INSTITUTO FEDERAL DO ESPÍRITO SANTO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE CONTROLE E AUTOMAÇÃO

DANIEL PINHEIRO SCARDINI

DIGITAL TWIN DE UM SEPARADOR TRIFÁSICO EM UNIDADE OFFSHORE UTILIZANDO REDES NEURAIS NARX E LSTM

Serra
2023

DANIEL PINHEIRO SCARDINI

DIGITAL TWIN DE UM SEPARADOR TRIFÁSICO EM UNIDADE OFFSHORE UTILIZANDO REDES NEURAIS NARX E LSTM

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Controle e Automação do Instituto Federal do Espírito Santo, como requisito para a obtenção do título de Mestre em Engenharia de Controle e Automação.
Orientador: Prof. Dr Gustavo Maia de Almeida
Co-orientador: Prof. Dr Leonardo Azevedo Scárdua

Serra
2023

Dados Internacionais de Catalogação na Publicação (CIP)

S285d 2023

Scardini, Daniel Pinheiro
Digital twin de um separador trifásico em unidade offshore utilizando redes neurais NARX e LSTM / Daniel Pinheiro Scardini. - 2023.
69 f.; il.; 30 cm

Orientador: Prof. Dr. Gustavo Maia de Almeida.
Coorientador: Prof. Dr. Leonardo Azevedo Scárdua.

Dissertação (mestrado) - Instituto Federal do Espírito Santo, Programa de Pós-graduação em Engenharia de Controle e Automação, 2023 .

1. Redes neurais. 2. Petróleo. 3. Long short-term memory (LSTM). 4. Nonlinear Autoregressive with Exogenous Inputs (NARX). I. Almeida, Gustavo Maia de. II. Instituto Federal do Espírito Santo. III. Título.

CDD 629.895

Bibliotecário: Valmir Oliveira de Aguiar - CRB6/ES 566

MINISTÉRIO DA EDUCAÇÃO

INSTITUTO FEDERAL DO ESPÍRITO SANTO

PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE CONTROLE E AUTOMAÇÃO

DANIEL PINHEIRO SCARDINI

"DIGITAL TWIN DE UM SEPARADOR TRIFÁSICO EM UMA UNIDADE OFFSHORE UTILIZANDO REDES NEURAIS NARX E LSTM"

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Controle e Automação do Instituto Federal do Espírito Santo, como requisito parcial para obtenção de título de Mestre em Engenharia de Controle e Automação.

Aprovado em 14 de Novembro de 2023

COMISSÃO EXAMINADORA

Documento assinado digitalmente
gov.br
GUSTAVO MAIA DE ALMEIDA
Data: 21/02/2024 09:08:38-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. Gustavo Maia de Almeida
Instituto Federal do Espírito Santo
Orientador

Documento assinado digitalmente
gov.br
LEONARDO AZEVEDO SCARDUA
Data: 22/02/2024 13:08:40-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. Leonardo Azevedo Scárdua
Instituto Federal do Espírito Santo
Co-Orientador

Documento assinado digitalmente
gov.br
DANIEL CRUZ CAVALIERI
Data: 22/02/2024 13:18:29-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. Daniel Cruz Cavalieir
Instituto Federal do Espírito Santo
Membro Interno

Documento assinado digitalmente
gov.br
VITOR FAICAL CAMPANA
Data: 05/03/2024 14:05:33-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. Vitor Faiçal Campana
Instituto Federal do Espírito Santo
Membro Externo

AGRADECIMENTOS

Agradeço, sobretudo, a Deus pela mão forte e poderosa com a qual tem sustentado a mim e minha família, pela direção segura que nunca faltou e pela misericórdia que se renova a cada dia, sem a qual nada subsistiria.

À minha esposa Larissa por todo o apoio e paciência, que mesmo com a chegada do nosso pequeno Isaque e todos os novos desafios que se apresentaram, foi sempre meu ânimo e motivação diária. Ao meu filho Isaque cujo sorriso fácil e alegria em descobrir o mundo são o maior combustível para minha jornada.

Aos meus pais por todos os esforços em me dar as condições para alcançar os meus objetivos profissionais e acadêmicos, e por serem sempre referências para mim.

Aos professores Gustavo Maia e Leonardo Scárdua pela paciência e orientação, me dando a direção e conhecimento indispensáveis ao alcance dos objetivos.

RESUMO

A extração de petróleo *offshore* é um processo complexo que enfrenta diversos desafios, especialmente no que se refere à limitação de espaço físico para a construção de vasos com capacidade suficiente para absorver as variações de carga a que o processo está sujeito. Além disso, a natureza turbulenta do escoamento nas tubulações de elevação do petróleo pode provocar golfadas ocasionais, tornando complexo o controle dos níveis de vasos de tratamento e separação, que são essenciais para garantir a qualidade dos produtos finais. Neste trabalho, é realizado o desenvolvimento de um *Digital Twin* de um separador trifásico utilizando redes neurais, especificamente as arquiteturas NARX e LSTM, para prever os níveis de vasos de tratamento e separação com 1, 3, 5 e 10 passos de antecedência. Cada modelo foi treinado por meio de otimização de hiperparâmetros com o método de Otimização Bayesiana e validado com dados de teste independentes obtendo resultados significativos.

O *Digital Twin* desenvolvido permite simulações de processo realistas, que utilizam técnicas de *machine learning* para identificar relações não lineares entre variáveis e replicar o comportamento do gêmeo físico com fidelidade. As simulações de processo são valiosas para testar estratégias de controle e ajustar a sintonia do modelo. A identificação do modelo de processo é realizada a partir das entradas e saídas conhecidas, utilizando os dados de processo disponíveis. Este trabalho evidencia a eficácia do *Digital Twin* como uma ferramenta que aprimora o controle e a operação dos sistemas de produção de petróleo *offshore*. Sua implementação representa um avanço significativo na gestão e controle desses sistemas, promovendo uma extração de petróleo mais eficiente e sustentável.

Palavras-chave: Redes Neurais, NARX, LSTM, *Digital Twin,* Tratamento de Petróleo, Separador Trifásico, Controle de Processos

ABSTRACT

Offshore oil extraction is a complex process that faces various challenges, especially regarding the limited physical space for constructing vessels with sufficient capacity to absorb the load variations to which the process is subjected. Additionally, the turbulent nature of flow in risers can cause occasional surges, making it complex to control the levels of treatment and separation vessels, which are essential to ensuring the quality of the final products. In this work, it is developed a Digital Twin of a three-phase separator using neural networks, specifically the NARX and LSTM architectures, to predict the levels of treatment and separation vessels with 1, 3, 5, and 10 steps ahead. Each model is trained using hyperparameter optimization with the Bayesian Optimization method and validated with independent test data, obtaining significant results.

The developed Digital Twin enables realistic process simulations, using machine learning techniques to identify non-linear relationships between variables and faithfully replicate the behavior of the physical twin. Process simulations are valuable for testing control strategies and fine-tuning the model. The process model identification is performed based on known inputs and outputs, using available process data. This study highlights the effectiveness of the Digital Twin as a tool that enhances control and operation of offshore oil production systems. Its implementation has the potential to be a significant advancement in the management and control of these systems, promoting more efficient and sustainable oil extraction.

Keywords: Neural Networks, NARX, LSTM, Digital Twin, Petroleum Treatment, Three-phase separator, Process Control

LISTA DE FIGURAS

Figura 1 – Produção de Petróleo no Brasil por localização.....10
Figura 2 – Visão esquemática de um sistema de produção *offshore*.....11
Figura 3 – Gêmeo Digital.....13
Figura 4 – Comparativo entre as arquiteturas de Redes LSTM e Redes NARX.....14
Figura 5 – Diagrama de implementação do código *Python* desenvolvido.....17
Figura 6 – *Plot* da autocorrelação da variável *Gas Flow*.....18
Figura 7 – Diagrama de Influência dos Parâmetros Ratio e Threshold na Quantidade de Entradas Defasadas Fornecidas à Rede NARX.....22
Figura 8 – Formação geológica de um reservatório típico.....26
Figura 9 – Interação entre poços injetores e produtores no reservatório.....27
Figura 10 – Formação de golfadas nos *risers* de produção.....29
Figura 11 – Fluxo de processo de produção e tratamento de petróleo e suas 3 fases.....30
Figura 12 – Visualização interna de um separador trifásico horizontal típico.....31
Figura 13 – Esquema de instrumentação e controle típicos de um separador trifásico horizontal.....33
Figura 14 – Abordagens de modelagem e simulação de processo.....35
Figura 15 – Evolução das citações ao termo '*Digital Twin*' ao longo do tempo.....38
Figura 16 – Arquitetura de uma Rede Neural NARX típica.....41
Figura 17 – Arquitetura de uma rede LSTM.....43
Figura 18 – Visualização do comportamento das variáveis de entrada do separador.....47
Figura 19 – Visualização do comportamento das variáveis de saída do separador.....49

LISTA DE TABELAS

Tabela 1 - Métricas de Erro MAE para cada modelo 55
Tabela 2 - Métricas de Erro MSE para cada modelo 56
Tabela 3 – Métricas de Erro MAE de um modelo NARX em quatro horizontes de previsão.....57

LISTA DE SIGLAS

ANP – Agência Nacional de Petróleo

LSTM – Memória de Curto Prazo Longa *(Long Short Term Memory)*

MAE – Erro Absoluto Médio *(mean absolut error)*

MAPE – Erro Percentual Absoluto Médio *(Mean Absolute Percentage Error)*

MSE – Erro Quadrático Médio (*Mean Squared Error*)

NARX – Rede Neural Auto-Regressiva *(Nonlinear Autoregressive Exogenous)*

NASA – Administração Nacional da Aeronáutica e Espaço (*National Aeronautics and Space Administration)*

PI – Sistema Operacional de Informação da Planta (*OSISoft Plant Information System)*

PID – Proporcional Integral Derivativo

PV – Variável de Processo *(Process Variable)*

RMSE – Raiz Quadrada do Erro Médio *(Root Mean Squared Error)*

RNN – Rede Neural Recorrente (*Recurrent Neural Networks*)

UEP – Unidade Estacionária de Produção

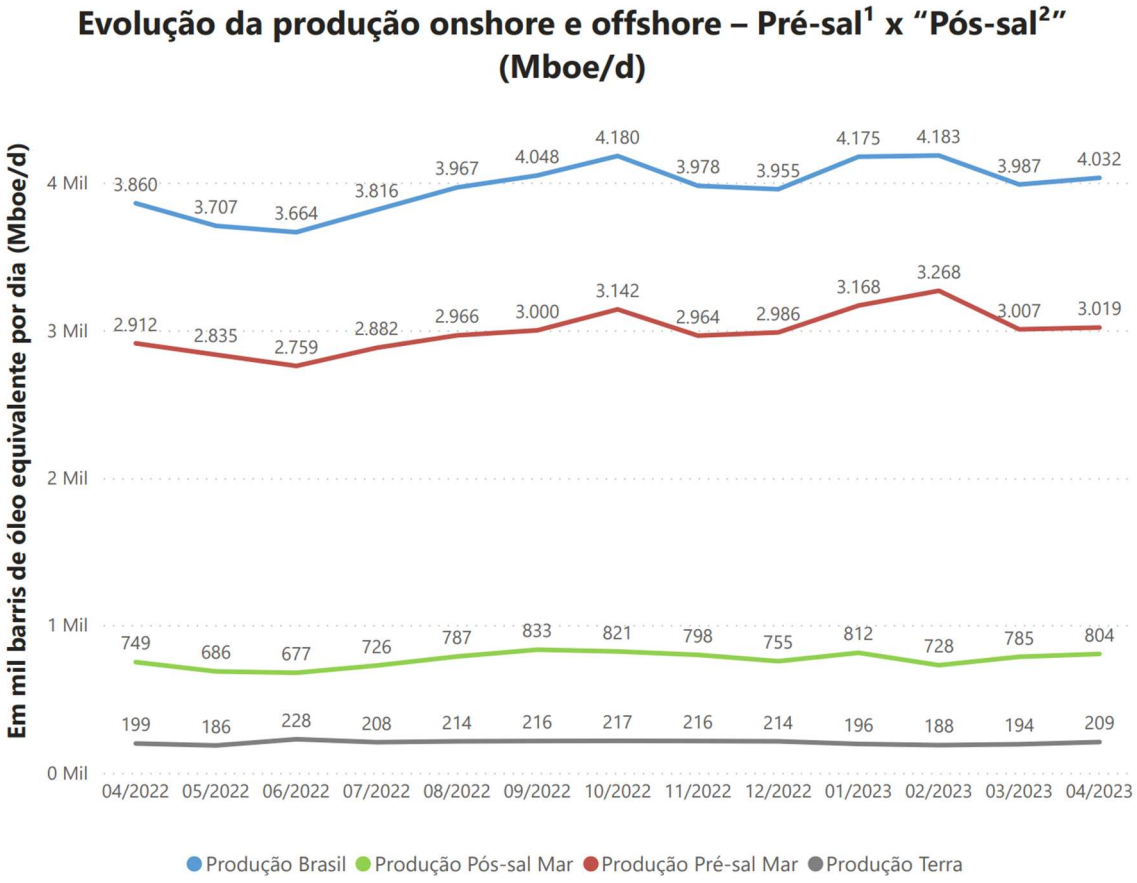
SUMÁRIO

1 INTRODUÇÃO 10
1.1 OBJETIVOS 16
1.1.1 Objetivos gerais 16
1.1.2 Objetivos específicos 16
1.2 METODOLOGIA 16
2 REFERENCIAL TEÓRICO 25
2.1 PRODUÇÃO DE PETRÓLEO *OFFSHORE* 25
2.1.1 Formação e migração do petróleo 25
2.1.2 Elevação e escoamento 26
2.1.3 Produção por golfadas 28
2.1.4 Separação e tratamento 29
2.2 MODELAGEM DINÂMICA DE PROCESSOS 33
2.3 IDENTIFICAÇÃO DE PROCESSOS 35
2.4 DIGITAL TWIN 36
2.5 REDES NARX 39
2.6 REDES LSTM 41
2.7 COMPARANDO REDES NARX E REDES LSTM 42
2.8 MÉTRICAS DE ERRO 43
2.9 OTIMIZAÇÃO BAYESIANA 44
3 DESENVOLVIMENTO 45
3.1 DETALHAMENTO DAS VARIÁVEIS DE ENTRADA E SAÍDA 45
3.2 PRÉ-PROCESSAMENTO DOS DADOS 50
3.3 TREINAMENTO DAS REDES 50
4 RESULTADOS E DISCUSSÃO 53
5 CONCLUSÃO 59
5.1 TRABALHOS FUTUROS 60

1 INTRODUÇÃO

Atualmente o setor de petróleo e gás é responsável por aproximadamente 13% do Produto Interno Bruto do Brasil, e de todo o petróleo e gás produzidos no país em Abril de 2023, 97% foram produzidos em plataformas de petróleo *offshore* (AGÊNCIA NACIONAL DO PETRÓLEO, 2023). A Figura 1 demonstra um gráfico comparativo entre a produção *offshore* e *onshore* durante o ano de 2021.

Figura 1 – Produção de Petróleo no Brasil

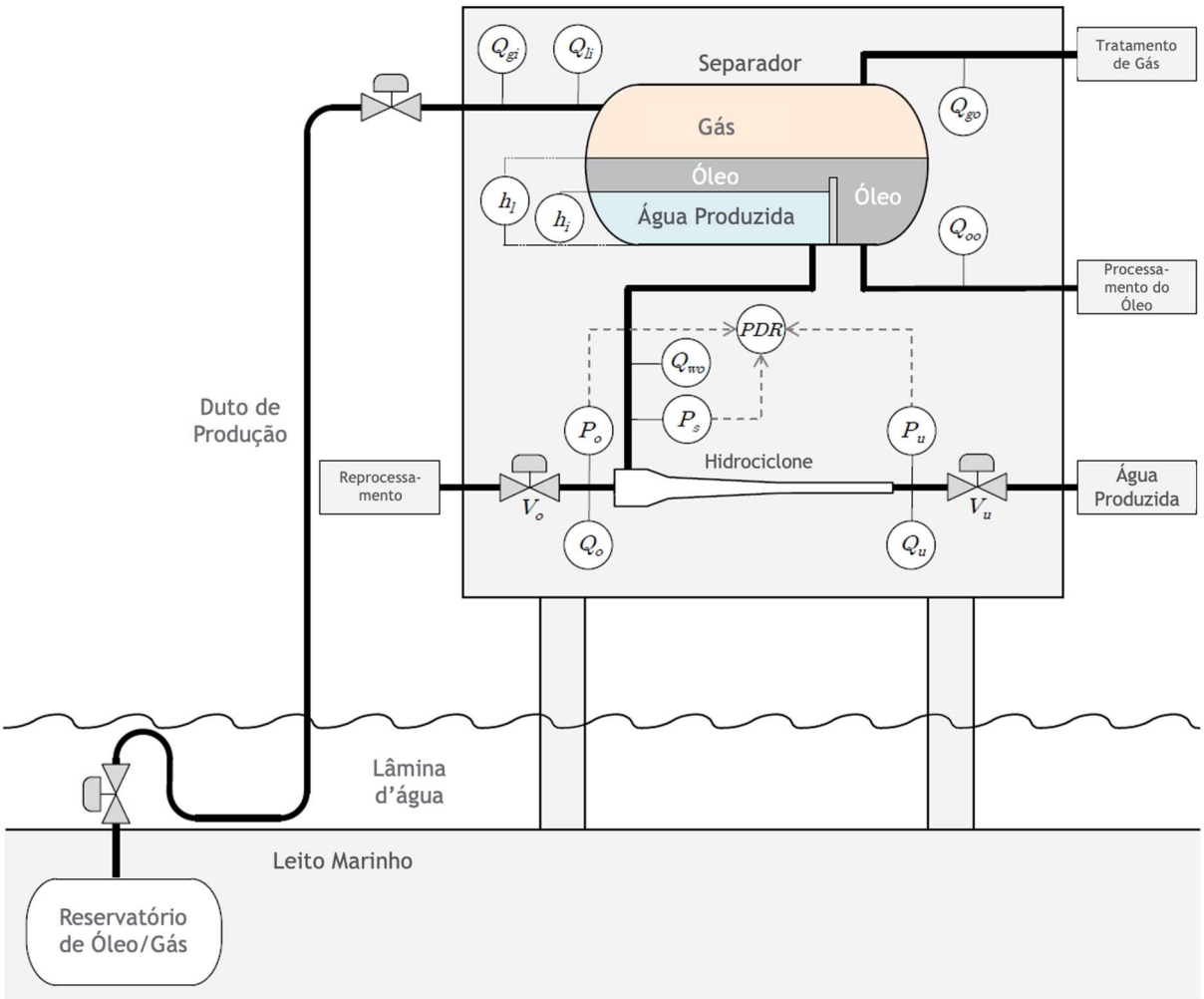


Fonte: (ANP, 2023)

O ambiente de produção de petróleo *offshor*e apresenta desafios significativos, incluindo o espaço físico limitado para o controle das variáveis do processo. A instabilidade do fluxo de produção, decorrente da variação de vazões e pressões, é um dos principais desafios enfrentados pelas instalações (CAMPOS et al., 2015). As causas dessa instabilidade são diversas e incluem transientes no fechamento e abertura de poços, características turbulentas do escoamento mul-

tifásico e golfadas. As golfadas são a perturbação mais comum e impactante na produção, causando sobrecarga em separadores, queima de gás no *flare*, perda de eficiência de separação e danos a equipamentos (NABUIFE; TANDOH; WHIDBORNE, 2022). A figura 2 apresenta um diagrama esquemático do processo de produção, desde a extração do reservatório até o tratamento no separador trifásico.

Figura 2 – Visão esquemática de um sistema de produção offshore



Fonte: Hansen et al. (2022)

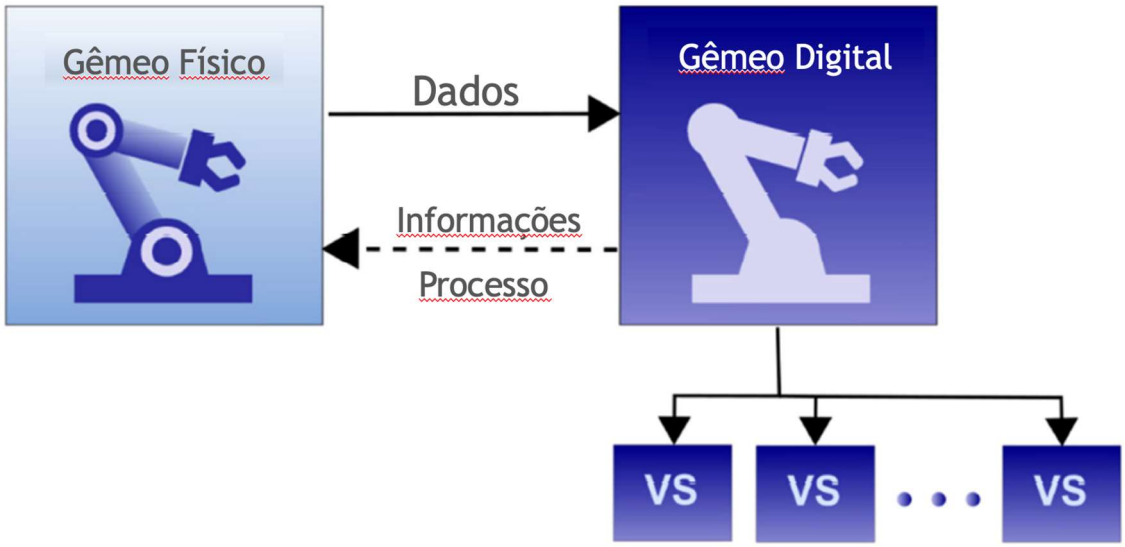
Diante dessa situação, é necessário estabelecer uma estratégia de controle que combine robustez para manter a estabilidade do processo mesmo sob perturbações severas com a capacidade de resposta para lidar com variações abruptas de carga devido ao fechamento ou abertura de poços. Controladores PID clássicos são amplamente utilizados na indústria *offshore* em geral, porém o seu desempenho não é satisfatório em casos em que a dinâmica do processo é altamente variável, como ocorre em plataformas de produção. Por exemplo, a dinâmica de uma plataforma com poucos poços abertos é totalmente diferente daquela com todos os poços abertos (DIEHL

et al., 2017), portanto durante a dinâmica de partida de planta os controladores trabalham com uma dinâmica muito variável, e as sintonias não são adaptadas para todas as dinâmicas diferentes.

