2023).

Tabela 5 – Funções dos controladores na coluna de destilação.

Controlador	Função
FC_1	Controla a vazão mássica de vapor de água
FC_2	Controla a vazão mássica do extrato
LC_1	Controla o nível do condensador
LC_2	Controle do nível do rebovedor
FM_BENZENO	Controle da composição molar do benzeno na corrente de topo
FC_3	Controle da vazão de vapor do rebovedor
B9_CondPC	Controle da carga térmica do rebovedor

Fonte: O autor (2023).

Os parâmetros de cada um dos controladores foram otimizados através da função *Tuning parameters* disponível no Aspen Dinâmico.

3.3 IMPLEMENTAÇÃO DE SENSORES PARA GERAÇÃO DE DADOS

No ambiente do Aspen Dinâmico, foram implementados sensores com o objetivo de medir vazões, pressões, temperaturas, composições e níveis de líquidos ao longo da coluna de destilação. Esses sensores desempenham um papel fundamental na obtenção de dados em tempo real para o controle e monitoramento do processo dinâmico.

A Tabela 6 oferece uma descrição detalhada de cada sensor inserido dentro do Aspen Dinâmico e sua respectiva função. As correntes mencionadas nesta tabela estão relacionadas à Figura 11, que ilustra a estratégia de controle associada à composição dos produtos. Esses sensores são essenciais para garantir a precisão das medidas e fornecer informações valiosas para a eficácia do controle dinâmico do sistema de destilação.

Tabela 6 – Sensores distribuídos ao longo da coluna de destilação.

Sensor	Variável monitorada
T_5Output_(°C)	Temperatura na corrente 5
P_5Output_(bar)	Pressão na corrente 5
TMF_5Output_(kg/h)	Vazão de alimentação
T_4Output_(C)	Temperatura na corrente 4
P_4Output_(bar)	Pressão na corrente 4
TMF_4Output_(kg/h)	Vazão na corrente 4
RROutput_	Razão de refluxo
LL_CONDOOutput_(m)	Nível de líquido no condensador
P_CONDOOutput_(bar)	Pressão no condensador
T_CONDOOutput_(°C)	Temperatura no condensador
T_11Output_(°C)	Temperatura do prato sensível (retificação)

P_11Output_(bar)	Pressão do prato sensível (retificação)
LL_11Output_(m)	Nível do prato sensível
REB_V.PFLOWOutput_(kmol/h)	Vapor empregado no refeedor
T_2Output_(C)	Temperatura na corrente 2
P_2Output_(bar)	Pressão na corrente 2
SUMPLEVELOutput_(m)	Nível do vaso de refluxo
LL_2Output_(m)	Nível de líquido no refeedor
FM_TOLUEOutput_(kg/kg)	Fração mássica de Tolueno
FM_HEPTOutput_(kg/kg)	Fração mássica de Heptano
FM_BENZOutput_(kg/kg)	Fração mássica de Benzeno
TFM_3Output_(kg/h)	Vazão na corrente 3
duty_condOutput_(Gj/h)	Calor latente de condensação
T_REFOutput_(°C)	Temperatura no refeedor
P_REFOutput_(bar)	Pressão no refeedor

Fonte: O autor 2024.

Os dados coletados pelos sensores foram obtidos através da aplicação de perturbações nas cargas de alimentação da coluna recuperadora de solvente, isto é, na vazão do vapor de água e na vazão e composição do extrato. Essas perturbações consideraram uma variação máxima de 20% nos valores padrão das variáveis.

3.4 TRATAMENTO DE DADOS POR MODELOS DE MACHINE LEARNING

A implementação de um algoritmo em Python, no ambiente Google Colab, utilizando o *Google Sheets* como base de dados, foi realizada com o propósito de conduzir uma análise exploratória, tratamento de dados e aplicação de modelos de

regressão linear multivariada, árvore de decisão, floresta aleatória e regressão logística.

Durante a análise exploratória, identificaram-se dados faltantes e médias invariantes ao longo do processo. As correlações entre as variáveis do processo foram examinadas por meio de matrizes de confusão, utilizando os métodos de Spearman, Kendall e Pearson.

Na implementação dos algoritmos de aprendizado de máquina, a base de dados foi dividida em um conjunto de teste, representando 70% da base original, e um conjunto de treinamento, correspondendo a 30% (Lopes e Lucena, 2023). As previsões foram realizadas pelos quatro modelos, utilizando os conjuntos de treinamento e teste como entrada. Os resultados obtidos compreenderam as saídas dos sensores e as frações molares previstas pelos modelos (Lopes e Lucena, 2023).

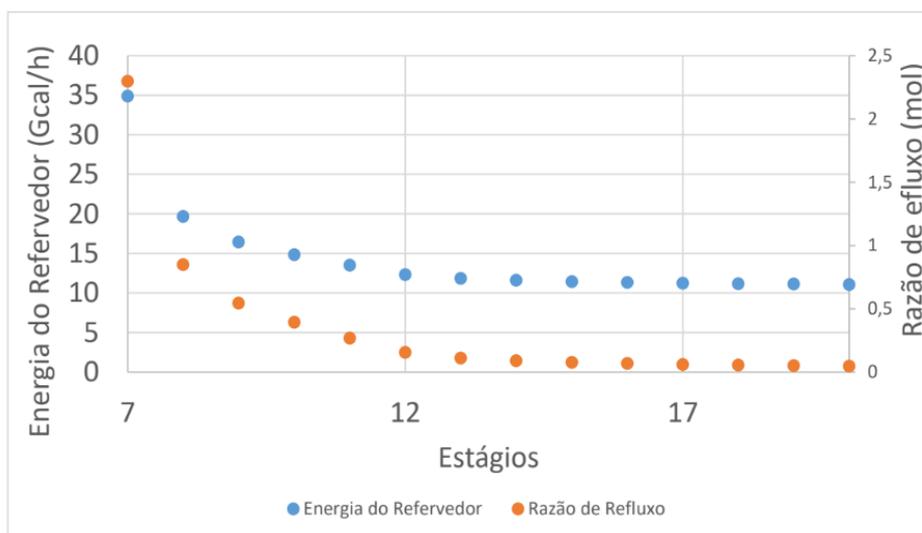
O algoritmo utilizado para todo o tratamento de dados está disponível no apêndice A.

4 RESULTADOS E DISCUSSÃO

4.1 SIMULAÇÃO EM ESTADO ESTACIONÁRIO

Com base nos resultados obtidos da coluna simplificada, gerou-se um gráfico que relacionou a energia consumida, a razão de refluxo e o número de estágios na coluna (Figura 11). O objetivo era determinar as melhores condições de operação para a recuperação eficiente do solvente. A análise desses dados levou à escolha da operação da coluna com 12 estágios, consumindo 12,322 Gcal/h de energia, com uma razão de refluxo de 0,1555/mol, e o sétimo prato foi selecionado como prato de alimentação. Essas escolhas foram fundamentadas nos resultados obtidos e nas considerações específicas do sistema estudado para posteriormente serem utilizadas na modelagem rigorosa da coluna de destilação (RadFrac).

Figura 11 –Razão de refluxo e energia do refervedor.



Fonte: O autor (2023).

Notou-se que, no módulo RadFrac, ao manter o prato de alimentação constante e aumentar a energia requerida pelo refervedor para 12,4 Gcal/h, juntamente com uma razão de refluxo de 0,156, toda a água inserida na coluna era expelida pela corrente de topo, resultando na recuperação do solvente com uma pureza de 99,99% na base. Entretanto, surgiu uma preocupação com a faixa de temperatura operacional, que

poderia propiciar a degradação do sulfolane, tornando inviável a condução do processo.

Para mitigar o risco de degradação do solvente, foram realizados ajustes na energia do refeedor e na razão de refluxo, visando alcançar o melhor projeto. Por meio desses ajustes, foi possível atingir uma razão de refluxo de 0,9 e uma carga térmica no refeedor de 9,42 Gcal/h, mantendo constante o número de estágios de separação. Além disso, observou-se que, com a otimização da razão de refluxo, a vazão de água utilizada poderia ser reduzida. Os resultados da otimização da vazão de vapor de água obtidos estão apresentados na Tabela 5.

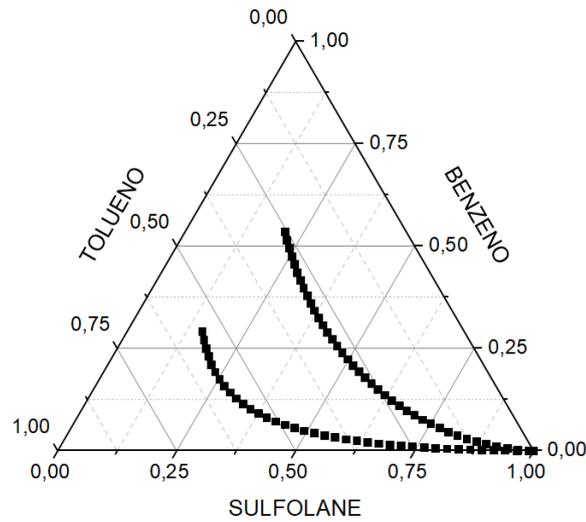
Tabela 7 – Composição da carga de alimentação da coluna recuperadora de solvente.

Composto	Vazão mássica (kg/h)	%, em massa	Ponto de ebulição (°C)
Sulfolane	80000	64,5	285,0
N-hexano	0	-	68,0
N-heptano	780,641	0,7	98,4
N-octano	0	-	125,6
Benzeno	28999,5	23,4	110,6
Tolueno	14174,7	11,4	80,1
Água	2600	-	100

Fonte: O autor (2023).

De acordo com as observações delineadas no mapa de curvas residuais presente na Figura 12, percebe-se que a destilação simples surge como uma opção termodinamicamente viável para a separação da mistura. No entanto, como discutido anteriormente, este método requer uma carga térmica substancial no refeedor, além de um elevado número de estágios na coluna, o que resulta em custos operacionais significativos.

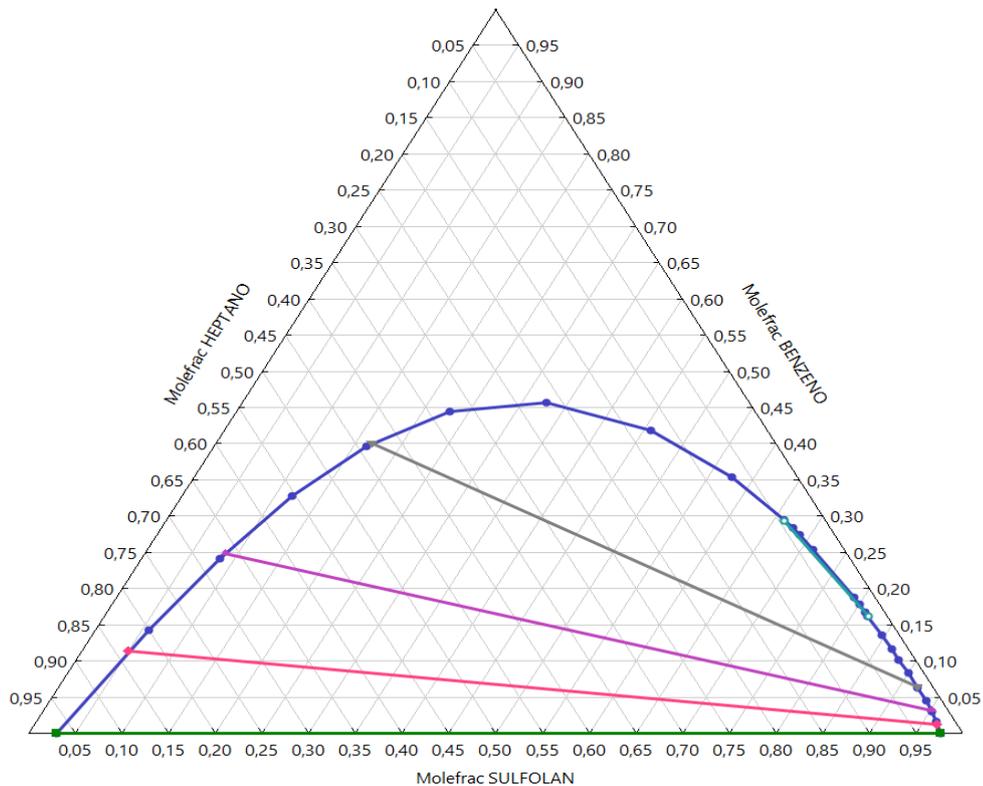
Figura 12 – Mapa de curva residual para mistura benzeno, tolueno e sulfolane.



Fonte: O autor (2023).

O consumo de energia está diretamente relacionado à predominância de uma fase homogênea na coluna de destilação, como indicado no diagrama ternário da Figura 13 e Tabela 5, onde observa-se a composição situada na região homogênea do diagrama.

