



UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO TECNOLÓGICO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE AUTOMAÇÃO E
SISTEMAS

Aline Evangelista Rubenich

**Técnicas de inteligência artificial para predição de temperatura da depleção de
hidrocarbonetos em áreas contaminadas**

Florianópolis
2023

Aline Evangelista Rubenich

Técnicas de inteligência artificial para predição de temperatura da depleção de hidrocarbonetos em áreas contaminadas

Dissertação submetida ao Programa de Pós-Graduação em Engenharia de Automação e Sistemas da Universidade Federal de Santa Catarina para a obtenção do título de Mestre em Engenharia de Automação e Sistemas.

Orientador: Prof. Rodolfo César Costa Flesch, Dr. Eng.

Coorientadores: Ahryman S. B. de Siqueira Nascimento, Dr. Eng. e Prof. Bernardo Barancelli Schwedersky, Dr. Eng.

Florianópolis
2023

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Rubenich, Aline Evangelista

Técnicas de inteligência artificial para predição de temperatura da depleção de hidrocarbonetos em áreas contaminadas / Aline Evangelista Rubenich ; orientador, Rodolfo César Costa Flesch, coorientador, Ahryman Seixas Busse de Siqueira Nascimento, coorientador, Bernardo Barancelli Schwdersky, 2023.

80 p.

Dissertação (mestrado) - Universidade Federal de Santa Catarina, Centro Tecnológico, Programa de Pós-Graduação em Engenharia de Automação e Sistemas, Florianópolis, 2023.

Inclui referências.

1. Engenharia de Automação e Sistemas. 2. Predição de temperatura. 3. Depleção da zona da fonte. 4. Derramamento. 5. Fase livre. I. Flesch, Rodolfo César Costa. II. Nascimento, Ahryman Seixas Busse de Siqueira. III. Schwdersky, Bernardo Barancelli IV. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Engenharia de Automação e Sistemas. V. Título.

Aline Evangelista Rubenich

Técnicas de inteligência artificial para predição de temperatura da depleção de hidrocarbonetos em áreas contaminadas

O presente trabalho em nível de Mestrado foi avaliado e aprovado por banca examinadora composta pelos seguintes membros:

Prof. Rodolfo César Costa Flesch, Dr. Eng.
Universidade Federal de Santa Catarina

Prof. Ademar Gonçalves da Costa Junior, Dr. Eng
Instituto Federal da Paraíba

Prof. Nestor Roqueiro, Dr. Eng.
Universidade Federal de Santa Catarina

Certificamos que esta é a **versão original e final** do trabalho de conclusão que foi julgado adequado para obtenção do título de Mestre em Engenharia de Automação e Sistemas.

Prof. Julio Elias Normey Rico, Dr. Eng
Coordenação do Programa de
Pós-Graduação

Prof. Rodolfo César Costa Flesch, Dr. Eng.
Orientador

Florianópolis, 2023.

AGRADECIMENTOS

A autora gostaria de expressar meus sinceros agradecimentos a todas as pessoas que contribuíram para a realização desta dissertação.

Ao meu orientador, Rodolfo César Costa Flesch, por sua orientação cuidadosa, apoio constante e valiosas contribuições ao longo deste processo. Sua sabedoria, conhecimento e orientação foram fundamentais para o meu crescimento acadêmico e para o sucesso deste trabalho.

Também expressar minha gratidão aos coorientadores, Ahryman Nascimento e Bernardo Schwedersky, pelo seu comprometimento e orientação adicional neste projeto. Que enriqueceram o meu trabalho, e sou grata por sua disponibilidade e apoio constante.

Agradeço sinceramente aos colegas do LIAE, Paulo Nishimoto, cuja colaboração, troca de ideias e discussões foram inestimáveis. Suas contribuições me ajudaram a expandir meus horizontes e a aprimorar minha compreensão do tema.

Estender o agradecimento à minha família, Abenil, Claudir, Gabriel e Ana Catarina, e expressando minha gratidão por todo o amor, apoio e encorajamento incondicionais que me deram ao longo de toda a minha jornada acadêmica.

Além disso, quero agradecer ao meu namorado Juarez, que esteve ao meu lado, apoiando-me nos momentos mais desafiadores e celebrando as vitórias junto comigo. Seu amor e apoio foram fundamentais para que eu pudesse concluir este trabalho com sucesso.

Agradeço a todos aqueles que, de alguma forma, contribuíram para o desenvolvimento deste trabalho e para a minha formação acadêmica. Seja através de palavras de encorajamento, recursos compartilhados, apoio moral ou outras formas de auxílio, cada um de vocês desempenhou um papel importante nessa conquista.

Por fim, a autora agradece à Universidade Federal de Santa Catarina, ao Programa de Pós-Graduação em Engenharia de Automação e Sistemas, à Petrobras e ao Programa de Recursos Humanos da Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – PRH-ANP, pela oportunidade de realizar este trabalho com recursos provenientes do investimento de empresas petrolíferas qualificadas na Cláusula de P, D&I da Resolução ANP n° 50/2015.