Apesar das técnicas avançadas de controle serem amplamente utilizadas em refinarias e indústrias petroquímicas, elas são pouco aplicadas na indústria de produção de petróleo *offshore.* De acordo com Simões et al. (2013), algumas razões prováveis incluem: grande quantidade de transientes que uma plataforma está sujeita, como paradas, alinhamentos de poços, e golfadas, bem como a presença de ruídos e oscilações na medição devido a fatores atípicos, como variação dos níveis de vasos devido ao balanço da embarcação; a produção de petróleo e gás nunca atinge um estado estacionário real, pois envolve a interação de muitas variáveis, como a relação entre gás, óleo e água, que varia com a vida útil do reservatório, além da pressão e vazão de *gas-lift*[1] utilizado para elevação artificial de poços; disponibilidade limitada de instrumentação devido a impossibilidade de acesso ou à natureza enxuta das instalações marítimas; falta de profissionais especializados a bordo para implementar e monitorar o funcionamento dos sistemas. A dificuldade de implementar técnicas avançadas de controle tem sido um obstáculo para muitas iniciativas de modernização. No entanto, é possível contornar essas dificuldades por meio da modelagem e simulação das plantas de processo, permitindo testes e avaliações comparativas que se aproximem do processo real. Nesse sentido, tem-se tornado cada vez mais popular a utilização de técnicas de implementação de *Digital Twins* ou "gêmeos digitais" (WANASINGHE et al., 2020). Um exemplo de configuração de *Digital Twin* pode ser visto na Figura 3.[1]

[1] Método de elevação artificial utilizando injeção de gás na coluna de produção a fim de viabilizar a produção de poços maduros

Figura 3 – Gêmeo Digital



Fonte: Singh et al. (2021)

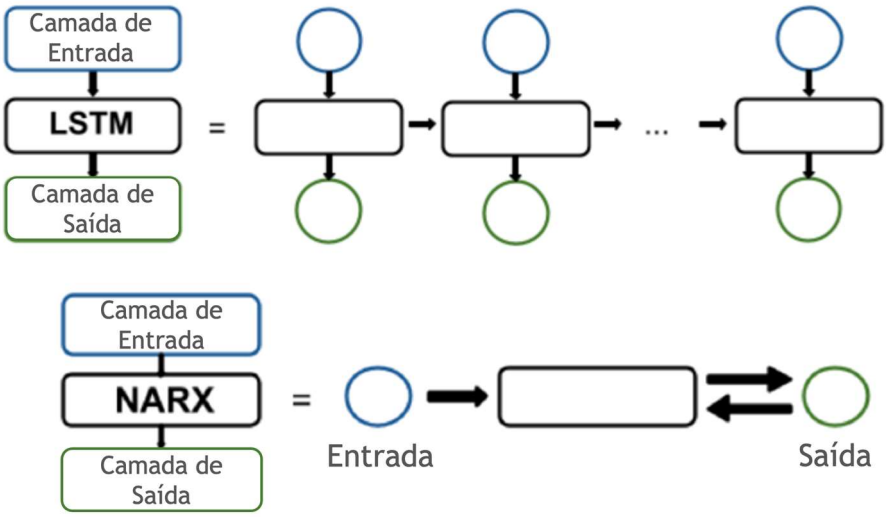
A implementação de um *digital twin* representa grande vantagem competitiva na indústria, com a possibilidade de utilização de métodos de controle mais modernos, permitindo a flexibilidade de rodar simulações muito próximas à realidade. A precisão e a especificidade do modelo obtido no desenvolvimento de um *Digital Twin* permitem a utilização de sintonias de controle mais arrojadas, baseadas na premissa de que a resposta do *digital twin* na simulação será igual à resposta do gêmeo físico, no momento da implantação de determinada estratégia, reduzindo os riscos de implementação (KRITZINGER et al., 2018).

Diversas técnicas de desenvolvimento de *Digital Twins* vêm sendo estudadas recentemente, sendo um campo de estudo que desperta muito interesse e é bastante atual na indústria. Dentre essas técnicas destaca-se o uso de redes neurais, que possuem grande capacidade de emular o comportamento de funções de complexidades diversas. Pela natureza dos processos, são muito utilizadas as redes que levam em conta os estados anteriores das variáveis de interesse, dividindo o campo de estudo em duas principais vertentes: redes neurais recorrentes e redes neurais autorregressivas. A principal diferença entre elas é a maneira com que aprendem as dependências de longo prazo. Nas redes neurais recorrentes comuns existem mecanismos de reutilização do sinal de saída que se propaga pela arquitetura da rede por um determinado prazo, influenciando as saídas futuras. Como ponto negativo, essa arquitetura sofre com o fenômeno de *vanishing gradient*, que atinge as redes neurais profundas e causa a diminuição da capacidade de aprendizagem da rede (KRITZINGER et al., 2018). O fenômeno de *vanishing gradient* é uma limitação que afeta a capacidade de aprendizagem das redes neurais profundas, e ocorre quando

os gradientes (derivadas parciais) dos pesos da rede se tornam muito pequenos à medida que as informações são propagadas pelas camadas, dificultando o ajuste dos pesos dos neurônios e, consequentemente, prejudicando o desempenho da rede (BENGIO et al., 1994).

As redes autorregressivas, por sua vez, utilizam uma abordagem ortogonal em relação às redes neurais recorrentes comuns, fornecendo os valores de entrada e saída de instantes passados como novas características de entrada para a rede. Essa arquitetura é composta por diversas camadas, sendo que cada uma delas está encarregada de estimar a distribuição condicional dos valores de saída para um número fixo de passos futuros, condicionado aos valores de entrada de um número fixo de passos anteriores. A principal vantagem desse tipo de rede é que ela não sofre do fenômeno de *vanishing gradient*, visto que as informações das camadas anteriores são explicitamente inseridas nas entradas das camadas posteriores, sem precisar passar pelo processo de retropropagação do erro. Isso permite que a rede seja treinada com um número maior de camadas, sem prejudicar o processo de aprendizagem. Essa característica tem sido explorada em diversas aplicações, como na previsão de séries temporais (LIU et al., 2020; LAGUNA et al., 2022). Uma visão comparativa entre as arquiteturas pode ser observada na Figura 4.

Figura 4 – Comparativo entre as arquiteturas de Redes LSTM e Redes NARX



Fonte: Heghedus; Shchipanov; Rong (2019)

Este trabalho tem como objetivo aplicar duas técnicas de Redes Neurais, NARX e LSTM, com o propósito de identificar o modelo de processo de um separador trifásico em uma plataforma de produção de petróleo *offshore* na Bacia de Campos. O intuito é criar um *Digital Twin* que possa simular realisticamente o processo e viabilizar testes e implementação de técnicas de

controle avançado. Os resultados obtidos pelas duas redes serão comparados considerando diferentes horizontes de previsão, visando avaliar a capacidade de ambas em replicar o comportamento do gêmeo físico, inclusive em condições limites. Este trabalho contribui para o avanço da aplicação de técnicas de modelagem de sistemas complexos e aprimoramento do controle avançado em processos de produção de petróleo *offshore*.

1.1 OBJETIVOS

1.1.1 Objetivos gerais

Este trabalho tem como objetivo geral o desenvolvimento de um *Digital Twin* do Separador Trifásico por meio da aplicação de técnicas de redes neurais, em especial as arquiteturas NARX e LSTM, utilizando dados históricos de leituras de entradas e saídas do sistema. Para tanto, serão realizados testes comparativos das duas arquiteturas considerando diferentes horizontes de predição, a fim de determinar a melhor abordagem para replicar com precisão o comportamento do gêmeo físico. Além disso, busca-se estabelecer e validar um método para obtenção de *Digital Twins* de sistemas de plataformas de produção *offshore*, possibilitando a implementação de técnicas avançadas de controle e tomada de decisão em ambientes operacionais desafiadores.

1.1.2 Objetivos específicos

- Obter os dados dos sensores do sistema de monitoramento da planta de processo (PI Processbook) para análise em ambiente de programação em *Python*;

- Implementar as redes neurais LSTM e NARX para identificar o modelo do processo de separação de fases do Separador Trifásico;

- Otimizar os parâmetros de treinamento das redes neurais usando a biblioteca Optuna, por meio da Otimização Bayesiana;

- Comparar os resultados de previsão das redes neurais NARX e LSTM para diferentes horizontes de previsão (1, 3, 5 e 10 passos) e determinar qual modelo tem melhor desempenho;

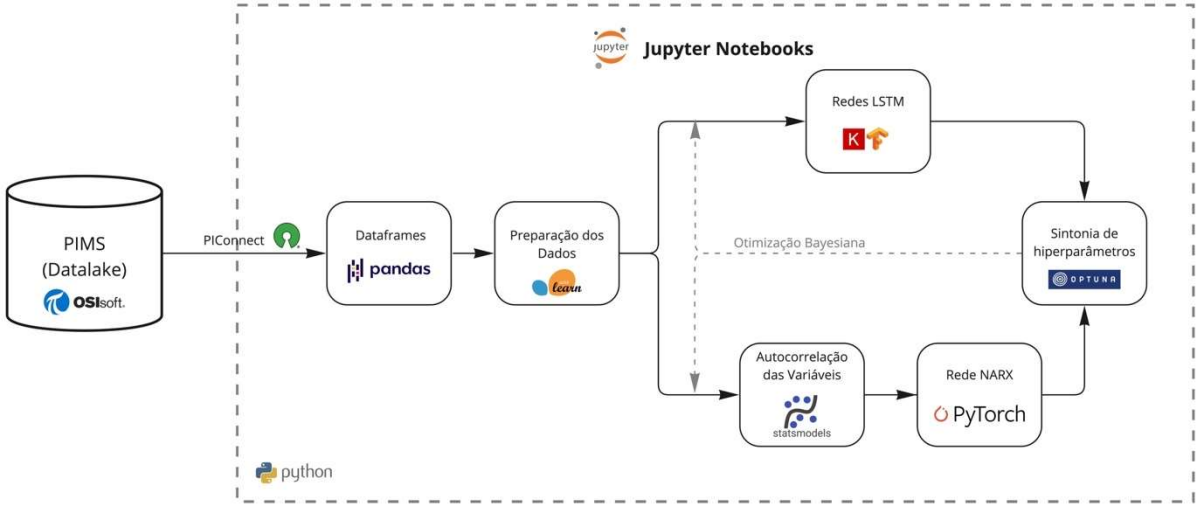
- Avaliar o desempenho do *Digital Twin* do Separador Trifásico por meio de métricas de erro amplamente aceitas e difundidas;

1.2 METODOLOGIA

Para obter os resultados esperados, foi realizada uma ampla revisão bibliográfica do Estado da Arte e das técnicas mais utilizadas para a obtenção de *Digital Twins* em plantas de processo. Após a revisão, verificou-se a adequação do uso de redes neurais autorregressivas do tipo NARX, assim como das redes recorrentes LSTM, e iniciou-se um estudo voltado às técnicas de identificação de sistemas dinâmicos.

Os dados da planta de processo foram coletados do servidor de informações da planta utilizado pela Petrobras, chamado *OSISoft Plant Information* (PI). Para a análise dos dados, utilizou-se *Python* por meio de *Jupyter Notebooks*, visando a simplicidade de utilização e a demonstração de etapas de desenvolvimento. Para realizar a conexão entre o servidor do PI e o *Python*, empregou-se a biblioteca PI-Connect, o que permitiu a coleta de dados por períodos estendidos e a utilização de dados mais atuais com facilidade. Assim, estabeleceu-se um *pipeline* de obtenção dos dados para as etapas subsequentes da pesquisa. Na Figura 5 pode ser observado um diagrama esquemático de bibliotecas e funções usados na construção do projeto e treinamento da rede NARX.

Figura 5 – Diagrama de implementação do código *Python* desenvolvido

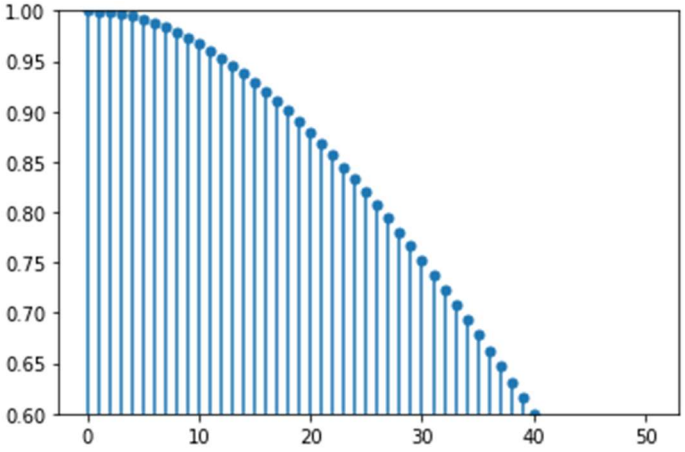


Fonte: Autor (2023)

Os dados coletados foram organizados em *dataframes* utilizando a biblioteca Pandas, na qual foram realizadas etapas de pré-processamento. De acordo com Wang (2019), as redes neurais aprendem com maior velocidade quando recebem dados normalizados, pois utilizam o método do gradiente descendente para atualização dos pesos de cada neurônio. Para isso, utilizou-se a biblioteca *Scikit-Learn* para efetuar a normalização, sendo o método escolhido o *RobustScaler,* por sua robustez a *outliers*. Seguindo as boas práticas de *machine learning*, o comando *.fit()* do *RobustScaler* foi realizado apenas nos dados de treinamento, enquanto que os dados de teste ou validação sofreram apenas o comando *.transform()*. Esse cuidado é importante para evitar o problema conhecido como *data leakage*, que ocorre quando são transformados os dados de teste juntamente com os dados de treinamento para efetuar transformações, inflando artificialmente a acurácia do modelo desenvolvido (KHALEGHI-MOGHADDAM, 2020).

Para o desenvolvimento do *Digital Twin* utilizando a rede neural NARX, é importante avaliar a autocorrelação das variáveis envolvidas no processo. A autocorrelação é uma medida de quão correlacionada uma variável é consigo mesma em diferentes momentos no tempo. Isso é importante para entender como o valor atual de uma variável pode estar correlacionado com seus valores passados. Uma maneira comum de avaliar a autocorrelação é por meio de *lags,* que são valores defasados no tempo. Os *lags* representam a quantidade de tempo entre o valor atual e o valor passado a ser avaliado. Na Figura 5, é possível visualizar um exemplo da curva de autorregressão da variável *Gas Flow*, que representa a correlação da variável com seus próprios valores passados. No eixo X, estão os *lags* selecionados, ou seja, a quantidade de tempo defasado que está sendo avaliado. No eixo Y, está a medida de correlação entre o valor atual e o valor passado. Essa curva de autorregressão é importante para identificar padrões e tendências que possam ajudar na previsão do comportamento futuro da variável em questão. Um exemplo de curva de autocorrelação pode ser observado na figura 6.

Figura 6 – *Plot* da autocorrelação da variável *Gas Flow*



Fonte: Autor (2023)

Para realizar o cálculo da autocorrelação e plotagem dos gráficos, foi utilizada a biblioteca *Statsmodels* em Python. A autocorrelação é um importante indicador da presença de correlação serial nos dados, ou seja, a dependência entre valores observados em diferentes momentos de tempo. O número de *lags* utilizados para treinamento da rede neural NARX foi estabelecido a partir da análise desses gráficos de autocorrelação, em que se define um valor mínimo para a correlação. Neste trabalho, optou-se por considerar esse valor mínimo como um hiperparâmetro da rede, o qual foi ajustado com o objetivo de obter uma melhor capacidade de predição.

Para a implementação da rede neural NARX e seu treinamento, optou-se pela utilização da biblioteca *Pytorch*, que é uma biblioteca *open-source* de aprendizado profundo para *Python*. Essa biblioteca é amplamente utilizada na indústria e na academia devido à sua flexibilidade e eficiência computacional, permitindo a construção de modelos complexos de forma relativamente fácil e rápida. Além disso, a biblioteca *Pytorch* oferece uma ampla gama de recursos para o desenvolvimento de modelos de aprendizado profundo, como a possibilidade de utilização de GPUs para acelerar o processamento, diversas funções de ativação e otimizadores, além de um sistema de autodiferenciação que simplifica o processo de retropropagação dos erros durante o treinamento da rede. A escolha da biblioteca *Pytorch* para a implementação da rede NARX neste trabalho foi baseada em sua popularidade e versatilidade, bem como em sua capacidade de lidar com grandes conjuntos de dados e modelos complexos de aprendizado de máquina. Para a implementação da rede LSTM, uma escolha foi feita em relação à tecnologia utilizada. Optou-se por utilizar a biblioteca Keras em conjunto com o *framework* TensorFlow. Essa combinação oferece uma plataforma poderosa e flexível para o desenvolvimento e treinamento de redes neurais, incluindo a arquitetura LSTM. A biblioteca Keras é conhecida por sua facilidade de uso e alto nível de abstração, o que permite uma implementação mais rápida e intuitiva da rede neural. Ela fornece uma ampla gama de camadas e funções de ativação, incluindo suporte específico para redes LSTM.

A divisão dos dados em conjuntos de treinamento e teste é uma prática comum em problemas de modelagem preditiva e aprendizado de máquina, sendo crucial para avaliar o desempenho do modelo. Para o presente estudo, a divisão foi realizada seguindo a premissa estabelecida por Ng (2018) [NG, A. Machine Learning Yearning: Technical Strategy for AI Engineers in the Era of Deep Learning. Deeplearning.ai, 2018], que sugere a alocação de 95% dos dados para treinamento e 5% para teste. Tal abordagem tem como objetivo garantir um conjunto de teste suficientemente grande para avaliar a capacidade de generalização do modelo, e simultaneamente, um conjunto de treinamento adequado para permitir a aprendizagem dos padrões presentes nos dados. É importante destacar que a escolha do tamanho do conjunto de treinamento e teste deve ser feita considerando a complexidade do problema em questão, o tamanho do *dataset* disponível e outros fatores específicos do problema em análise.

O treinamento de modelos para predição com diferentes números de passos de antecedência é uma prática comum em aplicações de séries temporais. De acordo com Azzouni et al. (2020), o número de passos de antecedência é um hiperparâmetro importante no treinamento de modelos de séries temporais, e sua escolha deve ser baseada em considerações específicas do problema em análise. Além disso, a escolha do número de passos de antecedência também pode

ter implicações na interpretação dos resultados obtidos, como destacado por Ruiz et al. (2019). Esses passos são definidos pela frequência de amostragem do processo, que se dá a cada 1 segundo, portanto cada passo representa a passagem de 1 segundo no tempo. Nesse sentido, a escolha de treinar o modelo proposto com 1, 3, 5 e 10 passos de antecedência se mostrou adequada para avaliar a capacidade de predição em diferentes horizontes temporais e analisar a robustez do modelo diante de diferentes configurações. Essa abordagem também é consistente com a literatura, que aponta a importância da avaliação do desempenho do modelo em diferentes horizontes de previsão (HYNDMAN; ATHANASOPOULOS, 2018). Os prazos escolhidos são de importância por permitirem tempo de atuação de sistemas de controle que possam ser desenvolvidos utilizando o Di*gital Twin* aqui proposto.

Para a análise de desempenho dos modelos, foram utilizadas métricas comuns na avaliação de modelos de previsão, como o erro quadrático médio (MSE) e o erro médio absoluto (MAE) (MONTGOMERY et al., 2015). A escolha dessas métricas é fundamentada na sua capacidade de avaliar a precisão e a capacidade de explicação do modelo, respectivamente. Além disso, a utilização de múltiplas métricas também permite uma avaliação mais abrangente e completa do desempenho do modelo (CHEN et al., 2018). Dessa forma, a escolha dos passos de antecedência e das métricas de avaliação se mostrou fundamentada na literatura e adequada para a construção e avaliação do modelo de *Digital Twin* proposto neste trabalho.