Figura 13 – Diagrama ternário para o sulfolane, heptano e benzeno.

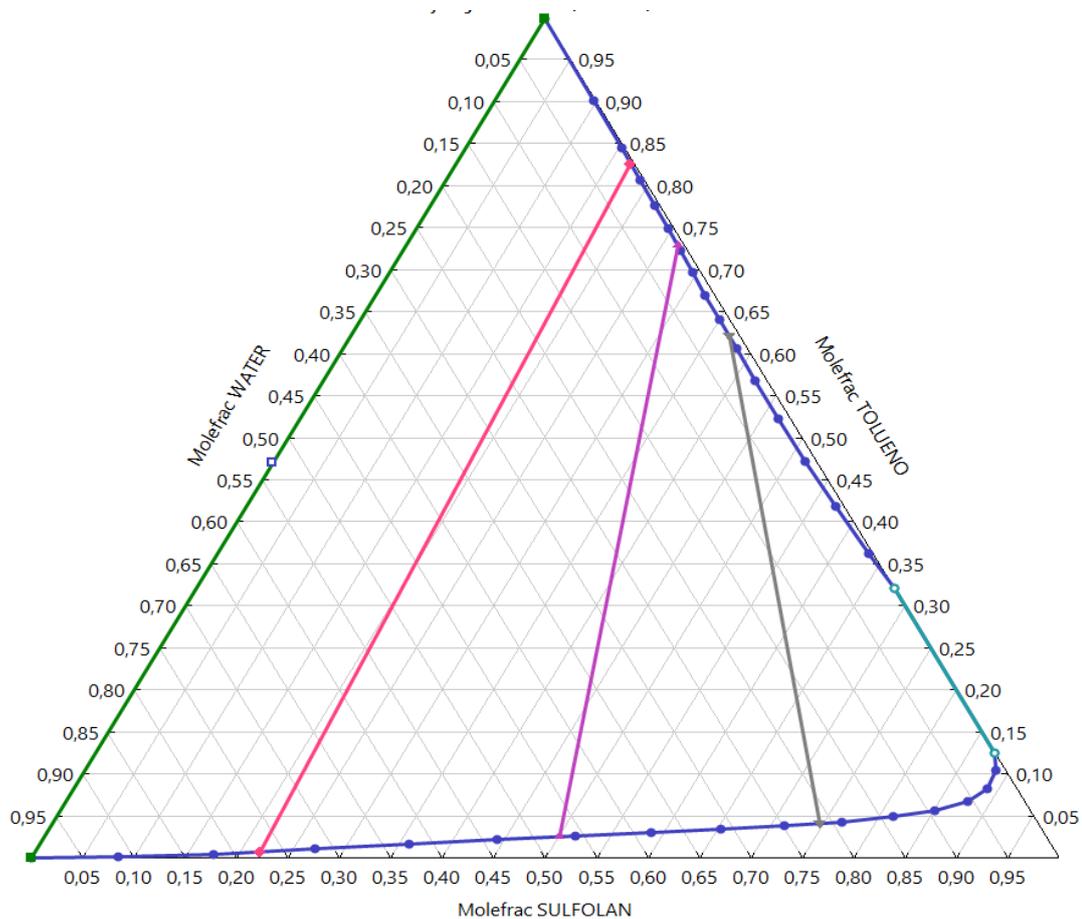


Fonte: O autor (2023).

A introdução controlada de vapor de água na coluna desempenhou um papel crucial no processo, atuando como um *entrainer*, pois quebrou as interações entre as moléculas de sulfolane (polar) e os hidrocarbonetos (apolares), facilitando a recuperação do solvente. Isso ocorreu devido à maior afinidade entre os compostos polares favorecendo sua separação durante o processo de recuperação do solvente.

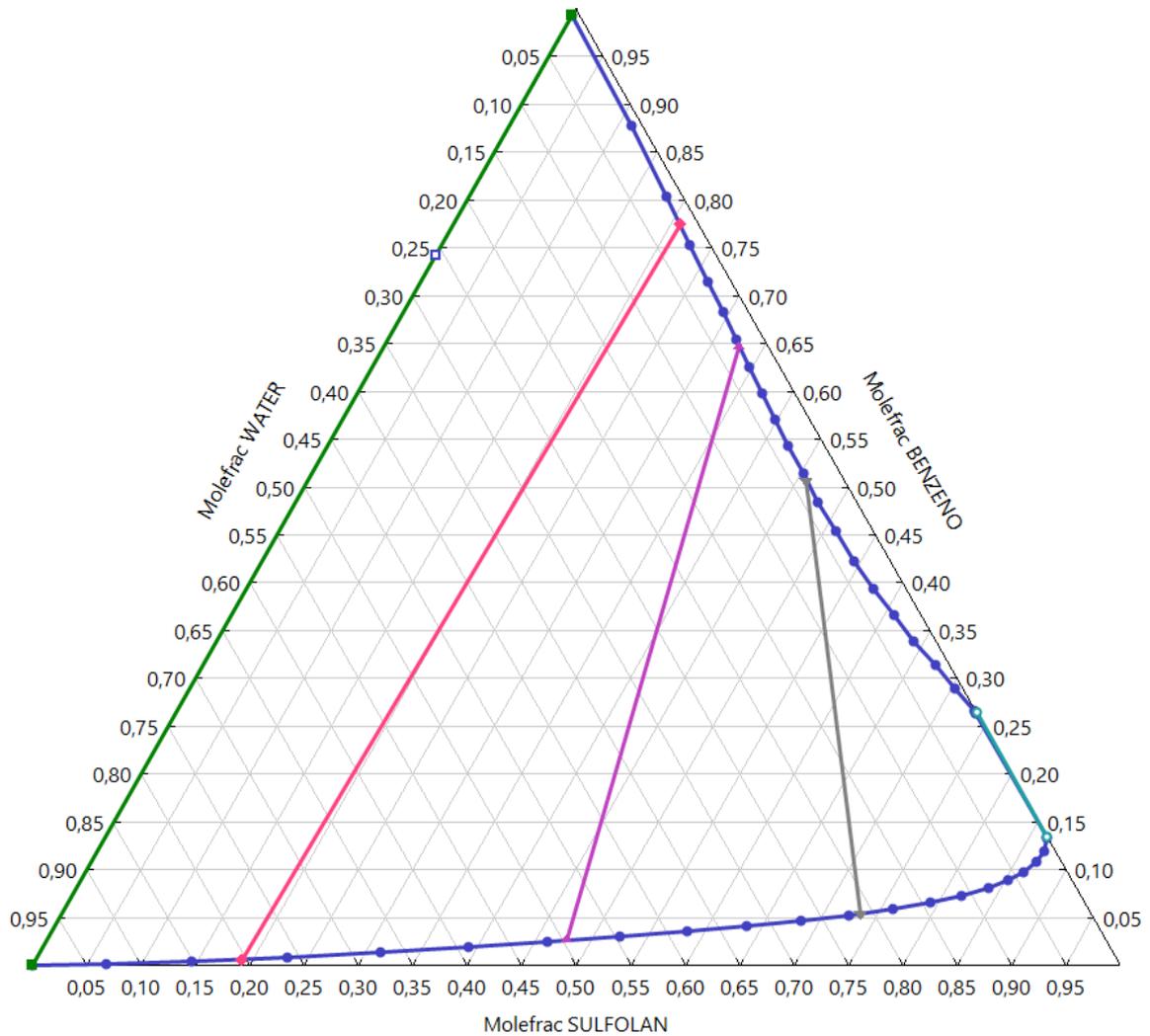
As Figuras 14 e 15 apresentam os diagramas ternários dos compostos presentes na coluna recuperadora de solvente. Através delas, evidencia-se que a incorporação de vapor saturado de água no sistema promove a formação de um sistema heterogêneo na coluna. Esse sistema compreende uma fase polar (água/sulfolane) e uma fase apolar (água/hidrocarbonetos). Portanto, dentro da própria coluna, ocorre uma separação parcial sem a necessidade do emprego de energia.

Figura 14 – Diagrama ternário: água, tolueno e sulfolane.



Fonte: O autor (2023).

Figura 15 – Diagrama ternário: água, benzeno e sulfolano.

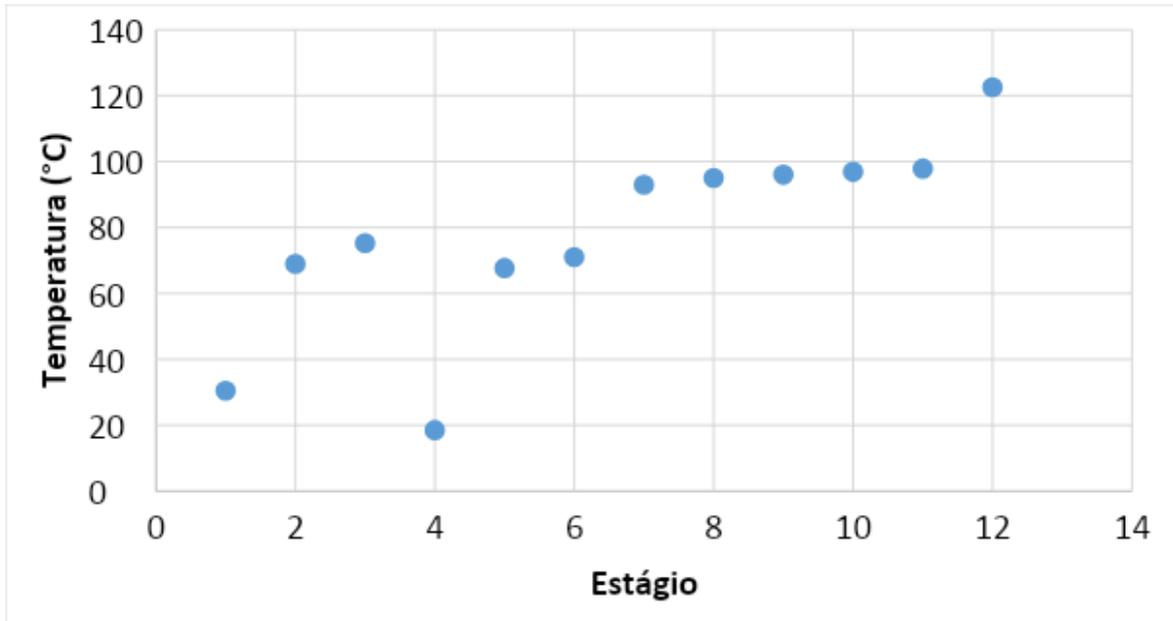


Fonte: O autor (2023).

A Figura 16 exibe o perfil de temperatura ao longo da coluna de destilação, destacando-se por uma diminuição na temperatura observada no 4º estágio. Este fenômeno sugere uma variação significativa nas condições termodinâmicas da coluna nesse ponto específico.

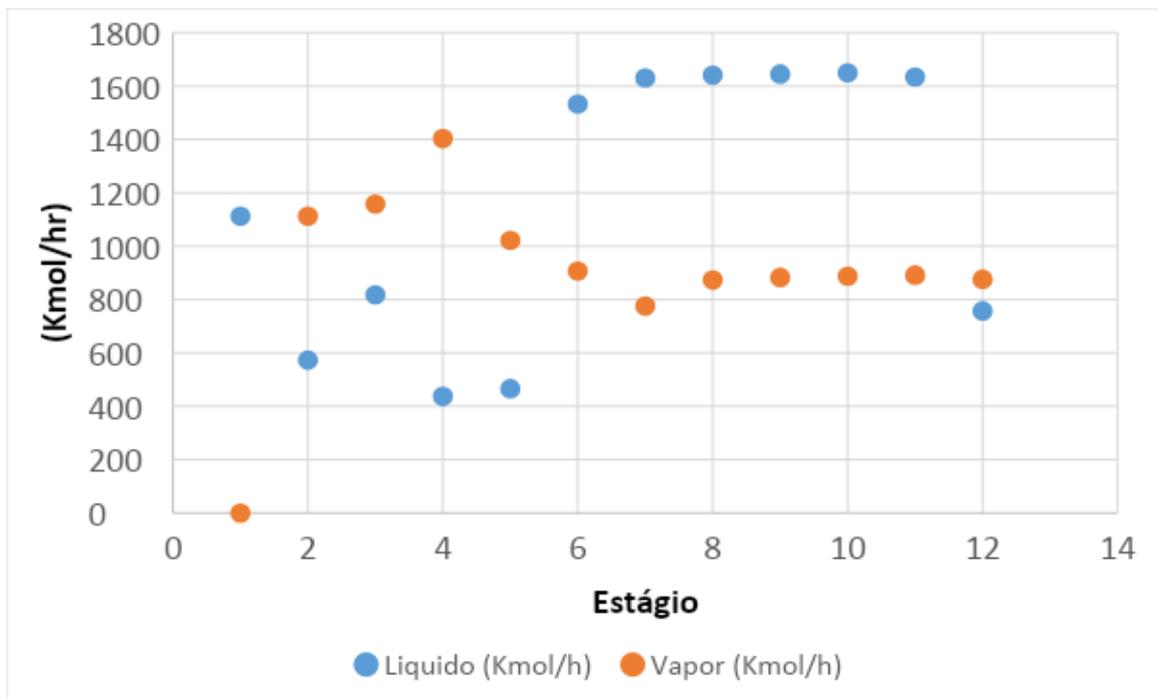
Analisando a Figura 17, que detalha a vazão molar de líquido e vapor em diferentes estágios da coluna, observa-se que essa queda de temperatura está diretamente relacionada a um aumento substancial na vazão molar de vapor neste estágio.