*“A Inteligência Artificial não é apenas
uma ferramenta para solucionar problemas humanos,
mas também uma ponte para uma maior
compreensão e preservação do nosso
precioso meio ambiente.” (Fei-Fei Li, 2021)*

RESUMO

O derramamento de hidrocarbonetos no solo apresenta riscos significativos ao meio ambiente e à saúde humana. Diante dessa situação, é importante realizar uma avaliação cuidadosa das áreas afetadas para determinar a necessidade de aplicar métodos de remediação. No entanto, os métodos tradicionais de remediação muitas vezes acarretam impactos ambientais consideráveis e altos custos. Portanto, há um grande esforço para buscar alternativas mais eficientes e sustentáveis. Uma abordagem promissora é o monitoramento da taxa de depleção dos hidrocarbonetos na área afetada, o que permite avaliar a taxa de biodegradação dos contaminantes. Estudos recentes mostram que uma maneira eficaz de avaliar essa taxa é por meio do acompanhamento do perfil de temperatura do solo, já que a biodegradação dos hidrocarbonetos resulta em um aumento na temperatura da região afetada. Nesse contexto, este estudo propõe o desenvolvimento de técnicas de predição para estimar o comportamento futuro dos perfis de temperatura do solo, a fim de estimar o processo de biodegradação em momentos posteriores. Para isso, são utilizadas técnicas de modelagem dinâmica do perfil de temperatura nas áreas afetadas, utilizando dados históricos para prever os valores futuros da série de dados. Com base em informações reais de temperatura coletadas por um sistema termodinâmico de monitoramento de fontes, foram implementados um modelo autorregressivo e modelos recentes de redes neurais recorrentes para estimar o comportamento futuro do perfil de temperatura. Os resultados obtidos a partir da comparação desses métodos indicam a viabilidade da modelagem proposta, fornecendo informações valiosas que podem auxiliar na tomada de decisões relacionadas às técnicas de remediação a serem empregadas. Ao fornecer uma predição do comportamento futuro do perfil de temperatura do solo, é possível obter uma compreensão mais clara do processo de biodegradação dos hidrocarbonetos, permitindo uma alocação mais eficiente de recursos e uma escolha mais assertiva das técnicas de remediação adequadas para cada situação. Além disso, a abordagem proposta contribui para a preservação do meio ambiente e para a proteção da saúde humana, mitigando parte dos impactos ambientais negativos associados a técnicas tradicionais de remediação de áreas contaminadas.

Palavras-chave: Predição de temperatura. Depleção da zona da fonte. Derramamentos. Fase livre.

ABSTRACT

The spillage of hydrocarbons in soil poses a serious problem that entails significant risks to the environment and human health. In light of this reality, it is important to conduct a thorough assessment of the affected areas to determine the need for remediation methods. However, traditional remediation approaches often come with considerable environmental impacts and high costs. Therefore, there are many active researches to explore more efficient and sustainable alternatives. One promising approach is monitoring the depletion rate of hydrocarbons in the affected zone, which allows for the assessment of the biodegradation rate of contaminants. Recent studies have shown that an effective way to evaluate this rate is by monitoring the soil temperature profile, as the biodegradation of hydrocarbons results in an increase in the temperature of the affected region. In this context, this study proposes the development of prediction techniques to estimate the future behavior of soil temperature profiles, aiming to predict the biodegradation process at later time instants. To achieve this, dynamic modeling techniques of the temperature profile in the affected areas are employed, using historical data to predict future values in the data series. Based on real temperature data collected by a thermodynamic source monitoring system, an autoregressive model and recent models of recurrent neural networks were implemented to predict the future behavior of the temperature profile. The results obtained from the comparison of these methods indicate the feasibility of the proposed modeling approach, providing valuable information that can assist in decision-making regarding the employed remediation techniques. By providing a prediction of the future behavior of the soil temperature profile, a clearer understanding of the hydrocarbon biodegradation process can be obtained, enabling more efficient resource allocation and informed selection of suitable remediation techniques for each situation. Moreover, this approach contributes to environmental preservation and the protection of human health by avoiding some of the limitations associated with traditional methods in terms of costs and environmental impacts.

Keywords: Temperature prediction. Source Zone Depletion. Spills. Free phase.

LISTA DE FIGURAS

Figura 1 – Hierarquia dos modelos paramétricos e não-paramétricos de uma ST	23
Figura 2 – Estrutura típica de um neurônio artificial.	29
Figura 3 – Arquitetura ESN. As conexões fixas entre os neurônios são representadas por linhas sólidas, enquanto os pesos treináveis são representados por linhas tracejadas.	32
Figura 4 – Arquitetura LSTM. Fluxo de informação de um bloco de memória LSTM no tempo t	35
Figura 5 – Representação algoritmo genético	37
Figura 6 – Representação do processo de recombinação	38
Figura 7 – Modelo conceitual de fluxo de carbono associado com NSZD em locais com a presença de LNAPL: (1) dinâmica de distribuição do LNAPL; (2) processo de depleção natural das fontes de contaminação.	41
Figura 8 – Influência da variação da temperatura atmosférica na temperatura do solo.	42
Figura 9 – Esquemático de instalação para monitoramento de temperatura. . .	44
Figura 10 – Sistema responsável pela coleta e envio dos dados adquiridos pelo STMF.	45
Figura 11 – Processo de predição de uma série temporal	46
Figura 12 – Esquemático de instalação para monitoramento de solo com o STMF.	48
Figura 13 – Séries temporais resultante das medições	49
Figura 14 – Perfis de temperatura empregados para realizar o modelamento . .	50
Figura 15 – Resultados para modelos ARIMA avaliados na tarefa da predição a temperatura em três profundidades distintas.	57
Figura 16 – Resultados obtidos com o modelo LSTM avaliado no conjunto de testes para a tarefa de predição de múltiplos instantes à frente. . . .	61
Figura 17 – Resultados obtidos com o modelo ESN avaliado no conjunto de testes para a tarefa de predição de múltiplos instantes à frente	65
Figura 18 – Resultados dos diferentes modelos avaliados no conjunto de testes para a tarefa de predição de múltiplos instantes à frente na profundidade de 17 cm	66
Figura 19 – Resultados dos diferentes modelos avaliados no conjunto de testes para a tarefa de predição de múltiplos instantes à frente na profundidade de 96 cm	67
Figura 20 – Resultados dos diferentes modelos avaliados no conjunto de testes para a tarefa de predição de múltiplos instantes à frente na profundidade de 117 cm	68