Com o objetivo de obter modelos com hiperparâmetros otimizados, foi implementado um código que utiliza a biblioteca Optuna para a sintonia dos hiperparâmetros do modelo. A biblioteca Optuna é amplamente utilizada em competições de *machine learning* por sua facilidade em realizar um grande número de tentativas de maneira persistente, permitindo a realização de centenas de iterações e o armazenamento de todo o processo de treinamento, possibilitando a continuação da etapa de otimização em caso de interrupções indesejadas. A otimização de hiperparâmetros com o Optuna foi realizada utilizando a técnica de Otimização Bayesiana, que apresenta melhores resultados de convergência (DENG, 2019). A otimização de hiperparâmetros é uma etapa crucial no processo de modelagem, pois afeta diretamente o desempenho do modelo. A biblioteca Optuna foi selecionada para esta tarefa, uma vez que demonstrou eficácia comprovada em competições de *machine learning* (KOMER et al., 2014). Além disso, a abordagem de Otimização Bayesiana adotada neste trabalho é uma técnica estatística amplamente reconhecida por sua eficácia na otimização de hiperparâmetros em diversas aplicações de *machine learning* (SNOEK et al., 2012). A utilização da biblioteca Optuna permitiu a sintonia dos

hiperparâmetros de forma automatizada e eficiente, levando a resultados mais robustos e confiáveis. O processo de otimização pode ser bastante oneroso, mas com o uso desta biblioteca foi possível reduzir significativamente o tempo de sintonia dos hiperparâmetros, tornando a modelagem mais eficiente e prática (KOMER et al., 2014). No presente estudo, foram realizados 500 ciclos de otimização de hiperparâmetros para cada modelo de rede treinado. Durante esses ciclos, foram ajustados diferentes parâmetros para otimizar o desempenho do modelo.

Parâmetros otimizados para as redes NARX:

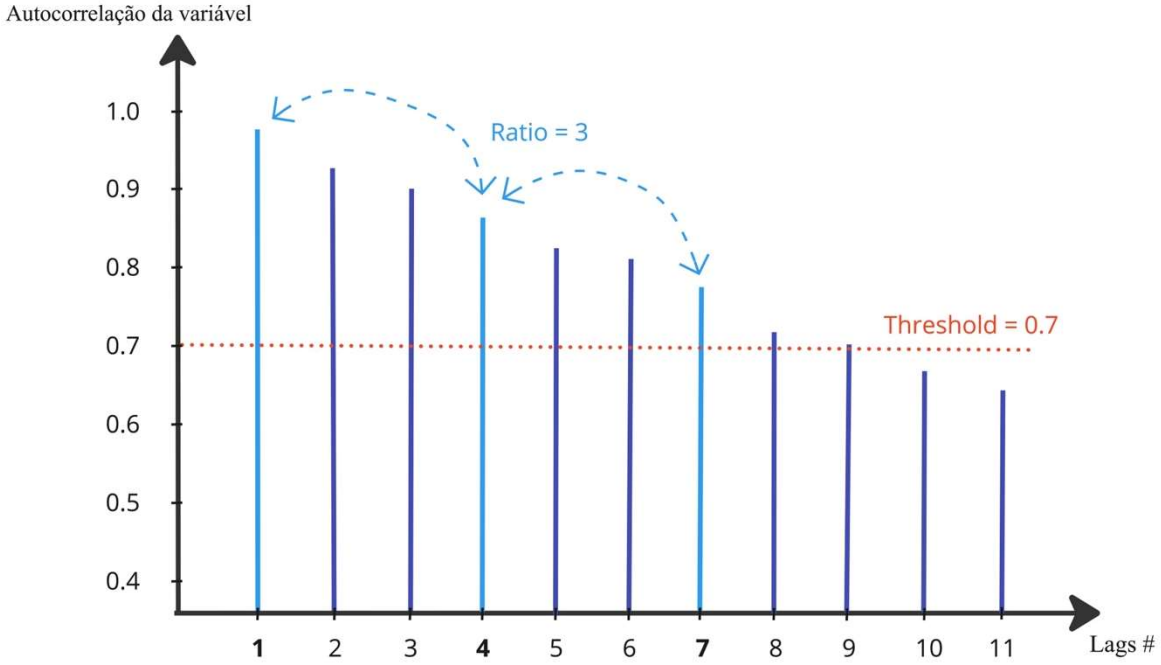
- Fator de autocorrelação: a definição deste parâmetro resulta na quantidade de *lags* a serem levados em consideração para cada variável de entrada ou saída. Esse valor aqui estabelecido é um *threshold* para incluir ou excluir *lags* do modelo. Valores entre 0,6 e 1,0;
- *Learning rate*: É a taxa de aprendizado da rede, parâmetro extremamente importante para que a rede alcance convergência satisfatória ao mínimo da função custo. Valores entre 1 x 10^{-1} e 1 x 10^{-6};
- *Dropout rate*: probabilidade de cada um dos neurônios internos ter seu peso considerado como peso 0 a cada rodada de treinamento. Objetiva evitar o *overfit*. Valores entre 0 e 0,5;
- *Hidden Size Factor*: Fator que define o número de neurônios na camada escondida. Como a quantidade de entradas varia com a alteração da quantidade de *lags* a serem considerados em cada rede, o valor otimizado foi um fator conforme a equação abaixo:

$$(Nentradas + Nsaídas) * fator$$

- *Ratio*: para evitar que as redes NARX tivessem um número de entradas muito grande e com dados de elevada autocorrelação, foi estabelecido um valor que representa o tamanho da janela entre os *lags* selecionados. A figura 7 ilustra esse fator que foi considerado como hiperparâmetro do modelo. Valores entre 1 e 40;
- *Optimizer:* o tipo de otimizador a ser utilizado na aprendizagem e ajuste dos pesos dos neurônios é um dos principais hiperparâmetros de uma rede neural. Nesta etapa foram testados os otimizadores: Adam, Adadelta, Adagrad, RMSProp e SGD;

A Figura 7 ilustra a influência dos parâmetros "*ratio*" e "fator de autocorrelação" nos dados a serem fornecidos para as redes NARX.

Figura 7 – Diagrama de Influência dos Parâmetros Ratio e Threshold na Quantidade de Entradas Defasadas Fornecidas à Rede NARX.



Fonte: Autor (2023)

No caso das redes LSTM, a biblioteca utilizada para implementação foi *Keras* rodando sobre o *Tensorflow*. Para esta arquitetura, os hiperparâmetros otimizados foram os seguintes:

- *Batch size:* é o tamanho do lote de dados de treinamento que é processado pela rede em cada iteração. Seu valor influencia tanto a velocidade de convergência do treinamento quanto a acurácia do modelo. Valores possíveis entre 8 e 64 dados por *batch*.
- *Dropout rate*: probabilidade de cada um dos neurônios internos ter seu peso considerado como peso 0 a cada rodada de treinamento. Objetiva evitar o *overfit*. Valores entre 0 e 0,5;
- *Learning rate*: É a taxa de aprendizado da rede, parâmetro extremamente importante para que a rede alcance convergência satisfatória ao mínimo da função custo. Valores entre 1×10^{-1} e 1×10^{-6};
- *N_past*: Número de leituras passadas a serem alimentadas à rede. Valores entre 1 e 60;
- Função de ativação: utilizada para introduzir não-linearidades na rede, é um dos hiperparâmetros que mais impacta no desempenho do modelo. Para este estudo, foram testadas as funções ReLU, tanh, ELU, e SeLU;
- Unidades LSTM: trata-se de um hiperparâmetro crucial, que define a capacidade da rede de aprender e generalizar informações. Profundidade usada entre 1 e 30;

Foram realizadas 500 iterações para cada modelo a fim de se obter a melhor combinação de hiperparâmetros. No total, foram gerados oito modelos, sendo quatro implementados com a arquitetura LSTM e quatro com a arquitetura NARX, cada um deles projetado para realizar predições com 1, 3, 5 e 10 passos de antecedência. Esse processo de otimização tem como objetivo a obtenção de modelos mais eficientes e precisos, o que pode contribuir significativamente para aprimorar a acurácia das previsões realizadas.

Na fase de testes realizada, foram coletados quatro novos conjuntos de dados em diferentes períodos de operação da planta de processo e em diferentes níveis de estabilidade, para avaliar a robustez e capacidade de generalização do *Digital Twin*. Cada conjunto de dados de teste foi coletado pelo menos três meses após o conjunto utilizado para o treinamento da rede, sendo o pior caso um conjunto de dados coletado 11 meses após o conjunto de treinamento. A avaliação da capacidade de generalização é fundamental para garantir a qualidade do modelo e sua aplicação em cenários reais (ZHU et al., 2020; SHI et al., 2018). Os resultados dos oito modelos (quatro LSTM e quatro NARX), foram gerados e comparados entre si para cada um dos quatro conjuntos de teste. A análise dos resultados obtidos permitiu verificar a capacidade dos modelos em prever com precisão o comportamento das variáveis de interesse em diferentes situações de operação da planta.

A avaliação dos resultados é conduzida por meio da análise das métricas de erro, mais especificamente o Erro Médio Absoluto (MAE) e o Erro Médio Quadrático (MSE). Essas métricas são amplamente utilizadas na área de previsão para medir a discrepância entre os valores previstos e os valores reais. O MAE é calculado como a média das diferenças absolutas entre as previsões e os valores observados. Ele fornece uma medida direta do erro médio sem considerar a direção da discrepância, o que o torna uma métrica intuitiva e fácil de interpretar. Por outro lado, o MSE calcula a média dos quadrados das diferenças entre as previsões e os valores reais. Ele penaliza erros maiores de forma mais significativa devido ao uso dos quadrados, o que o torna sensível a valores discrepantes. Ao analisar essas métricas de erro, é possível ter uma noção da qualidade das previsões realizadas pelos modelos em diferentes horizontes de previsão. A variação dessas métricas à medida que o horizonte de previsão aumenta oferece informações sobre o desempenho relativo dos modelos e sua capacidade de lidar com dependências de longo prazo nos dados.

Neste trabalho em particular, o uso do MAE e do MSE permite uma avaliação abrangente dos modelos, considerando tanto a magnitude quanto a distribuição dos erros. Com base nessas

métricas, é possível identificar qual modelo apresenta um desempenho consistente e confiável na previsão de eventos futuros, levando em conta as dependências temporais de longo prazo presentes nos dados analisados. A escolha dessas métricas específicas foi feita com base em sua relevância e aplicabilidade na área de previsão, mas é importante ressaltar que existem outras métricas disponíveis que podem ser adequadas para diferentes contextos e objetivos de estudo. A análise das métricas de erro contribui para uma compreensão mais aprofundada do desempenho dos modelos e auxilia na seleção daquele que melhor se adequa às necessidades da tarefa de previsão em questão.

No Capítulo 2, será apresentado o referencial teórico do trabalho, abordando sistemas de produção de petróleo *offshore*, estratégias de controle, modelagem e identificação de processos com redes neurais (NARX e LSTM), e otimização bayesiana para sintonização de hiperparâmetros. No Capítulo 3, serão detalhados os métodos e dados utilizados na implementação do *Digital Twin*, incluindo a escolha dos métodos de implementação, a aquisição e preparação dos dados, e as etapas de desenvolvimento do sistema. No Capítulo 4, serão apresentados os resultados e discussões do trabalho, abordando dados, gráficos, métricas de desempenho e outros resultados relevantes. Será avaliada a eficácia e robustez do *Digital Twin* desenvolvido, discutindo sua capacidade de atingir os objetivos propostos. Serão destacadas as contribuições e conclusões obtidas a partir das análises realizadas, assim como possíveis direcionamentos para trabalhos futuros. O Capítulo 5 trará as conclusões do trabalho, destacando os principais resultados e descobertas obtidos. Serão discutidos aspectos positivos e limitações do estudo, além de sugerir possíveis direções para pesquisas posteriores na área de digital twins utilizando redes neurais. O objetivo deste capítulo é fornecer uma visão geral das conclusões do trabalho e apontar caminhos para a continuidade da pesquisa nesse campo.

2 REFERENCIAL TEÓRICO

2.1 PRODUÇÃO DE PETRÓLEO *OFFSHORE*

A produção de petróleo offshore é uma atividade complexa que envolve a extração de fluidos compostos por misturas de gás natural, petróleo cru e água salgada (BROOKS et al., 2015). Essas misturas são de difícil manuseio, medição e transporte, além de representarem riscos ambientais e inviabilidade econômica (LIMA et al., 2014). Dessa forma, para que a exportação de petróleo e gás dos locais de extração, seja de plataformas offshore ou poços terrestres, para os locais de armazenamento ou refino seja possível, é necessário que sejam realizadas etapas de tratamento, sendo a primeira delas o processamento primário do petróleo (ABDEL-AAL; AGGOUR; FAHIM, 2003).

O processamento primário do petróleo consiste em uma série de operações unitárias que têm como objetivo separar os componentes do petróleo, como gás natural, óleo e água, para que possam ser tratados individualmente (LIMA et al., 2014). Essas operações envolvem desde a separação inicial das misturas até a remoção de contaminantes, como areia e sais (BROOKS et al., 2015). Vale ressaltar que a produção de petróleo offshore apresenta desafios únicos em relação à produção terrestre. Além das dificuldades já mencionadas, a produção offshore envolve a operação em ambientes hostis, como mar aberto, que pode levar a acidentes e impactos ambientais (GOLDBERG; GRUPO DE TRABALHO DE CONTROLE DE DERRAMAMENTOS DE PETRÓLEO, 2001).

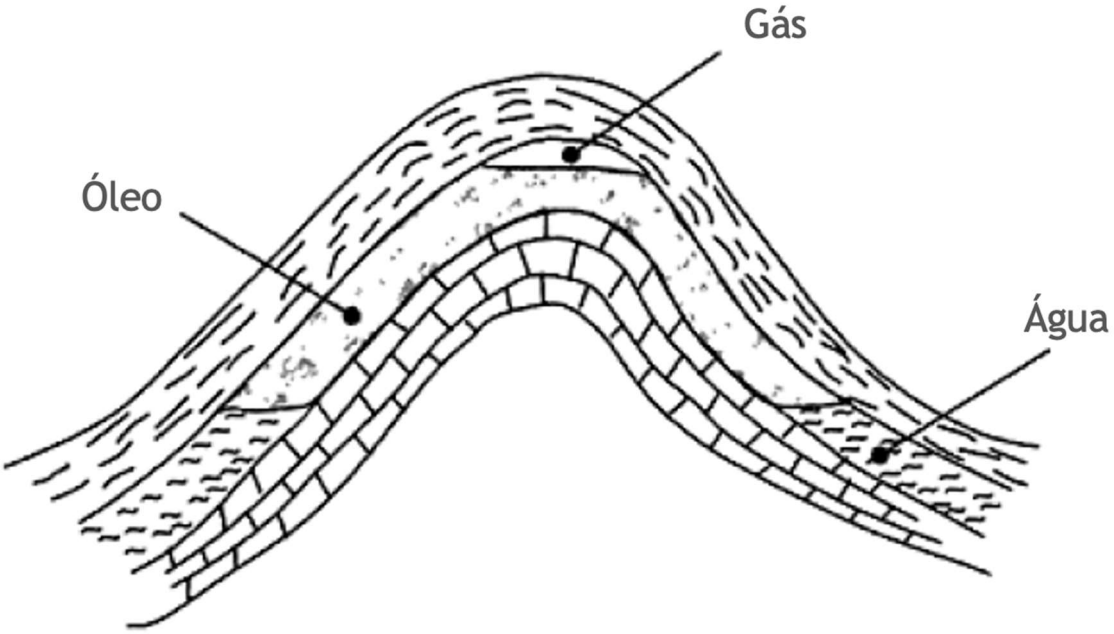
Diante dessas dificuldades, torna-se necessário o desenvolvimento de tecnologias e estratégias para otimizar o processo de produção *offshore*, reduzindo os riscos ambientais e os custos envolvidos.

2.1.1 Formação e migração do petróleo

De acordo com Arene e Kitwood (1979), o petróleo é uma substância oleosa que ocorre naturalmente em rochas sedimentares. A formação do petróleo tem origem na matéria orgânica acumulada sob camadas de lama, sal, rochas e sedimentos que, após milhões de anos sob pressões e temperaturas elevadas, sofreu processos de transformação química, originando o petróleo. Para que sejam formados reservatórios de petróleo, além das condições já descritas, é necessária a presença de uma rocha capeadora (ou trapa), que impede que o óleo formado e em migração para regiões superiores seja perdido para o mar. Essa rocha é um dos principais fatores para a

prospecção de petróleo, pois é uma condição essencial para a existência de um acúmulo significativo e para a viabilidade econômica de um bloco exploratório (MILANI et al., 2018). A migração do petróleo também é um processo importante a ser considerado. Conforme explicado por Gary et al. (2008), a migração ocorre devido à diferença de densidade entre o petróleo e a rocha geradora. Quando a pressão interna da rocha atinge um ponto crítico, o petróleo é expelido e migra para outras camadas de rochas porosas e permeáveis, onde se acumula formando reservatórios. A compreensão desses processos é fundamental para o entendimento da formação e exploração dos reservatórios de petróleo *offshore*. Um corte lateral de um reservatório está representado na Figura 8, permitindo melhor compreensão.

Figura 8 – Formação geológica de um reservatório típico



Fonte: Abdel-Aal; Aggour; Fahim (2003)

2.1.2 Elevação e escoamento

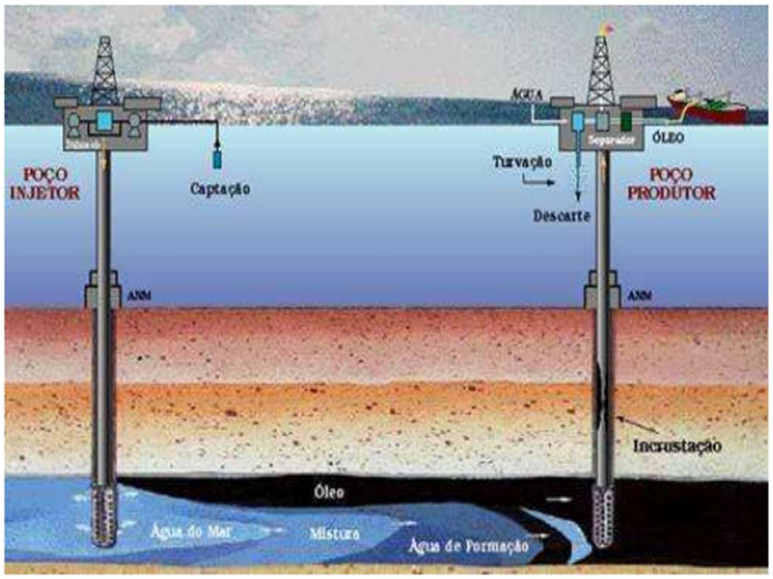
De acordo com Abdel-Aal, Aggour e Fahim (2003), o petróleo em reservatórios se encontra em condição trifásica, composta por óleo, água e gás, que podem se misturar em emulsões complexas durante o processo de extração e produção. A presença de gás no reservatório é um fator determinante para a produtividade e a pressão de reservatório é uma das principais características que determinam a vida útil e recuperabilidade do óleo contido no poço (ARAUJO; BERNARDO; JORGE, 2012). Com a extração do petróleo, ocorre uma migração de óleo e água de outras partes do reservatório para compensar a retirada, o que pode resultar em uma maior concentração de água produzida em conjunto com o petróleo (REIS et al., 2018). A relação custo-

benefício da exploração de um poço pode ser gradualmente deteriorada devido à maior concentração de água produzida em conjunto com o petróleo, o que pode levar à inviabilidade econômica de sua exploração e ao abandono do poço. A fim de aumentar o fator de recuperação de petróleo, é comum a utilização de técnicas de injeção de água do mar, gás ou reinjeção de água produzida, por meio de poços de injeção (PETROBRAS, 2007). Estes poços podem ser instalados em partes mais profundas do reservatório para injeção de água ou na chamada capa de gás, que é a parte de maior elevação do reservatório, para injeção de gás como método de recuperação.

Em resumo, a produção de petróleo envolve uma série de processos e fatores que afetam sua viabilidade econômica e seu fator de recuperação, como a presença de gás no reservatório, a pressão de reservatório, a migração de óleo e água, e a utilização de técnicas de injeção de água e gás. Por isso, é importante que os profissionais envolvidos na exploração e produção de petróleo estejam atualizados sobre as melhores práticas e tecnologias disponíveis para otimizar o processo e reduzir os custos envolvidos.

Na Figura 9 estão representados os poços injetores e produtores e uma interação desses fluidos no reservatório.

Figura 9 – Interação entre poços injetores e produtores no reservatório



Fonte: Thomas (2001)

A produção de petróleo é influenciada por diversos fatores, sendo a pressão de reservatório uma das principais características de um poço de produção, pois ela determina sua vida útil e a recuperabilidade do óleo contido (WANG et al., 2021). A existência de gás no reservatório é

essencial para a produtividade, já que é o principal responsável pela surgência do poço. No entanto, um poço que produz uma grande quantidade de gás associado ao petróleo pode apresentar maior surgência, mas a recuperação desse poço será comprometida em pouco tempo, pois o gás contido no reservatório é a energia produtiva natural de um poço (KHALIL et al., 2016).

Com o tempo, os poços maduros perdem surgência, tornando-se inviáveis sem o uso de métodos de elevação artificial. Alguns desses métodos incluem o uso de bombas centrífugas submersas, ou o uso de gas lift (THOMAS, 2001). A técnica de elevação via *gas lift* do tipo contínuo é especialmente relevante, sendo amplamente utilizada e tendo grande impacto no escoamento do petróleo. No entanto, ela também é uma das principais causas do fenômeno conhecido como "golfada", que ocorre devido à formação de "bolhas" no fluxo de produção (PEDERSEN; DURDEVIC; YANG, 2015).