Figura 16 – Temperatura ao longo da coluna.



Fonte: O autor (2023).

Figura 17 – Distribuição molar das fases.

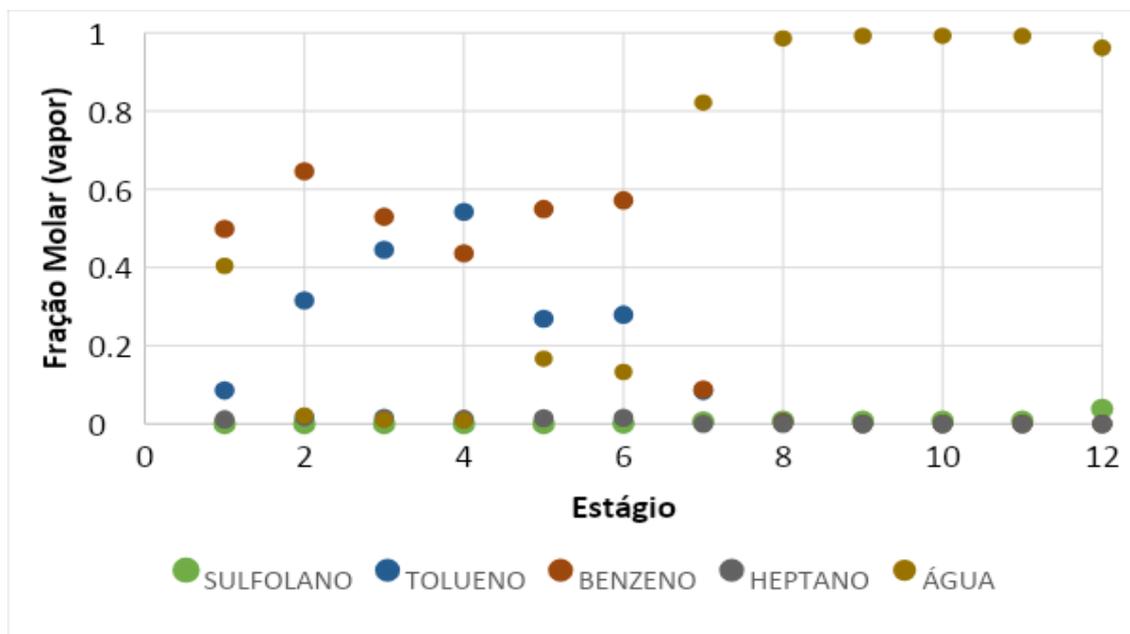


Fonte: O autor (2023).

Além disso, a análise da Figura 18, que demonstra a fração molar de vapor ao longo dos estágios da coluna de destilação, indica que a fração molar do sulfolane é

nula antes do sexto estágio. Isso significa que não se espera encontrar qualquer quantidade do solvente no produto de topo.

Figura 18 – Fração molar dos componentes ao longo da coluna recuperadora.



Fonte: O autor (2023).

A configuração adotada para a coluna de destilação, conforme documentado na Tabela 5, revelou-se altamente eficaz no contexto deste estudo. Uma análise comparativa entre os resultados alcançados utilizando a coluna recuperadora de solvente neste trabalho (conforme detalhado na Tabela 8) e os dados previamente publicados na literatura (Tabela 9) revela uma notável melhoria na eficiência do processo.

Os dados demonstram uma redução significativa de 74% na vazão mássica de água utilizada anteriormente, em conjunto com uma diminuição de 50% no número de estágios de separação necessários. Esses resultados indicam uma otimização substancial do processo de destilação, contribuindo para uma operação mais econômica e ambientalmente sustentável.

Adicionalmente, a aplicação de uma queda de pressão específica resultou em uma diminuição na temperatura dos produtos destilados no topo da coluna. Como resultado dessa otimização, os produtos agora são obtidos em temperatura ambiente, eliminando a necessidade de empregar um fluido refrigerante no condensador. Essa mudança não apenas simplifica o processo, mas também reduz os custos

operacionais associados à refrigeração, além de promover uma eficiência energética superior no sistema de destilação.

Tabela 8 – Valores gerados pela simulação na coluna de destilação sem degradação do solvente.

Parâmetros	Vazão mássica (kg.h ⁻¹)			
	Extrato	Água	Topo-D	Fundo-D
Sulfolane	80000	—	0	80000
N-hexano	0	—	0	0
N-heptano	780,841	—	780,841	0
N-octano	0	—	0	0
Benzeno	28999,5	—	28999,5	0
Tolueno	14174,7	—	14174,7	0
Água	—	2600	944,882	1655,12
Total	123955	2600	53954,245	80000,585
Temperatura (°C)	69	100	30,5	122,53

Fonte: Autor (2023).

Tabela 9 – Valores da coluna recuperadora de solvente disponíveis na literatura.

Parâmetros	Vazão mássica (kg.h ⁻¹)			
	Extrato	Água	Topo-D	Fundo-D
Sulfolane	80000	—	6,15E-25	80000
N-hexano	0	—	—	—
N-heptano	808,72	—	808,72	4,65E-29
N-octano	0	—	—	—
Benzeno	28875,32	—	28875,32	9,52E-16
Tolueno	14301,09	—	14301,09	6,56E-14
Água	—	10000	2805,97	7194,03
Total	123985,09	10000	46791,09	87194,03
Temperatura (°C)	69	100	55	112

Fonte: Adaptado de Lima (2019).

4.2 DINÂMICA E CONTROLE CONVENCIONAL

No contexto do ambiente de simulação dinâmica do software Aspen, foram conduzidos testes destinados à otimização dos parâmetros dos controladores empregados. A Tabela 10 apresenta os valores ótimos dos parâmetros, obtidos através de processos de otimização individual para cada controlador utilizado no processo de recuperação do solvente. Este procedimento visa aprimorar a eficiência e a estabilidade do sistema, garantindo um desempenho otimizado durante a operação do processo industrial.

Tabela 10 – Parâmetros ótimos obtidos para os controladores.

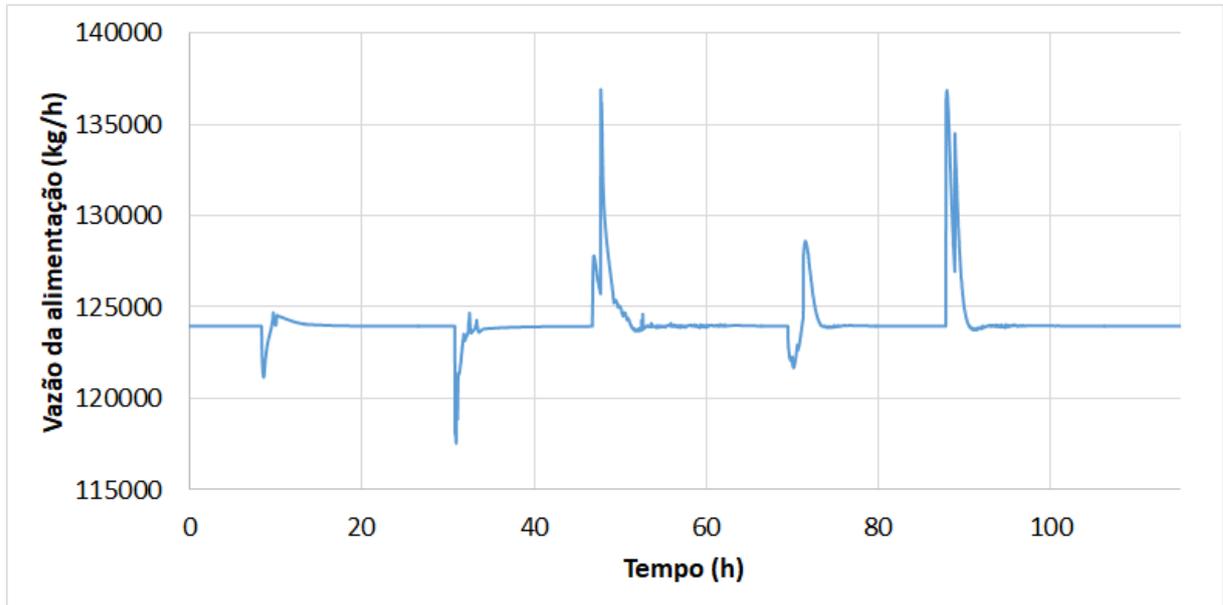
Controlador	Ganho proporcional (%%)	Ganho integral (min)	Tempo derivativo (min)
FC_1	1,034902	3,995995	--
FC_2	0,536515	0,948205	--
FC_3	0,676618	0,564341	--
LC_1	12797,06	43,29506	--
LC_2	0,397523	23,01309	--
FM_BENZENO	0,239179	0,893406	--
B9_CondPC	2,482577	1,407552	0,21361

Fonte: O autor(2023).

Com base nos parâmetros ótimos identificados, foram gerados os gráficos representativos das Figuras 19 e 20, os quais ilustram as perturbações impostas nas variáveis e o tempo necessário para que o sistema retorne ao setpoint. Os resultados evidenciam que os parâmetros obtidos promovem uma resposta robusta às perturbações, assegurando a rápida convergência do sistema ao setpoint.

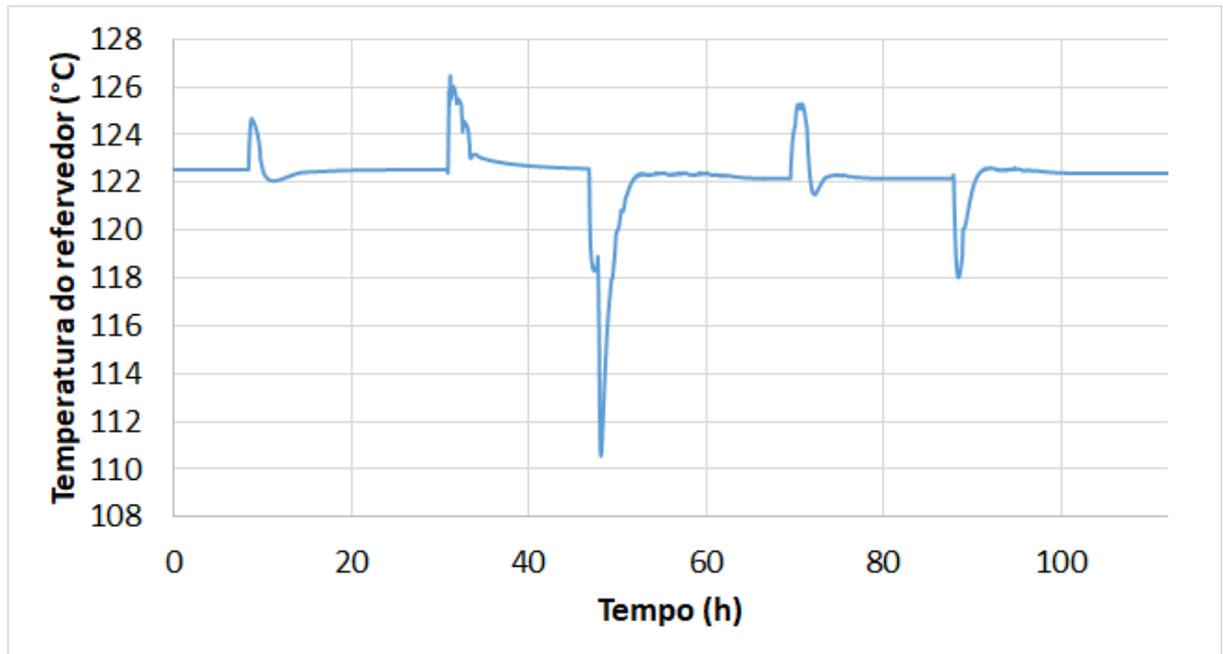
Ademais, ao realizar uma análise dinâmica detalhada da coluna, mediante a avaliação das figuras mencionadas, verifica-se que o modelo obtido no Aspen Dinâmico é altamente representativo. Observa-se um comportamento consistente com as expectativas teóricas, onde a diminuição da vazão resulta em um aumento da temperatura do refeedor, e vice-versa. Este padrão de resposta sugere uma concordância satisfatória entre o modelo simulado e o comportamento esperado do sistema físico em questão.

