



## TÍTULO DA TESE

Antonioni Barros Campos

Exame de Qualificação de Mestrado  
apresentado ao Programa de Pós-graduação em  
Engenharia Elétrica, COPPE, da Universidade  
Federal do Rio de Janeiro, como parte dos  
requisitos necessários à obtenção do título de  
Mestre em Engenharia Elétrica.

Orientadores: José Manoel de Seixas  
Natanael Nunes de Moura  
Junior

Rio de Janeiro  
Março de 2024

## TÍTULO DA TESE

Antonioni Barros Campos

EXAME DE QUALIFICAÇÃO SUBMETIDO AO CORPO DOCENTE DO  
INSTITUTO ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA  
DE ENGENHARIA DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO  
COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO  
GRAU DE MESTRE EM CIÊNCIAS EM ENGENHARIA ELÉTRICA.

Orientadores: José Manoel de Seixas

Natanael Nunes de Moura Junior

Aprovada por: Prof. Nome do Primeiro Examinador Sobrenome

Prof. Nome do Segundo Examinador Sobrenome

Prof. Nome do Terceiro Examinador Sobrenome

Prof. Nome do Quarto Examinador Sobrenome

Prof. Nome do Quinto Examinador Sobrenome

RIO DE JANEIRO, RJ – BRASIL

MARÇO DE 2024

*A alguém cujo valor é digno  
desta dedicatória.*

# Agradecimentos

Gostaria de agradecer a todos.

Resumo do Exame de Qualificação apresentado à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

## TÍTULO DA TESE

Antonioni Barros Campos

Março/2024

Orientadores: José Manoel de Seixas

Natanael Nunes de Moura Junior

Programa: Engenharia Elétrica

Apresenta-se, nesta tese, ...

Abstract of Qualifying Exam presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

THESIS TITLE

Antonioni Barros Campos

March/2024

Advisors: José Manoel de Seixas

Natanael Nunes de Moura Junior

Department: Electrical Engineering

In this work, we present ...

# Sumário

|   |             |
|---|-------------|
| <b>Lista de Figuras</b>   | <b>viii</b> |
| <b>Lista de Tabelas</b>   | <b>ix</b>   |
| <b>Lista de Abreviaturas</b>  | <b>x</b>    |
| <b>1 Introdução</b>   | <b>1</b>    |
| 1.1 Contexto . . . . .  | 1           |
| 1.2 Motivação . . . . .   | 3           |
| 1.3 Objetivo . . . . .  | 4           |
| <b>2 Equilíbrio de Fase no Contexto da Inteligência Computacional</b> | <b>5</b>    |
| 2.1 Simulação Composicional de Reservatórios . . . . .                | 6           |
| 2.2 Equilíbrio de Fases Vapor-Líquido . . . . .                       | 6           |
| 2.2.1 Razões de Equilíbrio . . . . .                                  | 7           |
| 2.2.2 Cálculo de <i>flash</i> . . . . .                               | 8           |
| <b>3 Trabalhos Relacionados</b>                                       | <b>14</b>   |
| 3.1 Cálculo de <i>flash</i> NPT . . . . .                             | 14          |
| 3.2 Cálculo de <i>flash</i> NVT . . . . .                             | 22          |
| 3.3 Cálculo de <i>flash</i> para compostos limpos . . . . .           | 23          |
| <b>4 Método proposto</b>  | <b>25</b>   |
| <b>Referências Bibliográficas</b>                                     | <b>27</b>   |

# Lista de Figuras

|     |  |    |
|-----|--|----|
| 2.1 | Fluxograma de inferência de propriedades de fluidos . . . . .  | 6  |
| 2.2 | Fluxograma da determinação das razões de equilíbrio através da equação de estado . . . . .   | 10 |
| 3.1 | Arquitetura proposta para o classificador . . . . .  | 15 |
| 3.2 | Arquitetura proposta para o inicializador . . . . .  | 16 |
| 3.3 | Diagrama esquemático do funcionamento da PINN. Equações diferenciais parciais da saída da rede neural em função das entradas são utilizadas na função custo, juntamente com condições iniciais e de contorno . . . . . | 17 |
| 3.4 | Rede MLP utilizada para a análise de estabilidade . . . . .  | 18 |
| 3.5 | MLP utilizada para estimativa inicial dos $K_i$ . . . . .  | 19 |
| 3.6 | PINN utilizada para estimativa inicial dos $K_i$ . . . . .   | 19 |
| 3.7 | Diagrama esquemático comparando a modelagem DL-KF e o cálculo de <i>flash</i> convencional . . . . .   | 21 |
| 3.8 | Arquitetura proposta para as Redes Neurais informadas pela Termodinâmica . . . . .   | 24 |
| 4.1 | Pipeline do experimento . . . . .  | 26 |



# Lista de Tabelas

|     |   |    |
|-----|---|----|
| 3.1 | Hiper parâmetros da rede MLP utilizada para a análise de estabilidade | 17 |
| 3.2 | Hiper parâmetros da rede PINN utilizada para a estimativa do $K_i$    | 18 |
| 3.3 | Resultados do caso 1 para diferentes modelos . . . . .                | 22 |
| 3.4 | Resultados do caso 2 para diferentes modelos . . . . .                | 22 |

# Lista de Abreviaturas

|                     |  |
|---------------------|--|
| $K_i$               | Razão de Equilíbrio para o componente $i$ , p. 7                 |
| $M$                 | Número de componentes em uma mistura, p. 24                      |
| $MW$                | Peso Molecular, p. 16  |
| $N$                 | Número de amostras utilizadas para treinamento de modelos, p. 24 |
| $T_c$               | Temperatura crítica, p. 16                                       |
| $VSP$               | Volume Shift Parameter, p. 16                                    |
| $W$                 | Pesos de uma Rede Neural, p. 24                                  |
| $\varepsilon_{p,i}$ | Profundidade potencial do poço de petróleo, p. 16                |
| $k_{ij}$            | Coefficiente de interação binária, p. 12                         |
| $n$                 | Número de mols, p. 16  |
| $p_c$               | Pressão crítica, p. 16   |
| $x_i$               | Fração molar da fase líquida para o componente $i$ , p. 7        |
| $y_i$               | Fração molar da fase gasosa para o componente $i$ , p. 7         |
| $z_i$               | Fração molar total para o componente $i$ , p. 7                  |
| BP                  | British Petroleum, empresa multinacional de petróleo e gás, p. 2 |
| EOS                 | Equation of State (Equação de Estado), p. 4                      |
| MLP                 | Multi-layer Perceptron, p. 16                                    |
| ML                  | <i>Machine Learning</i> , ou Aprendizado de Máquina, p. 5        |
| NPT                 | Composição (N), Pressão (P) e Temperatura (T), p. 14             |

|       |   |
|-------|---|
| NVT   | Composição (N), Volume (V) e Temperatura (T), p. 22                                       |
| PINN  | Physics-informed Neural Network, p. 16  |
| PR    | Equação de Estado Peng-Robinson, p. 5   |
| R     | Constante Universal dos Gases, p. 12  |
| SRK   | Equação de Estado Soave-Redlich-Kwong, p. 5   |
| $T_r$ | Temperatura reduzida, razão entre a temperatura do sistema e a temperatura crítica, p. 11 |
| VLE   | Vapor-Liquid Equilibrium (Equilíbrio Vapor-Líquido), p. 4                                 |
| V     | Volume, p. 12   |
| Z     | Fator de Compressibilidade, p. 12   |

# Capítulo 1

## Introdução

A indústria de petróleo e gás desempenha um papel fundamental no contexto mundial, impulsionando o crescimento econômico e atendendo à demanda global por energia. Essa indústria abrange todas as atividades relacionadas à exploração, produção, refino, distribuição e comercialização de petróleo e gás natural. Para uma prospecção e operações de futuras produções de petróleo, a exploração e produção de petróleo necessita de uma compreensão e gerenciamento de reservatórios de petróleo.

O presente trabalho pretende explorar o contexto da Simulação Composicional de Reservatórios com um olhar da Inteligência Computacional, mais especificamente do Equilíbrio de Fase através do cálculo de *flash* de fluidos de reservatório.

O texto divide-se em uma descrição mais detalhada do problema a ser atacado com algumas definições fundamentais associadas a equilíbrio de fases e estruturação do problema (Capítulo 2) e uma revisão sistemática dos principais trabalhos relacionados ao tema (Capítulo 3).

### 1.1 Contexto

No cenário global, a indústria de petróleo e gás tem enfrentado diversos desafios ao longo dos anos. Um desses desafios é a **volatilidade dos preços do petróleo** [1, 2], que é influenciada por fatores geopolíticos, econômicos e ambientais. Essa volatilidade pode afetar os investimentos e as decisões estratégicas das empresas do setor.

Outro desafio importante é a **transição para uma matriz energética mais sustentável** [1, 2]. Com o aumento das preocupações ambientais e a necessidade de reduzir as emissões de gases de efeito estufa, a indústria de petróleo e gás tem buscado investir em fontes de energia renovável e tecnologias de baixo carbono. Essa transição é fundamental para garantir a sustentabilidade do setor no longo prazo.

Além disso, a indústria de petróleo e gás tem uma importância estratégica para muitos países, especialmente aqueles que possuem grandes reservas de petróleo e

gás natural. Essas reservas podem ser uma fonte de receita significativa e podem contribuir para o desenvolvimento econômico desses países.

No contexto mundial, alguns países se destacam na indústria de petróleo e gás. Os Estados Unidos, por exemplo, é o maior produtor de petróleo do mundo com uma produção crescente de petróleo de xisto. A Arábia Saudita também é um importante *player* no setor e desempenha um papel crucial na estabilidade do mercado global [3].

Outros países, como Rússia, Irã e Iraque, também possuem grandes reservas de petróleo e gás natural e desempenham um papel significativo na produção e exportação desses recursos [3]. Além disso, empresas multinacionais do setor, como a ExxonMobil, Shell e BP, têm uma presença global e desempenham um papel importante na indústria.

No Brasil, a indústria de petróleo e gás desempenha um papel estratégico, sendo uma das principais fontes de receita e desenvolvimento econômico do país. O Brasil possui vastas reservas de petróleo e gás natural, localizadas principalmente na camada pré-sal, o que coloca o país como um dos principais produtores e exportadores desses recursos [4].

Entre 2007 e 2017, a arrecadação do setor de Petróleo e Gás totalizou R\$ 1,4 trilhão e espera-se que, nos próximos dez anos, esse valor alcance mais R\$ 1 trilhão. Essa arrecadação é proveniente de tributos, participação especial, bônus e royalties. Além disso, o setor tem um impacto significativo na economia do país, gerando empregos, renda e estimulando a pesquisa e desenvolvimento local. Também contribui para o superavit na balança comercial e para a formação de um fundo social importante. Para manter essa contribuição e atrair grandes investimentos nos próximos anos, é essencial que o Brasil adote uma agenda positiva, promovendo um ambiente favorável e garantindo segurança jurídica para os investidores [5].