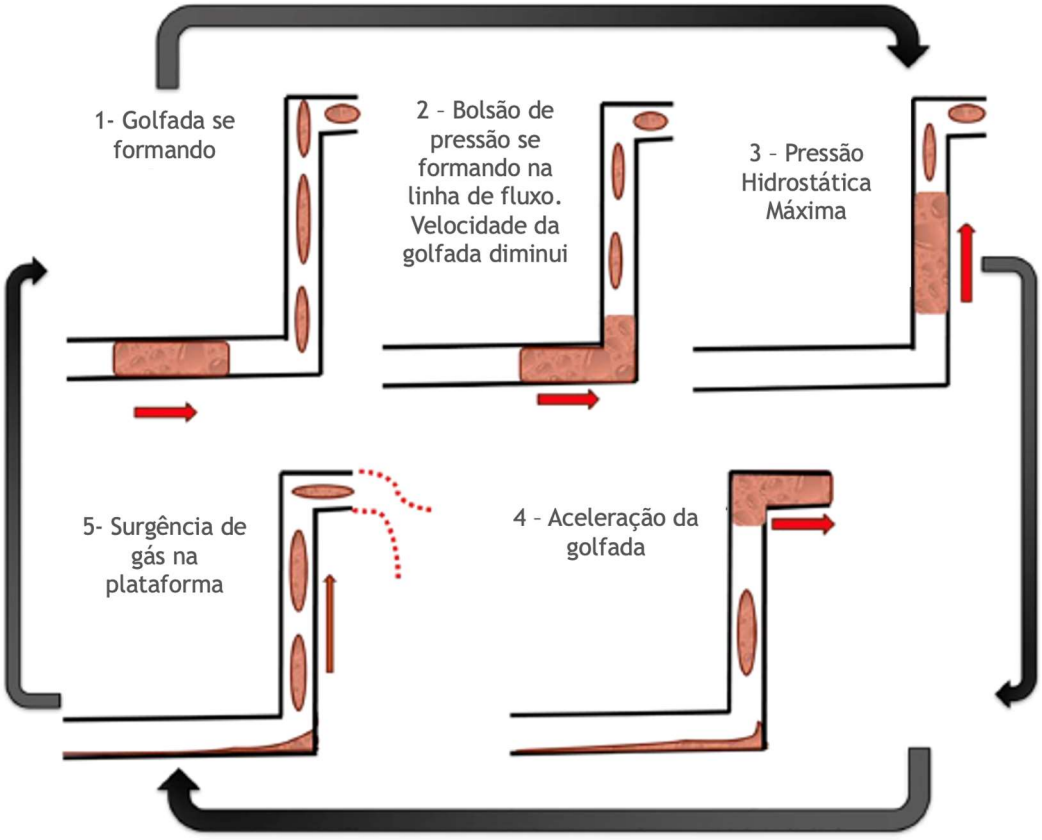
2.1.3 Produção por golfadas

A produção de petróleo é um processo complexo, que envolve a extração de hidrocarbonetos do subsolo e seu transporte até as unidades de processamento. Um dos problemas que pode afetar esse processo é a formação de golfadas, que pode causar diversos efeitos nocivos à produção, como paradas de unidade, perdas de produção, queima de gás no *flare,* corrosão acelerada de linhas e equipamentos, entre outros (BERTANI et al., 2019). As golfadas são um fenômeno cíclico que ocorre quando um bloco de líquido impede o fluxo de gás na linha de produção. Com o acúmulo de gás a montante desse bloqueio, a pressão se eleva até que o bloqueio é rompido, permitindo que uma grande quantidade de gás chegue à unidade de processamento de uma só vez. Isso pode causar uma grande variação na carga e pressão do processo, afetando a produção de forma significativa (BORDALO; MOROOKA, 2018).

Diversos fatores podem contribuir para a formação de golfadas, como a geometria da linha de fluxo, a batimetria do *riser* no fundo do oceano e a vazão de *gas lift* inadequada. O *gas lift* é um dos métodos mais utilizados para a elevação artificial de petróleo, sendo capaz de aumentar a eficiência da produção, mas também pode ser uma das causas da formação de golfadas se a vazão for muito alta ou muito baixa (BENITO et al., 2019). Para mitigar os efeitos das golfadas na produção de petróleo, é necessário o uso de técnicas de controle de produção, como o monitoramento da pressão e da vazão na linha de produção, o ajuste da vazão de *gas lift* e a instalação de equipamentos de separação de gás e líquido. O objetivo é evitar a formação de blocos de

líquido na linha de fluxo, reduzindo a frequência e a severidade das golfadas e mantendo a eficiência da produção (BERTANI et al., 2019). Uma representação gráfica da golfada está demonstrada na Figura 10.

Figura 10 – Formação de golfadas nos risers de produção



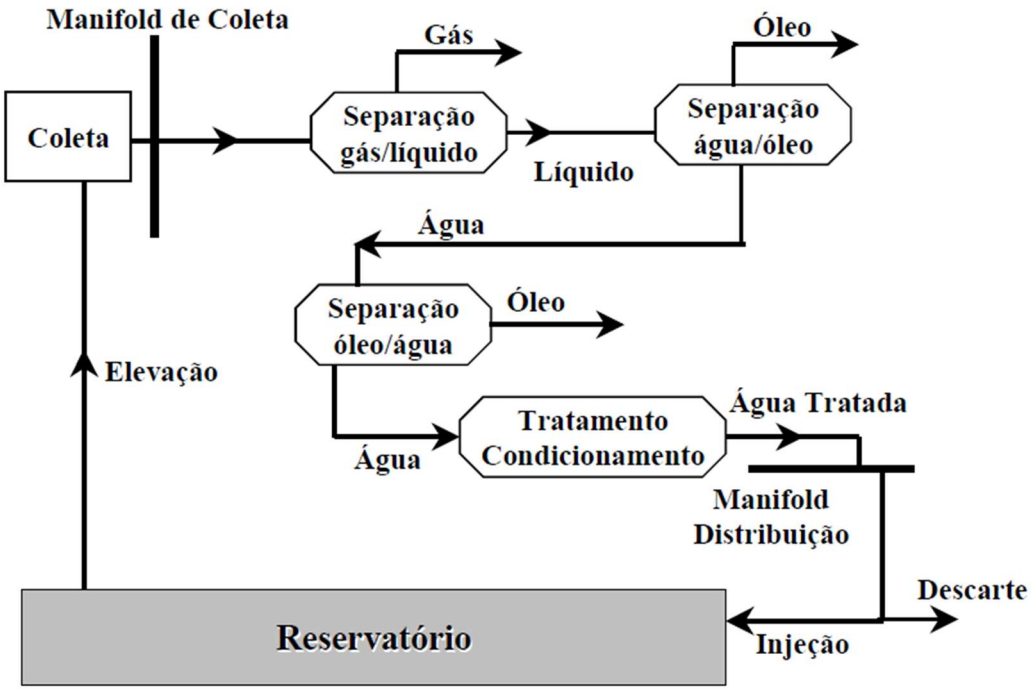
Fonte: Nnabuife; Tandoh; Whidborne (2022)

2.1.4 Separação e tratamento

Realizada a coleta do petróleo, sendo este uma mistura entre gás, óleo e água, o fluxo é direcionado ao sistema de processamento primário onde é realizado o tratamento e separação por gravidade.

A Figura 11 demonstra um fluxograma completo de um tipo de sistema de produção bastante utilizado.

Figura 11 – Fluxo de processo de produção e tratamento de petróleo e suas 3 fases

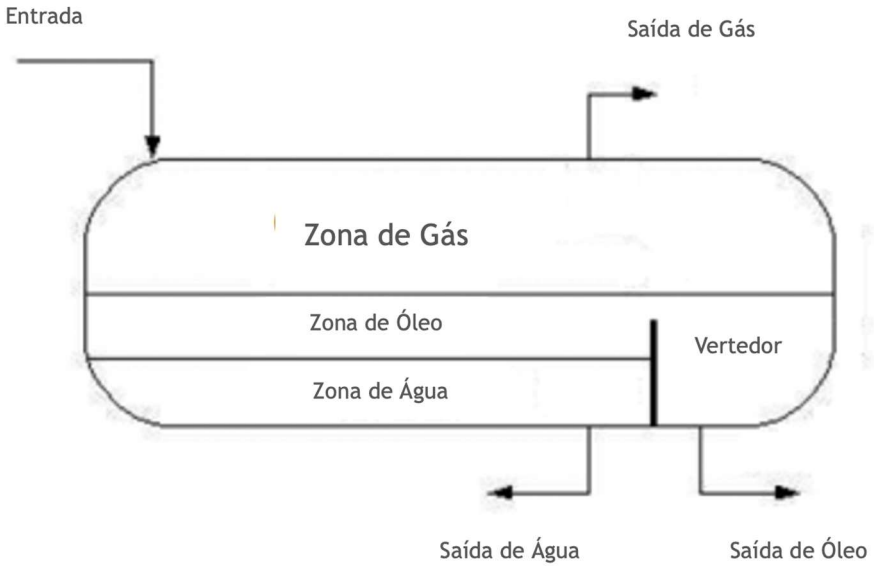


Fonte: Petrobras (2007)

O processo de separação e tratamento do petróleo *offshore* inicia-se com o uso do separador trifásico, que é responsável por separar o gás, o óleo e a água. Após a separação, cada corrente segue para um tratamento específico a fim de garantir que os produtos estejam de acordo com as especificações requeridas pela indústria. É comum a realização de tratamentos com o intuito de controlar a umidade do gás, o teor de óleos e graxas da água descartada, o percentual de água e sedimentos do óleo, dentre outros aspectos (THOMAS, 2001). O óleo segue então para os tratadores eletrostáticos, que utilizam a eletrostática para aglutinar gotículas de água, acelerando assim o processo de separação. Esses tratadores são extremamente importantes, uma vez que a presença de água no óleo pode causar problemas sérios no processo de refino e, consequentemente, reduzir a qualidade do produto final (VASCONCELOS et al., 2020). Após o tratamento pelo separador, a água é direcionada para os hidrociclones, onde ocorre a remoção de óleo por turbilhonamento. Essa etapa é essencial para garantir que a água esteja livre de resíduos oleosos antes de ser descartada no mar ou reinjetada nos poços de injeção. Vale ressaltar que esse processo é potencializado pela diferença de densidade entre água e óleo (ZHOU et al., 2019). Por fim, a água passa pelos flotadores, onde ocorre a remoção final de resíduos oleosos por flotação a gás dissolvido e injeção de produtos coagulantes/flotantes, como o polieletrólito.

O gás, por sua vez, é direcionado para a compressão, onde sua pressão é elevada para atender às pressões de trabalho das plantas de gás combustível, *gas lift* e exportação (THOMAS, 2001). Diante disso, é possível perceber que o processo de separação e tratamento do petróleo *offshore* é fundamental para garantir a qualidade do produto final e evitar danos ao meio ambiente. É necessário, portanto, que as empresas do setor invistam em tecnologias cada vez mais avançadas a fim de aprimorar esse processo e torná-lo mais eficiente e seguro (BENTLEY et al., 2016). Na Figura 12 está representado um corte lateral de um separador trifásico.

Figura 12 – Visualização interna de um separador trifásico horizontal típico



Fonte: Ghaffarkhah; Shahrabi; Moraveji (2018)

O controle de nível nos separadores e tratadores eletrostáticos é crucial para garantir a eficiência do processo de separação de óleo, água e gás. O controle de nível é responsável por manter o líquido em um nível constante dentro do equipamento, garantindo uma eficiente separação dos componentes. De acordo com Al-Anazi e Fahim (2019), o controle de nível em separadores trifásicos é importante porque o nível do líquido dentro do equipamento pode afetar a separação adequada de líquido e gás, além de afetar o fluxo e distribuição de líquidos no equipamento. Já em tratadores eletrostáticos, o controle de nível é ainda mais importante, uma vez que a eficiência do tratamento eletrostático depende da presença de uma camada uniforme e estável de líquido dentro do equipamento (SUNDARAMOORTHY et al., 2013). Além disso, um controle de nível inadequado pode levar a uma série de problemas operacionais, tais como a formação de emulsões, acúmulo de óleo e aumento do tempo de retenção, reduzindo a eficiência da separação e aumentando os custos operacionais (AL-ANAZI E FAHIM, 2019). Para evitar tais

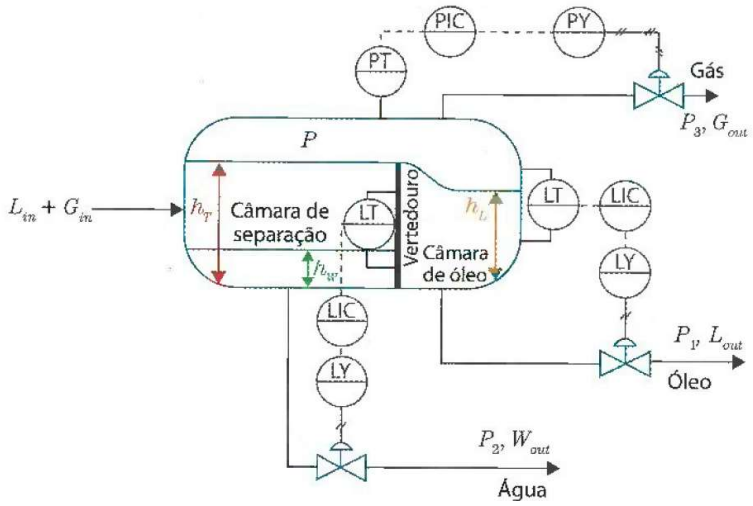
problemas, os separadores e tratadores eletrostáticos são equipados com instrumentos de controle de nível, como válvulas de controle, sensores de nível e transmissores de pressão. Esses equipamentos são capazes de controlar o nível de líquido dentro dos equipamentos, garantindo uma separação adequada dos componentes e maximizando a eficiência operacional.

O controle típico realizado em um sistema de separação trifásico tem três controladores principais:

- Controlador de pressão: responsável pelo controle da pressão de separação, importante para garantir o desprendimento de gás à quantidade desejada enquanto mantém ainda pressão suficiente para realizar o escoamento dos produtos da separação. A abertura da válvula controladora de pressão libera o gás para as etapas de tratamento e compressão, e também afeta a taxa de produção de líquido.
- Controlador de nível da câmara principal: é o principal controlador do processo, responsável por estabelecer o ponto de corte entre o fluido que está na câmara principal e o fluido que verte para a câmara de óleo. Os sensores de nível trabalham com medição de interface óleo-água e, embora não possuam alta precisão suficiente para garantir um ajuste de *setpoint* ideal permanente, permitem um processo contínuo de ajuste do *setpoint*. Esse ajuste depende de vários fatores, além da qualidade dos produtos, como vazão, viscosidade e temperatura dos fluidos. A estabilidade desse controlador é primordial para garantir um processo de separação adequado por todo o processo e evitar o acúmulo de bolsões de água na câmara de óleo ou o envio de produto oleoso para as saídas de água.
- Controlador de nível da câmara de óleo: responsável por manter uma coluna de líquido entre o sistema de separação trifásico e o restante do processo de tratamento de óleo, evitando a passagem de gás para os tratadores a jusante. Esse controlador também afeta diretamente a oscilação de fluxo para os equipamentos a jusante do sistema, por ser responsável pelo controle de carga do tratador eletrostático. O nível de líquido na câmara de água também influencia o controle de carga desse equipamento.

Um esquema típico de controle está descrito na Figura 13.

Figura 13 – Esquema de instrumentação e controle típicos de um separador trifásico horizontal



Fonte: Nunes; Medeiros; Araújo (2010)

2.2 MODELAGEM DINÂMICA DE PROCESSOS

A modelagem dinâmica de processos tem se mostrado fundamental para o aprimoramento da eficiência de processos e a mitigação de desperdícios na indústria, assim como para a diminuição dos impactos ambientais. Para atingir esses objetivos, é necessário validar a integridade do projeto e sua operabilidade prática por meio da simulação de toda a planta, utilizando modelos rigorosos.

De acordo com Aragão (2011), a modelagem matemática é uma ferramenta importante para a análise de processos e sistemas, permitindo a simulação de diferentes cenários e a identificação de pontos críticos. Além disso, ela pode ser utilizada para otimizar o uso de recursos e aprimorar a eficiência do processo produtivo. Diversas técnicas e pacotes computacionais para simulação e modelagem de processos têm sido desenvolvidos e comercializados no mercado. Cada uma dessas soluções apresenta técnicas diferentes para a modelagem, levando em consideração as características físico-químicas dos produtos, características construtivas dos equipamentos, modelos matemáticos e equações que regem as relações do processo, dentre outras.

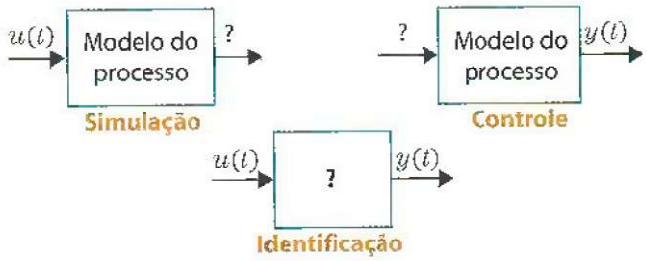
Conforme o objetivo da modelagem, capacidade computacional e rigor da simulação necessário para a finalidade do projeto, podem ser implementados modelos de maior ou menor complexidade, simplificando algumas relações ou desprezando certas especificidades enquanto ainda

mantendo uma boa generalização (BELAUD et al., 2019). Dessa forma, a modelagem dinâmica de processos é uma ferramenta essencial para aprimorar a eficiência da produção industrial e minimizar impactos ambientais, permitindo uma análise mais precisa e aprimoramento de projetos, além de permitir a identificação de pontos críticos e o uso eficiente de recursos.

Em resumo, a modelagem matemática é uma ferramenta valiosa para a otimização de processos industriais, contribuindo para a redução de custos, aumento da eficiência, segurança e minimização de impactos ambientais (NUNES; MEDEIROS; ARAÚJO, 2010). A modelagem pode ser realizada por meio de diferentes abordagens, que se adequam às características e objetivos específicos de cada processo. A modelagem fenomenológica, por exemplo, é uma abordagem que se baseia nos princípios da conservação de massa, energia e momento, e descreve o processo por meio das leis que regem os fenômenos físicos e químicos. Esse tipo de modelagem é capaz de fornecer informações detalhadas sobre o processo e pode ser utilizada para otimizar o desempenho de sistemas complexos (CHEN, 2017). Já a identificação de processo utiliza equações empírico-diferenciais que não possuem relação direta com as equações de balanço, mas que são geradas a partir de dados experimentais e ajuste de parâmetros. Essa abordagem é bastante útil quando não se dispõe de informações suficientes sobre o processo ou quando as informações disponíveis são limitadas (PINTO; SOUZA; ARAÚJO, 2020). Por fim, a modelagem híbrida utiliza uma combinação das abordagens empírica e fenomenológica, buscando unir as vantagens de cada uma. Essa abordagem tem sido utilizada com sucesso em diversos processos industriais, permitindo uma descrição mais precisa do comportamento do sistema e uma melhor compreensão das interações entre as variáveis (PÉREZ; MIRANDA; LÓPEZ, 2017).

Em suma, a escolha da abordagem de modelagem mais adequada para cada processo depende das informações disponíveis, do nível de detalhamento necessário e dos objetivos específicos de cada aplicação. Independentemente da abordagem utilizada, a modelagem matemática é uma ferramenta valiosa para a otimização de processos industriais e tem sido amplamente utilizada na indústria. Na Figura 14 pode ser observada uma representação destas 3 abordagens de modelagem.

Figura 14 – Abordagens de modelagem e simulação de processo



Fonte: Nunes; Medeiros; Araújo (2010)

2.3 IDENTIFICAÇÃO DE PROCESSOS

A identificação de processos é uma etapa fundamental para compreender e modelar o comportamento de sistemas complexos em diversas áreas da engenharia e ciências aplicadas. Com o avanço das tecnologias computacionais, têm surgido métodos e algoritmos que facilitam e aprimoram esse processo, proporcionando benefícios como maior precisão, eficiência e flexibilidade (HUANG et al., 2017). Existem diversas abordagens para a identificação de processos. Os modelos paramétricos são amplamente utilizados e baseiam-se na suposição de uma estrutura funcional conhecida. Esses modelos buscam minimizar o erro entre os dados medidos e os valores simulados, ajustando parâmetros específicos. Entre os métodos paramétricos, destacam-se a Regressão Linear e a Estimação de Máxima Verossimilhança (KRISHNA; SEKAR, 2020; LIU et al., 2018). Por outro lado, os modelos não paramétricos não pressupõem uma estrutura funcional específica, o que permite uma maior flexibilidade na modelagem de sistemas complexos. Técnicas como Redes Neurais Artificiais, Máquinas de Vetores de Suporte e Processos Gaussianos são utilizadas para mapear as relações entre as variáveis de entrada e saída (RAO et al., 2019; SAUL; JORDAN, 2016).

A utilização de meios computacionais na identificação de processos traz diversas vantagens. Esses métodos são capazes de lidar com sistemas complexos e não lineares, os quais são desafiadores de serem modelados por abordagens analíticas tradicionais. Além disso, a capacidade de processamento dos computadores permite trabalhar com grandes volumes de dados e realizar análises mais robustas (HUANG et al., 2017). A automação do processo de identificação por meio dos métodos computacionais torna-o mais rápido e eficiente. A flexibilidade desses métodos também permite ajustar os modelos de acordo com as necessidades específicas de cada aplicação, adaptando-se a diferentes contextos e problemas (KRISHNA; SEKAR, 2020). Atualmente, as tendências na área de identificação de processos utilizando meios computacionais

incluem o uso crescente de algoritmos de Aprendizado de Máquina, como Redes Neurais Profundas, que se destacam pela capacidade aprimorada de modelagem e predição (DELLI et al., 2021). Além disso, técnicas de Processamento de Sinais e Análise de Dados em tempo real têm sido exploradas para o monitoramento e controle de processos industriais (LIU et al., 2019). Diversos trabalhos têm sido desenvolvidos nessa área, abordando diferentes aplicações e aprimoramentos nos métodos de identificação de processos utilizando meios computacionais. Por exemplo, WANG et al. (2020) propuseram um método baseado em algoritmos genéticos para otimização de modelos de identificação de processos, enquanto DELLI et al. (2021) aplicaram Redes Neurais Profundas na identificação de sistemas dinâmicos complexos.