Figura 19 – Resposta do controlador FC_1 a perturbações na carga.



Fonte: O autor (2023).

Figura 20 –Resposta da temperatura no refervedor a perturbações na carga.



Fonte: O autor (2023).

4.3 APLICAÇÃO DE MÉTODOS DE INTELIGÊNCIA ARTIFICIAL NA COLUNA RECUPERADORA DE SOLVENTE

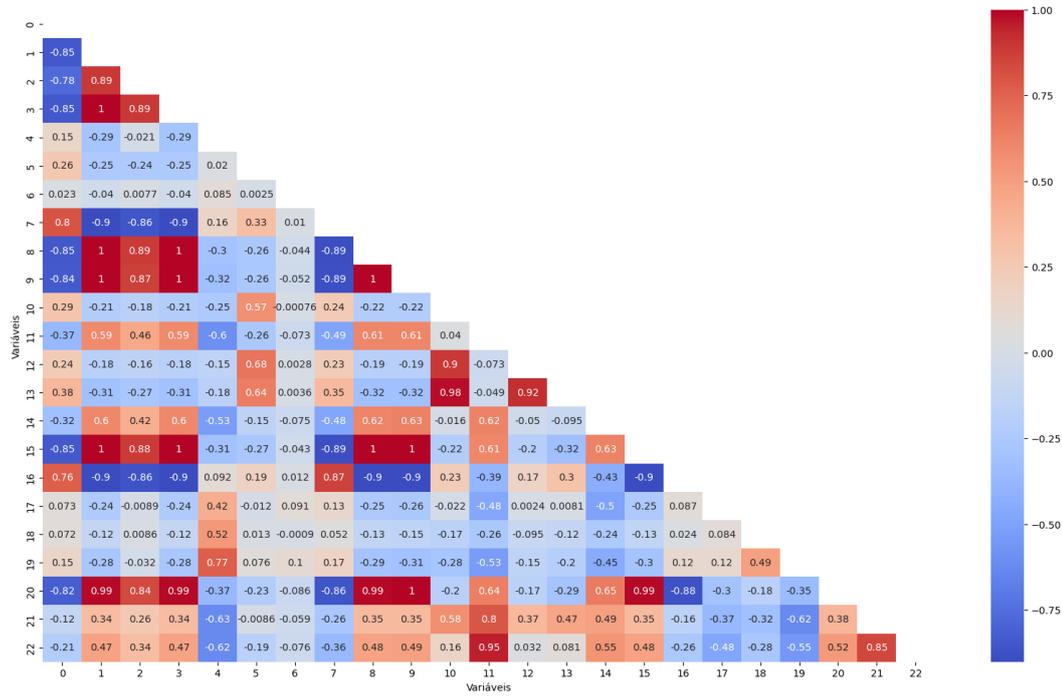
A análise exploratória de dados revelou que as variáveis relacionadas às temperaturas das correntes 4 e 5, a fração mássica de sulfolane e o nível de líquido no prato 2 não exibiram variações diante das perturbações introduzidas no sistema. Em face dessa inobservância, tais variáveis foram excluídas do conjunto de dados, com o intuito de otimizar a eficiência computacional.

Em seguida, procedeu-se com a redução de dimensionalidade do sistema em estudo, empregando a geração de matrizes de confusão por meio dos métodos de Spearman (Figura 22), Kendall (Figura 23) e Pearson (Figura 24).

A redução de dimensionalidade, realizada através das matrizes de confusão, (Figuras 21, 22, 23 e 24) evidencia uma correlação substancial entre grupos específicos de variáveis, medidas por sensores correspondentes, a saber: P_5Output, TFM_3Output, P_2Output, P_CONDOOutput, T_CONDOOutput, P_4Output e TM5_Output; P_REFOutput, P_11Output, REB_VPFLOWOutput, T_11Output e LL_11Output.

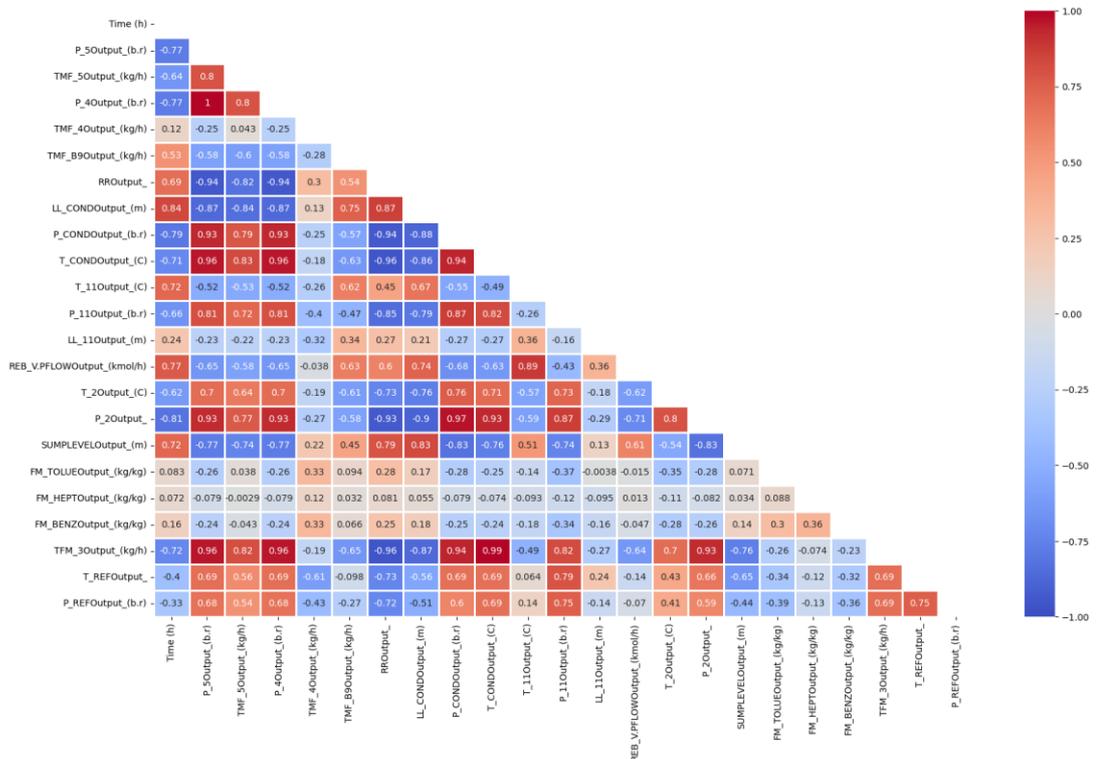
Logo, para o treinamento de modelos de inteligência artificial e considerando a praticidade na mensuração das variáveis no processo, a opção recaiu sobre o monitoramento da temperatura do condensador em detrimento de outras variáveis, da pressão do refeedor e da temperatura do prato 11 (prato sensível em relação a temperatura). Essa escolha visa mitigar desafios potenciais inerentes a implementação de modelos de aprendizagem de máquinas, incluindo *overfitting* (sobre ajuste), sensibilidade a pequenas alterações e redundância nas informações, contribuindo para a eficácia do processo de modelagem.

Figura 21: Matriz de confusão



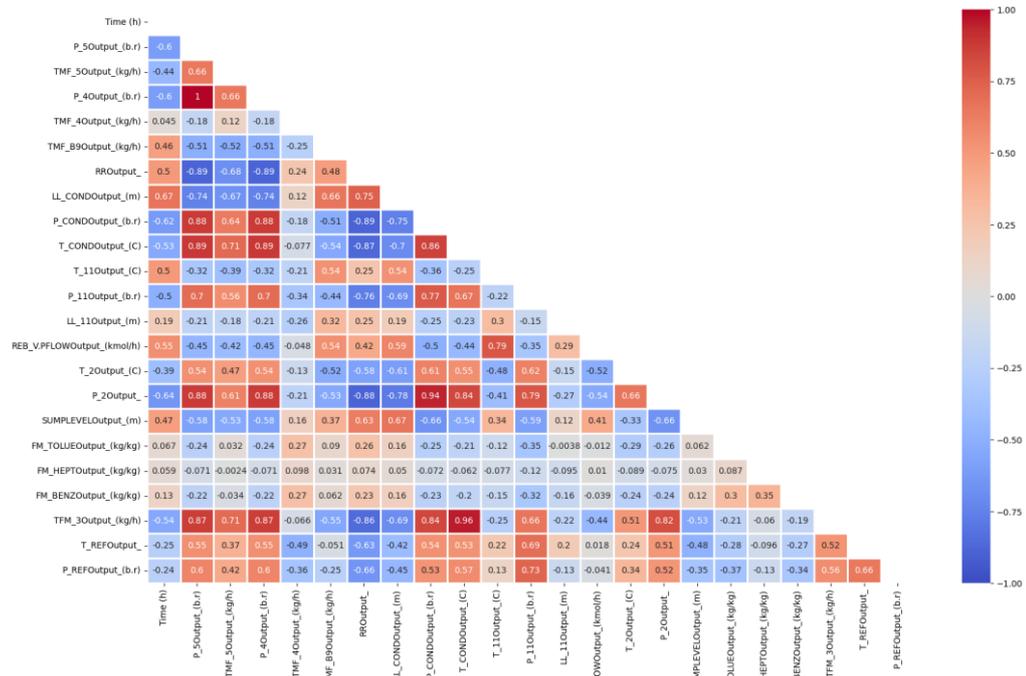
Fonte: O autor (2024).

Figura 22 –Matriz de confusão- método de Spearman.



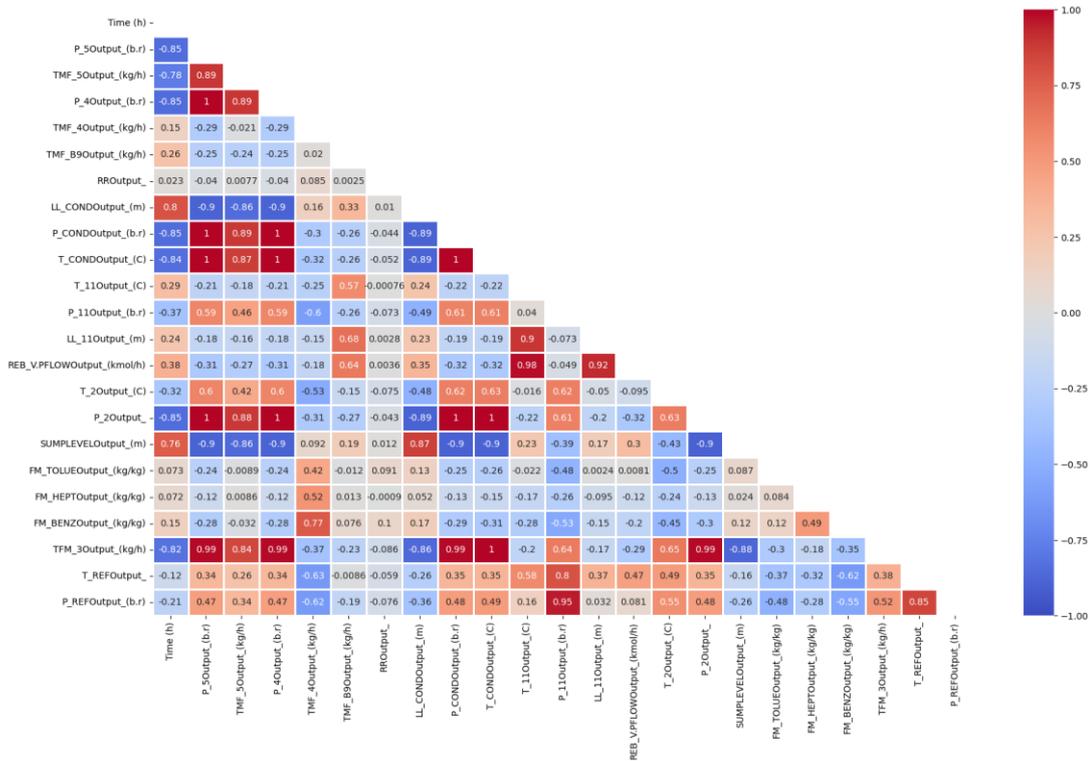
Fonte: O autor (2024).

Figura 23 –Matriz de confusão- método de Kendall



Fonte: O autor (2024).

Figura 24 –Matriz de confusão- método de Pearson



Fonte: O autor (2024).

A Tabela 11 apresenta os resultados de diferentes modelos de aprendizado de máquina aplicados ao conjunto de dados obtido na coluna recuperadora de solvente. Os modelos incluídos são Regressão Linear Múltipla, Árvore de Decisão, Floresta Aleatória e Regressão logística com avaliação através de diversas métricas, incluindo Score de Treino e Teste, além de erros absolutos médios (MAE) e erros quadráticos médios (MSE) para diferentes componentes químicos (heptano, benzeno e tolueno).