No entanto, a indústria de petróleo e gás no Brasil também enfrenta desafios significativos. A volatilidade dos preços do petróleo, as questões ambientais e as demandas por fontes de energia mais limpas são alguns dos desafios enfrentados pelo setor. A transição para uma matriz energética mais sustentável é uma preocupação crescente e tem impulsionado investimentos em energias renováveis e tecnologias de baixo carbono [5].

A Petrobras, principal empresa de óleo e gás brasileira e uma das principais do mundo, reconhece a importância de intensificar a descarbonização de suas operações, ao mesmo tempo em que garante a segurança energética do Brasil e do mundo. Para isso, a empresa busca repor suas reservas e produzir petróleo, enquanto reduz suas emissões de gases de efeito estufa e desenvolve produtos mais sustentáveis [6].

## 1.2 Motivação

A simulação de reservatórios é uma área da engenharia de reservatórios em que modelos de computador são usados para prever o fluxo de fluidos através de meios porosos representando o comportamento ao longo do tempo. Esses modelos consideram diversos aspectos, como a geometria do reservatório, as propriedades dos fluidos presentes (óleo, gás e água), a permeabilidade das rochas e as condições de pressão e temperatura. Os modelos são usados por empresas de petróleo e gás para o desenvolvimento de campos novos ou desenvolvidos onde as previsões de produção são necessárias para ajudar a tomar decisões de investimento [7]. Uma maneira similar de definir o que é uma simulação de reservatórios é a habilidade de utilizar formulas matemáticas para prever a performance de um reservatório de petróleo ou gás [8].

Existem várias razões pelas quais a simulação de reservatórios é tão relevante nos dias de hoje. Primeiramente, ela ajuda a entender como o petróleo é armazenado e flui dentro do reservatório, permitindo prever a produção futura de petróleo e gás. Com base nessas previsões, as empresas podem tomar decisões estratégicas sobre a viabilidade econômica de um campo de petróleo e planejar o desenvolvimento de poços e infraestrutura.

Além disso, a simulação de reservatórios auxilia no aumento da eficiência da produção. Com a modelagem computacional, é possível testar diferentes estratégias de produção, como a injeção de água ou gás para aumentar a pressão do reservatório e melhorar a recuperação de petróleo. Esses testes virtuais permitem otimizar a operação do campo, maximizando a produção e reduzindo os custos.

Outra relevância da simulação de reservatórios é a avaliação de riscos e incertezas. Os modelos permitem a análise de diferentes cenários e a consideração de incertezas geológicas e operacionais. Isso possibilita uma melhor compreensão dos riscos envolvidos na exploração e produção de petróleo, ajudando a mitigar problemas potenciais e a tomar decisões mais informadas.

Por fim, a simulação de reservatórios também contribui para a sustentabilidade e a preservação do meio ambiente. Ao prever a produção de petróleo e gás, é possível planejar o uso mais eficiente dos recursos naturais e reduzir o desperdício. Além disso, a modelagem computacional permite avaliar o impacto ambiental das atividades de exploração e produção, auxiliando no desenvolvimento de práticas mais seguras e sustentáveis.

Dentro da simulação de reservatórios, a **simulação composicional de reservatórios** é utilizada para avaliar o comportamento do fluxo composicional, levando em consideração todas as fases presentes, como função de pressão, tempo e posição. Também é capaz de estimar a recuperação de óleo para um determinado esquema

de produção existente e avaliar o impacto nas condições operacionais alteradas na recuperação [9].

Nos últimos anos, houve um crescente interesse na utilização de equações de estado (EOSs) para calcular o equilíbrio de misturas. A EOS de Peng-Robinson, considerada uma EOS realista, tem sido amplamente adotada devido ao seu histórico de aprimoramento e aplicabilidade em diversas áreas. Os procedimentos de cálculo usando EOSs têm sido alvo de estudos e modificações constantes. Os parâmetros da EOS, que descrevem a concentração e interação entre as misturas binárias, desempenham um papel crucial na precisão da correlação do equilíbrio líquido-vapor (VLE). Geralmente, esses parâmetros são obtidos por meio do ajuste de dados experimentais à temperatura desejada para o VLE. No entanto, é importante ressaltar que, devido à necessidade de iterações, a utilização de EOSs pode ser menos adequada em aplicações que requerem resultados rápidos [10].

Em essência, o Aprendizado Profundo pode atuar como um aproximador de funções, sendo capaz de estimar de forma implícita o processo físico da VLE (Equilíbrio Líquido-Vapor). Além disso, a maturidade tanto do hardware quanto do desenvolvimento de software nessa área pode ser de grande auxílio para acelerar os cálculos complexos de *flash* [10].

## 1.3 Objetivo

O presente trabalho tem como objetivo investigar o cálculo de *flash* de equilíbrio vapor-líquido para reservatórios de petróleo sob o olhar da inteligência computacional. Ou seja, utilizar modelos baseado em dados para auxiliar ou substituir os cálculos de *flash* convencionais.

Uma parte do trabalho tem o intuito de analisar se os modelos baseado em dados tem consistência fenomenológica quando comparados com os métodos convencionais. Sendo isso avaliado, um grande impacto que os modelos baseado em dados podem trazer é no tempo computacional. O trabalho também pretende fazer uma análise comparativa entre os tempos de execução dos dois métodos.

Para o cálculo de *flash* convencional, será utilizado uma ferramenta de código aberto desenvolvida pela Equinor chamada NeqSim <sup>1</sup>. NeqSim é uma biblioteca para estimativa de comportamento e propriedades dos fluidos.

---

<sup>1</sup><https://equinor.github.io/neqsimhome/>

## Capítulo 2

# Equilíbrio de Fase no Contexto da Inteligência Computacional

O Aprendizado de Máquina está em alta na pesquisa de propriedades termo-físicas, principalmente devido pelo progresso dos algoritmos de ML e a disponibilidade [11].

Os dados são geralmente utilizados no desenvolvimento de modelos matemáticos permeando entre dois extremos: Num extremo, modelos que não possuem nenhuma informação física e são obtidos puramente com dados. No outro extremo, modelos baseados em teorias físicas mas que ainda contém parâmetros que precisam ser ajustados utilizando alguns dados. Permeando nesse espectro, a mais importante característica do modelo é sua capacidade preditiva [11].

Uma equação de estado, como a PR [12] e a SRK [13, 14], é comumente usada para descrever o comportamento de fases fluidas em uma composição. Esses modelos são amplamente utilizados na indústria para prever propriedades termodinâmicas de misturas de fluidos. No entanto, os cálculos de equilíbrio de fases baseados em equações de estado geralmente envolvem a resolução de sistemas de equações não lineares, o que pode tornar os cálculos mais complexos e demorados em comparação com outros modelos de fluidos [15].

Dentro da inteligência computacional, aplicações vem demonstrando melhora na performance na predição, estimação e otimização de diferentes funções objetivo em problemas na indústria de óleo e gás (como exemplo, pressões mínimas de miscibilidade, taxa de produção de petróleo, precipitação de asfalto ao redor do poço, colocação de poço e caracterização de reservatório) [16].

Com o ganho de poder computacional, principalmente na última década, algoritmos de Aprendizado de Máquina e Aprendizado Profundo ganharam força em várias aplicações práticas. O aprendizado de máquina pode realizar previsões a um custo computacional muito menor em comparação com o tempo inicial necessário para realizar o cálculo, sem comprometer a precisão [17]. Além disso, em comparação



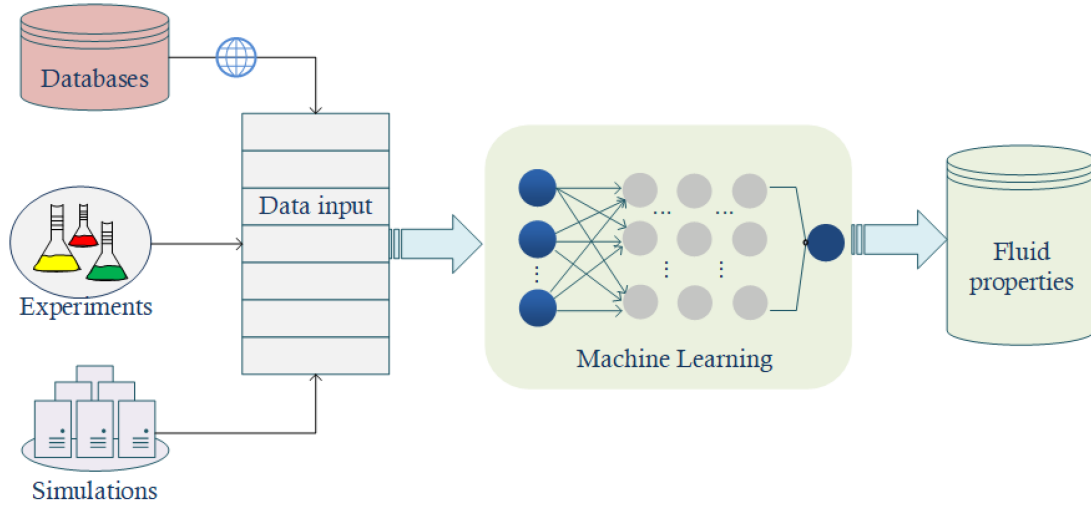


Figura 2.1: Fluxograma de inferência de propriedades de fluidos [17]

com simulações numéricas baseadas em física, o aprendizado de máquina pode ser mais eficiente computacionalmente [17].

Num primeiro momento, utilizando a figura 2.1 como referência, o presente trabalho dispõe, para formar os dados de entrada, de dados de uma base de dados (oriundos de análises de amostras de petróleo) e de simulação, e pretende retornar as composições para as fases existente em equilíbrio.

## 2.1 Simulação Composicional de Reservatórios

Durante a simulação de um reservatório, o reservatório é dividido em blocos e cada bloco possui propriedades como permeabilidade, porosidade, pressão, temperatura e composição do fluido. Essas propriedades podem variar de bloco para bloco e ao longo do tempo durante a simulação. O processo de deslocamento do fluido é descrito por equilíbrios termodinâmicos aplicados a cada bloco, permitindo uma análise detalhada do comportamento do reservatório e do seu desempenho sob diferentes condições operacionais [9].

## 2.2 Equilíbrio de Fases Vapor-Líquido

Uma fase é uma parte do sistema que apresenta propriedades físicas e químicas uniformes, composição homogênea e está separada de outras fases por superfícies definidas. Na produção de petróleo, as fases mais relevantes são a fase líquida de hidrocarboneto e a fase gasosa. Essas fases podem coexistir em equilíbrio quando as variáveis que descrevem as mudanças no sistema permanecem constantes ao longo do

tempo e da posição. As principais variáveis que determinam o estado de equilíbrio são **temperatura**, **pressão** e **composição** do sistema [8].