Em resumo, a utilização de métodos não paramétricos computacionais, como redes neurais, na identificação de processos oferece uma abordagem promissora e eficaz para lidar com sistemas complexos e não lineares. Essa abordagem permite uma modelagem mais precisa e acurada, impulsionada pela capacidade de processamento dos computadores modernos. A automação do processo e a flexibilidade desses métodos também são vantagens relevantes, proporcionando eficiência, velocidade e adaptabilidade às necessidades específicas de cada aplicação.(HUANG et al., 2017)

2.4 DIGITAL TWIN

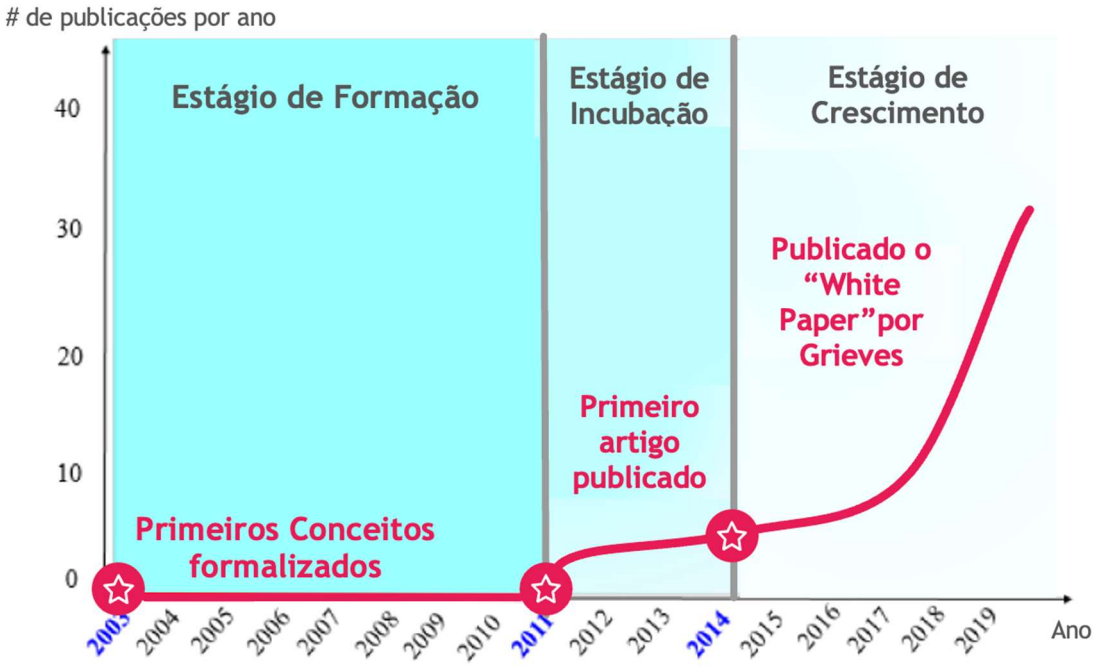
A evolução da ciência da informação e sua aplicação na indústria têm transformado a forma como as organizações abordam seus processos produtivos. O *Digital Twin* (Gêmeo Digital) é uma abordagem que tem sido amplamente utilizada na indústria para integrar sistemas virtuais e físicos, visando aprimorar a performance do sistema real por meio de informações geradas e testadas no ambiente virtual projetado para replicar o estado e funcionamento do equipamento físico (OCHÔA; PEIXOTO; SOARES, 2021). A aplicação de *Digital Twins* é vasta e sua implementação varia de acordo com as necessidades específicas do problema a ser solucionado, seja em logística, física, arquitetônica ou de processo.

Os *Digital Twins* têm se destacado como ferramentas importantes na transformação da indústria para os novos paradigmas da indústria 4.0. Sua popularidade pode ser observada pelo grande número de publicações e pesquisas desenvolvidas nos últimos anos (TAO et al., 2019). De fato, a implementação dos *Digital Twins* tem sido considerada como o principal gatilho para uma nova onda no ramo das simulações. O termo "*Digital Twin*" foi proposto pela primeira vez em

2003 por Michael Grieves, mas durante alguns anos não houve publicações sobre o assunto. Em 2011, foi publicado um artigo que especificava preliminarmente os conceitos de um *Digital Twin*, o que foi seguido por um aumento significativo na quantidade de publicações sobre o tema, incluindo sua adoção pela National Aeronautics and Space Administration (NASA).
A NASA definiu *Digital Twin* como uma "simulação multi-escala, multifísica, probabilística e de alta fidelidade que reflete, num intervalo de tempo observado, o estado do gêmeo físico correspondente baseado em dados históricos, dados de sensores em tempo real e modelo físico" (GLAESSGEN; STARGEL, 2012, p.23). A partir de 2017, após a publicação de Grieves e Vickers (2016), houve uma verdadeira explosão de interesse no tema, com um aumento significativo no número de citações e aplicações de *Digital Twins* na indústria. De fato, a aplicação de *Digital Twins* tem se mostrado uma estratégia valiosa para melhorar o desempenho de equipamentos e processos industriais, além de permitir a identificação precoce de possíveis falhas e permitir uma manutenção preditiva mais eficiente (ZHANG et al., 2019). Outros benefícios incluem a redução de custos de produção e manutenção, bem como a redução do tempo de inatividade do equipamento (TAN et al., 2020). Além disso, o uso de *Digital Twins* também tem sido estudado em áreas como saúde e agricultura, o que destaca o potencial dessa tecnologia em diferentes setores (TARATOUKHINE et al., 2020; WANG et al., 2021).

A Figura 15 demonstra os estágios de crescimento na quantidade de citações do termo ‘*Digital Twin’* em artigos acadêmicos, denotando a relevância do tema atualmente.

Figura 15 – Evolução das citações ao termo '*Digital Twin*' ao longo do tempo



Fonte: Tao et al. (2019)

O *Digital Twin*, abordagem que integra sistemas virtuais e físicos, tem transformado a indústria e redefinido os paradigmas de desenvolvimento organizacional. Dentre as categorias de *Digital Twin*, a Modelagem, Simulação, Verificação, Validação e Acreditação (MSVV) é uma das mais relevantes, pois engloba todas as etapas necessárias para a criação de um gêmeo digital. A modelagem física é uma etapa crucial na MSVV, pois é através dela que se extrai, define e descreve as características-chave de uma entidade física, permitindo a criação de uma representação virtual fiel da mesma (TAO et al., 2019). A modelagem virtual, por sua vez, é uma teoria útil para construir essa representação virtual, que deve ser capaz de reproduzir as mesmas características e comportamentos no espaço virtual. No entanto, é importante que a modelagem realizada seja avaliada em relação ao gêmeo físico quanto à sua fidelidade e representatividade, e validada para atestar sua adequação como solução ao problema em questão. Smarslok, Culler e Mahadevan (2012) propõem um método de quantificação de erros e avaliação de confiança de modelos baseado na teoria Bayesiana.

Apesar da grande quantidade de estudos e publicações sobre *Digital Twins*, ainda não há uma padronização estabelecida sobre a metodologia de modelagem e validação de um gêmeo digital. É importante ressaltar que a falta de uma padronização pode levar a diferentes abordagens e

resultados, tornando difícil a comparação e avaliação dos resultados obtidos por diferentes estudos. Portanto, é necessário que haja uma padronização para a metodologia de modelagem e validação de *Digital Twins*, para que os resultados possam ser comparados e avaliados de forma mais precisa (SUN et al., 2021). Outra questão importante é a segurança dos *Digital Twins*, que podem ser vulneráveis a ataques cibernéticos. Por isso, é fundamental que sejam adotadas medidas de segurança eficazes para proteger tanto o gêmeo digital quanto o sistema físico correspondente. Um estudo recente de Zhou et al. (2021) propõe um método para avaliar e mitigar vulnerabilidades de segurança em *Digital Twins*.

Dessa forma, fica claro que o *Digital Twin* é uma abordagem com grande potencial para aprimorar a eficiência e eficácia dos processos produtivos, mas é preciso que haja uma padronização das metodologias de modelagem e validação, bem como medidas de segurança eficazes, para que sua aplicação possa ser realizada de forma segura e confiável.

2.5 REDES NARX

As redes neurais NARX (*Nonlinear Autoregressive with eXogenous inputs*) têm sido amplamente utilizadas na modelagem e previsão de séries temporais em diversas áreas, como engenharia, finanças, medicina e meteorologia. Esses modelos se destacam por sua capacidade de capturar a dinâmica complexa e não linear das séries temporais, levando em consideração entradas exógenas, ou seja, variáveis externas que podem influenciar o comportamento da série temporal de interesse (LIU et al., 2020; TANG et al., 2021). Essas redes são compostas por uma camada de entrada, camadas ocultas e uma camada de saída. A camada de entrada recebe os dados da série temporal e os dados exógenos, que são processados pelas camadas ocultas para, então, produzir a saída da rede. O processo de treinamento dessas redes é geralmente realizado por meio de algoritmos de otimização, como o gradiente descendente estocástico, que ajustam os pesos da rede para minimizar a diferença entre a saída da rede e os valores reais da série temporal (LIU et al., 2020; TANG et al., 2021). Alguns estudos têm mostrado que as redes NARX podem superar outras técnicas de previsão de séries temporais, como modelos autorregressivos lineares, em termos de acurácia e robustez (TANG et al., 2021; LI et al., 2018). Além disso, essas redes têm sido utilizadas em aplicações práticas, como na previsão de demanda de energia elétrica, previsão de preços de ações e previsão de chuvas (ZHANG et al., 2020; YANG et al., 2020). No entanto, é importante destacar que o desempenho dessas redes pode ser afetado por diversos fatores, como o tamanho da janela de dados utilizada para a previsão, o número de neurônios nas camadas ocultas e a seleção das variáveis exógenas relevantes (LIU et al., 2020;

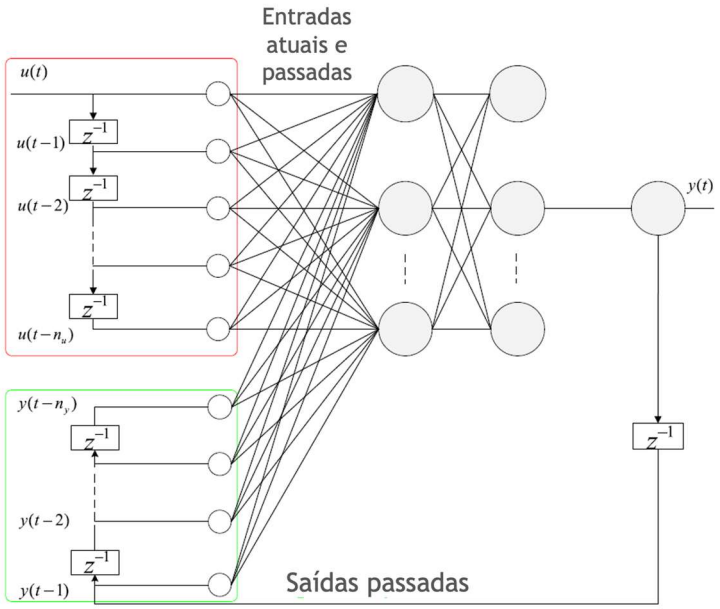
TANG et al., 2021). Portanto, é fundamental realizar uma análise cuidadosa dos dados e da estrutura da rede para obter resultados confiáveis e precisos.

Em resumo, as redes NARX apresentam uma abordagem diferenciada em relação às RNNs convencionais e são capazes de lidar com problemas de retenção de memória de longo prazo. Além disso, possuem aplicações em previsão de séries temporais e modelagem de sistemas dinâmicos complexos, sendo uma ferramenta importante para a análise e previsão de dados sequenciais. Dentre as aplicações das redes NARX, destaca-se o uso em previsão de séries temporais, onde os dados exógenos são incorporados na rede para melhorar a precisão das previsões. Por exemplo, Liu et al. (2020) utilizaram uma rede NARX para prever a produção de petróleo em um campo de petróleo, incorporando dados exógenos, como temperatura, pressão e umidade.

As redes NARX têm a capacidade de realizar uma modelagem do tipo caixa-preta, o que significa que elas podem inferir relações não-lineares e desconhecidas inicialmente, simplificando a modelagem de sistemas complexos sem necessariamente simplificar o próprio modelo (LIU et al., 2020). Essa abordagem é particularmente útil em problemas complexos, onde os modelos analíticos podem ser difíceis ou impossíveis de formular. Além disso, a implementação de modelos NARX é relativamente simples, e não requer um conhecimento especializado aprofundado. Em vez disso, é necessário apenas a capacidade de avaliar criticamente os resultados e determinar se eles são adequados para resolver o problema em questão (HENAFF; SZLAM; LECUN, 2017). O uso de redes NARX tem sido aplicado em diversas áreas, como em previsão de demanda de energia (JAMIL et al., 2018), previsão de preços de ações (HUANG et al., 2019) e previsão de consumo de água (WANG et al., 2021). Além disso, as redes NARX também são utilizadas em sistemas de controle de processos industriais, como em um estudo que propôs o uso de uma rede NARX para controlar a temperatura de um reator químico (MEDEIROS et al., 2017).

Uma configuração típica de rede neural NARX está representada na Figura 16.

Figura 16 – Arquitetura de uma Rede Neural NARX típica



Fonte: Liang et al. (2021)

2.6 REDES LSTM

As Redes Neurais Recorrentes (*Recurrent Neural Networks* ou RNN) são amplamente utilizadas na modelagem de dados sequenciais, pois possuem a capacidade de processar entradas de comprimentos variáveis e de lidar com dependências temporais em dados. No entanto, as RNNs têm uma desvantagem significativa em relação a outros tipos de redes neurais, que é a dificuldade em lidar com a perda de informação de longo prazo, conhecida como problema *vanishing gradient*, que se manifesta à medida que a rede fica mais profunda (BENGIO et al., 1994). Para solucionar esse problema, a LSTM (*Long Short-Term Memory*) foi proposta por Hochreiter e Schmidhuber (1997), sendo uma variante da RNN que utiliza portões de by-pass de camadas para permitir que os sinais passem por mais tempo, conferindo maior capacidade de retenção de memória. A LSTM é composta por células de memória, portões de entrada, portões de esquecimento e portões de saída. Essa arquitetura permite que a LSTM capture padrões de longo prazo em sequências, mantendo a relevância de informações anteriores (GREFF et al., 2016).

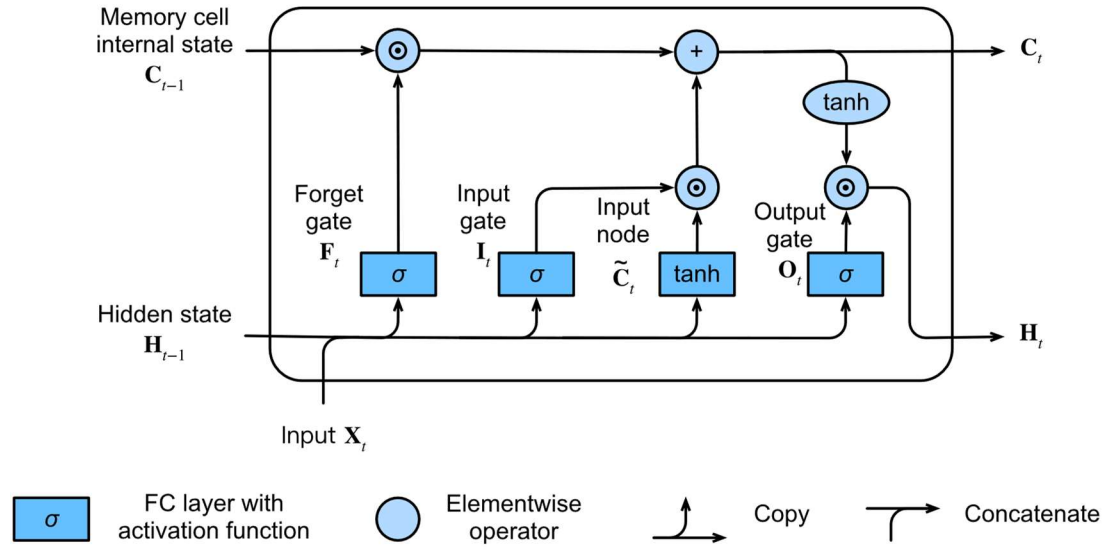
Uma das aplicações mais comuns das LSTMs é na modelagem e previsão de séries temporais, em que a rede é treinada com séries históricas de dados para prever os valores futuros. Além

disso, as LSTMs também têm sido amplamente utilizadas em diversas áreas, como processamento de linguagem natural, reconhecimento de fala e visão computacional (HAN et al., 2021). Portanto, a utilização da LSTM tem sido considerada uma solução eficaz para o problema de retenção de informações de longo prazo em dados sequenciais, permitindo o desenvolvimento de modelos robustos para uma variedade de aplicações em diferentes áreas.

2.7 COMPARANDO REDES NARX E REDES LSTM

As redes NARX e LSTM são duas abordagens distintas de Redes Neurais Recorrentes (RNNs) que possuem como objetivo a modelagem de séries temporais. Enquanto as redes NARX adotam uma abordagem paralela, alimentando a rede com sinais antigos e atuais simultaneamente, as redes LSTM adotam uma abordagem sequencial, alimentando a rede com um sinal por vez. Ambas possuem vantagens e desvantagens em relação ao desempenho e à capacidade de modelagem. Uma das principais vantagens das redes NARX é a sua capacidade de modelar dados não-lineares de forma mais eficiente do que outras RNNs, como as redes LSTM (LAMBERTS; SONTAG, 2013). Além disso, as redes NARX apresentam um bom desempenho na predição de séries temporais de longo prazo (HE; WANG, 2019). No entanto, um dos principais desafios das redes NARX é a sua complexidade computacional, o que pode tornar a sua implementação mais difícil em alguns casos. Por outro lado, as redes LSTM têm sido amplamente utilizadas em problemas de séries temporais devido à sua capacidade de reter informações de longo prazo, o que é essencial para a modelagem de séries temporais complexas (HOCHREITER; SCHMIDHUBER, 1997). Além disso, as redes LSTM são menos propensas ao problema do *vanising gradient* do que outras RNNs, o que permite que elas possam ser treinadas com sucesso em redes profundas (BENGIO et al., 1994). No entanto, as redes LSTM podem apresentar dificuldades na modelagem de séries temporais que possuem múltiplas escalas de tempo (GERS et al., 2002). Portanto, a escolha entre as redes NARX e LSTM para a modelagem de séries temporais depende das características do problema em questão, incluindo a natureza dos dados, a escala temporal, o desempenho desejado e a complexidade computacional permitida. A Figura 17 contém a representação gráfica da arquitetura de uma rede LSTM típica.

Figura 17 – Arquitetura de uma rede LSTM



Fonte: Zhang, Lipton, Smola (2021)

2.8 MÉTRICAS DE ERRO

A previsão de séries temporais é uma tarefa desafiadora em diversas áreas, como finanças, economia, energia e clima. Com o avanço das redes neurais artificiais, especialmente as redes NARX e redes LSTM, essas técnicas têm se mostrado promissoras para modelar e prever séries temporais complexas. No entanto, a avaliação adequada do desempenho desses modelos é fundamental para garantir a precisão e a confiabilidade das previsões. Ao avaliar o desempenho dos modelos de previsão de séries temporais, é comum utilizar métricas de erro para medir a discrepância entre as previsões e os valores reais. Duas métricas amplamente utilizadas são o Erro Médio Absoluto (MAE) e o Erro Médio Quadrático (MSE). O MAE, definido como a média das diferenças absolutas entre as previsões e os valores observados, fornece uma medida direta do erro médio sem levar em consideração a direção da discrepância. Essa métrica tem sido amplamente aplicada na avaliação de modelos de previsão de séries temporais. Segundo Santos et al. (2018), "o MAE é uma métrica robusta e de interpretação intuitiva, sendo especialmente adequada para séries com valores discrepantes" (p. 403). Além disso, Sohail et al. (2020) destacam que o MAE é uma métrica amplamente utilizada em redes LSTM para avaliar o desempenho na previsão de séries temporais financeiras. O MSE, por sua vez, calcula a média dos quadrados das diferenças entre as previsões e os valores reais. Ele penaliza erros maiores

de forma mais significativa devido ao uso dos quadrados. Essa métrica é frequentemente utilizada em conjunto com o MAE para obter uma visão mais abrangente do desempenho do modelo. De acordo com Gharibi et al. (2020), "o MSE é uma métrica útil para avaliar o desempenho geral de modelos de previsão de séries temporais, pois destaca a importância de prever valores extremos" (p. 303). Além disso, Hajizadeh et al. (2021) ressaltam que o MSE é amplamente utilizado em modelos NARX para avaliar a precisão da previsão de séries temporais climáticas.