Tabela 11 -Métricas de desempenho para os modelos de aprendizagem de máquinas aplicados.

	Modelo Linear	Árvore de Decisão	Floresta Aleatória	Regressão Logística
Score de Treino	0,4488	1,0	0,983841	1,0
Score de Teste	-230111	0,8339	0,899750	0,939379
MAE Treino do Heptano	0,0001	1,1482E-16	0,000007	3,531E-06
MAE Treino do Benzeno	0,0005	4,5009E-15	0,000027	3,531E-06
MAE Treino do Tolueno	0,0010	1,4696E-15	0,000039	3,531E-06
MAE Teste do Heptano	1,1359	0,000017	0,000024	0,000063
MAE Treino do Benzeno	1,8726	0,000058	0,000073	3,531E-06
MAE Teste do Tolueno	2,8519	0,000069	0,000097	0,000071

MSE Treino do Heptano	3,0331E-07	1,7502E-32	1,1600E-08	6,26E-08
MSE Teste do Benzeno	1,9359E-06	2,9630E-29	6,2657E-08	1,16E-08
MSE Treino do Tolueno	4,8712E-06	3,2080E-30	8,7359E-08	5,75E-07
MSE Teste do Heptano	2243	1,7251E-07	9,9246E-08	0,000009
MSE Treino do Benzeno	6094	5,7504E-07	3,9644E-07	1,72E-32
MSE Teste do Tolueno	14133	6,9005E-07	4,3858E-07	3,331E-07

Fonte: O autor (2024).

Na análise das métricas de desempenho dos modelos de aprendizagem de máquina aplicados, os resultados fornecem uma visão abrangente do comportamento de cada abordagem.

O modelo de regressão linear múltipla demonstrou-se altamente ineficaz, conforme antecipado, devido ao seu desempenho insatisfatório nos conjuntos de treinamento e teste, assim como aos valores significativamente elevados observados nos cálculos dos erros.

A Árvore de Decisão destaca-se com um score de treino perfeito de 1.0, sugerindo uma adaptação excepcional aos dados de treino. No entanto, seu score de teste de 0.8339 revela uma possível dificuldade em generalizar para dados não vistos, indicando um caso de overfitting (sobre ajuste).

Contrastando com a Árvore de Decisão, a Floresta Aleatória demonstra uma performance notável em ambas as métricas de treino e teste, com scores de 0.983841 e 0.899750, respectivamente. Este equilíbrio sugere uma capacidade razoável de generalização e resistência a overfitting.

Por sua vez, a Regressão Logística apresenta resultados consistentes, com scores de treino e teste de 1.0 e 0.939379, respectivamente. Isso indica uma boa

adaptação aos dados de treino e uma capacidade sólida de generalização para dados não vistos.

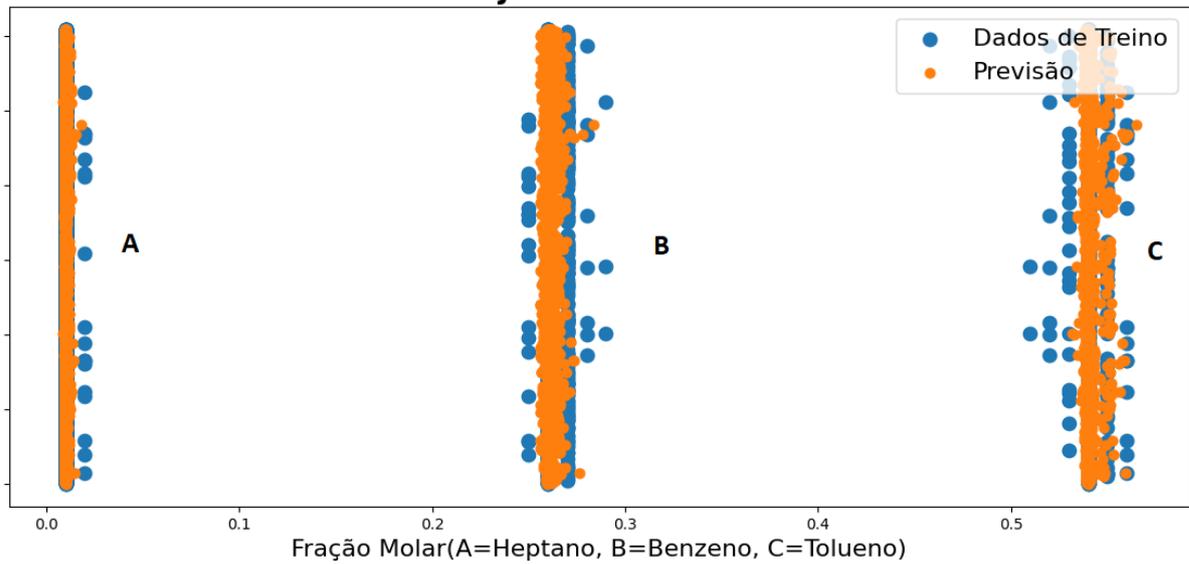
Ao analisar as métricas de erro, nota-se que os valores de MAE para os modelos são geralmente baixos, indicando boa precisão em prever as variáveis alvo (heptano, benzeno, tolueno). Destaca-se que a Árvore de Decisão e a Floresta Aleatória têm MAE de treino muito próximos de zero, sugerindo *overfitting* nesses casos.

Os valores de MSE reforçam as observações anteriores. A Árvore de Decisão, com MSE próximo de zero, reforça a ideia de *overfitting*. A Floresta Aleatória, por sua vez, apresenta MSE razoáveis, indicando uma boa capacidade de generalização. A Regressão Logística também mostra resultados competitivos, enquanto o modelo linear exibe MSE elevado, sugerindo desafios significativos na adaptação aos dados.

Em resumo, os dados destacam a Regressão Logística e Floresta aleatória como as opções mais equilibradas, apresentando bom desempenho tanto no treino quanto no teste. A Árvore de Decisão evidencia o *overfitting* e o modelo linear não é representativo devido ao score de teste negativo e MSE elevado.

Além disso, nas Figuras 24 a 28, são apresentados os dados obtidos no Aspen Dinâmico, juntamente com as previsões correspondentes dos modelos. O objetivo é destacar de maneira mais intuitiva as métricas de desempenho alcançadas. Nestas figuras, observa-se que a fração mássica de heptano (estimada e prevista) é próxima de zero, a fração mássica de benzeno varia entre 0,2 e 0,3, e a fração mássica de tolueno é superior a 0,5.

Figura 25 –Dados do conjunto de treinamento e previsões realizadas pelo modelo Linear.



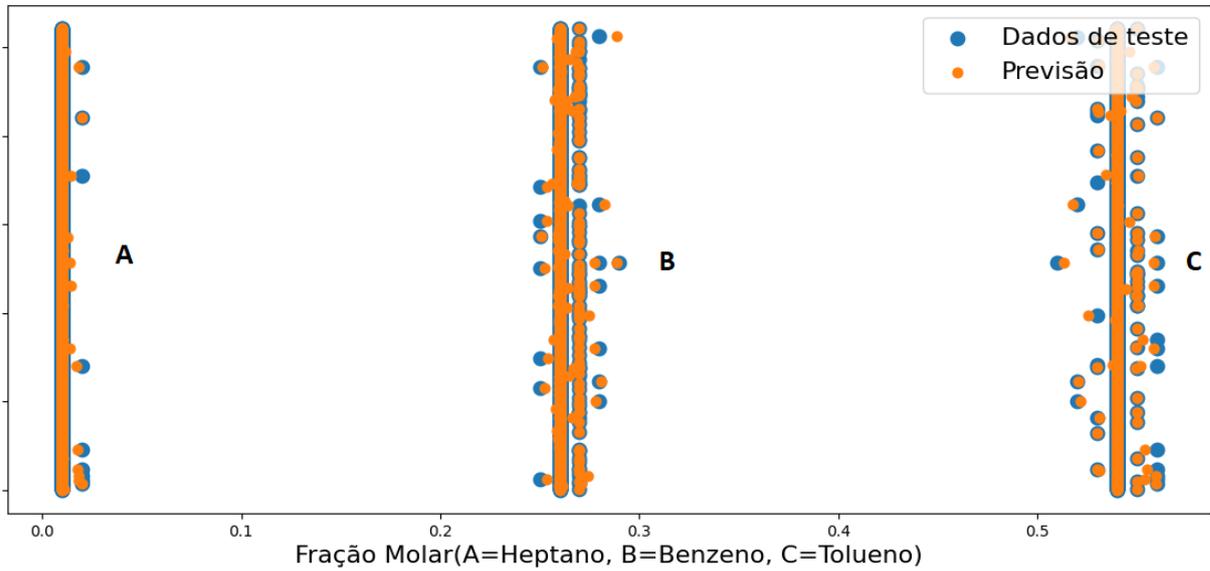
Fonte: O autor (2024).

Figura 26 –Dados do conjunto de treinamento e previsões realizadas pelo modelo de árvore de decisão.



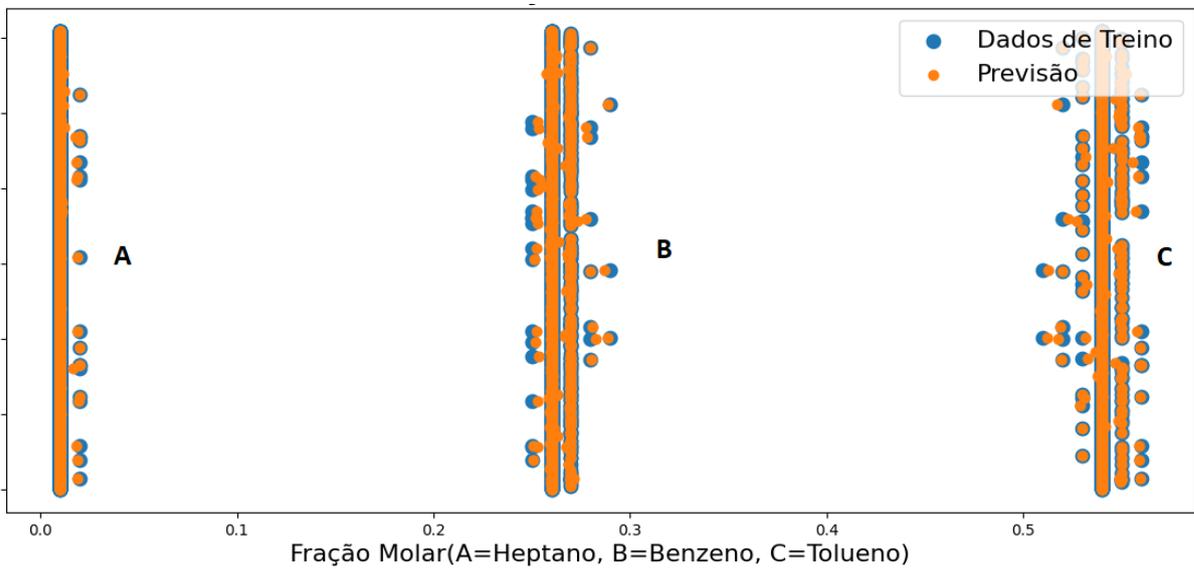
Fonte: O autor (2024).

Figura 27 –Dados do conjunto de testes e previsões realizadas pelo modelo de árvore de decisão.



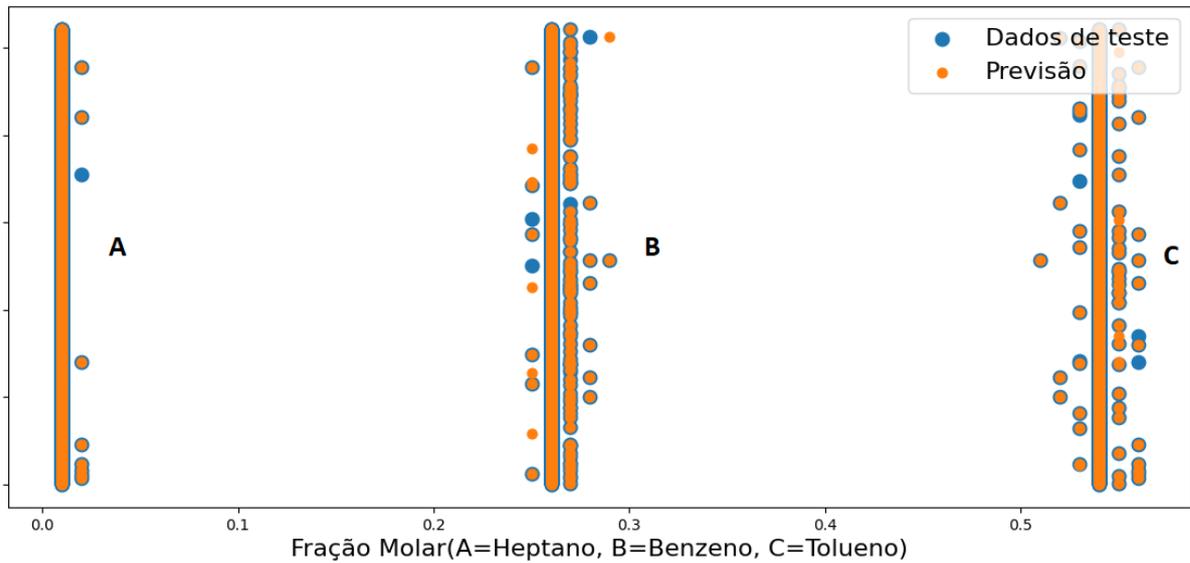
Fonte: O autor (2024).