### 2.2.1 Razões de Equilíbrio

Em sistemas com vários componentes, a razão de equilíbrio  $K_i$  de um componente é a relação entre sua fração molar na fase gasosa ( $y_i$ ) e sua fração molar na fase líquida ( $x_i$ ). Essa relação é expressa matematicamente como:

$$K_i = \frac{y_i}{x_i} \quad (2.1)$$

Onde o  $K_i$  representa a Razão de Equilíbrio,  $y_i$  a fração molar da fase gasosa e  $x_i$  a fração molar da fase líquida para o componente  $i$ .

Além das Razões de Equilíbrio, outras relações básicas são definidas. De maneira conveniente, todos os cálculos de equilíbrio de fases são baseados em 1 mol de mistura de hidrocarbonetos, resultando nas seguintes relações.

$$z_i = x_i n_L + y_i n_V \quad (2.2)$$

$$\sum_i z_i = 1 \quad (2.3)$$

$$\sum_i x_i = 1 \quad (2.4)$$

$$\sum_i y_i = 1 \quad (2.5)$$

Combinando equações 2.1 e 2.2 para eliminar  $y_i$ , temos:

$$z_i = x_i n_L + (x_i K_i) n_V \quad (2.6)$$

E resolvendo para  $x_i$ :

$$x_i = \frac{z_i}{n_L + n_V K_i} \quad (2.7)$$

A equação 2.2 pode também ser resolvida para  $y_i$ , combinando com a equação 2.1 para eliminar  $x_i$ , obtendo:

$$y_i = \frac{z_i K_i}{n_L + K_i n_V} \quad (2.8)$$

Combinando as equações 2.4 e 2.7 e equações 2.5 e 2.8 chegamos ao seguinte resultado:

$$\sum_i x_i = \sum_i \frac{z_i}{n_L + n_V K_i} \quad (2.9)$$

$$\sum_i y_i = \sum_i \frac{z_i K_i}{n_L + K_i n_V} \quad (2.10)$$

Como  $\sum_i y_i - \sum_i x_i = 0$ :

$$\sum_i \frac{z_i K_i}{n_L + K_i n_V} - \sum_i \frac{z_i}{n_L + n_V K_i} = 0 \quad (2.11)$$

$$\sum_i \frac{z_i (K_i - 1)}{n_L + K_i n_V} = 0 \quad (2.12)$$

$$f(n_V) = \sum_i \frac{z_i (K_i - 1)}{n_V (K_i - 1) + 1} = 0 \quad (2.13)$$

O conjunto de equações acima são suficientes para representar as relações de fase para realizar os cálculos de composição de um sistema de hidrocarbonetos. Esses cálculos são conhecidos como cálculos de *flash* [8].

### 2.2.2 Cálculo de *flash*

O cálculo de *flash* é fundamental em todas as análises de engenharia de reservatórios e processos. Eles são necessários para determinar as quantidades de líquido e gás de hidrocarbonetos que coexistem em um reservatório, levando em consideração a pressão e temperatura específicas. Esses cálculos também são realizados para determinar a composição das fases de hidrocarbonetos presentes. Com base na composição geral de um sistema de hidrocarbonetos em uma determinada pressão e temperatura, os cálculos de *flash* são realizados com o objetivo de determinar:

- Fração da fase gasosa  $n_V$
- Fração da fase líquida  $n_L$
- Composição da fase gasosa  $y_i$
- Composição da fase líquida  $x_i$

A sequencia de passos para realizar o cálculo de *flash* é descrito a seguir.

1. **Cálculo de  $n_V$ :** A equação 2.13, conhecida como equação de Rachford-Rice, pode ser resolvida para  $n_V$  usando o método iterativo de Newton-Raphson. Aplicando o método:

- Assume-se um valor para  $n_V$  entre 0 e 1 (por exemplo  $n_V = 0,5$ ). Um bom valor inicial pode ser calculado pela seguinte relação:

$$n_V = \frac{A}{A - B} \quad (2.14)$$

Onde,

$$A = \sum_i [z_i (K_i - 1)] \quad (2.15)$$

$$B = \sum_i \left[ \frac{z_i (K_i - 1)}{K_i} \right] \quad (2.16)$$

- Avalia a função  $f(n_V)$  dada pela equação 2.13 usando o valor inicial de  $n_V$  dado pelo passo anterior.
- Se o valor absoluto de  $f(n_V)$  for menor que uma tolerância preestabelecida  $\varepsilon$  (exemplo  $\varepsilon = 10^{-15}$ ), então  $n_V$  é a solução desejada.
- Caso o valor absoluto seja maior do que a tolerancia, aplique um passo de iteração do método de Newton-Raphson a partir da seguinte relação:

$$n_V^{i+1} = n_V^i - \frac{f(n_V^i)}{f'(n_V^i)} \quad (2.17)$$

Onde,

$$f'(n_V) = - \sum_i \left[ \frac{z_i (K_i - 1)^2}{[n_V (K_i - 1) + 1]^2} \right] \quad (2.18)$$

O valor de  $n_V^{i+1}$  é o novo valor de  $n_V$  que será utilizado na próxima iteração.

- O procedimento acima é repetido até o valor absoluto de  $f(n_V)$  atingir valores menores do que a tolerância  $\varepsilon$  (convergência atingida).

2. **Cálculo de  $n_L$ :** Por estarmos trabalhando em termos de fração (considerando 1 mol), a fração da fase líquida pode ser determinada a partir da equação:

$$n_L = 1 - n_V \quad (2.19)$$

3. **Cálculo de  $x_i$ :** A composição da fase líquida pode ser obtida pela equação 2.7.

$$x_i = \frac{z_i}{n_L + n_V K_i} \quad (2.20)$$

4. **Cálculo de  $y_i$ :** A composição da fase gasosa pode ser obtida pela equação 2.8.

$$y_i = \frac{z_i K_i}{n_L + K_i n_V} \quad (2.21)$$

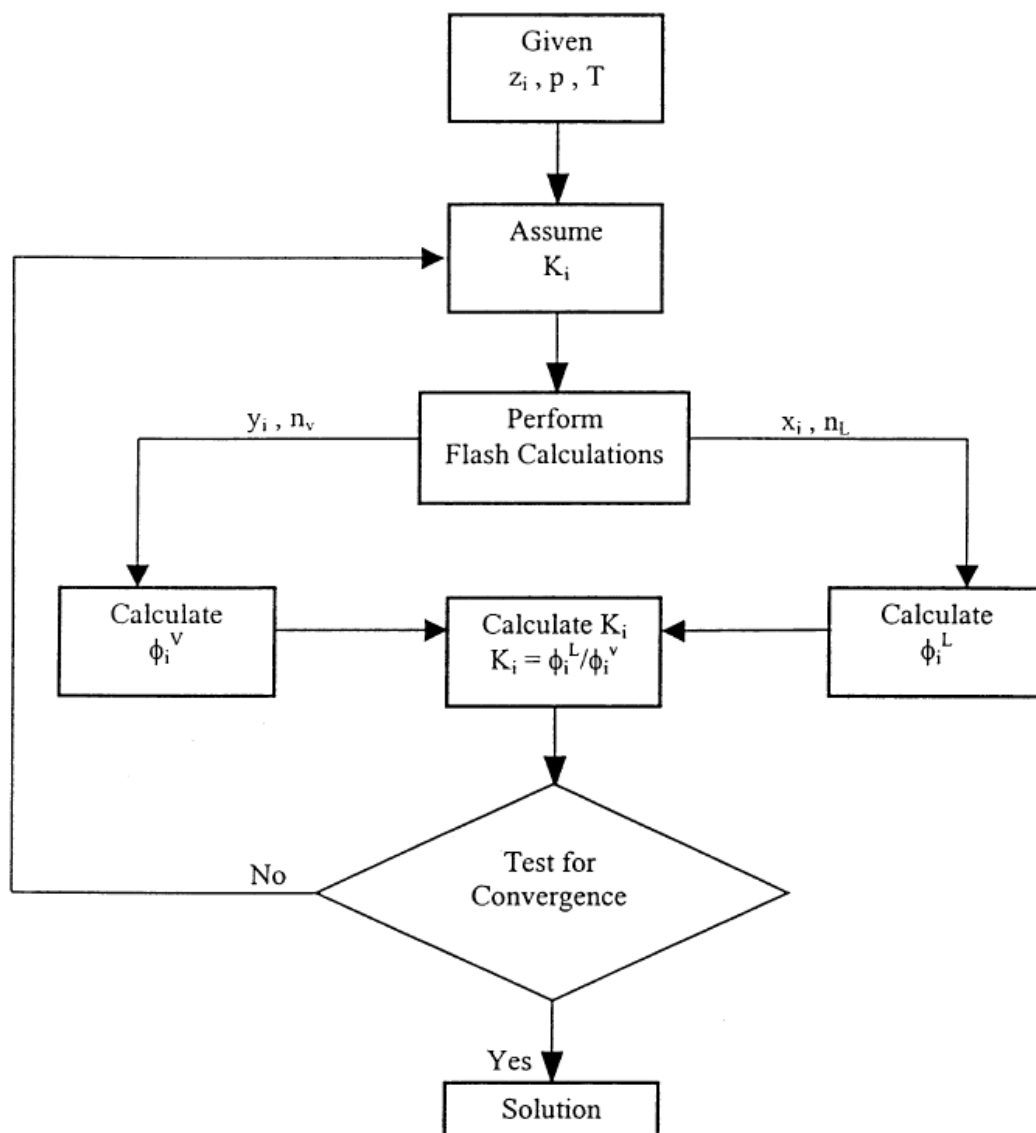


Figura 2.2: Fluxograma da determinação das razões de equilíbrio através da equação de estado [8]

Para a obtenção das razões de equilíbrio necessárias para os cálculos de *flash*, é preciso ser utilizada equação de estado para descrever as relações entre as propriedades das misturas.

O fluxograma da figura 2.2 detalha o cálculo para obtenção da convergência das razões de equilíbrio para misturas reais, com a utilização de uma equação de estado, que fornece os resultados oriundos do cálculo de *flash* ( $n_V$ ,  $n_L$ ,  $y_i$  e  $x_i$ ).

Os passos são os seguintes:

1. Assume-se uma premissa de valor inicial para as razões de equilíbrio para cada

componente da mistura. A equação de Wilson pode fornecer esses  $K_i$  iniciais.

$$K_i^A = \frac{p_{ci}}{p} \exp \left[ 5,37 (1 + \omega_i) \left( 1 - \frac{T_{ci}}{T} \right) \right] \quad (2.22)$$

Onde  $K_i^A$  = Valor inicial da razão de equilíbrio do componente i.