A escolha entre o MAE e o MSE como métricas de erro depende do contexto e dos objetivos da previsão. Enquanto o MAE é mais adequado para avaliar o erro médio absoluto, o MSE é sensível a valores discrepantes e permite uma avaliação mais detalhada da variação dos erros ao longo do tempo. É importante ressaltar que a escolha da métrica de erro apropriada depende do contexto específico da previsão e dos objetivos do estudo. Por exemplo, em algumas aplicações financeiras, a ênfase pode ser dada à minimização do erro absoluto, enquanto em outras, como previsão de demanda de energia, a precisão percentual pode ser mais relevante.

2.9 OTIMIZAÇÃO BAYESIANA

A otimização bayesiana é uma abordagem poderosa para resolver problemas de otimização em diversas áreas de pesquisa. Ela combina técnicas estatísticas e de aprendizado de máquina para encontrar o conjunto ideal de parâmetros de um modelo, levando em consideração a função objetivo e as restrições definidas. Uma das vantagens da otimização bayesiana é sua capacidade de lidar com espaços de busca complexos e de alta dimensionalidade. Em um estudo realizado por Snoek, Larochelle e Adams (2012), a otimização bayesiana foi aplicada para ajustar hiperparâmetros em redes neurais convolucionais, resultando em melhorias significativas no desempenho de classificação de imagens. Outra aplicação relevante da otimização bayesiana é na otimização de algoritmos. Em um estudo conduzido por Swersky, Snoek e Adams (2014), a otimização bayesiana foi utilizada para ajustar os parâmetros de um algoritmo de busca em larga escala. Os resultados mostraram que a otimização bayesiana superou os métodos tradicionais de otimização, melhorando significativamente o desempenho do algoritmo. Comparado a outros métodos de otimização, a otimização bayesiana oferece vantagens distintas. Por exemplo, em um estudo de Shahriari et al. (2016), foi observado que a otimização bayesiana superou os métodos clássicos em problemas de otimização com orçamento limitado de avaliações do modelo. A capacidade da otimização bayesiana de fazer escolhas inteligentes sobre quais pontos explorar no espaço de busca levou a uma melhor eficiência computacional.

A biblioteca Optuna, desenvolvida em Python, é uma ferramenta popular para a implementação da otimização bayesiana. Ela oferece uma interface simples para definir espaços de busca, funções objetivo e restrições, além de utilizar algoritmos eficientes de amostragem, como o algoritmo de árvore de decisão Parzen Estimator (TPE). Em um estudo recente de Kandasamy et al. (2018), o Optuna foi utilizado para otimizar hiperparâmetros em modelos de aprendizado de máquina aplicados em detecção de anomalias, resultando em melhorias significativas no desempenho dos modelos.

3 DESENVOLVIMENTO

3.1 DETALHAMENTO DAS VARIÁVEIS DE ENTRADA E SAÍDA

A coleta de dados desempenha um papel crucial na condução de pesquisas científicas, especialmente no desenvolvimento de modelos preditivos. Nesse contexto, a frequência de amostragem desempenha um papel importante na obtenção de informações detalhadas sobre o comportamento dinâmico do sistema em estudo. No presente trabalho, a frequência de amostragem de 1 Hz foi escolhida com base em sua capacidade de capturar variações rápidas nas variáveis do processo, considerando que o período de amostragem é significativamente menor em relação aos tempos característicos das variáveis em questão. Essa escolha visa assegurar que as amostras coletadas forneçam informações precisas e atualizadas sobre o estado do sistema em tempo real.

Os dados coletados consistem em leituras obtidas de diferentes instrumentos de medição ao longo do tempo. É importante destacar que alguns desses instrumentos são considerados redundantes, sendo utilizado um conceito conhecido como "reservas de outros". Esse conceito é comumente aplicado em instrumentação que requer calibração periódica. A utilização de múltiplos instrumentos para medir a mesma variável permite a detecção de possíveis desvios e a calibração cruzada, garantindo maior confiabilidade nos dados coletados.

Além disso, é importante considerar a configuração dos dispositivos de controle durante o processo de coleta de dados. No caso específico deste estudo, os dispositivos operam em uma configuração *split-range*, o que significa que em determinadas situações um dos dispositivos pode estar bloqueado para manutenção. Nesse sentido, é necessário selecionar qual dos dispositivos será considerado na análise, levando em conta critérios específicos relacionados às condições operacionais e aos objetivos do estudo.

No processo de desenvolvimento do modelo preditivo, foi realizada uma cuidadosa seleção das variáveis a serem utilizadas. Essa seleção baseou-se tanto na experiência do pesquisador no campo em estudo quanto na observação da interação dessas variáveis com os resultados desejados. Essa abordagem visa identificar as variáveis mais relevantes que possuem maior influência no comportamento do sistema, a fim de garantir a eficácia e a precisão do modelo final. (LEE et al., 2019).

Foram elas:

- Pressão de entrada do vaso separador (*Inlet Pressure*): desempenha um papel crucial na inferência do fluxo de líquido e gás dentro do sistema. Essa variável está diretamente relacionada às variações na vazão de saída do vaso separador. Ao monitorar a pressão de entrada, é possível estimar se há um aumento no fluxo de líquido ou gás entrando no vaso.

- Vazão de saída de gás (*Gas Flow*): O monitoramento da vazão de saída de gás é de extrema importância para a detecção de golfadas, que são caracterizadas pela passagem de bolsões de gás de maneira não homogênea. Essas golfadas podem impactar negativamente a eficiência do processo de separação e a qualidade dos produtos finais. Portanto, é essencial identificar esses eventos e tomar medidas corretivas para garantir uma operação estável e eficiente.

- Percentual de abertura da válvula de saída de água (*Water Valve Opening*): A válvula de saída de água é responsável por ajustar o fluxo de água que deixa o separador, influenciando diretamente o nível de líquido na câmara principal. Através da variação do percentual de abertura da válvula, é possível controlar a taxa de escoamento da água e manter o nível desejado.

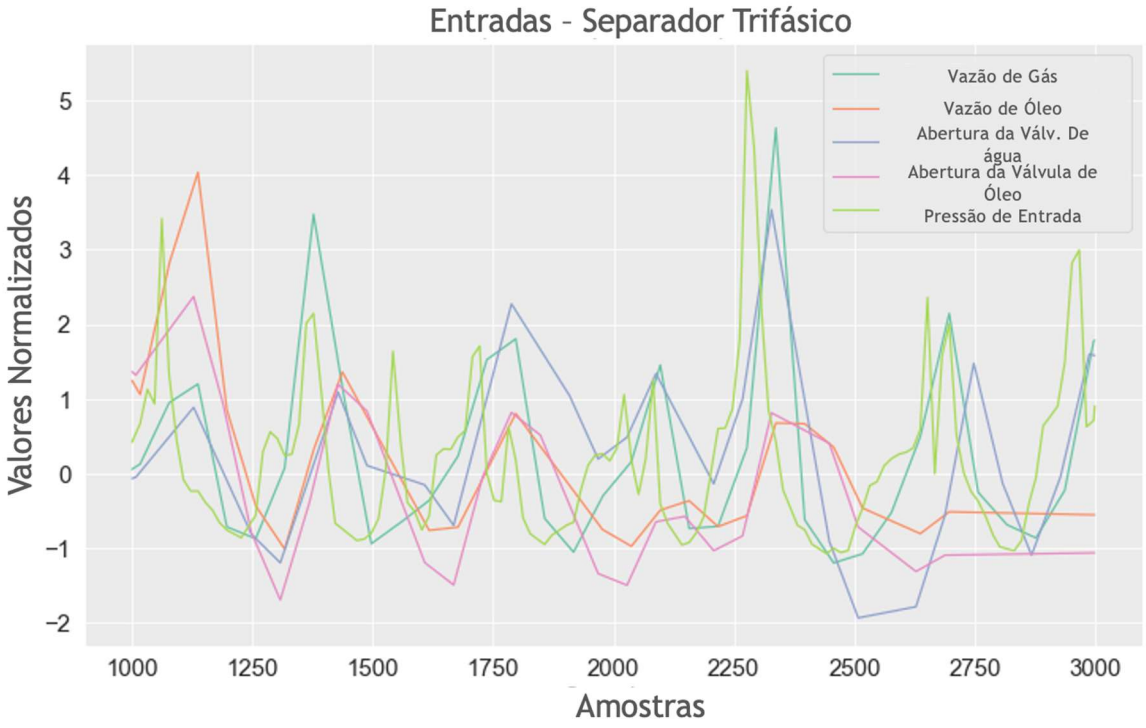
- Percentual de abertura da válvula de saída de óleo (*Oil Valve Opening*): papel crucial como um elemento final de controle no controlador de nível da câmara de óleo de um separador trifásico. Esse parâmetro é responsável por regular a vazão de saída do óleo e influenciar diretamente o nível de óleo dentro do separador. A abertura da válvula afeta a taxa de escoamento do óleo, o que por sua vez afeta a separação eficiente dos componentes do fluido no processo de separação trifásica.

- Vazão de saída de óleo (*Oil Flow*): a medição da vazão de saída de óleo permite uma melhor compreensão do comportamento do sistema, pois fornece informações sobre a taxa de escoamento do óleo, mesmo sob condições variáveis de pressão. Essa informação é essencial para

avaliar o desempenho do sistema, identificar possíveis variações indesejadas e tomar medidas corretivas adequadas.

Uma visualização de exemplo do comportamento das variáveis de entrada já normalizadas consta na Figura 18.

Figura 18 –Visualização do comportamento das variáveis de entrada do separador



Fonte: Autor (2022)

As variáveis de saída serão:

- Nível da Câmara principal do Separador trifásico: Essa variável representa a altura do líquido dentro da câmara e é de extrema importância no controle e operação eficiente do separador. O nível da câmara principal é uma medida direta da quantidade de líquido presente no separador e está intimamente relacionado à eficiência da separação dos componentes (gás, óleo e água). O controle adequado do nível é essencial para garantir que cada fase seja direcionada corretamente para sua saída correspondente, evitando misturas indesejadas. Diversos estudos têm abordado a importância do controle preciso do nível da câmara principal em separadores trifásicos. Por exemplo, em um estudo realizado por Li et al. (2020), foi evidenciado que um controle inadequado do nível pode levar a uma má separação das fases e, consequentemente, com-

prometer a qualidade dos produtos finais. Além disso, o nível da câmara principal também desempenha um papel crucial na segurança operacional. Um nível muito baixo pode resultar em entradas de ar indesejadas na câmara, causando instabilidades no processo. Por outro lado, um nível muito alto pode levar a transbordamentos e perda de líquido valioso. Portanto, o monitoramento e controle preciso do nível da câmara principal do separador trifásico são fundamentais para garantir a eficiência do processo de separação, a qualidade dos produtos finais e a segurança operacional. Ao considerar essa variável de saída em nossa análise, buscamos entender como as redes neurais NARX e LSTM podem contribuir para um controle mais efetivo e confiável do nível, proporcionando benefícios significativos na operação de separadores trifásicos.

- Nível da Câmara de Óleo do Separador trifásico: O nível da câmara de óleo em um separador trifásico desempenha um papel fundamental na ocorrência do fenômeno conhecido como "efeito de selagem líquida" ou "selagem líquida". Esse efeito refere-se à formação de uma camada de líquido (geralmente óleo) que atua como uma barreira para impedir a passagem de gás através do equipamento.

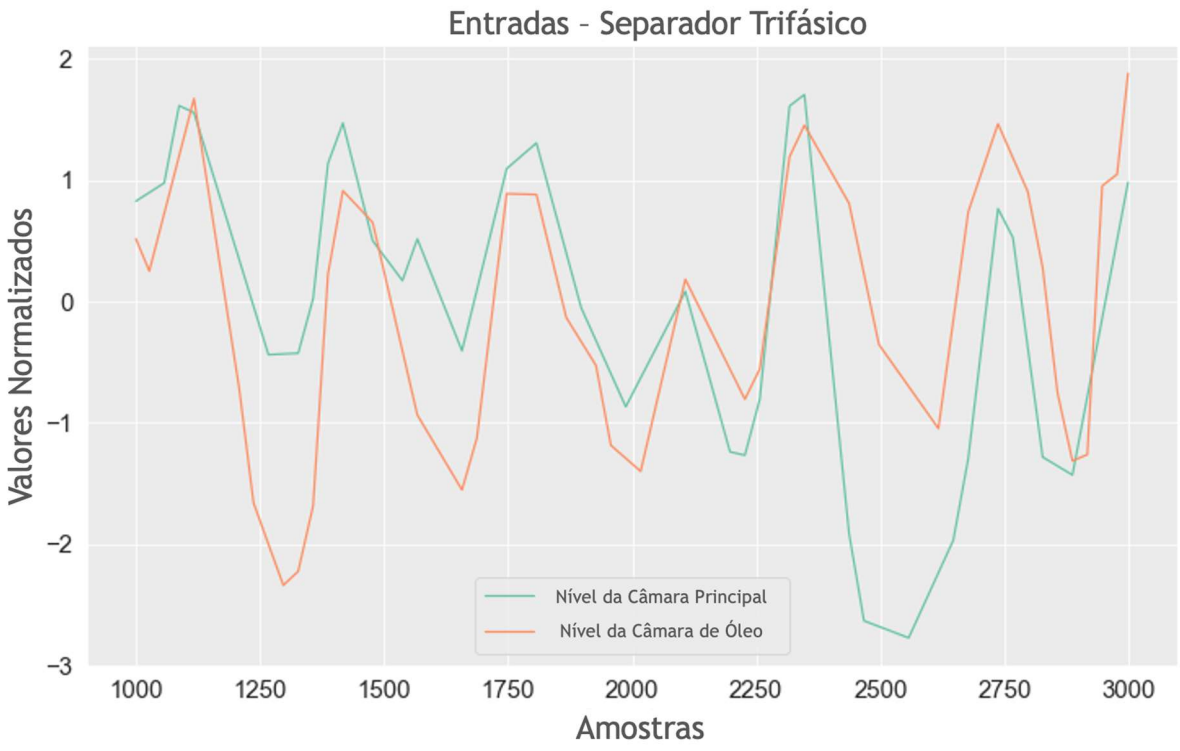
Quando o nível da câmara de óleo está adequadamente controlado e mantido em uma posição ótima, ocorre a formação de uma camada líquida que cobre a saída de gás do separador. Essa camada de óleo cria uma barreira física que evita a passagem do gás para a linha de saída, garantindo que apenas o líquido seja escoado.

Esse efeito de selagem líquida é essencial para a operação eficiente do separador trifásico. Ele garante que o gás seja retido dentro do equipamento, permitindo que as fases líquidas (óleo e água) sejam separadas adequadamente. Caso contrário, sem a presença da camada líquida, o gás poderia ser arrastado juntamente com as fases líquidas, comprometendo a eficiência do processo de separação.

Diversos estudos têm destacado a importância do controle preciso do nível da câmara de óleo para o efeito de selagem líquida. Por exemplo, em um trabalho realizado por Bezerra et al. (2019), foi investigado experimentalmente o efeito do nível da câmara de óleo na eficiência da separação e no arraste de gás. Os resultados mostraram que um nível inadequado da câmara de óleo pode levar a uma diminuição na eficiência da separação e ao aumento do arraste de gás para a linha de saída.

Uma demonstração do comportamento das saídas pode ser vista na Figura 19, após normalização.

Figura 19 – Visualização do comportamento das variáveis de saída do separador



Fonte: Autor (2023)

Os dados utilizados neste trabalho foram selecionados com o objetivo de representar situações de grande variação de carga no separador, ou seja, períodos em que o poço alinhado estava produzindo em regime instável. Essa escolha se baseia na premissa de que esses períodos representam os piores casos para o processo, sendo, portanto, mais desafiadores para a modelagem. Além disso, a escolha desses dados permite uma melhor visualização das curvas das variáveis de processo, que se mostram bastante pronunciadas nessa situação, possibilitando uma melhor comparação e análise. Os dados foram coletados no período entre 12:00 do dia 23 de maio de 2020 e 06:00 do dia 24 de maio de 2020.

3.2 PRÉ-PROCESSAMENTO DOS DADOS

Após a preparação dos dados, uma etapa crucial no processo de modelagem é o pré-processamento dos dados. O pré-processamento envolve várias etapas, incluindo a limpeza, transformação e normalização dos dados. A normalização dos dados é particularmente importante, pois garante que as variáveis tenham a mesma escala e distribuição, evitando assim que algumas variáveis dominem outras durante a modelagem (ROUSSEEUW; CROUX, 1993).

Nesse contexto, a normalização dos dados desempenha um papel fundamental na obtenção de resultados precisos e confiáveis. Uma abordagem comumente utilizada é a normalização robusta, que lida efetivamente com a presença de valores discrepantes ou *outliers* nos dados (PEDREGOSA et al., 2011). Nesse sentido, o uso da biblioteca Scikit-learn é uma escolha adequada, pois essa esta biblioteca possui funções capazes de aplicar tais transformações aos dados, reduzindo a influência dos *outliers* na normalização. É importante destacar que, ao realizar a normalização dos dados, é fundamental evitar o problema de *data leakage*, ou seja, vazamento de informações dos dados de treinamento para os dados de teste e validação. (KAUFMAN, ROSSET e PERLICH, 2011)

3.3 TREINAMENTO DAS REDES

Para cada um dos horizontes de previsão estabelecidos neste estudo, foram desenvolvidas e otimizadas redes neurais autoregressivas com entradas exógenas (NARX) e redes de memória de longo prazo (LSTM). A sintonização dos hiperparâmetros foi realizada por meio de uma abordagem de otimização Bayesiana utilizando a biblioteca Optuna. No total, foram construídos e otimizados oito modelos, correspondendo aos horizontes de previsão de 1, 3, 5 e 10 passos à frente.

Cada tipo de rede neural foi submetido a um algoritmo de otimização específico, com a finalidade de ajustar e otimizar os hiperparâmetros relevantes para cada modelo. Essa abordagem permitiu explorar diferentes configurações e buscar os conjuntos ideais de hiperparâmetros para cada horizonte de previsão em cada arquitetura de rede. Neste trabalho já foram citados os parâmetros para cada arquitetura no capítulo introdutório, aqui estes parâmetros serão explicados com maior nível de detalhe.

Para as redes NARX, os parâmetros otimizados por meio do algoritmo foram os seguintes:

- Fator de Autocorrelação: O fator de autocorrelação desempenha um papel significativo na estruturação de uma rede neural autoregressiva com entradas exógenas (NARX), pois influencia diretamente a capacidade da rede em analisar informações passadas representadas por variáveis com atraso. A definição do fator de autocorrelação estabelece a extensão pela qual a rede considera observações passadas ao avaliar as relações entre as variáveis. Em outras palavras, o fator de autocorrelação determina a importância atribuída às variáveis com atraso, onde um valor mais elevado indica que a rede prioriza informações mais recentes, enquanto um valor mais baixo indica uma maior consideração às observações com atraso mais longo. A determinação cuidadosa do fator de autocorrelação é essencial para a modelagem adequada e precisa da dinâmica temporal dos dados em uma rede NARX, permitindo que a rede capture as dependências temporais relevantes e, assim, obtenha resultados mais precisos e confiáveis.

- *Droupout Rate*: A taxa de *dropout* é um hiperparâmetro essencial no treinamento de redes neurais e desempenha um papel significativo na regularização e no controle do *overfitting.* Essa técnica introduz aleatoriamente a exclusão temporária de unidades (neurônios) durante o treinamento, o que impede que a rede dependa excessivamente de um subconjunto específico de unidades e ajuda a evitar o superajuste aos dados de treinamento. A definição adequada da taxa de *dropout* é crucial para garantir um equilíbrio entre a capacidade de aprendizado da rede e a capacidade de generalização para dados não vistos. Ao definir a taxa de *dropout*, é importante considerar o *trade-off* entre a capacidade de aprendizado e a regularização da rede. Uma taxa de *dropout* muito alta pode levar a uma perda excessiva de informações importantes durante o treinamento, prejudicando a capacidade de aprendizado da rede. Por outro lado, uma taxa de *dropout* muito baixa pode não ser eficaz o suficiente para regularizar a rede e controlar o *overfitting.*

- *Learning rate*: é um parâmetro fundamental no treinamento de redes neurais, pois influencia diretamente a convergência da rede para o mínimo global da função custo. Esse parâmetro determina o tamanho dos passos que são dados durante o processo de otimização dos pesos da rede, afetando a rapidez e a estabilidade da convergência. Uma taxa de aprendizado adequada é essencial para garantir que a rede neural alcance uma convergência satisfatória sem ficar presa em mínimos locais indesejados ou oscilar em torno do mínimo global. Valores muito altos de taxa de aprendizado podem fazer com que a

rede salte de um mínimo para outro sem atingir a convergência, resultando em instabilidade no processo de treinamento. Por outro lado, valores muito baixos de taxa de aprendizado podem levar a um treinamento lento, exigindo um número maior de iterações para alcançar a convergência.