Figura 28–Dados do conjunto de treinamento e previsões realizadas pelo modelo floresta aleatória.



Fonte: O autor (2024).

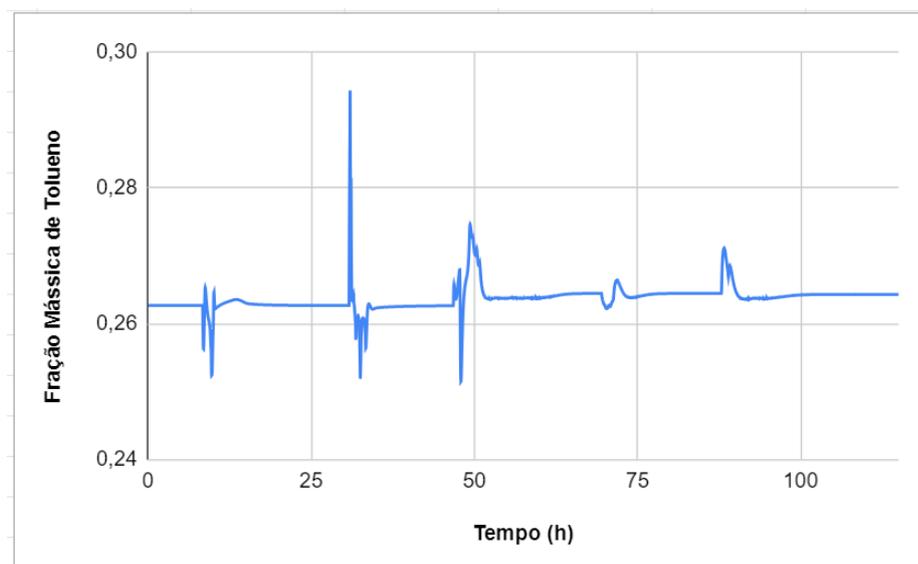
Figura 29 –Dados do conjunto de testes e previsões realizadas pelo modelo de floresta aleatória.



Fonte: O autor (2024).

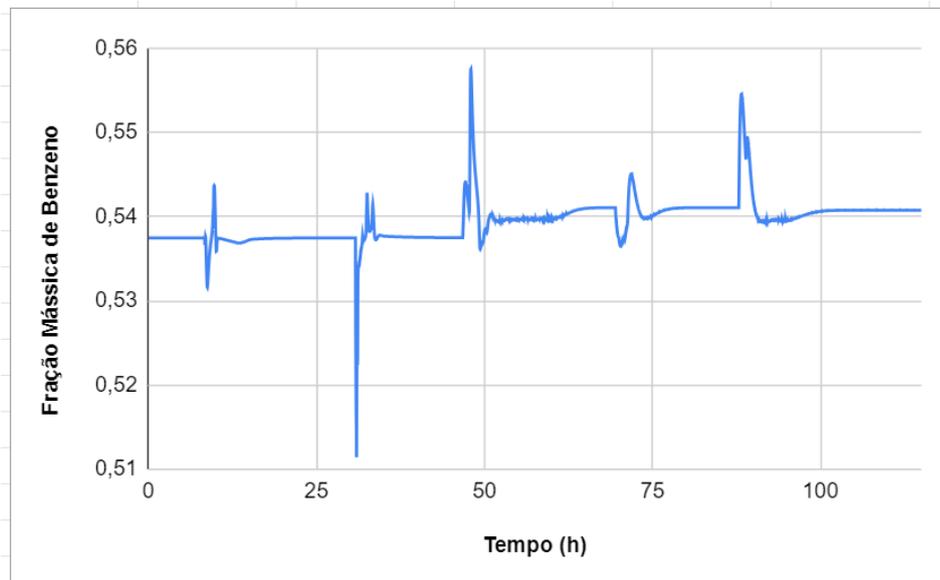
A análise aprofundada dos gráficos das Figuras 30 e 31, exibindo as frações molares de tolueno e benzeno ao longo do tempo, revela a robustez e a precisão dos modelos de IA empregados no controle do processo. Estes gráficos não apenas corroboram a eficácia do controle realizado IA, mas também destacam a habilidade dos modelos em lidar com flutuações e perturbações que possam surgir durante a operação da coluna.

Figura 30 – Fração mássica de tolueno controle por IA.



Fonte: O autor (2024).

Figura 30 – Fração mássica de benzeno controle por IA.



Além disso, a análise dos gráficos revela a consistência do desempenho dos modelos ao longo do tempo, evidenciando uma estabilidade e confiabilidade cruciais para garantir a eficiência e a qualidade do processo.

Portanto, os resultados obtidos por meio destes gráficos reforçam a importância e a viabilidade das técnicas de inteligência artificial como estratégia de controle em processos industriais. A habilidade destes modelos em lidar com a complexidade e as incertezas do ambiente de processo demonstra o significativo potencial que a inteligência artificial pode oferecer para otimizar e aprimorar operações industriais em diversas áreas.

5 CONCLUSÃO

Ao longo desta análise abrangente, foram percorridas diversas etapas do processo de separação de hidrocarbonetos e recuperação do sulfolane, identificando nuances e otimizações significativas. Inicialmente, a destilação simples foi reconhecida como uma escolha termodinamicamente adequada para a separação da mistura, mas com a ressalva de demandar uma carga térmica considerável e um alto número de estágios na coluna, resultando em custos operacionais substanciais.

Para mitigar esses custos, implementou-se a injeção controlada de vapor d'água na coluna, observando-se melhorias na eficiência de separação entre o solvente e os hidrocarbonetos. A formação de um sistema heterogêneo na coluna, composto por fases polar e apolar, indicou uma separação parcial sem a necessidade de energia adicional.

Explorando os mapas de curvas residuais de destilação para o processo extrativo com a introdução de vapor de água, foi identificada a formação de um azeótropo instável entre a água e os hidrocarbonetos. A presença de água na corrente dos produtos de topo, contrariamente à expectativa de exclusividade dos compostos hidrocarbonetos, destaca a complexidade dessa interação.

A análise específica da coluna de recuperação de solvente revelou uma redução abrupta na temperatura em um estágio específico, indicando uma taxa de evaporação tão pronunciada que apenas a água permanece no estado líquido. Além disso, a fração molar do solvente foi observada como nula antes do sexto estágio, reforçando a eficiência da configuração adotada.

A transição para o Aspen Dinâmico envolveu testes para otimização dos parâmetros dos controladores, evidenciando respostas robustas a perturbações e fornecendo parâmetros ótimos para cada controlador. A escolha criteriosa das variáveis monitoradas, como a temperatura do condensador, visou mitigar desafios associados à implementação de modelos de aprendizado de máquina.

Finalmente, a aplicação de modelos de aprendizado de máquina revelou que a Floresta Aleatória se destacou como a opção mais equilibrada, apresentando bom desempenho em treino e teste. A Árvore de Decisão evidenciou overfitting, a Regressão Logística mostrou consistência, e o modelo linear demandou uma avaliação mais detalhada devido a resultados menos favoráveis.

Finalmente, esta análise forneceu uma compreensão aprofundada do processo, desde as operações termodinâmicas até a implementação de técnicas avançadas de controle e modelagem, destacando melhorias significativas na eficiência operacional. Essa abordagem integrada permite uma visão holística do sistema, proporcionando insights valiosos para futuras otimizações e aprimoramentos.

Neste sentido, estudos futuros podem considerar: a investigação e otimização de outras variáveis de controle utilizando métodos convencionais de análise dinâmica; a integração de novos modelos de inteligência artificial; a implementação do controle de temperatura com base em modelos de inteligência artificial; e a adoção da versão 12 do Aspen Plus para incorporar modelos in-line e realizar uma comparação estatística entre os modelos de inteligência artificial e o controle convencional.

REFERÊNCIAS

AL KALBANI, F.; ZHANG, J. **Inferential Composition Control of a Distillation Column Using Active Disturbance Rejection Control with Soft Sensors.** *Sensors*, v. 23, n. 2, p. 1019, 16 jan. 2023.

AL-MALAH, KAMAL I. M. **ASPEN PLUS®: CHEMICAL ENGINEERING APPLICATIONS.** 1ED. NEW JERSEY: WILEY-BLACKWELL, 2017. 623P.

BRAY, N. **INTRODUCTION TO MULTIVARIATE DATA ANALYSIS IN CHEMICAL ENGINEERING.** CONTROL ENGINEERING, 2013.

BELGIU, M.; DRĂGUȚ, L. **RANDOM FOREST IN REMOTE SENSING: A REVIEW OF APPLICATIONS AND FUTURE DIRECTIONS.** *ISPRS JOURNAL OF PHOTOGRAMMETRY AND REMOTE SENSING*, V. 114, N. 114, P. 24–31, ABR. 2016.

CARLSON, E. C. **Don't Gamble With Physical Properties For Simulations.** *Chemical engineering progress*. 1996, Vol 92, Num 10, pp 35-46

CHEREMISINOFF, P. N. **WASTE MINIMIZATION AND COST REDUCTION FOR THE PROCESS INDUSTRIES.** PARK RIDGE, N.J., U.S.A.: NOYES PUBLICATIONS, 1995.

DANTAS, S.; PEREIRA BRITO, R. **AVALIAÇÃO DO COMPORTAMENTO DINÂMICO DE UMA COLUNA DE DESTILAÇÃO AZEOTRÓPICA DE ALTA PUREZA.** VI CONGRESSO DE INICIAÇÃO CIENTÍFICA DA UNIVERSIDADE FEDERAL DE CAMPINA GRANDE.

FAHIM. M. A; AL-SAHHAF. T.A; ELKILANI. A. E. **INTRODUÇÃO AP REFINO DE PETRÓLEO.** ELSEVIER EDITORA LTDA, 2012.

FERNÁNDEZ, EMILSE BELÉN. 56F. 2018. **SIMULAÇÃO DE UMA UNIDADE DE EXTRAÇÃO DE BTX NO ASPEN PLUS E AVALIAÇÃO DA VIABILIDADE**

ECONÔMICA PRELIMINAR. Anais do 9º 9º CONGRESSO BRASILEIRO DE PESQUISA E DESENVOLVIMENTO EM PETRÓLEO E GÁS. Maceió, Al, 2017.

FIGUEIRA, C. V. **MODELOS DE REGRESSÃO LOGÍSTICA.** DISSERTAÇÃO DE MESTRADO. UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL, PORTO ALEGRE, 2006.

FOULADVAND, M. T.; ASADI, J.; LOTFOLLAHI, M. N. **SIMULATION AND OPTIMIZATION OF AROMATIC EXTRACTION FROM LUBE OIL CUTS BY LIQUID-LIQUID-LIQUID EXTRACTION.** CHEMICAL ENGINEERING RESEARCH AND DESIGN, V. 165, P. 118–128, JAN. 2021.

FOUST, ALAN S.; CLUMP, CURTIS W.; WENZEL, LEONARD A.; MAUS, LOUIS; ANDERSEN, BRYCE L. **PRINCÍPIOS DAS OPERAÇÕES UNITÁRIAS.** TRADUÇÃO DA 2ED. RIO DE JANEIRO: LTC, 1982. 670P.

GERBAUD, V. ET AL. **REVIEW OF EXTRACTIVE DISTILLATION. PROCESS DESIGN, OPERATION, OPTIMIZATION AND CONTROL.** CHEMICAL ENGINEERING RESEARCH AND DESIGN, V. 141, P. 229–271, JAN. 2019.

GÉRON, A. MÃOS A OBRA: **APRENDIZADO DE MÁQUINAS COM SKLEARN & TENSORFLOW.** 2. ED. RIO DE JANEIRO: ATLAS BOOKS, 2019.

GLADSON EULER LIMA JÚNIOR, LUÍS GONZAGASALES VACONCELOS, KAROLINE DANTAS BRITO, ROMILDO PEREIRA BRITO, **OPTIMIZATION OF OF DISTILLATION COLUMN CONTROL USING Q-LEARNING+PID: A DYNAMIC DYNAMIC APPROACH,** 2023. ISSN: 2178-9010.