2. Utilizando a composição geral  $\mathbf{z}$  e os  $K$  iniciais, realiza-se o cálculo de *flash* para determinar  $n_V$ ,  $n_L$ ,  $y_i$  e  $x_i$ .
3. Usando a composição calculada para ambas as fases, determinar os coeficientes de fugacidade  $\Phi_i^L$  e  $\Phi_i^V$ .

O conceito do coeficiente de fugacidade, juntamente com as equações de estado, surgem quando se quer tratar misturas reais de componentes e está intrinsecamente ligadas aos cálculos de equilíbrio de fases. Fisicamente, a fugacidade de um componente i em uma fase em comparação a fugacidade desse componente na outra fase é a medida do potencial de transferência desse componente entre fases. A fase com menor fugacidade recebe do componente da fase com maior fugacidade. Fugacidades iguais indicam zero transferência. Fugacidades iguais para todos os componentes da mistura indicam um equilíbrio termodinâmico.

O conceito da equação de estado surgiu da necessidade de uma expressão analítica para relacionar pressão  $p$ , temperatura  $T$  e volume  $V$ . As mais famosas equações de estado para descrição de misturas reais são as equações de estado empíricas cúbicas. Uma das mais significativas é a equação de Soave-Redlich-Kwong [14] que é descrita pelas equações a seguir.

$$p = \frac{RT}{V - b} - \frac{a\alpha}{V(V + b)} \quad (2.23)$$

$$\alpha = \left[ 1 + m \left( 1 - \sqrt{T_r} \right) \right]^2 \quad (2.24)$$

$$m = 0,480 + 1,574\omega + 0,176\omega^2 \quad (2.25)$$

Onde T é a temperatura do sistema em °R.

Introduzindo o fator de compressibilidade, a equação de estado SRK torna-se:

$$Z^3 - Z^2 + (A - B - B^2) Z - AB = 0 \quad (2.26)$$

Onde,

$$A = \frac{(a\alpha)p}{(RT)^2} \quad (2.27)$$

$$B = \frac{bp}{RT} \quad (2.28)$$

Onde,  $p$  é a pressão em psia,  $T$  é a temperatura em °R e  $R = 10,730 \text{ psia} \cdot \text{ft}^3 / \text{lb} \cdot \text{mol} \cdot ^\circ \text{R}$

Na prática, para uma mistura de hidrocarbonetos, o coeficiente de fugacidade de um componente  $i$  foi definido por Reid, Prausnitz e Sherwood (1977) como:

$$\ln \Phi_i = \left( \frac{1}{RT} \right) \left[ \int_V^\infty \left( \frac{\partial p}{\partial n_i} - \frac{RT}{V} \right) dV \right] - \ln Z \quad (2.29)$$

E, utilizando como exemplo a equação de estado SRK, para a fase líquida:

$$\ln \Phi_i^L = \frac{b_i (Z^L - 1)}{b_m} - \ln (Z^L - B) - \left( \frac{A}{B} \right) \left[ \frac{2\Psi_i}{(a\alpha)_m} - \frac{b_i}{b_m} \right] \ln \left[ 1 + \frac{B}{Z^L} \right] \quad (2.30)$$

Onde,

$$\Psi_i = \sum_j [x_j \sqrt{a_i a_j \alpha_i \alpha_j} (1 - k_{ij})] \quad (2.31)$$

$$(a\alpha)_m = \sum_i \sum_j [x_i x_j \sqrt{a_i a_j \alpha_i \alpha_j} (1 - k_{ij})] \quad (2.32)$$

E para fase gasosa:

$$\ln \Phi_i^V = \frac{b_i (Z^V - 1)}{b_m} - \ln (Z^V - B) - \left( \frac{A}{B} \right) \left[ \frac{2\Psi_i}{(a\alpha)_m} - \frac{b_i}{b_m} \right] \ln \left[ 1 + \frac{B}{Z^V} \right] \quad (2.33)$$

Onde,

$$\Psi_i = \sum_j [y_j \sqrt{a_i a_j \alpha_i \alpha_j} (1 - k_{ij})] \quad (2.34)$$

$$(a\alpha)_m = \sum_i \sum_j [y_i y_j \sqrt{a_i a_j \alpha_i \alpha_j} (1 - k_{ij})] \quad (2.35)$$

4. Calcula-se o novo conjunto de razões de equilíbrio através da relação:

$$K_i = \frac{\Phi_i^L}{\Phi_i^V} \quad (2.36)$$

5. Verifica-se se a solução convergiu através da seguinte comparação entre  $K_i$  e uma tolerância  $\varepsilon$ :

$$\sum_{i=1}^M \left[ \frac{K_i}{K_i^A} - 1 \right]^2 \leq \varepsilon \quad (2.37)$$

Se as condições da etapa 5 forem satisfeitas, a solução foi alcançada. Caso não, uma nova iteração é realizada utilizando os  $K_i$  calculados na iteração anterior.

Devido a característica não linear das equações que regem o equilíbrio de fases de misturas reais conjugado ao método de cálculo em si, que possui dois processos iterativos, tornando o cálculo mais lento, vários trabalhos vem sendo feito na criação de modelos de rede neural e aprendizado profundo para auxiliar nos cálculos de equilíbrio de fase.

Para uma variedade de problemas na engenharia de petróleo, as Redes Neurais têm sido a escolha usual para treinamento e inferência [17]. Tal fato pode ser constatado nas publicações feitas nos últimos anos no contexto dos cálculos de equilíbrio de fases.



# Capítulo 3

## Trabalhos Relacionados

Os trabalhos revisados a seguir foram retirados dos sites de busca de artigos científicos **ScienceDirect** e **Google Scholar** nos períodos entre o ano de 2020 até publicações no ano de 2024 utilizando as palavras-chave "machine", "learning", "deep", "learning", "flash", "calculations", "phase" e "equilibrium".

### 3.1 Cálculo de *flash* NPT

Na área de cálculos de *flash* utilizando Composição, Pressão e Temperatura, o trabalho de QU *et al.* [18] visa acelerar o cálculo de *flash* utilizando redes neurais para resolver duas tarefas distintas:

- Classificar se um conjunto  $P$ ,  $T$  e  $\mathbf{z}$  é estável ou não, ou seja, possui mais de uma fase ou não. Esse problema é de classificação binária e esse modelo substitui o teste de estabilidade normalmente empregado antes da divisão de fases para acelerar o processo.
- Inicializar as razões de equilíbrio  $K_i$  para o problema de divisão das fases convencional. Uma rede neural inicializa os  $K_i$  de maneira mais precisa reduzindo o número de iterações na etapa de divisão das fases.

Na geração dos dados, foram usadas como referência quatro tipos de petróleo (*Wet Gas*, *Gas Condensate*, *Volatile Oil* e *Black Oil*), ajustados distribuições de Dirichlet baseados nas composições características desses petróleos e gerados um conjunto de amostras utilizando a técnica de amostragem do Hipercubo Latino [19]. Os dados foram divididos em três estudos de caso: um primeiro com apenas dois componentes ( $CH_4$  e  $C_6H_{14}$ , um segundo caso envolvendo quatro componentes ( $CH_4$ ,  $C_2H_6$ ,  $C_3H_8$  e  $C_4H_{10}$ ). O terceiro estudo de caso considera nove componentes.

O classificador possui três camadas ocultas com 32 neurônios e foi utilizado a função de ativação SiLU. A camada de saída possui um único neurônio com uma

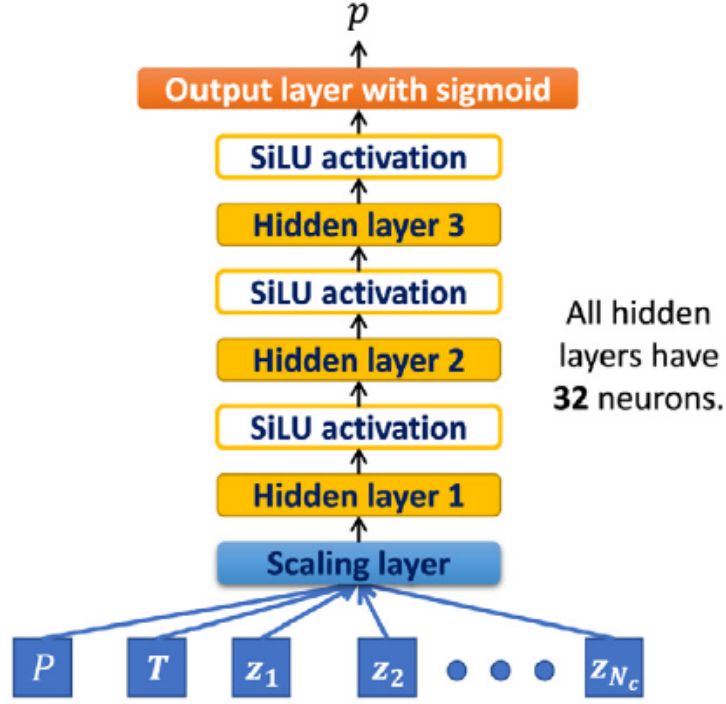


Figura 3.1: Arquitetura proposta para o classificador [18]

função de ativação sigmoide. A função custo utilizada foi a entropia cruzada. A rede neural retorna a probabilidade do fluido, com composição  $\mathbf{z}$  e nas condições de temperatura  $T$  e pressão  $P$  ser estável (única fase). A arquitetura está ilustrada na Figura 3.1. A performance do classificador no conjunto de teste foi de  $bce = 0,002$  e acurácia de 99,93%.

O inicializador por sua vez recebe as mesmas *features* como entrada mas a saída são as razões de equilíbrio iniciais no formato  $\ln K_i$  devido ao fato dos valores de  $K_i$  estarem em diferentes ordens de magnitude. A arquitetura da rede tem uma camada oculta e três blocos residuais. Cada bloco residual possui duas camadas ocultas e uma conexão somando a entrada da primeira camada a saída da segunda camada. Todas as camadas ocultas possuem 64 neurônios e função de ativação SiLU. A camada de saída não tem função de ativação. Após os blocos residuais, um bloco concatena a saída do ultimo bloco com as entradas da rede. Segundo CHENG *et al.* [20], essa abordagem permite a rede aprender regras simples. No caso do *flash* NPT, as variáveis de pressão  $P$  e temperatura  $T$  tem grande impacto nos valores de  $K_i$ . A função custo empregada no treinamento da rede inicializadora é o erro médio absoluto.

A performance da rede inicializadora no conjunto de teste é de  $mae = 9,66e^{-4}$  em termos de  $\ln K_i$  e  $mae = 1,86e^{-3}$  em termos de  $K_i$ .