LISTA DE TABELAS

Tabela 1 – Configuração da ferramenta de otimização baseada em AG para o modelo ARIMA	55
Tabela 2 – Configurações ótimas para o modelo ARIMA	56
Tabela 3 – Métricas de desempenho para o modelo ARIMA	57
Tabela 4 – Parâmetros considerados para o AG para o modelo LSTM	59
Tabela 5 – Métricas de desempenho para o modelo LSTM	61
Tabela 6 – Parâmetros considerados para o AG para o modelo ESN	63
Tabela 7 – Valores dos parâmetros para o modelo ESN	64
Tabela 8 – Métricas de desempenho para o modelo ESN	65
Tabela 9 – Métricas de desempenho para todos os modelos para profundidade de 17 cm	67
Tabela 10 – Métricas de desempenho para todos os modelos para profundidade de 96 cm	68
Tabela 11 – Métricas de desempenho para todos os modelos para profundidade de 117 cm	69

LISTA DE ABREVIATURAS E SIGLAS

ACF	Função de autocorrelação – do inglês <i>Autocorrelation Function</i>
ADAM	Estimativa de Momento Adaptativo – do inglês <i>Adaptive Moment Estimation</i>
AG	Algoritmo Genético
ARIMA	Modelo Auto-Regressivo Integrado de Médias Móveis – do inglês <i>Autoregressive Integrated Moving Average</i>
ESN	Rede de estado de eco – do inglês <i>Echo State Network</i>
LNAPL	Líquido imiscível menos denso que a água – do inglês <i>Liquid Non-Aqueous Phase Liquids</i>
LSTM	Memória longa de curto prazo – do inglês <i>Long Short-Term Memory</i>
MAE	Erro médio absoluto – do inglês <i>Mean Absolute Error</i>
MAPE	Erro médio percentual absoluto – do inglês <i>Mean Absolute Percentage Error</i>
MSE	Erro quadrático médio - do inglês <i>Mean Squared Error</i>
NSZD	Depleção natural da zona da fonte – do inglês <i>Natural Source Zone Depletion</i>
PACF	Função de autocorrelação parcial – do inglês <i>Partial Autocorrelation Function</i>
QTA	Queda da Taxa de Aprendizado
RC	Computação de reservatório – do inglês <i>Reservoir Computing</i>
RMSE	Raiz do erro quadrático médio – do inglês <i>Root Mean Squared Error</i>
RNA	Rede Neural Artificial
RNN	Rede neural recorrente – do inglês <i>recurrent neural networks</i>
SSM	modelo de espaço de estado - do inglês <i>model in state space</i>
ST	Série temporal
STMF	Sistema Termodinâmico de Monitoramento de Fontes
TA	Taxa de Aprendizado
UFSC	Universidade Federal de Santa Catarina
WMSE	Erro quadrático médio ponderado – do inglês <i>Weighted Mean Squared Error</i>

LISTA DE SÍMBOLOS

x	Variável de estacionalidade
c	Constante
p	Ordem modelo autoregressivo
Φ	Coefficientes de autocorrelação na defasagem
ε	Resíduo
μ	Expectativa das variáveis de estacionalidade
q	Ordem do modelo de média móvel
θ	Pesos dos termos aplicados
k	Ativação de neurônio
W	Pesos
α	Taxa de vazamento
u	Entrada
y	Saída predita
$f(\cdot)$	Função de ativação das unidades internas

SUMÁRIO

1	INTRODUÇÃO	14
1.1	OBJETIVOS	17
1.1.1	Objetivo geral	17
1.1.2	Objetivos específicos	18
1.2	ESTRUTURA DO DOCUMENTO	18
2	REVISÃO DA LITERATURA	20
2.1	SÉRIES TEMPORAIS	20
2.2	MODELAGEM DINÂMICA PARA PREDIÇÃO EM SÉRIES TEMPORAIS	22
2.3	MODELO AUTORREGRESSIVO INTEGRADO DE MÉDIAS MÓVEIS	25
2.4	MODELOS DE APRENDIZADO DE MÁQUINA	27
2.4.1	Visão geral sobre redes neurais	28
2.4.2	Computação de reservatório	30
2.4.3	Redes neurais LSTM	34
2.5	ALGORITMOS GENÉTICOS	36
2.6	COMENTÁRIOS FINAIS	39
3	MÉTODO E ESTUDO DE CASO	40
3.1	DEPLEÇÃO NATURAL NA ZONA DA FONTE E SUA RELAÇÃO COM A TEMPERATURA DO SUBSOLO	40
3.2	ESTUDO DE CASO: PREDIÇÃO DE TEMPERATURA DO STMF	43
3.2.1	Sistema de aquisição de dados	43
3.2.2	Método para predição de séries temporais de temperaturas de subsolo	45
3.2.3	Séries de dados utilizadas	47
3.2.4	Avaliação da qualidade da predição	49
3.2.4.1	Erro médio percentual absoluto	51
3.2.4.2	Erro médio absoluto	51
3.2.4.3	Erro quadrático médio	51
3.2.4.4	Raiz do erro quadrático médio	51
3.2.4.5	Erro médio quadrático ponderado	52
3.2.4.6	Coeficiente de determinação	52
3.3	CONSIDERAÇÕES FINAIS	52
4	RESULTADOS DA PREDIÇÃO DE TEMPERATURAS DE DEPLEÇÃO DE HIDROCARBONETOS	54
4.1	PREDIÇÃO COM MODELO ARIMA	54
4.1.1	Otimização do modelo ARIMA com algoritmo genético	54
4.1.2	Resultados para o modelo ARIMA	56
4.2	PREDIÇÃO COM A REDE NEURAL LSTM	58