JUANCARLOS MENDEZ-B ET AL. **MODELING AND SIMULATION OF AN ENERGY INTEGRATED DISTILLATION COLUMN IN A BOND GRAPH APPROACH.** ENTROPY, V. 24, N. 9, P. 1191–1191, 25 AGO. 2022.

JUNQUEIRA, TÁSSIA L. 187F. 2010. **SIMULAÇÃO DE COLUNAS DE DESTILAÇÃO CONVENCIONAL, EXTRATIVA E AZEOTRÓPICA NO PROCESSO DE PRODUÇÃO DE BIOETANOL ATRAVÉS DA MODELAGEM DE NÃO EQUILÍBRIO E DA MODELAGEM DE ESTÁGIOS DE EQUILÍBRIO COM EFICIÊNCIA.** DISSERTAÇÃO DE MESTRADO. CAMPINAS, SP: UNICAMP, 2010.

KISTER, H. Z. **DISTILLATION OPERATION**. 1ST ED. NEW YORK: MCGRAW-HILL, INC., 1990.

KWON, H. et al. **Distillation column temperature prediction based on machine-learning model using wavelet transform**. Em: Computer Aided Chemical Engineering. [s.l.] Elsevier, 2022. p. 1651–1656.

LEI, Z.; LI, C.; CHEN, B. **EXTRACTIVE DISTILLATION: A REVIEW**. SEPARATION & PURIFICATION REVIEWS, V. 32, N. 2, P. 121–213, 12 JAN. 2003.

LIMA, A. J. R. C. **EFEITO DA INJEÇÃO DE VAPOR DE ÁGUA NA DESTILAÇÃO DE UMA PLANTA SIMPLIFICADA DE EXTRAÇÃO DE AROMÁTICOS A PARTIR DA NAFTA**. TRABALHO DE CONCLUSÃO DE CURSO. RECIFE, PE: UFPE, 2019.

LUYBEN, W. L. **Distillation Design and Control Usign Aspen Simulation**. 2 nd ed. Hoboken: John Wiley & Sons, Inc., 2013.

MA, X. C. ET AL. **COMPUTER-AIDED NAPHTHA LIQUID–LIQUID EXTRACTION: MOLECULAR RECONSTRUCTION, SUSTAINABLE SOLVENT DESIGN AND MULTISCALE PROCESS OPTIMIZATION**. FUEL, V. 334, P. 126651–126651, 1 FEV. 2023.

MA, Y. ET AL. **A REVIEW OF EXTRACTIVE DISTILLATION FROM AN AZEOTROPIC PHENOMENON FOR DYNAMIC CONTROL**. CHINESE JOURNAL OF JOURNAL OF CHEMICAL ENGINEERING, V. 27, N. 7, P. 1510–1522, JUL. 2019.

2019.

MAARTEN, R. ET AL. **MACHINE LEARNING IN CHEMICAL ENGINEERING: STRENGTHS, WEAKNESSES, OPPORTUNITIES, AND THREATS**. ENGINEERING, THREATS. ENGINEERING, V. 7, N. 9, JUL. 2021.

MCCABE, W. L.; SMITH, J. C.; HARRIOTT, P. **UNIT OPERATIONS OF CHEMICAL ENGINEERING**. 5TH ED. NEW YORK: MCGRAW-HILL CHEMICAL ENGINEERING SERIES, 1993.

MEADOWCROFT, T. A. **PROCESS CONTROL DESIGN AND PRACTICE – A NEW APPROACH TO TEACHING CONTROL TO CHEMICAL ENGINEERS**. DISPONÍVEL EM: <[HTTPS://PEER.ASEE.ORG/PROCESS-CONTROL-DESIGN-AND-PRACTICE-](https://peer.asee.org/process-control-design-and-practice-)

A-NEW-APPROACH-TO-TEACHING-CONTROL-TO-CHEMICAL-ENGINEERS>.
ACESSO EM: 13 FEV. 2024.

MEIRELLES, L. B.; SILVA, S. M. C. DA; RAJAGOPAL, K. **CARACTERIZAÇÃO DA NAFTA PETROQUÍMICA PARA A PRODUÇÃO DE AROMÁTICOS**. Anais do XX Congresso Brasileiro de Engenharia Química. Anais...São Paulo: Editora Edgard Blücher, 2015.

MEIRELLES, L. B. **CARACTERIZAÇÃO DA NAFTA PETROQUÍMICA PARA PRODUÇÃO DE AROMÁTICOS**. DISSERTAÇÃO DE MESTRADO. RIO DE JANEIRO, RJ: UFRJ, 2014.

MESQUITA, F.M. R. **ESTUDO DE SOLVENTES ALTERNATIVOS PARA EXTRAÇÃO DE AROMÁTICOS**. TESE DE DOUTORADO. FORTALEZA, CE: UFCE, UFCE, 2015.

MEZZOMO, H. **DESTILAÇÃO EXTRATIVA DE ETANOL UTILIZANDO GLICEROL - MODELAGEM TERMODINÂMICA, OTIMIZAÇÃO E DETERMINAÇÃO DE UMA CONFIGURAÇÃO ÓTIMA**. DISSERTAÇÃO DE MESTRADO. PORTO ALEGRE, RS: UFRGS, 2014.

NEMETH, B.; LANG, P.; LASZLO HEGELY. **Optimisation of solvent recovery in two batch distillation columns of different size**. Journal of Cleaner Production, v. 275, p. 122746–122746, 1 dez. 2020.

OCHOA-ESTOPIER, L. M. et al. **Prediction of flooding in distillation columns using machine learning**. Digital Chemical Engineering, v. 7, n. 100098, p. 100098, 2023.

OÑA, R. D.; EBOLI, L.; MAZZULLA, G. **KEY FACTORS AFFECTING RAIL SERVICE QUALITY IN THE NORTHERN ITALY: A DECISION TREE APPROACH**. TRANSPORT, V. 29, N. 1, P. 75–83, 25 MAR. 2014.

ORAVEC, J.; BAKOSOVA, M.; ARTZOVA, P. **Advanced process control design for a distillation column using UniSim design**. 1 jun. 2017.

RAPPEL, R.; DE GÓIS, L. M. N.; MATTEDI, S. **LIQUID-LIQUID EQUILIBRIA DATA FOR SYSTEMS CONTAINING AROMATIC + NONAROMATIC + SULFOLANE AT 308.15 AND 323.15 K.** FLUID PHASE EQUILIBRIA, V. 202, N. 2, P. 263–276, NOV. 2002.

SANTOS, N. S. **SIMULAÇÃO DINÂMICA DE UMA COLUNA DE DESTILAÇÃO.** TRABALHO DE CONCLUSÃO DE CURSO. SALVADOR, BA: UCSAL, 2019.

SEADER, J. D.; HENLEY, E. J.; ROPER, D. K. **SEPARATION PROCESS PRINCIPLES CHEMICAL AND BIOCHEMICAL OPERATIONS.** HOBOKEN, NJ: NJ: JOHN WILEY & SONS, INC, 2011.

SHI, J. ET AL. STUDY OF AROMATIC EXTRACTION FROM LIGHT CYCLE OIL FROM THE VIEWPOINT OF INDUSTRIAL APPLICATIONS. FUEL, V. 357, P. 130023–130023, 1 FEV. 2024.

Shiras, R.N., D.N. Hanson, and C.H. Gibson, Ind. Eng. Chem., 42, 871–876 (1950).

SIEMENS. **USE OF PROCESS ANALYTICS IN AROMATICS (BTX AND PHENOL) PRODUCTION PLANTS.** ESTUDO DE CASO. HOUSTON, TEXAS, 2015.

SILVA, M. G. L.; LUCENA, S. **APLICAÇÃO DE TÉCNICAS APRENDIZAGEM DE MÁQUINAS PARA PREDIÇÃO DE FRAÇÕES MOLARES DO DESTILADO EM UMA COLUNA DE DESTILAÇÃO BINÁRIA.** IN: ANAIS DO WEB ENCONTRO NACIONAL DE ENGENHARIA QUÍMICA. ANAIS...DIAMANTINA(MG) ONLINE, 2023.

SILVA, R. J. M. C. L. **ESTUDO DO PROCESSO DE EXTRAÇÃO DE BENZENO, TOLUENO E XILENO (BTX) E DA VIABILIDADE ECONÔMICA PRELIMINAR UTILIZANDO O ASPEN PLUS® E MATLAB® .** TRABALHO DE CONCLUSÃO DE CURSO. RECIFE, PE: UFPE, 2016.

SILVEIRA, M. B. G. DA ET AL. **APLICAÇÃO DA REGRESSÃO LOGÍSTICA NA ANÁLISE DOS DADOS DOS FATORES DE RISCO ASSOCIADOS À HIPERTENSÃO ARTERIAL.** RESEARCH, SOCIETY AND DEVELOPMENT, V. 10, V. 10, N. 16, P. E20101622964–E20101622964, 4 DEZ. 2021.

SIMONELLI, G. ET AL. **SIMULAÇÃO DO CONTROLE DE UMA COLUNA DE DESTILAÇÃO DESCONTÍNUA UTILIZANDO O SCILAB.** ENGEVISTA, V. 19, N. 2, P. 498, P. 498, 8 MAIO 2017.

SISHI, M.; ARNESH TELUKDARIE. THE APPLICATION OF DECISION TREE REGRESSION TO OPTIMIZE BUSINESS PROCESSES. 5 ABR. 2021.

SMITH, C. A.; CORRIPIO, A. B. PRINCÍPIOS E PRÁTICA DO CONTROLE AUTOMÁTICO DE PROCESSO. 3ª ED. RIO DE JANEIRO, RJ: LTC - LIVROS TÉCNICOS E CIENTÍFICOS EDITORA LTDA., 2008.

WANG, Q. ET AL. OPTIMAL DESIGN OF A NEW AROMATIC EXTRACTIVE DISTILLATION PROCESS AIDED BY A CO-SOLVENT MIXTURE. ENERGY PROCEDIA, V. 105, P. 4927–4934, MAIO 2017.

YAN, R. ET AL. A DECISION TREE BASED DATA-DRIVEN DIAGNOSTIC STRATEGY FOR AIR HANDLING UNITS. ENERGY AND BUILDINGS, V. 133, P. 37–P. 37–45, DEZ. 2016.

ZHAO, J., LASTERNAS, B., LAM, K.P., YUN, R., AND LOFTNESS, V., OCCUPANT BEHAVIOR AND SCHEDULE MODELING FOR BUILDING ENERGY SIMULATION THROUGH OFFICE APPLIANCE POWER CONSUMPTION DATA MINING, ENERGY AND BUILDINGS, VOL. 82, PP. 341-355, 2014.

ZHAO, N.; LU, J. Neural network algorithm and its application in temperature control of distillation tower. 2021.

APÊNDICE A – CÓDIGO EM PYTHON DESENVOLVIDO PARA ANÁLISE EXPLORATÓRIA DE DADOS E IMPLEMENTAÇÃO DE MÉTODOS DE IA

```

import pandas as pd
import random
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
!pip install --upgrade gspread
# Criando conexão entre o colab e o google Sheet
from google.colab import auth
auth.authenticate_user()

import gspread
from google.auth import default
creds, _ = default()

gc = gspread.authorize(creds)

#Abrindo o arquivo
spreasheet= gc.open('Dados_IA_1')
#Abrindo as abas do arquivo
page=spreasheet.sheet1
#olhando os dados
page.row_values(1)
#Escrever na planilha
#page.update('CELULA')
dataset= pd.DataFrame(page.get_all_records())
dataset.head()
dataset.tail()
dataset.describe()
dataset.shape
dataset.isnull().sum()
#Verificando a presença de colunas com um valor invariável
for col in dataset.columns:
    if len(dataset[col].unique()) == 1:
        print( {col} )
#Apagando as colunas
dataset=dataset.drop(['T_5Output_(C)'], axis=1)
dataset=dataset.drop(['T_4Output_(C)'], axis=1)
dataset=dataset.drop(['FM_SULFOutput_(kg/kg)'], axis=1)
dataset=dataset.drop(['LL_2Output_(m)'], axis=1)
dataset=dataset.drop(['duty_condOutput_(Gj/h)'], axis=1)
dataset.head()
from sklearn.preprocessing import StandardScaler

# Cria uma instância do StandardScaler
scaler = StandardScaler()

# Ajuste os dados
scaled_data = scaler.fit_transform(dataset)
import numpy as np
import seaborn as sns