O trabalho de CHEN *et al.* [21] e CHEN *et al.* [22] segue a mesma linha de

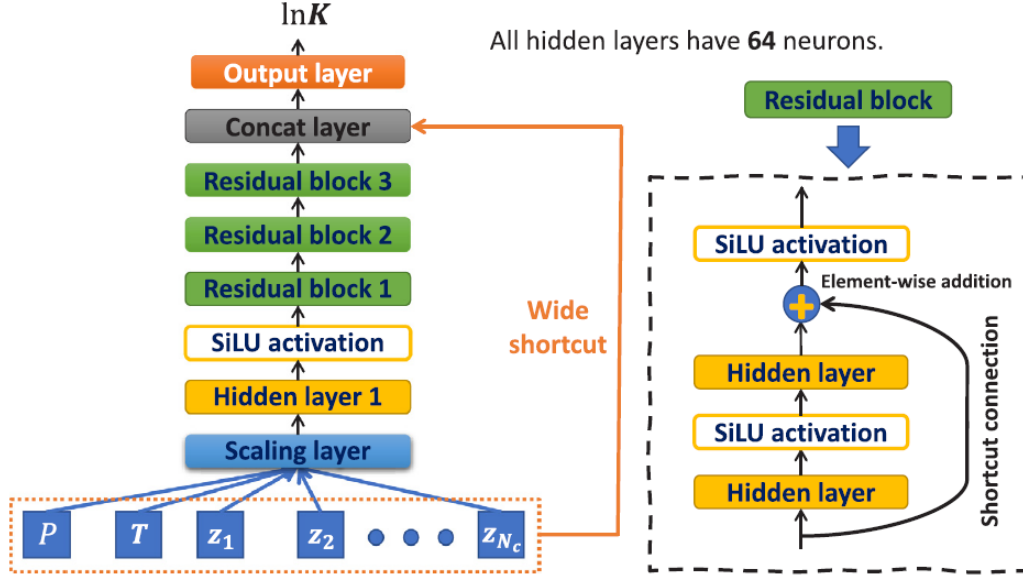


Figura 3.2: Arquitetura proposta para o inicializador [18]

substituir os métodos do teste de estabilidade e o cálculo de *flash* propriamente por modelos baseados em dados. Especificamente nesse trabalho, foi proposto uma metodologia de utilizar a equação de estado de Peng-Robinson modificada (PR-C), para fluidos de hidrocarbonetos aonde a pressão de capilaridade tem um efeito considerável, para geração dos dados. Para a tarefa de análise de estabilidade da mistura, foi utilizado um modelo *multi-layer perceptron* para resolver um problema de classificação binária, onde as classes indicam se a mistura é de fase única ou bifásica. No cálculo de *flash*, são comparadas as performances de uma MLP e uma *physics-informed neural network* para uma estimativa inicial de  $K_i$ . Um esquema ilustrativo da PINN é mostrado na figura 3.3.

Para a geração do modelo, foram selecionados seis pseudo-componentes para representar diferentes misturas de hidrocarbonetos ( $C_1-N_2$ ,  $C_2-CO_2-C_3$ ,  $C_{4-6}$ ,  $C_{7-12}$ ,  $C_{13-24}$ ,  $C_{25+}$ ). São utilizados os parâmetros da equação de estado PR-C ( $n_i$ ,  $p_{c,i}$ ,  $T_{c,i}$ ,  $MW_i$ ,  $\omega_i$ ,  $VSP_i$ ,  $k_{ij}$  e  $\varepsilon_{p,i}$ ) como entrada do modelo ao invés da composição do fluido, além das condições do reservatório ( $p$ ,  $T$  e  $r_p$ ). Cinco dos seis parâmetros da equação de estado possuem seis valores (um para cada pseudo-componente) e o parâmetro  $k_{ij}$  é uma matriz 6x6 simétrica somados aos três parâmetros inerentes ao reservatório fornece uma entrada para as redes neurais com dimensão 60.

Nesse trabalho, para incluir várias configurações de misturas de hidrocarbonetos, uma grande quantidade de amostras foram geradas num total de 123.050 amostras utilizando o método de amostragem Hipercubo Latino [19].

A estrutura otimizada da rede MLP para a análise de estabilidade é mostrada na figura 3.4. Os hiper parâmetros utilizados no treinamento dos modelos estão

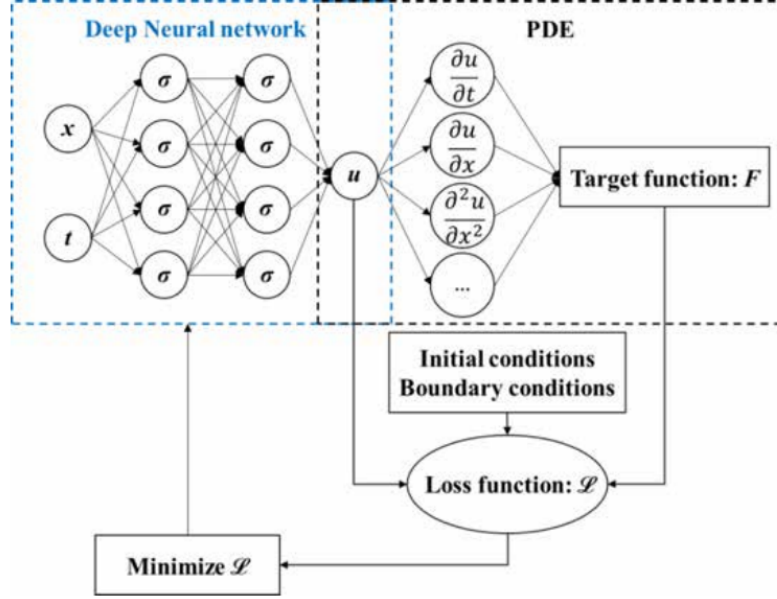


Figura 3.3: Diagrama esquemático do funcionamento da PINN. Equações diferenciais parciais da saída da rede neural em função das entradas são utilizadas na função custo, juntamente com condições iniciais e de contorno [22]

| Hiper parâmetro      | Valor            |
|----------------------|------------------|
| <i>Batch size</i>    | 64               |
| <i>Epochs</i>        | 80               |
| <i>Learning Rate</i> | 0,001            |
| Otimizador           | Adam             |
| Função custo         | Entropia cruzada |

Tabela 3.1: Hiper parâmetros da rede MLP utilizada para a análise de estabilidade [22]

listados na tabela 3.1.

Para validação da performance dos modelos gerados, foi utilizado uma estratégia de validação cruzada utilizando 10 *folds*. Para o melhor modelo, a acurácia no conjunto de teste foi de 96,96%.

Em relação ao tempo computacional para a análise de estabilidade, o cálculo de equilíbrio convencional teve médias de 269,56 segundos enquanto a rede neural 0,28 segundos.

Para a estimativa dos  $K_i$  foram comparados modelos MLP e PINN. A arquitetura da rede MLP utilizada é mostrada na figura 3.5. Utilizou-se os mesmos parâmetros de entrada da MLP da análise de estabilidade e a saída é um vetor de seis dimensões, uma para o  $K_i$  de cada pseudo-componente. Foi utilizada a mesma estratégia de validação cruzada 10-*fold* e uma quantidade de 52.102 amostras. A tabela 3.2 mostra

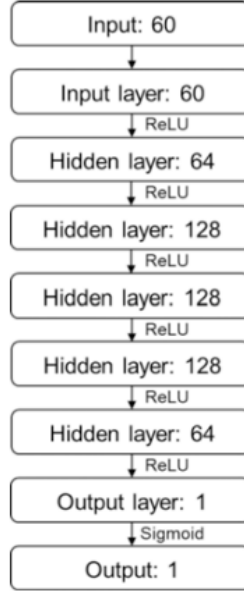


Figura 3.4: Rede MLP utilizada para a análise de estabilidade [22]

| Hiper parâmetro      | Valor            |
|----------------------|------------------|
| <i>Batch size</i>    | 128              |
| <i>Epochs</i>        | 40               |
| <i>Learning Rate</i> | 0,001            |
| Otimizador           | Adam             |
| Função custo         | Entropia cruzada |

Tabela 3.2: Hiper parâmetros da rede PINN utilizada para a estimativa do  $K_i$  [22]

os hiper parâmetros da MLP com melhores resultados na validação cruzada.

A arquitetura da rede PINN utilizada é mostrada na figura 3.6. A arquitetura da rede está representado na área tracejada da figura. Para a PINN, a função custo não necessita dos valores reais de  $K_i$ . Nesse caso, as estimativas iniciais de  $K_i$  geradas pela rede são utilizadas para calcular  $A_{total}$  (soma das energias livres de Helmholtz das duas fases. Minimizando  $A_{total}$ , o sistema está em equilíbrio) utilizando a equação de estado PR-C. A razão da função de  $A_{total}$  estar sendo dividida por  $10^7$  é para que  $A_{total}$  e os  $K_i$  estejam na mesma ordem de grandeza. Sem o escalonamento, uma pequena mudança nos  $K_i$  pode alterar significativamente o  $A_{total}$  e se o  $A_{total}$  variar significativamente, o gradiente muda significativamente também tornando o treinamento instável.

Para avaliar a performance dos modelos na estimativa dos  $K_i$ , foi utilizado o erro médio absoluto, como descrito na equação abaixo.

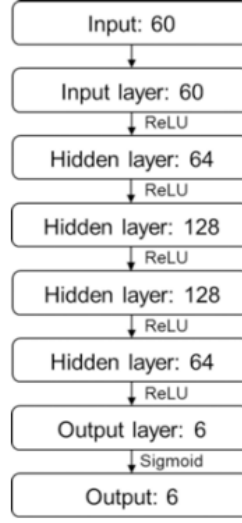


Figura 3.5: MLP utilizada para estimativa inicial dos  $K_i$  [22]

$$\mathcal{L} = \frac{1}{ns} \sum_{j=1}^{ns} (A_{total,j}/10^7)$$

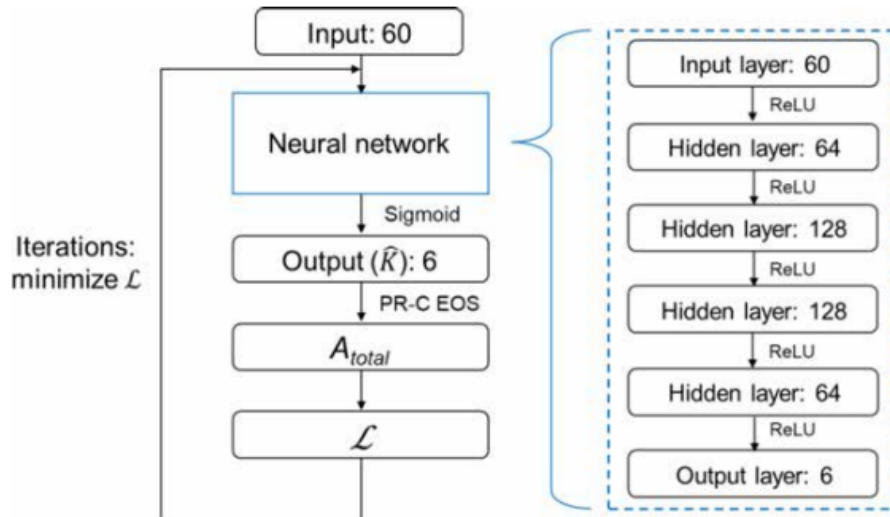


Figura 3.6: PINN utilizada para estimativa inicial dos  $K_i$  [22]

$$MAE = \frac{1}{ns \cdot nc} \sum_{j=1}^{ns} \sum_{i=1}^{nc} \left| \frac{\hat{K}_{i,j} - K_{i,j}}{K_{i,j}} \right| \quad (3.1)$$

Onde  $nc$  é a quantidade de pseudo-componentes e  $\hat{K}_{i,j}$  representa o valor predito da razão de equilíbrio do pseudo-componente  $j$  na amostra  $i$ .