4.2.1	Otimização da rede LSTM com Algoritmo Genético	58
4.2.2	Resultados para o modelo LSTM	60
4.3	PREDIÇÃO COM REDE ESN	62
4.3.1	Otimização da rede neural ESN com Algoritmo Genético	62
4.3.2	Resultados para rede ESN	64
4.4	COMPARAÇÃO DOS RESULTADOS	66
5	CONSIDERAÇÕES FINAIS	70
5.1	CONCLUSÕES	70
5.2	SUGESTÕES PARA TRABALHOS FUTUROS	71
	REFERÊNCIAS	73

1 INTRODUÇÃO

A indústria do petróleo e gás ainda tem grande participação na matriz energética brasileira, apesar de haver uma tendência de aumento da participação de energias renováveis. O plano nacional de energia para 2050 (EMPRESA DE PESQUISA ENERGÉTICA, 2018) mostra a participação dos derivados de petróleo como principal fonte de energia para o consumo final, com previsão de ser dominante até o final do plano. Atualmente no Brasil, o setor de petróleo e gás é responsável por 34,9% da oferta interna de energia, sendo a maior fonte de contribuição (EMPRESA DE PESQUISA ENERGÉTICA, 2021).

Com as previsões de exploração do setor petrolífero sendo promissoras, a comunidade científica e a opinião pública têm se preocupado cada vez mais com a investigação e a pesquisa para melhoria do processo e para amenizar o impacto ambiental dessa atividade. Diante disso, observa-se que as atividades que estão ligadas à indústria petrolífera, como exploração, refino, armazenamento e transporte, estão sujeitas à ocorrência de derramamento de hidrocarbonetos no meio ambiente (PERES, 2019). Os efeitos adversos desses derramamentos sobre a biodiversidade e os ecossistemas em geral podem persistir por muitas décadas, até mesmo séculos (GAUTAM *et al.*, 2023). Além disso, a exposição prolongada para seres humanos a esses compostos pode resultar em problemas respiratórios, neurológicos e até câncer (ADIPAH, 2019). A limpeza de um derramamento de petróleo para a recuperação de áreas afetadas pode ser extremamente cara e frequentemente é suportada pelo governo ou pela indústria petrolífera (AL-MAJED; ADEBAYO; HOSSAIN, 2012). Além disso, muitas vezes processos tradicionais de remediação apresentam eficiência limitada, em especial para aplicação em contaminantes residuais (STOCKWELL, 2015).

Depois do derramamento, o petróleo ou derivado infiltra-se no solo e espalha-se lateralmente devido às características geológicas locais, influenciado pela gravidade. Ele permanece sorvido no solo, nas águas subterrâneas e no lençol freático como uma solução não homogênea, resultando em duas fases distintas. A fase do contaminante é conhecida como fase livre ou líquido imiscível menos denso que a água (LNAPL, do inglês *Liquid Non-Aqueous Phase Liquids*) (STOCKWELL, 2015). O local que apresenta LNAPL é denominado zona da fonte e ele passa por um processo de depleção natural da zona da fonte (NSZD, do inglês *Natural Source Zone Depletion*) (PERES, 2019). Durante o processo de depleção, ocorre a transferência física de massa através da dissolução e vaporização dos constituintes químicos para as fases aquosa (água subterrânea) e gasosa (gás do solo), além da biodegradação dos constituintes do LNAPL por meio de reações microbianas (SMITH *et al.*, 2021).

Existe uma vasta gama de tecnologias para lidar com a descontaminação do solo e da água subterrânea. A seleção da abordagem mais apropriada exige uma

compreensão detalhada das propriedades físico-químicas do contaminante, do volume do vazamento, do tempo decorrido desde o incidente, das características geológicas e hidrogeológicas, da análise do ambiente físico na superfície e abaixo do solo no local específico, bem como da extensão da pluma de contaminação. Uma revisão das principais técnicas de remediação e de seus usos de acordo com os parâmetros listados acima é apresentada em Mariano (2006).

No Brasil, ainda predominam as técnicas de remoção do solo contaminado e de bombeamento de contaminantes. Entretanto, em alguns casos essas técnicas podem não ser eficientes ou podem até mesmo ser desnecessárias. Por exemplo, a remoção física do solo pode não retirar totalmente a pluma do contaminante e o LNAPL continua agindo como fonte de contaminação (ASKARANI; SALE, 2020). Além disso, o bombeamento tem eficiência limitada quando a contaminação residual tem baixa concentração (PERES, 2019).

De maneira geral, NSZD também é considerada uma estratégia de remediação, baseada em processos biológicos. Essa estratégia pode ser empregada com eficiência significativamente maior na remoção da contaminação residual persistente, chegando a ter efeito melhor que técnicas que envolvem processos de extração hidráulicos, pneumáticos ou por injeção de remediadores químicos nesse caso. Além disso, as técnicas que são mais invasivas apresentam custos elevados. Apesar de ser promissor em algumas situações, o uso de NSZD como método de remediação depende de uma quantificação da taxa efetiva depleção de massa de contaminantes, de modo a garantir a eficiência do método (PERES, 2019).