```

```

import matplotlib.pyplot as plt

# Calcula a matriz de correlação usando numpy.corrcoef()
corr_matrix = np.corrcoef(scaled_data, rowvar=False)

# Cria uma máscara para remover a diagonal principal
mask = np.triu(np.ones_like(corr_matrix, dtype=bool))

# Configura o tamanho da figura
plt.figure(figsize=(20, 12))

# Plota a matriz de correlação usando seaborn
sns.heatmap(corr_matrix, annot=True, cmap='coolwarm', mask=mask)

# Define os rótulos dos eixos
plt.xlabel('Variáveis')
plt.ylabel('Variáveis')

# Exibe o gráfico
plt.show()

a=dataset.corr()
from tabulate import tabulate
tabela = tabulate(a, headers='keys', tablefmt='fancy_grid')

print(tabela)
nome_do_arquivo_excel = 'Tabela1.xlsx'
a.to_excel(nome_do_arquivo_excel, index=False)
# Método de Spearman
corr = dataset.corr(method='spearman') # Matriz de correlação
lower_triangle = np.tril(corr, k = -1) # seleciona o menor triangula
da matriz
mask = lower_triangle == 0 # Apaga a diagonal do mapa de calor, pois a
diagonal sempre será 1
#coolwarm
#rocket
plt.figure(figsize = (20,12))
sns.heatmap(lower_triangle, center = 0, cmap = 'coolwarm', annot= True,
xticklabels = corr.index, yticklabels = corr.columns,
cbar= True, linewidths= 1, mask = mask, vmin = -1, vmax =
+1) # mapa de calor
plt.show()
# Método de Kendall
corr = dataset.corr(method='kendall') # Matriz de correlação
lower_triangle = np.tril(corr, k = -1) # seleciona o menor triangula
da matriz
mask = lower_triangle == 0 # Apaga a diagonal do mapa de calor, pois a
diagonal sempre será 1
#coolwarm
#rocket
plt.figure(figsize = (20,12))
sns.heatmap(lower_triangle, center = 0, cmap = 'coolwarm', annot= True,
xticklabels = corr.index, yticklabels = corr.columns,

```

```

        cbar= True, linewidths= 1, mask = mask, vmin = -1, vmax =
+1) # mapa de calor
plt.show()
# Método de Pearson
corr = dataset.corr(method='pearson') # Matriz de correlação
lower_triangle = np.tril(corr, k = -1) # seleciona o menor triangula
da matriz
mask = lower_triangle == 0 # Apaga a diagonal do mapa de calor, pois a
diagonal sempre será 1
#coolwarm
#rocket
plt.figure(figsize = (20,12))
sns.heatmap(lower_triangle, center = 0, cmap = 'coolwarm', annot= True,
xticklabels = corr.index, yticklabels = corr.columns,
        cbar= True, linewidths= 1, mask = mask, vmin = -1, vmax =
+1) # mapa de calor
plt.show()
dataset.corr(method='pearson')
colunas_para_excluir = [
    'P_5Output_(b.r)',
    'TFM_3Output_(kg/h)',
    'P_2Output_',
    'P_CONDOOutput_(b.r)',
    'TMF_5Output_(kg/h)',
    'P_11Output_(b.r)',
    'SUMPLEVELOutput_(m)',
    'REB_V.PFLOWOutput_(kmol/h)',
    'LL_11Output_(m)',
    'FM_TOLUEOutput_(kg/kg)',
    'FM_BENZOutput_(kg/kg)',
    'FM_HEPTOutput_(kg/kg)',
    'Time (h)'
]

X = dataset.drop(colunas_para_excluir, axis=1)

Colunas=
dataset[["FM_HEPTOutput_(kg/kg)", "FM_BENZOutput_(kg/kg)", "FM_TOLUEOutpu
t_(kg/kg)"]]
y= Colunas

from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size =
0.30, random_state = 0)
def calcular_mae(y, y_pred):
    erro = y - y_pred
    erro_absoluto = np.absolute(erro)
    return erro_absoluto.mean()
def calcular_mse(y, y_pred):
    erro = y - y_pred
    erro_quadratico = erro ** 2
    return erro_quadratico.mean()
from sklearn.linear_model import LinearRegression

```

```

modelo_linear = LinearRegression()
modelo_linear

modelo_linear.fit(X_train, y_train)
modelo_linear
modelo_linear.score(X_train, y_train)
y_pred_train = modelo_linear.predict(X_train)
y_pred_test = modelo_linear.predict(X_test)
y_train.shape
y_pred_train.shape
# Cálculo de MAE
mae_train = calcular_mae(y_train, y_pred_train)
mae_test = calcular_mae(y_test, y_pred_test)
print("MAE treino:", mae_train)
print("MAE teste:", mae_test)
# Cálculo de MSE
mse_train = calcular_mse(y_train, y_pred_train)
mse_test = calcular_mse(y_test, y_pred_test)
print("Estatística para o Modelo LiNEAR", "\n")
print("Score de Treino:", modelo_linear.score(X_train, y_train), "\n")
print("Score de Teste:", modelo_linear.score(X_test, y_test), "\n")
print("MSE treino:", mse_train, "\n")
print("MSE teste:", mse_test, "\n")
fig, variav=plt.subplots(figsize=(14,6))
leituras = np.linspace(0, len(y_train_a), len(y_train_a))
variav.scatter(np.asarray(y_train_a).reshape(-1),leituras,label='Dados
de Treino',s=60);
variav.scatter(np.asarray(y_pred_train).reshape(-
1),leituras,label='Previsão',s=40);
variav.set_xlabel('Fração Molar(Heptano, Benzeno, Tolueno)')
variav.set_yticklabels([])

plt.legend(fontsize=16, frameon=True, framealpha=0.8,
facecolor='white');
variav.set_title('Modelo Linear- Conjunto de dados de Treinamento',
fontsize=20, fontweight='bold');
coeficientes = modelo_linear.coef_
interceptacao = modelo_linear.intercept_

# Exibir a equação
print(f"Equação da Regressão Linear: y = {coeficientes[0]} * x +
{interceptacao}")
y_test_a = y_test.values.flatten().tolist()
fig, variav=plt.subplots(figsize=(14,6))
leituras = np.linspace(0, len(y_test_a), len(y_test_a))
variav.scatter(np.asarray(y_test_a).reshape(-1),leituras,label='Dados
de teste',s=60);
variav.scatter(np.asarray(y_pred_test).reshape(-
1),leituras,label='Previsão',s=40);
variav.set_xlabel('Fração Molar(Heptano, Benzeno, Tolueno)')
variav.set_yticklabels([])

plt.legend(fontsize=16, frameon=True, framealpha=0.8,
facecolor='white');

```

```

variav.set_title('Modelo Linear- Conjunto de dados de teste',
fontsize=20, fontweight='bold');
print(y_train)
print(y_pred_train)
from sklearn.tree import DecisionTreeRegressor
modelo_flor = DecisionTreeRegressor()
modelo_flor

modelo_flor.fit(X_train, y_train)
modelo_flor
modelo_flor.score(X_train, y_train)
y_pred_train = modelo_flor.predict(X_train)
y_pred_test = modelo_flor.predict(X_test)
# Cálculo de MAE
mae_train = calcular_mae(y_train, y_pred_train)
mae_test = calcular_mae(y_test, y_pred_test)
print("MAE treino:", mae_train)
print("MAE teste:", mae_test)
# Cálculo de MSE
mse_train = calcular_mse(y_train, y_pred_train)
mse_test = calcular_mse(y_test, y_pred_test)
print("Estatística para a Árvore de decisão","\n")
print("Score de Treino:", modelo_flor.score(X_train, y_train),"\n")
print("Score de Teste:", modelo_flor.score(X_test, y_test),"\n")
print("MSE treino:", mse_train,"\n")
print("MSE teste:", mse_test,"\n")
y_train_a = y_train.values.flatten().tolist()
y_test_a = y_test.values.flatten().tolist()
from sklearn.tree import export_text
from sklearn.tree import DecisionTreeRegressor

# Substitua DecisionTreeRegressor pelo seu modelo de regressão
modelo_regressao = DecisionTreeRegressor()

# Substitua X_treino e y_treino pelos seus dados de treinamento
modelo_regressao.fit(X_train, y_train)

# Exporta a árvore de decisão como texto
tree_rules = export_text(modelo_regressao,
feature_names=list(X_train.columns))
print(tree_rules)

from sklearn.tree import export_graphviz
import graphviz

# Substitua DecisionTreeRegressor pelo seu modelo de regressão
modelo_regressao = DecisionTreeRegressor()

# Substitua X_treino e y_treino pelos seus dados de treinamento
modelo_regressao.fit(X_train, y_train)

# Exporta a árvore de decisão em formato DOT
dot_data = export_graphviz(modelo_regressao, out_file=None,
feature_names=list(X_train.columns),

```

```

        filled=True, rounded=True,
special_characters=True)

# Cria e exibe a figura
graph = graphviz.Source(dot_data)
graph.render("arvore_de_decisao") # Isso irá gerar um arquivo PDF
chamado "arvore_de_decisao.pdf"
graph.view("arvore_de_decisao") # Isso irá abrir o visualizador de PDF
padrão

fig, variav=plt.subplots(figsize=(14,6))
leituras = np.linspace(0, len(y_test_a), len(y_test_a))
variav.scatter(np.asarray(y_test_a).reshape(-1),leituras,label='Dados
de teste',s=60);
variav.scatter(np.asarray(y_pred_test).reshape(-
1),leituras,label='Previsão',s=40);
variav.set_xlabel('Fração Molar(Heptano, Benzeno, Tolueno)')
variav.set_yticklabels([])

plt.legend(fontsize=16, frameon=True, framealpha=0.8,
facecolor='white');
variav.set_title('Árvore de decisão- Conjunto de dados de Teste',
fontsize=20, fontweight='bold');
from sklearn.ensemble import RandomForestRegressor
modelo_rf = RandomForestRegressor()
modelo_rf

modelo_rf.fit(X_train, y_train)
modelo_rf
modelo_rf.score(X_train, y_train)
y_pred_train = modelo_rf.predict(X_train)
y_pred_test = modelo_rf.predict(X_test)
# Cálculo de MAE
mae_train = calcular_mae(y_train, y_pred_train)
mae_test = calcular_mae(y_test, y_pred_test)
print("MAE treino:", mae_train)
print("MAE teste:", mae_test)
# Cálculo de MSE
mse_train = calcular_mse(y_train, y_pred_train)
mse_test = calcular_mse(y_test, y_pred_test)
print("Estatística para Floresta aleatória","\n")
print("Score de Treino:", modelo_rf.score(X_train, y_train),"\n")
print("Score de Teste:", modelo_rf.score(X_test, y_test),"\n")
print("MSE treino:", mse_train,"\n")
print("MSE teste:", mse_test,"\n")
y_train_a = y_train.values.flatten().tolist()
fig, variav=plt.subplots(figsize=(14,6))
leituras = np.linspace(0, len(y_train_a), len(y_train_a))
variav.scatter(np.asarray(y_train_a).reshape(-1),leituras,label='Dados
de Treino',s=60);
variav.scatter(np.asarray(y_pred_train).reshape(-
1),leituras,label='Previsão',s=40);
variav.set_xlabel('Fração Molar(Heptano, Benzeno, Tolueno)')
variav.set_yticklabels([])

```

```

plt.legend(fontsize=16, frameon=True, framealpha=0.8,
facecolor='white');
variav.set_title('Floresta aleatória- Conjunto de dados de
Treinamento', fontsize=20, fontweight='bold');
fig, variav=plt.subplots(figsize=(14,6))
leituras = np.linspace(0, len(y_test_a), len(y_test_a))
variav.scatter(np.asarray(y_test_a).reshape(-1),leituras,label='Dados
de teste',s=60);
variav.scatter(np.asarray(y_pred_test).reshape(-
1),leituras,label='Previsão',s=40);
variav.set_xlabel('Fração Molar(Heptano, Benzeno, Tolueno)')
variav.set_yticklabels([])

plt.legend(fontsize=16, frameon=True, framealpha=0.8,
facecolor='white');
variav.set_title('Floresta Aleatória- Conjunto de dados de Teste',
fontsize=20, fontweight='bold');
from sklearn.ensemble import GradientBoostingRegressor
modelo_gbr= GradientBoostingRegressor()
modelo_gbr

y_train_a = y_train[:, 0]
y_test_a = y_test[:, 0]

modelo_gbr.fit(X_train, y_train_a)
modelo_gbr

modelo_gbr.score(X_train, y_train_a)
y_pred_train = modelo_gbr.predict(X_train)
y_pred_test = modelo_gbr.predict(X_test)
# Cálculo de MAE
mae_train = calcular_mae(y_train_a, y_pred_train)
mae_test = calcular_mae(y_test_a, y_pred_test)
print("MAE treino:", mae_train)
print("MAE teste:", mae_test)
# Cálculo de MSE
mse_train = calcular_mse(y_train_a, y_pred_train)
mse_test = calcular_mse(y_test_a, y_pred_test)
print("Estatística para GBR","\n")
print("Score de Treino:", modelo_gbr.score(X_train, y_train_a),"\n")
print("Score de Teste:", modelo_gbr.score(X_test, y_test_a),"\n")
print("MSE treino:", mse_train,"\n")
print("MSE teste:", mse_test,"\n")

```