Foram comparados quatro métodos de estimativa de  $K_i$  iniciais. Os valores obtidos pela análise de estabilidade convencional (método A), pelo modelo MLP com o conjunto de dados pequeno de 1.000 amostras (método B), pelo modelo MLP com o conjunto de dados grande com as 52.102 amostras (método C), pela PINN com 1.000 amostras (método D). O MAE do método A foi relativamente alto enquanto pelos métodos C e D reduziram o erro em 77,26% e 82,12% respectivamente, comparando com o método A. Já o método B piorou os resultados do método A.

O objetivo principal do trabalho era de diminuir o tempo computacional gasto pelos métodos de cálculo de equilíbrio de fase líquido-vapor convencionais. Na última seção, o trabalho compara os tempos de execução entre o método convencional e o assistido por ML. O tempo médio gasto no convencional foi de 269,80 segundos enquanto o assistido por ML de 0,43 segundos.

O trabalho de ZHENG *et al.* [23] propõe-se a investigar a modificação do cálculo de *flash* convencional inserindo três redes neurais. O modelo ANN-STAB (Rede neural para teste de estabilidade), o modelo ANN-KV (Rede neural para cálculo das razões de equilíbrio  $K$ ) e o modelo ANN-FUG (Rede neural para cálculo de fugacidade). O nome proposto para a solução foi modelagem DF-KF (*Deep Learning for K-Values and Fugacity Calculation*).

Devido o teste de estabilidade, a estimativa dos valores  $K$  e o cálculo dos coeficientes de fugacidade afetarem a eficiência computacional do cálculo de *flash*, foi proposto um cálculo de *flash* baseado em dados ilustrado na figura 3.7.

Para a preparação dos dados para os modelos, foram criados conjuntos separados de dados de treinamento (80% do total), validação (10% do total) e teste (10% do total). Os dados de treinamento foram usados para ajustar os pesos das redes neurais, os dados de validação para ajustar os hiper parâmetros, como o número de camadas ocultas e neurônios das redes, por exemplo e os dados de teste para avaliar a performance dos modelos.

Para a arquitetura dos modelos, chegou-se ao valor de duas camadas ocultas e de 30 neurônios que trouxeram um melhor equilíbrio entre erro médio relativo quadrático e tempo computacional.

Para a saída dos modelos, são retornados a fração de vapor, os valores das razões de equilíbrio  $K_i$  e os coeficientes de fugacidade dos componentes para os modelos ANN-STAB, ANN-KV e ANN-FUG respectivamente.

O método foi comparado com outros quatro modelos: a modelagem original base-

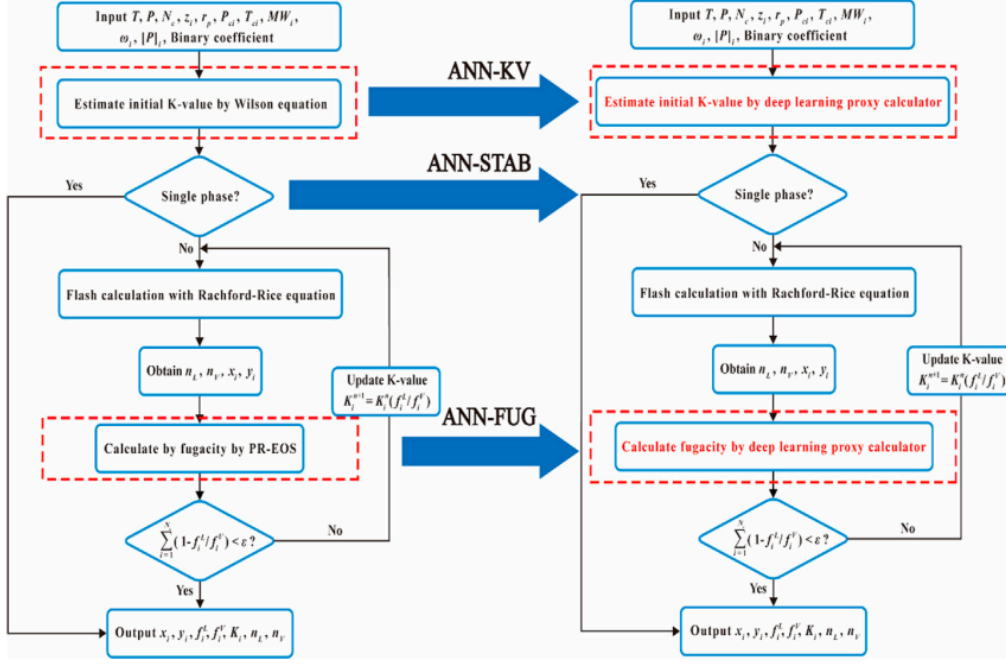


Figura 3.7: Diagrama esquemático comparando a modelagem DL-KF e o cálculo de *flash* convencional [23]

ada na equação de estado, modelo substituto (*surrogate*), DL-FUG (*Deep Learning for Fugacity Calculation*) e DL-KV (*Deep Learning for K-Values Calculation*).

Foram estudados dois casos:

- **Caso 1:** Três componentes ( $C_1/C_6/C_{10}$ ) amostrando pressões entre 0 e 200 bar com passo de 0,5 bar e composições de zero a um com intervalos de 0,1 totalizando 1960 amostras geradas. A avaliação dos modelos foi feita a partir do erro médio percentual (MPE, Mean Percentage Error) calculado pela formula abaixo.

$$MPE = \frac{1}{n} \sum \left| \frac{o - \hat{o}}{\hat{o}} \right| \times 100\% \quad (3.2)$$

Resultados para os modelos resumidos na tabela 3.3.

- **Caso 2:** Seis componentes ( $C_1/C_{2-3}/C_{4-6}/C_{7-15}/C_{16-27}/C_{28+}$ ) amostrando utilizando a mesma estratégia do caso 1, fornecendo um total de 65.990 amostras. Resultados para os modelos resumidos na tabela 3.4.

De maneira geral, o DL-KF teve resultados equiparáveis com os melhores modelos e com a modelagem original trazendo uma considerável eficiência computacional.

O DL-KF foi testado em composições reais de petróleo sempre estando dentro dos 5% de desvio em relação ao método de cálculo de fase convencional considerado pelo autores aceitável. E, em relação ao tempo computacional gasto, o DL-KF trás uma redução de 50,60% no número de iterações em relação a modelagem utilizando



| Modelos          | Acurácia (%) | Tempo de CPU (s) | Aceleração |
|------------------|--------------|------------------|------------|
| EOS original     | 100,0        | 35,28            | 1          |
| <i>Surrogate</i> | 97,32        | 0,51             | 69,18      |
| DL-FUG           | 99,50        | 3,35             | 10,53      |
| DL-KV            | 99,45        | 1,62             | 21,78      |
| DL-KF            | 99,18        | 0,82             | 43,02      |

Tabela 3.3: Resultados do caso 1 para diferentes modelos [23]

| Modelos          | Acurácia (%) | Tempo de CPU (s) | Aceleração |
|------------------|--------------|------------------|------------|
| EOS original     | 100,0        | 1250,34          | 1          |
| <i>Surrogate</i> | 97,23        | 6,38             | 195,98     |
| DL-FUG           | 99,84        | 29,87            | 41,86      |
| DL-KV            | 99,58        | 26,55            | 47,08      |
| DL-KF            | 99,29        | 7,80             | 160,30     |

Tabela 3.4: Resultados do caso 2 para diferentes modelos [23]

equação de estado mostrando que é uma melhor ferramenta para aceleração de cálculos de *flash*.

## 3.2 Cálculo de *flash* NVT

O trabalho de LI *et al.* [24] utilizou redes neurais profundas densas para modelar reservatórios abordando o problema de *flash* utilizando a Composição, Volume e Temperatura (NVT *flash*) como parâmetros de entrada. A ideia é acelerar os cálculos equilíbrio de fase substituindo os algoritmos convencionais, que são iterativos. O modelo proposto recebe, como entrada, propriedades composicionais (pressão crítica  $p_c$ , temperatura crítica  $T_c$  e fator acêntrico  $\omega$ ), a fração molar total  $\mathbf{z} = [z_1, \dots, z_M]^T$ , a concentração molar total  $C$  e temperatura  $T$ . As saídas são as frações molares do vapor  $X_i$ , frações molares do líquido  $Y_i$  e o número de fases  $N$ . Essa abordagem difere-se das pesquisas mais recentes na aceleração dos cálculos de *flash* utilizando modelos de aprendizado de máquina porque geralmente o teste de estabilidade para avaliar se existe mais de uma fase é preservado. Para o treinamento do modelo, foi utilizado o erro médio quadrático com regularização pela norma  $L^2$  como função custo e uma arquitetura com 5 camadas ocultas com 100 neurônios cada e função de ativação do tipo ReLU. Nesse trabalho, três fluidos de reservatórios reais foram investigados, incluindo petróleos com 5 componentes, 8 componentes e 14 compo-

nentes. Para esses fluidos, foram discretizados uniformemente a densidade molar e temperatura. Para investigar o efeito do tamanho do conjunto de dados, essas discretizações foram feitas num grid de 51 x 51, 71 x 71, 101 x 101, 151 x 151, 201 x 201 e 301 x 301 para um dos três fluidos. Para os outros dois fluidos, foram utilizados grids de 301 x 301. Através do trabalho, a abordagem mostra bons resultados tanto na performance geral do modelo quanto as previsões das composições nas fases.

O trabalho de ZHANG *et al.* [25] propõe-se a resolver o mesmo problema de *flash* NVT utilizando-se de duas redes neurais. Todo o experimento descrito nesse artigo está bem alinhado com o trabalho de LI *et al.* [24], diferindo-se na arquitetura da(s) rede(s) neural(is). Nesse trabalho, a primeira rede neural tem o propósito de extrair características dos valores de entrada e a saída dessa rede sofre um *padding* de zeros ou um *cropping* dependendo da quantidade de componentes do petróleo em relação a quantidade de componentes predefinida na entrada da segunda rede. Então os dados de treinamento e teste são uniformizados para uma certa dimensão, beneficiando a segunda rede. Foram avaliados vários hiper parâmetros da rede neural visando diminuição dos erros no conjunto de teste.