Os métodos tradicionais de quantificar a estimativa do processo são baseados em análise de emissões de gases e são influenciáveis pelas condições ambientais (terrestres e subterrânea). Adicionalmente, os métodos exigem algumas análises de amostras em laboratórios, o que torna o processo dispendioso (VAN DE VEN *et al.*, 2021). Pesquisas mais recentes exploram um novo método de avaliação, que emprega a medição do perfil vertical da temperatura na zona da fonte. O calor liberado pela biodegradação é diretamente relacionado com a taxa de NSZD por meio de relações termodinâmicas, sendo um método mais fácil de medição e que não sofre tanto com alteração por fatores ambientais (SWEENEY; RIRIE, 2014). Assim, uma vez que se tenha uma relação confiável entre o perfil de temperatura e a taxa de biodegradação, uma instrumentação fixa na região contaminada permite obter uma estimativa de forma praticamente on-line da taxa de NSZD. Essa estimativa pode ser empregada para acompanhamento da eficiência do método na depleção de hidrocarbonetos de petróleo e avaliação da adequação da técnica como método de remediação.

Apesar de o modelo ser parte importante para a remediação por NSZD, não há uma abordagem amplamente empregada na literatura. Além disso, a obtenção dos modelos não é algo trivial, pois eles envolvem diferentes domínios de conhecimento

e apresentam características não lineares. Para que se obtenha uma estimativa confiável da taxa de NSZD a partir das medições de temperatura, a escolha cuidadosa de parâmetros e técnicas de modelagem é necessária, o que pode ser um processo iterativo e exigente em termos computacionais (TIAN; CHEN; GANGULI, 2021). Além do emprego de técnicas tradicionais de modelagem, podem ser empregadas técnicas baseadas em dados e em modelos de aprendizado de máquina. Essas têm se mostrado ser um campo promissor por permitirem a solução de vários problemas de modelagem para sistemas com características bastante variadas (SCHWEDERSKY, 2017). Por se tratar de um problema dinâmico, as técnicas da área de aprendizado profundo, que englobam arquiteturas de redes neurais artificiais (RNA) recorrentes, são as escolhas naturais quando se opta pela abordagem de identificação com base em dados e modelos de aprendizado. A grande vantagem dessas técnicas frente à modelagem fenomenológica é que exigem pouco conhecimento da física que rege o processo (BAROT; BURGSTEINER; KOLLERITSCH, 2020).

Nos últimos anos, os modelos de aprendizado de máquina têm ganhado popularidade devido às suas estruturas específicas e seu desempenho notável em várias aplicações. Destacam-se dois modelos em particular: as redes convolucionais e as redes recorrentes, cada uma com características distintas que impulsionaram seu sucesso em tarefas específicas. As redes convolucionais são amplamente utilizadas em tarefas de visão computacional. Elas se destacam por sua capacidade de extrair recursos relevantes de imagens, permitindo a detecção de padrões complexos (SCHWEDERSKY, 2017; MENDES *et al.*, 2020). Outro modelo de destaque é a rede recorrente conhecida como memória longa de curto prazo (LSTM, do inglês *Long Short-Term Memory*). A LSTM tem a capacidade de aprender a partir de sequências de dados, capturando dependências de longo prazo e superando as limitações das redes neurais tradicionais. Isso permite a detecção de padrões temporais complexos e o processamento eficiente de informações sequenciais (YAMAK; YUJIAN; GADOSEY, 2019). Além desses modelos, outras arquiteturas recorrentes têm recebido crescente interesse. Um exemplo são as redes de estado de eco (ESN, do inglês *Echo State Network*), que têm sido aplicadas com sucesso na modelagem de problemas dinâmicos. As ESNs se destacam por sua capacidade de processar sequências temporais complexas, aprendendo a partir de estados internos e adaptando-se de forma eficiente a diferentes contextos dinâmicos (JAEGER, 2002).

Como forma de avaliar experimentalmente a eficiência do método de remediação por NSZD e de estudar diferentes formas de modelagem do problema, foram desenvolvidas estações de monitoramento da taxa de biodegradação baseadas na medição do perfil de temperatura no contexto de um projeto de pesquisa na Universidade Federal de Santa Catarina (UFSC). Essas estações, junto com o modelo de estimativa, são chamadas de Sistema Termodinâmico de Monitoramento de Fontes (STMF) e

estão sendo empregadas para o monitoramento remoto de áreas contaminadas *in situ*, possibilitando um acompanhamento contínuo da temperatura resultante da depleção das fontes de contaminação. As estações foram, inicialmente, avaliadas em ambiente de laboratório e, posteriormente, foram para campo avaliar áreas com contaminação controlada dentro do Núcleo Ressacada de Pesquisas em Meio Ambiente da UFSC. Atualmente, há um STMF em operação em área de contaminação na Refinaria de Paulínia (Replan), da Petrobras.

Com base nos dados reais medidos em áreas contaminadas em campo, propõe-se a criação de modelos dinâmicos de predição das temperaturas subsuperficiais, como forma de avaliar um método que permita estimar as temperaturas em momentos futuros. Os dados estimados de temperatura podem ser usados para avaliação das taxas esperadas de NSZD em momentos futuros, porém esse processo exige modelos que sejam capazes de entender as relações de longo prazo e as sazonalidades associadas aos dados. Dessa forma, não está no escopo deste trabalho criar um modelo que faça predições confiáveis da taxa de NSZD, mas propor e avaliar um método de estimativa das temperaturas em instantes futuros de tempo, considerando horizontes curtos, dada a disponibilidade limitada de dados obtidos durante a realização desta dissertação. Para essa tarefa, foram empregados modelos tradicionais, como modelos de média móvel integrada (ARIMA, do inglês *Autoregressive Integrated Moving Average*) (SHUMWAY *et al.*, 2017), e redes neurais recorrentes, em particular as arquiteturas LSTM e ESN. Para os modelos de predição é necessário realizar os ajustes dos hiperparâmetros, os quais precisam ser obtidos para cada série temporal. Como método de otimização dessa etapa, foi utilizada a técnica de algoritmo genético (AG) (DRÉO *et al.*, 2006b) para encontrar as combinações ótimas de hiperparâmetros para cada modelo. Essa abordagem é bastante utilizada na literatura para esse fim (NIKBAKHT; ANITESCU; RABCZUK, 2021; SILVA; LUDERMIR; ALMEIDA, 2021).