- *Hidden Size Factor*: O fator de tamanho oculto (*Hidden Size Factor*) é um parâmetro que determina o número de neurônios na camada oculta de uma rede neural. A escolha adequada desse fator é crucial para garantir que a rede neural tenha capacidade suficiente para capturar a complexidade e os padrões dos dados. O valor do fator de tamanho oculto neste trabalho foi determinado levando em consideração a quantidade de entradas e saídas da rede. Essa relação pode ser expressa pela seguinte equação:

$$(Nentradas + Nsaídas) * fator$$

O número de entradas refere-se à quantidade de variáveis de entrada utilizadas para alimentar a rede neural, enquanto o número de saídas se refere à quantidade de variáveis de saída que a rede deve prever. O fator é um valor que pode ser ajustado empiricamente ou por meio de técnicas de otimização para obter o melhor desempenho da rede. Ao determinar o tamanho da camada oculta com base no fator de tamanho oculto, é importante considerar a complexidade dos dados e a capacidade de generalização desejada. Um número muito pequeno de neurônios pode levar a um modelo subajustado, incapaz de aprender padrões complexos nos dados. Por outro lado, um número muito grande de neurônios pode resultar em *overfitting*, no qual o modelo se ajusta excessivamente aos dados de treinamento, prejudicando sua capacidade de generalização para novos dados.

- *Ratio*: O parâmetro *Ratio* foi utilizado para controlar o número de entradas das redes NARX e evitar que o modelo tivesse um número excessivamente grande de variáveis de entrada com alta autocorrelação. Esse valor representa o tamanho da janela entre os *lags* selecionados para a construção do modelo. A escolha adequada do valor do *Ratio* é importante para equilibrar a complexidade do modelo e a capacidade de capturar padrões relevantes nos dados. Um valor muito baixo pode resultar em um modelo simplificado demais, incapaz de capturar a dinâmica temporal dos dados. Por outro lado, um valor muito alto pode levar a um modelo com excesso de informações e maior complexidade computacional, sem trazer benefícios significativos para a precisão da previsão.

- *Optimizer*: O hiperparâmetro "*Optimizer*" refere-se ao tipo de algoritmo de otimização utilizado para ajustar os pesos dos neurônios em uma rede neural durante o processo de aprendizagem. A escolha do otimizador é crucial, pois diferentes algoritmos podem apresentar comportamentos distintos na convergência do modelo e na eficiência do treinamento. Neste estudo, foram testados diversos otimizadores, incluindo Adam, Adadelta, Adagrad, RMSProp e SGD. Cada um desses otimizadores possui características específicas que afetam o desempenho da rede neural. O Adam (Adaptive Moment Estimation) é um otimizador que combina os benefícios do método de gradiente descendente estocástico (SGD) com as vantagens do RMSProp. Ele ajusta a taxa de aprendizado de forma adaptativa para cada parâmetro individual, além de manter estimativas de momento adaptáveis para controlar a direção e a magnitude das atualizações dos pesos (KINGMA; BA, 2015). O Adadelta é um otimizador que ajusta a taxa de aprendizado adaptativamente com base no histórico recente dos gradientes. Ele utiliza uma técnica de média móvel ponderada dos gradientes anteriores para atualizar os pesos de forma eficiente (ZEILER, 2012). O Adagrad é um otimizador que adapta a taxa de aprendizado para cada parâmetro com base na frequência com que esse parâmetro é atualizado. Ele dá mais importância aos parâmetros menos frequentes, aplicando uma taxa de aprendizado menor a eles (DUCHI; HAZAN; SINGER, 2011). O RMSProp (*Root Mean Square Propagation*) é um otimizador que também adapta a taxa de aprendizado para cada parâmetro individual, mas utiliza uma técnica de média móvel ponderada dos gradientes ao quadrado. Isso permite que a taxa de aprendizado se ajuste de forma diferente para diferentes parâmetros, melhorando a convergência em direção ao mínimo global da função custo (TIELEMAN; HINTON, 2012). Por fim, o SGD (*Stochastic Gradient Descent*) é o otimizador clássico que atualiza os pesos de acordo com o gradiente da função custo calculado em cada exemplo de treinamento. Ele não utiliza informações de momentos adaptativos ou histórico dos gradientes, sendo considerado mais simples em comparação aos outros otimizadores mencionados. (BOTTOU, 2010).

4 RESULTADOS E DISCUSSÃO

Nesta seção, serão apresentados e discutidos os dados, gráficos, métricas de desempenho e outros resultados relevantes, com o objetivo de responder às perguntas de pesquisa e avaliar a eficácia e robustez dos métodos empregados. Inicialmente, serão apresentados os resultados da validação cruzada e da avaliação dos modelos em diferentes conjuntos de dados.

Os resultados obtidos pelos oito modelos, divididos em quatro baseados em redes NARX e quatro baseados em redes LSTM, são apresentados nas Tabelas 1 e 2. Cada modelo foi testado em quatro conjuntos de dados distintos para uma avaliação abrangente do desempenho.

Cada experimento representa diferentes períodos de operação do Separador Trifásico, com um intervalo mínimo de três meses entre eles e em relação aos dados de treinamento do modelo. O teste #2, por exemplo, utilizou dados de um período 11 meses posterior ao período utilizado para treinamento da rede neural. Essas condições desafiadoras foram escolhidas considerando as mudanças que ocorrem ao longo do tempo em uma planta de processo, como formação de incrustações, recalibração ou perda de calibração de instrumentos e alterações nas configurações dos poços e trens de produção.

As incrustações nas tubulações e válvulas são gradualmente formadas pelos elementos presentes no fluido processado, afetando o fluxo e a eficiência do separador. A recalibração ou perda de calibração dos instrumentos e válvulas é necessária para manter a precisão e confiabilidade das leituras e do controle de processo. Durante os testes, observou-se a ocorrência de perda de calibração, resultando em medições imprecisas e comprometendo o desempenho geral do separador.

Além disso, as alterações nas configurações dos poços e trens de produção são comuns ao longo do tempo para otimizar os processos industriais. Essas alterações podem influenciar a composição e as propriedades do fluido que passa pelo separador, afetando sua eficiência e capacidade de separação. Ao abordar essas situações desafiadoras nos testes, busca-se avaliar a capacidade dos modelos em lidar com condições operacionais variáveis e fornecer resultados precisos e confiáveis, mesmo diante das mudanças inerentes ao contexto industrial do Separador Trifásico.

Tabela 1 – Métricas de Erro MAE para cada modelo

# Teste	**MAE**	***1-step***	***3-steps***	***5-steps***	***10-steps***
Teste #1	NARX	0.0185	0.0188	0.0272	0.0737
	LSTM	0.0230	0.0272	0.0695	0.1256
Teste #2	NARX	0.0181	0.0188	0.0276	0.0723
	LSTM	0.0299	0.0446	0.0742	0.1213
Teste #3	NARX	0.0185	0.0203	0.0282	0.0786
	LSTM	0.0230	0.0272	0.0695	0.1256
Teste #4	NARX	0.0185	0.0203	0.0282	0.0785
	LSTM	0.0255	0.0372	0.0943	0.1561

Fonte: Autor (2023)

Ao analisar os resultados das métricas de erro MAE para os testes NARX e LSTM, observamos que os modelos NARX tendem a apresentar menores valores de erro em comparação com os modelos LSTM em todas as configurações de horizonte de previsão. No Teste #1, por exemplo, o modelo NARX possui um MAE de 0.0185 no horizonte de previsão de *1-step*, enquanto o modelo LSTM apresenta um MAE ligeiramente maior, de 0.0230. Essa tendência se repete nos demais horizontes de previsão (*3-steps, 5-steps e 10-steps*), onde os modelos NARX apresentam MAEs menores em comparação com os modelos LSTM.

Essa diferença de desempenho pode ser observada também nos demais testes. No Teste #2, por exemplo, os modelos NARX têm MAEs inferiores em todos os horizontes de previsão em comparação com os modelos LSTM. O mesmo padrão é observado nos Testes #3 e #4, onde os modelos NARX tendem a ter MAEs menores do que os modelos LSTM. Esses resultados sugerem que os modelos NARX têm uma capacidade superior em realizar previsões mais precisas em comparação com os modelos LSTM para os conjuntos de dados e configurações de teste utilizados. No entanto, é importante ressaltar que a escolha do modelo mais adequado depende do contexto específico, incluindo as características dos dados, a complexidade do problema e as necessidades da aplicação.

Tabela 2 – Métricas de Erro MSE para cada modelo

# Teste	MSE	*1-step*	*3-steps*	*5-steps*	*10-steps*
Teste #1	NARX	0.0022	0.0029	**0.0045**	**0.0134**
	LSTM	**0.0010**	**0.0013**	0.0097	0.0305
Teste #2	NARX	**0.0021**	**0.0027**	**0.0042**	**0.0125**
	LSTM	0.0045	0.0049	0.0146	0.0433
Teste #3	NARX	0.0023	0.0034	**0.0053**	**0.0160**
	LSTM	**0.0012**	**0.0015**	0.0098	0.0312
Teste #4	NARX	0.0023	0.0034	**0.0052**	**0.0160**
	LSTM	**0.0018**	**0.0030**	0.0218	0.0488

Fonte: Autor (2023)

A análise dos resultados apresentados na tabela indica as diferenças de desempenho entre os modelos NARX e LSTM em termos de métrica de erro MSE (*Mean Squared Error*) em diferentes horizontes de previsão.

No Teste #1, tanto o modelo NARX quanto o LSTM apresentam valores de MSE relativamente baixos no horizonte de previsão de 1-step e 3-steps. No entanto, nos horizontes de previsão de 5-steps e 10-steps, o modelo LSTM exibe MSEs maiores do que o modelo NARX. Essa tendência é observada também no Teste #3, onde o LSTM mostra MSEs menores aos do NARX nos horizontes de previsão de *1-step* e *3-steps*, e valores maiores para *5-steps* e *10-steps*.

Por exemplo, no Teste #2, o modelo LSTM apresenta MSEs maiores do que o modelo NARX em todos os horizontes de previsão. A diferença é mais notável no horizonte de previsão de 5-steps e 10-steps, onde o LSTM tem MSEs significativamente maiores.

Já no Teste #4, a rede LSTM apresentou MSEs ligeiramente mais baixos do que a rede NARX para os horizontes de *1-step* e *3-steps*, e significativamente maiores para horizontes de *5-steps e 10-steps.*

Os resultados obtidos nas métricas de erro MAE e MSE despertam curiosidade devido às diferenças observadas entre as arquiteturas de rede NARX e LSTM. Ao analisar a métrica MAE, fica evidente que a rede NARX apresentou consistentemente resultados melhores em compara-

ção com a rede LSTM. No entanto, ao considerar a métrica MSE, observamos um comportamento distinto, onde as redes LSTM mostraram resultados superiores para horizontes de previsão mais curtos, mas resultados significativamente piores para horizontes de previsão mais longos. As redes LSTM são conhecidas por sua habilidade em capturar informações sequenciais e padrões imediatos, tornando-as mais adequadas para previsões de curto prazo. No entanto, quando se trata de horizontes de previsão mais longos, as redes LSTM enfrentam dificuldades em manter a memória das dependências de longo prazo, o que pode levar a uma deterioração do desempenho.

Considerando a notável eficácia e consistência dos resultados obtidos pelas redes NARX em relação às redes LSTM, conduziram-se experimentos mais exigentes com o propósito de avaliar a robustez de uma rede pré-treinada. Nesses experimentos, o modelo foi inicialmente treinado apenas para o horizonte de *1-step* e, posteriormente, testado para os horizontes de *3-steps, 5-steps e 10-steps*, a fim de avaliar sua capacidade de previsão em diferentes horizontes temporais.

Tabela 3 – Métricas de Erro MAE de um modelo NARX em quatro horizontes de previsão

# Teste	**MAE**	***1-step***	***3-steps***	***5-steps***	***10-steps***
Teste #1	NARX	0.0022	0.0035	0.0062	0.0166
Teste #2	NARX	0.0021	0.0034	0.0079	0.0174
Teste #3	NARX	0.0023	0.0039	0.0078	0.0199
Teste #4	NARX	0.0023	0.0040	0.0077	0.0191

Fonte: Autor

É de grande interesse notar a consistência dos resultados obtidos em todos os testes realizados, mesmo levando em consideração que esses resultados são naturalmente inferiores em comparação com aqueles obtidos quando as redes foram treinadas de forma específica para os horizontes de previsão em que foram avaliadas.

Essa consistência pode indicar a capacidade da arquitetura NARX de generalizar os padrões e relações presentes nos dados de entrada, permitindo um desempenho razoável mesmo em horizontes de previsão que não foram alvo de treinamento direto. Isso sugere que as redes estão capturando informações relevantes durante o treinamento que podem ser aplicadas de maneira mais ampla para diferentes horizontes de previsão.

No entanto, é importante ressaltar que, apesar dessa consistência, os resultados obtidos nos testes específicos para os horizontes de previsão treinados mostraram-se superiores, indicando a importância de um treinamento direcionado para obter os melhores resultados possíveis. Esse resultado é esperado, uma vez que o treinamento voltado para um horizonte de previsão específico permite que a rede aprenda características mais refinadas e ajuste seus parâmetros de acordo com as necessidades específicas desse horizonte.

5 CONCLUSÃO

Com base nos resultados obtidos, foi possível implementar com sucesso um *Digital Twin* de um separador trifásico, demonstrando a capacidade de replicar com precisão o comportamento do sistema físico. Os modelos NARX utilizados no *Digital Twin* mostraram-se mais robustos e capazes de lidar com diferentes condições de previsão em comparação com os modelos LSTM nos conjuntos de dados e configurações de teste utilizados.

Ao avaliar as métricas de erro MAE e MSE, os modelos NARX apresentaram consistentemente valores menores em todos os horizontes de previsão analisados. Isso indica que esses modelos foram capazes de realizar previsões mais precisas e robustas, independentemente do horizonte de previsão considerado. Além disso, os modelos NARX mantiveram sua precisão mesmo em previsões de longo prazo, o que confirma sua capacidade de lidar com horizontes de previsão mais longos. Esses resultados sugerem que as redes NARX são mais confiáveis e adaptáveis para problemas semelhantes ou com dados de natureza semelhante, tornando-as uma opção promissora para o desenvolvimento de *Digital Twins* em uma variedade de aplicações práticas. A implementação bem-sucedida do *Digital Twin* do separador trifásico contribui para o avanço do conhecimento na área e fornece uma ferramenta poderosa para prever e controlar o comportamento do sistema físico.

Com o *Digital Twin*, é possível realizar simulações, testes e otimizações virtuais do separador trifásico, reduzindo a necessidade de intervenções físicas e os custos associados. Além disso, o *Digital Twin* permite um monitoramento contínuo e em tempo real do sistema físico, identificando possíveis falhas, otimizando o desempenho e auxiliando na tomada de decisões mais informadas. Portanto, os resultados obtidos com a implementação do *Digital Twin* do separador trifásico demonstraram não apenas a precisão e confiabilidade dos modelos NARX em previsões, mas também a viabilidade de replicar o comportamento do sistema físico. Essa conquista representa um avanço significativo no desenvolvimento de *Digital Twins* para sistemas complexos e oferece oportunidades para melhorar a eficiência, confiabilidade e segurança de diversas aplicações industriais.

5.1 TRABALHOS FUTUROS

Como sugestões para trabalhos futuros, recomenda-se a realização de testes e simulações de diferentes técnicas de controle de processo utilizando o *Digital Twin* desenvolvido. Essas técnicas podem incluir desde sintonias de controle clássico até abordagens mais avançadas, como controle *fuzzy*, *neuro-fuzzy*, controle preditivo, MPC, etc. Um estudo comparativo entre essas diversas técnicas pode ser conduzido para identificar aquela que melhor se adequa ao processo em questão, levando em consideração critérios como desempenho, simplicidade de implementação na planta real e custo.

Além disso, é interessante explorar outras técnicas de desenvolvimento de *Digital Twins* e comparar seus resultados com o trabalho atual. Isso permitirá avaliar a eficácia e as vantagens de diferentes abordagens na replicação precisa e confiável do comportamento de sistemas físicos. Essa comparação pode ajudar a identificar as melhores práticas e aprimorar ainda mais a implementação de *Digital Twins* em diferentes processos industriais.

Adicionalmente, uma extensão promissora desse trabalho é aplicar o desenvolvimento do *Digital Twin* a outros processos na indústria. Cada sistema apresenta características e desafios únicos, e replicar seu comportamento por meio de um *Digital Twin* pode trazer benefícios significativos, como otimização de desempenho, identificação de falhas e tomada de decisões mais informadas. Explorar a aplicação do *Digital Twin* em outros contextos industriais permitirá ampliar o conhecimento e as possibilidades de uso dessa tecnologia.

Dessa forma, essas sugestões de trabalhos futuros podem contribuir para o avanço do campo de *Digital Twins*, fornecendo *insights* valiosos sobre técnicas de controle, abordagens de desenvolvimento e aplicações em diferentes setores industriais.

REFERÊNCIAS

ABDEL-AAL, H. K.; AGGOUR, M.; FAHIM, M. A. **Petroleum and gas field processing**. New York: Marcel Dekker, 2003.

AGÊNCIA NACIONAL DO PETRÓLEO. **Especial ANP 20 Anos**. Disponível em: <https://www.gov.br/anp/pt-br/acesso-a-informacao/institucional/especial-anp-20-anos>.

AL-ANAZI, M.; FAHIM, M. A. Review on Level Control in Three-Phase Separators. **Journal of Engineering Research and Reports**, v. 6, n. 1, p. 1-13, 2019. Disponível em: https://doi.org/10.9734/jerr/2019/v6i117124. Acesso em: 15 abr. 2023.

ARAGÃO, A. P. **Modelagem e simulação computacional de processos produtivos**: o caso da cerâmica vermelha de Campos dos Goytacazes, RJ. 2011. Dissertação (Mestrado em Engenharia de Produção) – Centro de Tecnologia, Universidade Estadual do Norte Fluminense, Campos dos Goytacazes, 2011.

ARAGÃO, Márcio Antonio de. Introdução à modelagem matemática de sistemas dinâmicos. **Ed. da UFMG**, 2011.

ARENE, E.O.; KITWOOD, T. M. **An Introduction to the Chemistry of Carbon Compounds**, Londres: Longman, 1979.

AZZOUNI, Y.; LAHMER, M.; DESSOUKI, A.; et al. Optimal selection of the forecast horizon in time series forecasting. **International Journal of Forecasting**, v. 36, n. 4, p. 1478-1491, 2020.

BELAUD, Jean-Pierre et al. Process modeling: from conceptualization to implementation. **John Wiley & Sons**, 2019.

BENGIO, Y.; SIMARD, P.; FRASCONI, P. Learning long-term dependencies with gradient descent is difficult. **IEEE transactions on neural networks**, [s.l.], v. 5, n. 2, p. 157-166, mar. 1994. DOI: 10.1109/72.279181.

BENITO, J. L. et al. Analysis and modeling of slug flow and gas lift instability in vertical pipes. **Journal of Petroleum Science and Engineering**, v. 179, p. 345-358, 2019.

BENTLEY, R. W. et al. Global oil supply 2016: prospects to 2040. **Edinburgh: Uppsala University,** 2016.

BEZERRA, M. L. M. et al. Effect of Oil Level on Three-Phase Separator Performance in a Petroleum Refinery. **Journal of Petroleum Science and Engineering**, v. 182, p. 106360, 2019.

BERTANI, R. et al. Slugging and gas lift instability: from field observations to a comprehensive instability map. **Journal of Petroleum Science and Engineering**, v. 180, p. 665-677, 2019.

BORDALO, S. N.; MOROOKA, C. K. Slug flow induced oscillations on subsea petroleum pipelines. **Journal of Petroleum Science and Engineering**, v. 165, p. 535–549, jun. 2018.

BOTTOU, L. Large-scale machine learning with stochastic gradient descent. In: LEON BOTTOU, O.; CHAPLEAU, Y.; DECELLE, A.; DOUZE, M.; GROUSSET, F.; LIU, C. K. (Eds.).

COMPSTAT'2010: **Proceedings in Computational Statistics**. Berlin, Heidelberg: Springer, 2010. p. 177-186.

BROOKS, A., AKINJOGUNLA, O., MAHMOOD, T., & ODEY, O. (2015). Investigation of primary production and water injection in offshore reservoirs using reservoir simulation. **Journal of Petroleum Science and Engineering**, 135, 279-288.

CAMPOS, M. et al. Advanced Anti-slug Control for Offshore Production Plants. **IFAC-PapersOnLine**, v. 48, n. 6, p. 83–88, 2015.

CHEN, T.; GAO, J.; ZHU, Y.; et al. Multi-step ahead time series forecasting using residual analysis based decision trees. **Neurocomputing**, v. 282, p. 74-84, 2018.

CHEN, Y. Model-based process control. In: MCAVOY, T.J. (Ed.). Handbook of advanced process control and optimization. **Boca Raton: CRC Press**, 2017. p. 1-22.

DELLI, V. et al. Deep learning-based identification of a nonlinear dynamic system with real-time parameter estimation. **Applied Soft Computing**, v. 101, p. 107003, 2021.

DENG, Shuai. An Introduction to Optuna. 2019. **Medium.** Disponível em: https://towardsdatascience.com/an-introduction-to-optuna-4e92f785aad. Acesso em: 01 maio 2023.

DIEHL, F. C.; ANZAI, T. K.; ALMEIDA, C. S.; VON MEIEN, O. F.; NETO, S. S.; ROSA, V. R.; CAMPOS, M. C. M. M.; REOLON, F.; GEREVINI, G.; RANZAN, C.; FARENZENA, M.; TRIERWEILER, J. O. Fast Offshore Wells Model (FOWM): A practical dynamic model for multiphase oil production systems in deepwater and ultra-deepwater scenarios. **Computers & Chemical Engineering**, v. 99, p. 304-313, 6 abr. 2017. Disponível em: https://www.sciencedirect.com/science/article/abs/pii/S0098135417300443. Acesso em: 29 julho 2023.