Baseado nos trabalhos acima, ZHANG *et al.* [26] seguiu uma abordagem diferente alterando os valores preditos pela modelagem. Ao contrário dos trabalhos anteriores, a saída do modelo retornava, além da quantidade de fases, a fração molar de vapor  $\phi$  e a fração de cada componente na fase vapor somente, diminuindo a dimensão da saída em quase pela metade, impactando fortemente, segundo os autores, na eficiência do treinamento. Nesse trabalho, também foi avaliado os efeitos de capilaridade que ocorrem em reservatórios não convencionais. A ideia é de avaliar a alta aplicabilidade da configuração de rede neural obtida anteriormente no artigo.

Nesses trabalhos, os conjuntos de dados para treinamento e avaliação dos modelos foram gerados através do método iterativo por substituições sucessivas que foi introduzido anteriormente para resolver problemas de *flash* NVT.

### 3.3 Cálculo de *flash* para compostos limpos

Com a transição para energias renováveis, a aplicação da análise de equilíbrio de fases para misturas de hidrogênio natural surgiu como um tópico relevante e foi explorada pelo trabalho de ZHANG *et al.* [27] posteriormente aos trabalhos mostrados acima. Nesse trabalho, foram utilizadas duas composições de mistura de hidrogênio, o gás Robel e o gás Ramsay1, ambos sendo uma mistura de  $H_2$  em maior fração com  $O_2$ ,  $CO_2$ ,  $CH_4$  e  $N_2$ .

Também foram analisados modelos híbridos baseados em rede neurais com a inserção de regras termodinâmicas nos quais foram denominados Redes Neurais informadas pela Termodinâmica (*Thermodynamic-Informed Neural Networks*, TINN).

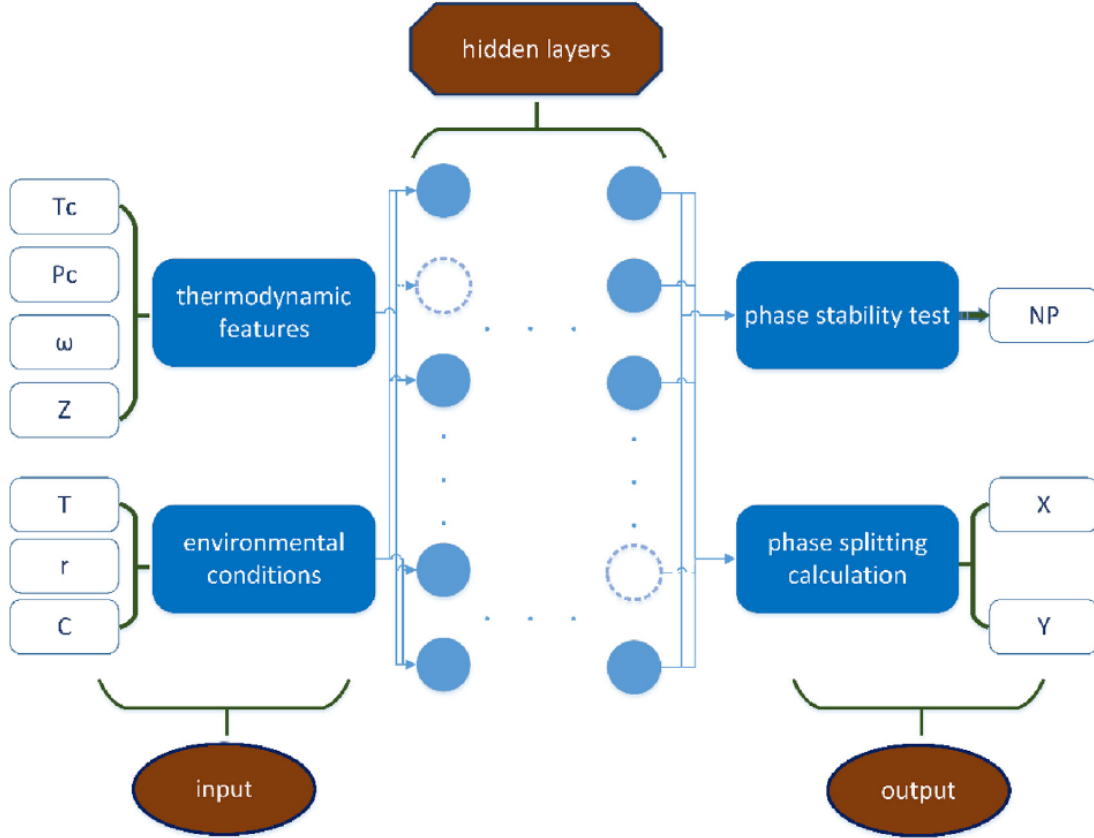


Figura 3.8: Arquitetura proposta para as Redes Neurais informadas pela Termodinâmica [27]

Foram utilizados as regras de conservação para as composições ( $\sum_i X_i = 1$  e  $\sum_i Y_i = 1$ ) para atingir uma consistência termodinâmica, na função custo utilizada no treinamento dos modelos. A função custo  $L$  final é a seguinte.

$$L = \frac{1}{N} \sum_{n=1}^N \|\mathbf{o} - \hat{\mathbf{o}}\|^2 + \lambda \|\mathbf{W}\|_2^2 + \sum_{i=1}^M X_i + \sum_{i=1}^M Y_i - 2 \quad (3.3)$$

As variáveis de entrada e saídas são similares aos trabalhos anteriores com a diferença da adição nas *features* do raio do poro para incorporar efeitos de capilaridade dos reservatórios.

As abordagens no treinamento dos modelos utilizadas nos trabalhos anteriores também foram usadas. Aplicação da regularização pela norma  $L_2$  dos pesos da rede, utilização de *dropout* e *batch normalization*. Foram realizadas análises comparativas de tempo computacional entre os métodos iterativos convencionais com os modelos de aprendizado profundo e de resultados dos cálculos de *flash* entre as abordagens com resultados satisfatórios no ponto de vista dos autores.

# Capítulo 4

## Método proposto

A pesquisa propõe-se avaliar o problema do equilíbrio vapor-líquido de mistura de hidrocarbonetos (maior parte dos componentes do petróleo) utilizando modelos baseado em dados para alterar ou substituir os métodos convencionais. Uma técnica de cálculo de *flash* bem estabelecida dados a composição do fluido ( $\mathbf{N}$ ), pressão ( $P$ ) e temperatura ( $T$ ) é o *flash* NPT. Esse método de cálculo de *flash* vai ser o utilizado no trabalho.

Na figura 4.1, tem-se um fluxograma esquemático das etapas do experimento que servirão de insumo para o trabalho.

O ponto de partida é de uma base de um pouco mais de 400 amostras de composições reais de petróleos de campos brasileiros. Essas composições possuem 24 componentes, na sua maioria hidrocarbonetos.

Com esses dados, um pré processamento deverá ser feito para garantir que as composições somem 1 (ou 100%) para garantia do balanço de matéria. Com a mesma base inicial, uma parte da análise é de avaliar se amostras sintéticas geradas a partir da base de composições trarão ganho nos modelos gerados para o cálculo do *flash*. Estimadores de Máxima Verosimilhança a partir da distribuição de Dirichlet [28] (abordagem utilizada em outros trabalhos), modelos Wasserstein GAN [29] e Diffusion Probabilistic Models [30] são possíveis candidatos a geradores de composições.

Com todas as composições, precisa-se amostras pressões e temperaturas que servirão, juntamente com as composições da base inicial e as geradas sinteticamente, a entrada para o cálculo de *flash*. E, tendo todas as entradas, necessita-se gerar os dados de saída do cálculo de *flash*, a partir de simulações (utilizando o NeqSim). Tais dados seriam as frações de líquido e vapor e suas respectivas composições, no caso de uma mistura de duas fases ou a indicação se o fluido naquelas condições de composição  $z_i$ , pressão  $p$  e temperatura  $T$  é somente líquido ou somente vapor. Finalizada essas primeiras etapas, tem-se todo o conjunto de dados necessários para treinar e avaliar os modelos gerados.

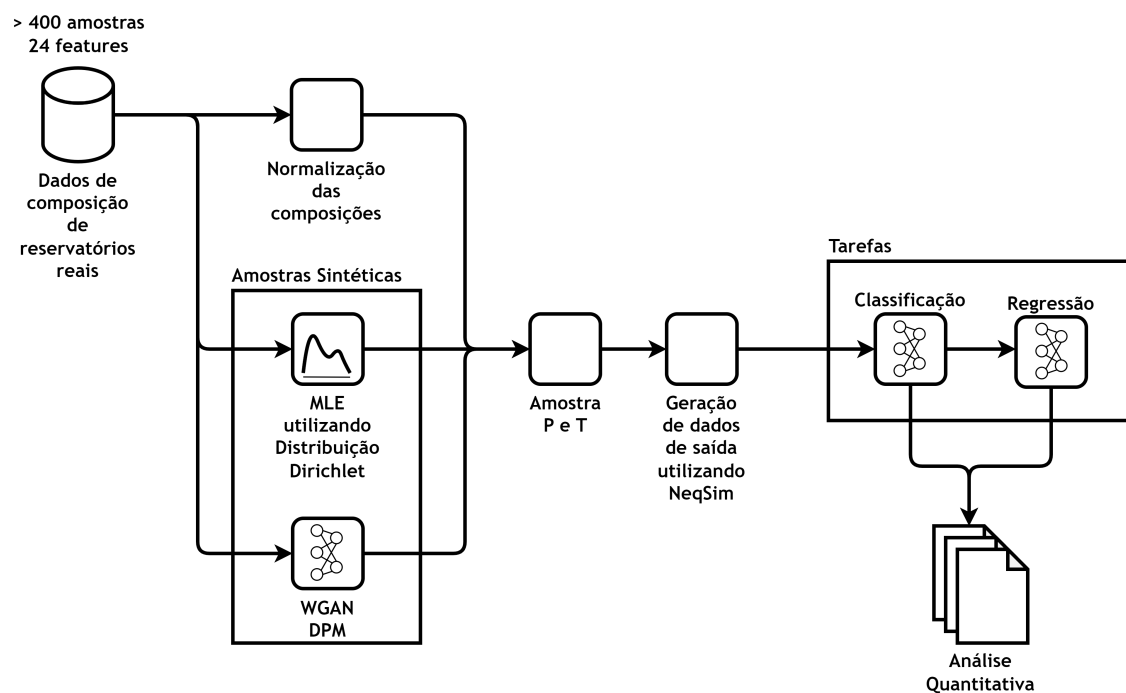


Figura 4.1: Pipeline do experimento

Partindo do método de cálculo de *flash* original, tem-se o teste de estabilidade que, no contexto da inteligência computacional pode ser tratado como um problema de classificação. Existem algumas maneiras de tratar esse problema de classificação: através de uma classificação com três classes, onde as classes seriam mistura, vapor e líquido, através de uma classificação binária aonde as classes seriam estável (vapor ou líquido) ou instável (mistura) ou ainda ser tratado como uma regressão aonde a saída do modelo seria a fração de vapor ou líquido onde, 0 ou 1 indicariam ausência ou presença total da fase especificada e algo entre 0 e 1 uma mistura com a fração já retornada do modelo.