1.1 OBJETIVOS

Nas próximas seções estão descritos o objetivo geral e os objetivos específicos.

1.1.1 Objetivo geral

O objetivo geral deste estudo é realizar uma avaliação experimental da utilização de diversas arquiteturas e modelos de redes neurais artificiais recorrentes, assim como modelos tradicionais, para prever valores de temperatura em alguns pontos do perfil de temperatura do solo resultante do processo de depleção natural de hidrocarbonetos de petróleo com base em medições de estações de monitoramento.

1.1.2 Objetivos específicos

O objetivo geral foi dividido em objetivos específicos, consistindo em:

- analisar e pré-processar os dados de campo para identificar quais dinâmicas devem ser modeladas e quais dados podem ser considerados como entrada dos modelos de predição;
- criar modelos de predição dos dados reais de temperatura utilizando um modelo autorregressivo e arquiteturas de redes neurais artificiais recorrentes;
- propor um método de ajuste dos hiperparâmetros dos modelos com base em algoritmos genéticos;
- comparar os resultados das predições dos modelos desenvolvidos em perfis de temperatura com dinâmicas distintas.

1.2 ESTRUTURA DO DOCUMENTO

No restante desta dissertação, no Capítulo 2 é apresentada uma revisão da literatura que aborda séries temporais, modelagem de sistemas dinâmicos e algoritmos genéticos. O capítulo é dividido em seis seções, cada uma explorando um aspecto específico. A Seção 2.1 aborda os conceitos fundamentais das séries temporais, fornecendo uma base teórica para a compreensão dessas estruturas de dados. A Seção 2.2 concentra-se nas técnicas de modelagem utilizadas para análise e predição de comportamento futuro de séries temporais. Além disso, aponta os desafios e complexidades envolvidos no ajuste de modelos para capturar adequadamente a estrutura e a dinâmica das séries temporais. Nas Seções 2.3 a 2.4.1, são apresentadas as arquiteturas específicas utilizadas para modelar as séries temporais neste trabalho. Isso inclui o modelo ARIMA, que é amplamente utilizado na literatura, e as redes neurais recorrentes, como LSTM e ESN. As vantagens e limitações de cada abordagem são discutidas em relação à análise de séries temporais. A Seção 2.5 destaca os conceitos e técnicas relacionados à implementação de algoritmos genéticos, empregados neste trabalho para ajuste dos parâmetros dos modelos. Por fim, na Seção 2.6, são apresentadas considerações finais sobre os métodos discutidos ao longo do capítulo, destacando suas contribuições e relevância para a pesquisa em questão.

O Capítulo 3 apresenta a descrição detalhada do processo dos dados e metodologias de métricas. O capítulo está dividido em três seções. Na Seção 3.1, são discutidas a dinâmica do LNAPL no solo após derramamentos de petróleo e a aplicação da NSZD como método de remediação, com monitoramento baseado na temperatura do solo. Na Seção 3.2, é apresentado o STMF e ela inclui informações sobre o sistema de aquisição de dados empregado, o método utilizado para a predição das séries temporais de temperaturas de subsolo e uma descrição das séries de dados empregadas.

Além disso, são apresentadas métricas de avaliação da qualidade das predições: erro médio percentual absoluto, erro médio absoluto, erro quadrático médio, raiz do erro quadrático médio, erro médio quadrático ponderado e coeficiente de determinação. Por fim, a Seção 3.3 apresenta as considerações finais sobre a metodologia de desenvolvimento do trabalho e sobre suas conexões com o restante dos desenvolvimentos da pesquisa.

O Capítulo 4 apresenta os resultados dos modelos utilizados e suas principais características de cada uma. A Seção 4.1 aborda o modelo ARIMA, com foco na configuração do modelo e sua validação. Além disso, são apresentados os resultados do treinamento do modelo, destacando as métricas de desempenho obtidas. A Seção 4.2 apresenta o modelo LSTM aplicado para a análise dos dados dos reservatórios. São descritos os detalhes da implementação desse modelo específico, incluindo as camadas, as funções de ativação e os parâmetros utilizados. Além disso, são discutidos os resultados. A Seção 4.3 faz a mesma análise para o modelo ESN. Por fim, a Seção 4.4 traz uma comparação dos resultados dos diferentes modelos em cenários de predição para temperaturas de solo medidas em diferentes profundidades.

O Capítulo 5 conclui esta dissertação, fornecendo as considerações finais do trabalho. Nesse capítulo, são apresentadas as principais conclusões alcançadas por meio do desenvolvimento do estudo, destacando os principais *insights* e descobertas obtidas. Além das conclusões, também são propostas sugestões para trabalhos futuros relacionados ao tema abordado. Essas propostas visam aprofundar a pesquisa, explorar lacunas identificadas durante o estudo e contribuir para o avanço do conhecimento na área.