DIPIETRO, R. et al. Analyzing and exploiting NARX recurrent neural networks for long-term dependencies. In**: WORKSHOP TRACK -** INTERNATIONAL CONFERENCE ON LEARNING REPRESENTATION 2018, 6., 2018, Vancouver, BC. **Anais**… Vancouver. Disponível em: https://arxiv.org/pdf/1702.07805.pdf. Acesso em: 5 jun. 2022.

DUCHI, J.; HAZAN, E.; SINGER, Y. Adaptive subgradient methods for online learning and stochastic optimization. **Journal of Machine Learning Research**, v. 12, n. Jul, p. 2121-2159, 2011.

GARCÍA, S.; HERRERA, F. An extension on "statistical comparisons of classifiers over multiple data sets" for all pairwise comparisons. **Journal of Machine Learning Research**, v. 9, p. 2677-2694, 2008.

GERS, F. A. et al. Learning to forget: Continual prediction with LSTM. **Neural computation**, v. 12, n. 10, p. 2451-2471, 2000.

GHAFFARKHAH, A.; SHAHRABI, M. A.; MORAVEJI, M. K. 3D Computational-Fluid-Dynamics Modeling of Horizontal Three-Phase Separators: An Approach for Estimating the Optimal Dimensions. **SPE Production & Operations**, v. 33, n. 04, p. 879–895, 1 nov. 2018.

GHARIBI, W. et al. Short-term wind power forecasting using long short-term memory neural network. **Electric Power Systems Research**, v. 185, 106523, 2020.

GLAESSGEN, W. H.; STARGEL, D. S. TheDigital Twin Paradigm forFuture NASA and U.S. Air Force Vehicles. In: **53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference - Special Session on the Digital Twin**, [s.l.: s.n.], 2012. Disponível em: https://ntrs.nasa.gov/citations/20120008178.

GOLDBERG, E. D., & GRUPO DE TRABALHO DE CONTROLE DE DERRAMAMENTOS DE PETRÓLEO. (2001). Derramamentos de petróleo no mar: ciência, estratégia e tecnologia. **Edusp**.

GRIEVES, M.; VICKERS, J. Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems. **Transdisciplinary Perspectives on Complex Systems**, p. 85–113, 17 ago. 2016.

HAJIZADEH, S. et al. Modeling monthly temperature using NARX neural network and autoregressive models: a comparative analysis. **Theoretical and Applied Climatology**, v. 145, n. 1-2, p. 235-246, 2021.

HANSEN, L. et al. Performance Comparison of Control Strategies for Plant-Wide Produced Water Treatment. **Energies**, v. 15, n. 2, p. 418, 1 jan. 2022.

HE, X.; WANG, C. Research on NARX neural network for long-term prediction of electricity consumption. **Journal of Physics: Conference Series**, v. 1168, n. 3, p. 032016, 2019.

HEGHEDUS, C.; SHCHIPANOV, A.; RONG, C. Advancing Deep Learning to Improve Upstream Petroleum Monitoring. **IEEE Access**, v. 7, p. 106248–106259, 2019.

HENAFF, M.; SZLAM, A.; LECUN, Y. Recurrent orthogonal networks and long-memory tasks. In: PROCEEDINGS OF THE 33 RD INTERNATIONAL CONFERENCE ON MACHINE LEARNING, 2017, Nova York. **Anais**… Nova York. DOI 10.48550/arXiv.1602.06662. Disponível em: https://arxiv.org/abs/1602.06662. Acesso em: 5 jun. 2022.

HUANG, B. et al. Nonlinear process identification and control using Gaussian processes. **Journal of Process Control**, v. 57, p. 30-44, 2017.

HUANG, J. et al. Stock price prediction with ensemble learning using NARX recurrent neural network. **International Journal of Machine Learning and Cybernetics**, v. 10, n. 1, p. 77-88, 2019.

HUANG, X.; ZHU, Y.; LIAO, X.; LI, X.; CHEN, X. A Comprehensive Survey on Long Short-Term Memory Neural Networks. **arXiv** preprint arXiv:2004.11553, 2020.

HYNDMAN, R. J.; ATHANASOPOULOS, G. Forecasting: principles and practice. 2. ed. Melbourne: **OTexts**, 2018.

IZENMAN, A. J. Modern multivariate statistical techniques: regression, classification, and manifold learning. **Springer Science & Business Medi**a, 2013.

JAMIL, M. et al. Load forecasting using NARX neural network: A comparative study with multilayer perceptron. In: 2018 **International Conference on Computing, Mathematics and Engineering Technologies** (iCoMET). IEEE, 2018. p. 1-5.

KARSOLIYA, S. **Approximating Number of Hidden layer neurons in Multiple Hidden Layer BPNN Architecture International Journal of Engineering Trends and Technology**. **International Journal of Engineering Trends and Technology** [s.l: s.n.]. Disponível em: http://www.ijettjournal.org/volume-3/issue-6/IJETT-V3I6P206.pdf. Acesso em: 5 jun. 2022.

KAUR, H.; SINGH, D. A review of artificial neural network modeling techniques and its applications. **International Journal of Computer Science and Information Technology Research**, v. 7, n. 1, p. 1-16, 2019.

KAUFMAN, S.; ROSSET, S.; PERLICH, C. Leakage in data mining: formulation, detection, and avoidance. In: **Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining, 11.**, 2011, [s.l.], p. 556-563. **Anais…** DOI 10.1145/2020408.2020496. Disponível em: https://dl.acm.org/doi/10.1145/2020408.2020496. Acesso em: 5 jun. 2022.

KHALIL, M., WANG, H., SAVAGE, D., & WELLS, J. (2016). Gas-oil relative permeability for compositional simulation of heavy oil reservoirs with gas injection. **Journal of Petroleum Science and Engineering**, 147, 221-231.

KHALEGHI-MOGHADDAM, J.; GHAFARI, M.; SHOKOUHI, S. B. Machine Learning in COVID-19 Fight: An Overview of Clinical Decision Support System. **Current Medical Imaging Reviews**, v. 16, n. 1, p. 21-27, 2020.

KINGMA, D. P.; BA, J. Adam: A method for stochastic optimization. In: **International Conference on Learning Representations** (ICLR), 2015.

KRITZINGER, W.; KARNER, M.; TRAAR, G.; HENJES, J.; SANKARAN, B. Digital Twin in manufacturing: A categorical literature review and classification. **IFAC-PapersOnLine**, v. 51, n. 11, p. 1016–1022, 1 jan. 2018. Disponível em: https://www.researchgate.net/publication/327491752_Digital_Twin_in_manufacturing_A_categorical_literature_review_and_classification. Acesso em: 29 jul. 2023.

KRISHNA, A.; SEKAR, R. An enhanced maximum likelihood estimator based identification of dynamic systems. **International Journal of Automation and Control**, v. 14, n. 3, p. 257-278, 2020.

KOMER, Brent et al. The Hyperspace Analogue to Language (HAL): A General Framework for Document Representation. **Journal of Artificial Intelligence Research**, v. 51, p. 659-697, 2014.

LAGUNA, J. et al. Digital twins for Industry 4.0: A survey. **Computers in Industry**, v. 135, p. 103-125, 2022. DOI: https://doi.org/10.1016/j.compind.2021.103125.

LAMBERTS, J. J.; SONTAG, E. D. Stability and time-delay in NARMAX systems. **IEEE Transactions on Automatic Control**, v. 58, n. 2, p. 266-281, 2013.

LEE, K. M. et al. A Comparative Study of Neural Network Training Methods for Process System Modeling. **IEEE Access,** v. 7, p. 45915-45926, 2019. DOI: 10.1109/ACCESS.2019.2904089.

LI, H. et al. Separation efficiency improvement of a three-phase separator using intelligent control. **IEEE Access**, v. 8, p. 206805-206814, 2020.

LI, Y.; WANG, N.; DUAN, Z.; PENG, Y. A survey on recurrent neural networks for sequence learning. **IEEE Transactions on Neural Networks and Learning Systems**, v. 31, n. 4, p. 1235-1255, 2020.

LI, H.; XIA, Y.; LIU, J. Online fault detection of rolling bearing based on improved deep recurrent neural networks. **Neurocomputing**, v. 226, p. 113-122, 2017.

LI, Y. et al. Improved short-term wind power prediction using a novel hybrid model based on wavelet transform and nonlinear autoregressive models with exogenous input. **Energy**, v. 145, p. 900-913, 2018.

LIANG, Y. et al. Nonlinear Autoregressive Model with Exogenous Input Recurrent Neural Network to Predict Satellites' Clock Bias. **IEEE Access**, p. 1–1, 2021.

LIU, G. et al. Industrial process monitoring and control based on deep learning: An overview. **IEEE Transactions on Industrial Informatics**, v. 15, n. 5, p. 3194-3205, 2019.

LIU, Y. et al. A review on regression algorithms for dynamic modeling of process systems. **Chemical Engineering Science**, v. 186, p. 147-164, 2018.

LIMA, R. S., GUIMARÃES, L. H. S., & CHAVES, T. H. (2014). Caracterização do petróleo e do condensado produzido no Campo de Urucu, Solimões, Brasil. **Química Nova**, 37

LIU, J. et al. NARX Prediction-Based Parameters Online Tuning Method of Intelligent PID System. **IEEE Access**, v. 8, p. 130922–130936, 2020.

LIU, Q. et al. An Optimal NARX Neural Network Identification Model for a Magnetorheological Damper With Force-Distortion Behavior. **Frontiers in Materials**, v. 7, 14 fev. 2020.

LIU, J. et al. Nonlinear autoregressive model with exogenous input based on recurrent neural network for prediction of chaotic time series. **Neurocomputing**, v. 381, p. 129-141, 2020.

LIU, Z. et al. Multi-step prediction of oil production rate based on NARX with data preprocessing. **IEEE Access**, v. 8, p. 134157-134168, 2020.

MEDEIROS, J. L. et al. Predictive control of chemical reactors using nonlinear autoregressive neural network models. **Chemical Engineering Transactions**, v. 57, p. 523-528, 2017.

Microsoft Power BI. Disponível em: https://app.powerbi.com/view?r=eyJrIjoiNzVmNzI1MzQtNTY1NC00ZGVhLTk5N2ItNzBkMDNhY2IxZTIxIiwidCI6IjQ0OTlmNGZmLTI0YTYtNGI0Mi1iN2VmLTEyNGFmY2FkYzkxMyJ9. Acesso em: 5 jun. 2022.

MILANI, E. J. et al. Petróleo na margem continental brasileira: Geologia, exploração, resultados e perspectivas. **Brazilian Journal of Geophysics**, v. 18, n. 3, p. 352–396, 6 ago. 2018. ISSN 1809-4511. Available at: https://sbgf.org.br/revista/index.php/rbgf/article/view/1418. Date accessed: 05 june 2022.

MONTGOMERY, D. C.; JENNINGS, C. L.; KULAHCI, M. Introduction to Time Series Analysis and Forecasting. 2. ed. **Hoboken: John Wiley & Sons**, 2015.

NG, A. **Machine Learning Yearning:** Technical Strategy for Al Engineers, In the Era of Deep Learning. Massachusetts: GitHub (MIT Lincensed), 2018, 118 p. ISBN – 13 978-1999579500

NNABUIFE, S. G.; TANDOH, H.; WHIDBORNE, J. F. Slug flow control using topside measurements: A review. **Chemical Engineering Journal Advances**, v. 9, p. 100204, 15 mar. 2022.

KANDASAMY, K. et al. Neural Architecture Search with Bayesian Optimisation and Optimal Transport. In: **International Conference on Learning Representations**, 2018.

NUNES, G. C.; MEDEIROS, J. L.; ARAÚJO, O. Q. F. **Modelagem e controle na produção de petróleo**: Aplicações em MatLab, São Paulo: Blucher, 2010. ISBN 978-85-212-0567-8.

OCHÔA, V. M.; PEIXOTO, J. A.; SOARES, A. B. Autogemini: Gêmeos digitais na indústria 4.0. In: SALÃO INTEGRADO DE ENSINO, PESQUISA E EXTENSÃO DA UERGS, 10., 2021, Rio Grande do Sul. **Anais**... Rio Grande do Sul: Uergs, 2021. Não paginado.

PEDERSEN, S.; DURDEVIC, P.; YANG, Z. Review of Slug Detection, Modeling and Control Techniques for Offshore Oil & Gas Production Processes∗∗Supported by the Danish National Advanced Technology Foundation through PDPWAC Project (J.nr. 95-2012-3). **IFAC-PapersOnLine**, v. 48, n. 6, p. 89–96, 2015.

PEDREGOSA, F. et al. Scikit-learn: Machine learning in Python. **Journal of machine learning research**, v. 12, p. 2825-2830, 2011.

PÉREZ, R.; MIRANDA, L.; LÓPEZ, R. A hybrid modeling approach for the simulation of a hydrogen production plant. **Energy Procedia**, v. 126, p. 101-108, 2017.

PETROBRAS. Unidade de Negócios de Exploração e Produção do Espírito Santo. Injeção de água no reservatório. **Vitória: Petrobras**, 2007.

PETROBRAS. **Pilotos de otimização da produção do E&P/SSE**: relatório técnico final. Centro de Pesquisas e Desenvolvimento Leopoldo A. Miguez de Mello, 2020.

PETROBRAS. **Processamento primário de petróleo**. Rio de Janeiro: Escola de Ciências e Tecnologia E&P, Universidade Petrobras, 2007.

RAO, C. V. et al. Nonlinear process identification using support vector machines with model predictive control. **Industrial & Engineering Chemistry Research**, v. 58, n. 14, p. 5713-5722, 2019.

REIS, D. M.; COSTA, R. S.; CARVALHO, A. M.; LIMA, E. L. S.; SANTOS, M. C. F. Avaliação da utilização de água do mar como fluido injetado em reservatório de petróleo. In: **Anais do Congresso Brasileiro de Engenharia de Petróleo e Gás**, 2018, Natal. Anais... Rio de Janeiro: IBP, 2018.

ROUSSEEUW, P. J.; CROUX, C. Alternatives to the median absolute deviation. **Journal of the American Statistical Association**, v. 88, n. 424, p. 1273-1283, 1993.

RUIZ, V. R.; FLORES, B. E.; CAMPO, M. A.; et al. An assessment of the impact of the forecast horizon on time series forecasting. In: **International Conference on Computational Science and Its Applications**. Springer, Cham, p. 204-216, 2019.

SANTOS, J. C. et al. A comparative analysis of methods to predict stock market indexes using nonlinear models. **Expert Systems with Applications**, v. 114, p. 402-412, 2018.

SAUL, L. K.; JORDAN, M. I. Exploiting boundedness and smoothness when using Gaussian priors. In: **Advances in Neural Information Processing Systems**, p. 683-690, 1996.

SCHRAMM, L. L. (ed.). **Emulsions**: Fundamentals and Applications in the Petroleum Industry. Washington: Advances in Chemistry - American Chemical Society, 5 maio 1992. ISBN 0-8412-2006-9

SHAHRIARI, B. et al. Taking the Human Out of the Loop: A Review of Bayesian Optimization. **Proceedings of the IEEE**, v. 104, n. 1, p

SHI, Q.; ZHU, C.; TAO, F. Digital twin and its potential application in the construction industry. **Journal of Cleaner Production**, v. 196, p. 1055-1066, 2018.

SIMÕES, S.; SANTOS, W. T.; CAMPOS, M. C. M. M.; PIMENTA, A.; TEIXEIRA, A. F.; VON MEIEN, O. F. Advanced Control Systems for Offshore Production Platforms. In: **OFFSHORE TECHNOLOGY CONFERENCE BRASIL, Rio de Janeiro, 2013**. Disponível em: https://onepetro.org/OTCBRASIL/proceedings-abstract/13OTCB/All-13OTCB/OTC-24286-MS/39775. Acesso em: 5 jun. 2022.

SINGH, M.; FUENMAYOR, E.; HINCHY, E. P.; QIAO, Y.; MURRAY, N.; DEVINE, D. Digital Twin: Origin to Future. **Applied System Innovation**, v. 4, n. 2, p. 36, 24 maio 2021. Disponível em: https://www.mdpi.com/2571-5577/4/2/36. Acesso em: 29 jul. 2023

SMARSLOK, B.; CULLER, A.; MAHADEVAN, S. Error Quantification and Confidence Assessment of Aerothermal Model Predictions for Hypersonic Aircraft. In: **53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference**, Hawaii, 2012.

SNOEK, Jasper et al. Practical bayesian optimization of machine learning algorithms. In: **Advances in neural information processing systems**. 2012. p. 2951-2959.

SOHAIL, A. et al. A deep learning approach for stock market prediction using long short-term memory neural networks. **Neural Computing and Applications**, v. 32, n. 22, p. 16921-16930, 2020.

SUN, X. et al. A review of digital twin technology: State-of-the-art and future directions. **Journal of Cleaner Production**, v. 321, p. 128879, 2021.

SUNDARAMOORTHY, S.; HOQUE, M. E.; ISLAM, M. R. et al. Review of Level Control Systems in Oil and Gas Production Separators. **Journal of Petroleum Science and Engineering**, v. 108, p. 59-66, 2013. Disponível em: https://doi.org/10.1016/j.petrol.2013.04.003. Acesso em: 15 abr. 2023.

SWERSKY, K.; SNOEK, J.; ADAMS, R. P. Multi-Task Bayesian Optimization. In: **Advances in Neural Information Processing Systems**, v. 27, p. 2004-2012, 2014.

TANG, J. et al. Nonlinear autoregressive models with exogenous inputs for hydrological prediction: A review. **Journal of Hydrology**, v. 592, p. 125663, 2021.

TAO, F. et al. Digital Twin in Industry: State-of-the-Art. **IEEE Transactions on Industrial Informatics**, v. 15, n. 4, p. 2405–2415, abr. 2019.

TAO, F. et al. Digital twin: a critical review and implications for future research. **International Journal of Production Research**, v. 57, n.

THOMAS, J. E. (org). **Fundamentos de Engenharia de Petróleo**. Rio de Janeiro: Interciência, 2001. ISBN 85-7193-046-5

TIELEMAN, T.; HINTON, G. RMSProp: Divide the gradient by a running average of its recent magnitude. COURSERA: **Neural Networks for Machine Learning**, v. 4, n. 2, p. 26-31, 2012.

VASCONCELOS, J. V. et al. Eletrostática em processos de tratamento de água e efluentes**.** **Química Nova**, São Paulo, v. 43, n. 9, p. 1203-1213, 2020

VUJIČIĆ, T. et al. **Comparative Analysis of Methods for Determining Number of Hidden Neurons in Artificial Neural Network**. In: CENTRAL EUROPEAN CONFERENCE ON INFORMATION AND INTELLIGENT SYSTEMS, [s.n.], 2016, Varaždin. **Anais**... Varaždin. Disponível em: http://archive.ceciis.foi.hr/app/public/conferences/1/ceciis2016/papers/IIS-4.pdf.

WANG, L. X.; YANG, Z. B.; HE, L. Y.; XU, J. H. Multi-objective optimization design for train weight forecasting model using deep learning algorithm. **Journal of Rail Transport Planning & Management**, v. 9, n. 2, p. 127-143, 2019.

WANG, Y. et al. A genetic algorithm-based method for identification of dynamic processes. **Chemical Engineering Science**, v. 211, p. 115299, 2020.

WANASINGHE, T. R.; WROBLEWSKI, L.; PETERSEN, B.; GOSINE, R. G.; JAMES, L. A.; DE SILVA, O.; MANN, G. K. I.; WARRIAN, P. J. Digital Twin for the Oil and Gas Industry: Overview, Research Trends, Opportunities, and Challenges. **IEEE Access**, v. 8, p. 104175-104207, June 2020. Disponível em: https://www.researchgate.net/publication/341812422_Digital_Twin_for_the_Oil_and_Gas_Industry_Overview_Research_Trends_Opportunities_and_Challenges. Acesso em: 29 jul. 2023.

YANG, Q. et al. A hybrid model of multiple regression analysis and nonlinear neural network for short-term wind speed forecasting. **Journal of Cleaner Production**, v. 225, p. 767-777, 2019. DOI: https://doi.org/10.1016/j.jclepro.2019.03.142.

YUAN, X.; JIANG, L.; ZHANG, Q. A survey on the LSTM networks. **Journal of Information Processing Systems**, v. 13, n. 4, p. 889-901, 2017.

ZEILER, M. D. ADADELTA: An adaptive learning rate method. **arXiv** preprint arXiv:1212.5701, 2012.

ZHANG, S. et al. Short-term wind power forecasting based on least squares support vector regression optimized by differential evolution algorithm and nonlinear autoregressive models with exogenous input. **Renewable Energy**, v. 158, p. 85-97, 2020.

ZHANG, Aston; LIPTON, Zachary C.; LI, Mu; SMOLA, Alexander J. 10.1. Long Short-Term Memory (LSTM). In: **DIVE INTO DEEP LEARNING**. 1.0.3 ed. [S.l.]: [s.n.], [2024]. Disponível em: https://d2l.ai/chapter_recurrent-modern/lstm.html. Acesso em: 02/01/2024.

ZHOU, D. et al. Evaluation of hydrocyclone as pre-treatment to dissolved air flotation (DAF) for the separation of oil from wastewater. **Separation Science and Technology**, v. 54, n. 5, p. 868-877,

ZHU, Y.; WANG, D.; CHEN, Z.; LIU, W. Digital Twin driven adaptive control of large-scale complex systems: A review. **Annual Reviews in Control**, v. 49, p. 265-279, 2020.