Para a obtenção das frações dos componentes de cada fase, pode-se trabalhar com um modelo regressão. Na literatura, geralmente trabalha-se com as razões de equilíbrio para cada componente do fluido na saída do modelo para uma estimativa inicial para utilização no cálculo de *flash*, no lugar da correlação de Wilson ou com os coeficientes de fugacidade. Pode-se também avaliar modelos que forneçam uma predição das razões de equilíbrio finais para cada componente.

Para a análise quantitativa, avaliações dos modelos baseados em figuras de mérito comparando com os resultados gerados pelo simulador NeqSim serão realizadas.

# Referências Bibliográficas

- [1] *OPEC Annual Report*. In: Report, Organization of the Petroleum Exporting Countries (OPEC), Vienna, Austria, 2022.
- [2] *World Energy Outlook*. In: Report, International Energy Agency (IEA), Paris, France, 2023.
- [3] *World Energy Review*. In: Report, Eni SpA, Rome, Italy, 2023.
- [4] BARBOSA, G. *Petróleo e Estado*. Rio de Janeiro, RJ, Agência Nacional do Petróleo, Gás Natural e Biocombustíveis - ANP, 2015.
- [5] *A Relevância do Petróleo & Gás para o Brasil*. Rio de Janeiro, RJ, Instituto Brasileiro de Petróleo, Gás e Biocombustível - IBP, 2019.
- [6] “Energia em Transformação”. [https://petrobras.com.br/energia-em-transformacao?gad\\_source=1&gclid=CjwKCAiA0PuuBhBsEiwAS7fsNTRxzBFw\\_qPJbRy\\_TGYNdRZOHdPDANYp\\_vKej4T\\_AiGS1IGrTVc1yRoCi-kQAvD\\_BwE#introducao](https://petrobras.com.br/energia-em-transformacao?gad_source=1&gclid=CjwKCAiA0PuuBhBsEiwAS7fsNTRxzBFw_qPJbRy_TGYNdRZOHdPDANYp_vKej4T_AiGS1IGrTVc1yRoCi-kQAvD_BwE#introducao). Accessed: 2024-02-28.
- [7] BEGUM, N., DADASHPOUR, M., KLEPPE, J. “Subchapter 1.6 - A case study of reservoir parameter estimation in Norne oil field, Norway by using Ensemble Kalman Filter (EnKF)”. In: Moitra, A., Kayal, J., Mukerji, B., et al. (Eds.), *Innovative Exploration Methods for Minerals, Oil, Gas, and Groundwater for Sustainable Development*, Elsevier, pp. 61–78, 2022. ISBN: 978-0-12-823998-8. doi: <https://doi.org/10.1016/B978-0-12-823998-8.00007-7>. Disponível em: <<https://www.sciencedirect.com/science/article/pii/B9780128239988000077>>.
- [8] AHMED, T. *Reservoir Engineering Handbook*. Houston, Texas, Elsevier Science, 2010. ISBN: 9780080966670. Disponível em: <[https://books.google.com.br/books?id=LXJcG\\_jwTHAC](https://books.google.com.br/books?id=LXJcG_jwTHAC)>.
- [9] WANG, P., STENBY, E. H. “Phase equilibrium calculation in compositional reservoir simulation”, *Computers & Chemical Engineering*, v. 16,

- pp. S449–S456, 1992. ISSN: 0098-1354. doi: [https://doi.org/10.1016/S0098-1354\(09\)80053-7](https://doi.org/10.1016/S0098-1354(09)80053-7). Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0098135409800537>>. European Symposium on Computer Aided Process Engineering—1.
- [10] SUN, S., ZHANG, T. “Chapter seven - Recent progress in accelerating flash calculation using deep learning algorithms”. In: Sun, S., Zhang, T. (Eds.), *Reservoir Simulations*, Gulf Professional Publishing, pp. 289–322, 2020. ISBN: 978-0-12-820957-8. doi: <https://doi.org/10.1016/B978-0-12-820957-8.00007-1>. Disponível em: <<https://www.sciencedirect.com/science/article/pii/B9780128209578000071>>.
- [11] JIRASEK, F., HASSE, H. “Perspective: Machine Learning of Thermophysical Properties”, *Fluid Phase Equilibria*, v. 549, pp. 113206, 2021. ISSN: 0378-3812. doi: <https://doi.org/10.1016/j.fluid.2021.113206>. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0378381221002697>>.
- [12] PENG, D.-Y., ROBINSON, D. B. “A New Two-Constant Equation of State”, *Industrial & Engineering Chemistry Fundamentals*, v. 15, n. 1, pp. 59–64, 1976. doi: [10.1021/i160057a011](https://doi.org/10.1021/i160057a011). Disponível em: <<https://doi.org/10.1021/i160057a011>>.
- [13] REDLICH, O., KWONG, J. N. S. “On the Thermodynamics of Solutions. V. An Equation of State. Fugacities of Gaseous Solutions.” *Chemical Reviews*, v. 44, n. 1, pp. 233–244, 1949. doi: [10.1021/cr60137a013](https://doi.org/10.1021/cr60137a013). Disponível em: <<https://doi.org/10.1021/cr60137a013>>. PMID: 18125401.
- [14] SOAVE, G. “Equilibrium constants from a modified Redlich-Kwong equation of state”, *Chemical Engineering Science*, v. 27, n. 6, pp. 1197–1203, 1972. ISSN: 0009-2509. doi: [https://doi.org/10.1016/0009-2509\(72\)80096-4](https://doi.org/10.1016/0009-2509(72)80096-4). Disponível em: <<https://www.sciencedirect.com/science/article/pii/0009250972800964>>.
- [15] WANG, K., LUO, J., WEI, Y., et al. “Practical application of machine learning on fast phase equilibrium calculations in compositional reservoir simulations”, *Journal of Computational Physics*, v. 401, pp. 109013, 2020. ISSN: 0021-9991. doi: <https://doi.org/10.1016/j.jcp.2019.109013>. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0021999119307193>>.
- [16] RAHMANIFARD, H., PLAKSINA, T. “Application of artificial intelligence techniques in the petroleum industry: a review”, *Artificial Intelligence*

*Review*, v. 52, n. 4, pp. 2295–2318, Dec 2019. ISSN: 1573-7462. doi: 10.1007/s10462-018-9612-8. Disponível em: <<https://doi.org/10.1007/s10462-018-9612-8>>.

- [17] SOFOS, F., STAVROGIANNIS, C., EXARCHOU-KOUVELI, K. K., et al. “Current Trends in Fluid Research in the Era of Artificial Intelligence: A Review”, *Fluids*, v. 7, n. 3, 2022. ISSN: 2311-5521. doi: 10.3390/fluids7030116. Disponível em: <<https://www.mdpi.com/2311-5521/7/3/116>>.
- [18] QU, J., FANEY, T., DE HEMPTINNE, J.-C., et al. “PTFlash : A vectorized and parallel deep learning framework for two-phase flash calculation”, *Fuel*, v. 331, pp. 125603, 2023. ISSN: 0016-2361. doi: <https://doi.org/10.1016/j.fuel.2022.125603>. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0016236122024334>>.
- [19] STEIN, M. “Large Sample Properties of Simulations Using Latin Hypercube Sampling”, *Technometrics*, v. 29, n. 2, pp. 143–151, 1987. ISSN: 00401706. Disponível em: <<http://www.jstor.org/stable/1269769>>.
- [20] CHENG, H., KOC, L., HARMSSEN, J., et al. “Wide & Deep Learning for Recommender Systems”, *CoRR*, v. abs/1606.07792, 2016. Disponível em: <<http://arxiv.org/abs/1606.07792>>.
- [21] CHEN, F., LUO, S., WANG, S., et al. “A generalized machine learning-assisted phase-equilibrium calculation model for shale reservoirs”, *Fluid Phase Equilibria*, v. 558, pp. 113423, 2022.
- [22] CHEN, F., LUO, S., WANG, S., et al. “A Novel Machine-Learning Assisted Phase-Equilibrium Calculation Model for Liquid-Rich Shale Reservoirs”. In: *SPE Reservoir Simulation Conference?*, p. D011S006R001. SPE, 2023.
- [23] ZHENG, Z., DI, Y., YU, E. “DL-KF modeling for acceleration of flash calculations in phase equilibrium using deep learning methods”, *Frontiers in Earth Science*, v. 10, pp. 1041589, 2023.
- [24] LI, Y., ZHANG, T., SUN, S. “Acceleration of the NVT Flash Calculation for Multicomponent Mixtures Using Deep Neural Network Models”, *Industrial & Engineering Chemistry Research*, v. 58, n. 27, pp. 12312–12322, 2019. doi: 10.1021/acs.iecr.9b00527. Disponível em: <<https://doi.org/10.1021/acs.iecr.9b00527>>.



- [25] ZHANG, T., LI, Y., LI, Y., et al. “A self-adaptive deep learning algorithm for accelerating multi-component flash calculation”, *Computer Methods in Applied Mechanics and Engineering*, v. 369, pp. 113207, 2020. ISSN: 0045-7825. doi: <https://doi.org/10.1016/j.cma.2020.113207>. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0045782520303923>>.
- [26] ZHANG, T., LI, Y., SUN, S., et al. “Accelerating flash calculations in unconventional reservoirs considering capillary pressure using an optimized deep learning algorithm”, *Journal of Petroleum Science and Engineering*, v. 195, pp. 107886, 2020. ISSN: 0920-4105. doi: <https://doi.org/10.1016/j.petrol.2020.107886>. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0920410520309438>>.
- [27] ZHANG, T., ZHANG, Y., KATTERBAUER, K., et al. “Deep learning-assisted phase equilibrium analysis for producing natural hydrogen”, *International Journal of Hydrogen Energy*, v. 50, pp. 473–486, 2024. ISSN: 0360-3199. doi: <https://doi.org/10.1016/j.ijhydene.2023.09.097>. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0360319923046633>>.
- [28] MINKA, T. “Estimating a Dirichlet Distribution”. September 2000. Disponível em: <<https://www.microsoft.com/en-us/research/publication/estimating-dirichlet-distribution/>>.
- [29] ARJOVSKY, M., CHINTALA, S., BOTTOU, L. “Wasserstein GAN”. 2017.
- [30] SOHL-DICKSTEIN, J., WEISS, E. A., MAHESWARANATHAN, N., et al. “Deep Unsupervised Learning using Nonequilibrium Thermodynamics”. 2015.