2 REVISÃO DA LITERATURA

A revisão da literatura é dividida em seções, cada uma explorando um aspecto específico de séries temporais e modelos para sua predição. A Seção 2.1 aborda os conceitos fundamentais das séries temporais, fornecendo uma base teórica para a compreensão dessas estruturas de dados e de suas principais características, como dinamicidade, tendências, padrões, sazonalidade e autocorrelação. A Seção 2.2 apresenta as principais técnicas de modelagem utilizadas para análise e predição de comportamento de séries temporais, bem com uma discussão sobre a complexidade de ajuste de modelos para capturar a estrutura e a dinâmica das séries temporais. A Seção 2.3 apresenta o modelo ARIMA e é seguida na seção 2.4 por uma descrição dos modelos de aprendizado de máquina, com destaque para os modelos LSTM e ESN, detalhados nas Seções 2.4.2 e 2.4.3, respectivamente. A Seção 2.5 aborda a implementação de algoritmos genéticos empregada para ajuste dos hiperparâmetros dos modelos. Na Seção 2.6, são apresentadas considerações finais sobre o capítulo.

2.1 SÉRIES TEMPORAIS

Os avanços tecnológicos na computação incentivaram a utilização de sistemas de dados para aquisição e gerenciamento em diversas áreas do conhecimento, permitindo armazenar grandes quantidades de dados em diferentes formatos, incluindo a informação temporal para organização cronológica (PARMEZAN, 2016). Graças ao componente temporal inerente à maioria dos dados do mundo real, é relativamente fácil gerar séries temporais. No entanto, o mesmo não pode ser dito sobre sua análise, classificação e predição, processos que ainda apresentam alguns desafios (BROCKWELL; DAVIS, 2002). A análise mais simples de uma série temporal (ST), tipicamente é realizada de forma gráfica, que permite avaliar o comportamento geral, procurar periodicidade e verificar tendências nos dados, o que abre possibilidade para se fazer predição de valores futuros. Além da análise gráfica, as STs permitem o uso de técnicas que modelam matematicamente o comportamento futuro com base em dados históricos sequenciais, incluindo seus próprios dados passados. Ao encontrar esse modelo matemático, é possível realizar predições dentro do contexto temporal (CHATFIELD; XING, 2019).

Na literatura, é comum utilizar o conceito de decomposição de uma ST em elementos independentes dos eventos representados para análise e compreensão aprimoradas. Os principais componentes abordados nessa decomposição são denominados tendência, sazonalidade e resíduo (GONÇALVES, 2017). Essa abordagem visa identificar e separar esses elementos para obter uma visão mais clara e interpretativa dos dados temporais.

A tendência refere-se à direção geral ou padrão de longo prazo dos dados ao

longo do tempo. Em outras palavras, a tendência representa a mudança sistemática e contínua no valor médio da série temporal ao longo de um período extenso. Caso a ST tenha uma tendência, ela pode ser classificada em ascendente (crescente), decrescente (decrecente) ou horizontal (constante) (PARMEZAN, 2016). A identificação e análise da tendência em uma série temporal podem fornecer informações valiosas sobre o comportamento subjacente dos dados e ajudar a prever padrões futuros.

Sazonalidade em STs refere-se a padrões repetitivos ou ciclos que ocorrem regularmente em determinados períodos de tempo. Esses padrões podem ser influenciados por fatores sazonais, como estações do ano, feriados e eventos culturais, ou flutuações regulares em atividades econômicas, por exemplo. A presença de sazonalidade em uma série temporal pode ter um impacto significativo na interpretação dos dados e na previsão de tendências futuras. A sazonalidade pode ser categorizada em dois tipos: aditiva e multiplicativa. Na sazonalidade aditiva, a flutuação sazonal é estável e não depende do nível global da série, sendo adicionada a tendência de forma consistente. Na sazonalidade multiplicativa, o tamanho da flutuação sazonal varia de acordo com o nível global da série, sendo multiplicada pela tendência (MUEEN *et al.*, 2009). O entendimento do tipo de sazonalidade presente em uma série temporal é importante para a modelagem e a predição adequadas.

O resíduo refere-se à parte não explicada ou imprevisível dos dados observados. É a diferença entre os valores observados e os valores estimados pela tendência e pela sazonalidade. Os resíduos representam as flutuações não sistemáticas presentes na série, que podem ser causadas por diversos fatores, como variações aleatórias, eventos imprevistos ou erros de medição e de modelo (PARMEZAN, 2016). O estudo dos resíduos é importante na análise de séries temporais, pois permite verificar se o modelo de decomposição utilizado é apropriado e se captura adequadamente os padrões sistemáticos da série. Resíduos aleatórios indicam uma boa captura da tendência e da sazonalidade, enquanto resíduos com padrões ou estruturas sugerem a presença de informações não consideradas no modelo (CHOU DHURY; MIDDYA; ROY, 2022). Além disso, a análise dos resíduos permite verificar a validade de pressupostos estatísticos, como a normalidade dos erros, a independência entre os resíduos e a homogeneidade da variância. Caso esses pressupostos sejam violados, é necessário revisar o modelo ou aplicar técnicas de correção.

A análise de séries temporais utiliza diferentes técnicas para descrever os componentes de uma série, permitir uma melhor utilização das informações disponíveis e atingir diferentes objetivos. Os quatro principais objetivos da análise de séries temporais são a descrição, explicação, predição e controle (BOURGUIGNON; SANTOS-NETO; CASTRO, 2021). A análise descritiva busca descrever os comportamentos da série, identificando tendências, variações sazonais, valores discrepantes e alterações estruturais. Para isso, são utilizadas ferramentas como peridiogramas, histogramas e

diagramas de dispersão. A análise de explicação envolve a identificação de relações entre duas ou mais séries temporais. É utilizada para compreender as variações em uma série com base em outra série relacionada. A predição tem como objetivo resumir as propriedades da série e caracterizar seu comportamento. Busca-se identificar ou sugerir um modelo que permita prever os valores futuros da série com base nos valores passados (CARLOS, 2022). A análise de controle utiliza os valores da série como dados de controle de um processo específico. Seu objetivo é medir a qualidade desse processo por meio da análise da série temporal (CHATFIELD; XING, 2019).

2.2 MODELAGEM DINÂMICA PARA PREDIÇÃO EM SÉRIES TEMPORAIS

A situação ideal para fazer previsões de um fenômeno que gera séries temporais seria ter conhecimento das equações regem o mesmo. No entanto, em muitos casos reais, tal conhecimento não está disponível e, portanto, não é possível saber as equações que definem o comportamento das variáveis de interesse de forma precisa. Quando isso ocorre, é comum utilizar uma abordagem baseada em modelos, na qual se tenta identificar o comportamento do fenômeno por meio de aproximação (GOMES, 2005).

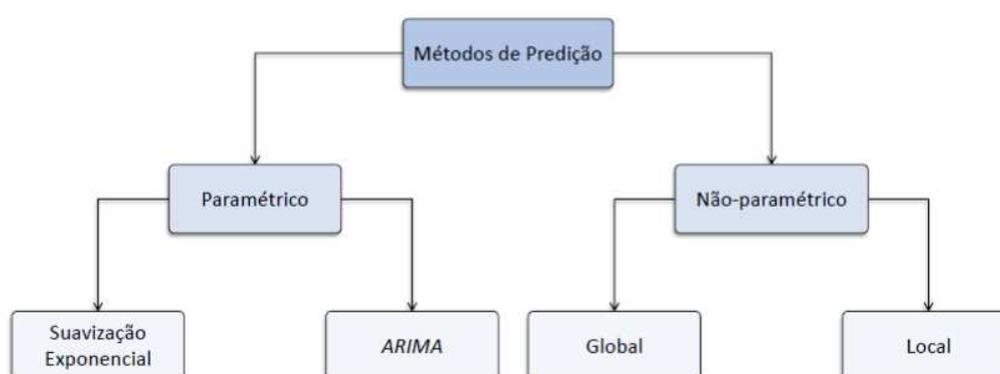
Ao avaliar os dados de STs, tipicamente é necessário assumir duas hipóteses na formulação dos modelos: a normalidade das informações em questão ou dos termos estocásticos relacionados aos modelos e a estacionariedade das séries temporais de interesse envolvidas (SALLES; LONDE, 2020). Embora tais hipóteses sejam consideradas desejáveis em muitos casos, existem métodos e técnicas disponíveis para contornar essas suposições. É importante avaliar a adequação dessas hipóteses e técnicas às características dos dados específicos em questão.

A utilização de modelos pode ser muito útil na análise de sistemas, pois permite uma melhor compreensão dos fenômenos envolvidos. No entanto, a qualidade do modelo geralmente determina um limite superior na qualidade da solução final do problema, tornando a modelagem um gargalo comum no desenvolvimento de sistemas mais complexos. Como resultado, há uma forte demanda por esquemas avançados de modelagem e identificação (NELLES, 2020).

Na área de identificação de sistemas, existem dois principais métodos utilizados: métodos paramétricos e métodos não paramétricos. Esses métodos são empregados para obter um modelo matemático que descreva o comportamento de um sistema dinâmico a partir de dados observados (LEVINE, 2018). Uma representação de tal divisão pode ser vista na Figura 1. Os métodos paramétricos assumem uma estrutura de modelo pré-definida com parâmetros desconhecidos a serem estimados. Essa abordagem busca encontrar os valores ótimos dos parâmetros que melhor se ajustem aos dados observados (BROCKWELL; DAVIS, 2002). Um exemplo de método paramétrico bastante popular de modelagem dinâmica para predição de STs é o modelo ARIMA,

um modelo estatístico que leva em consideração a autocorrelação e a sazonalidade dos dados para fazer previsões (BOX *et al.*, 2015). Por outro lado, os métodos não paramétricos não fazem suposições sobre a forma funcional do sistema e não se restringem a uma estrutura de modelo específica. Eles tentam inferir o comportamento do sistema diretamente dos dados observados, sem a necessidade de especificar um conjunto de parâmetros. Esses métodos são mais flexíveis, mas podem exigir uma maior quantidade de dados para obter resultados confiáveis (HASTIE *et al.*, 2009). Um exemplo desse último método é o modelo do tipo RNA.

Figura 1 – Hierarquia dos modelos paramétricos e não-paramétricos de uma ST



Fonte: Parmezan (2016)

Além dos exemplos apresentados anteriormente, outra abordagem comumente utilizada é a baseada em modelos de espaço de estado, como o modelo de espaço de estado linear (SSM, do inglês *State-Space Model*) ou o filtro de Kalman (SCHWEDERSKY, 2017). Eles dividem a série temporal em componentes latentes não observados, como tendência, sazonalidade e ruído, e estimam esses componentes usando informações históricas e observações presentes.

Existem, ainda, técnicas mais avançadas de modelamento dinâmico, como redes neurais recorrentes (RNNs, do inglês *Recurrent Neural Networks*) e modelos de aprendizado de máquina baseados em árvores, que são capazes de aprender relações complexas e não lineares nos dados. Esses modelos são especialmente úteis quando a série temporal possui padrões não lineares ou dependências de longo prazo (LIPTON; BERKOWITZ; ELKAN, 2015).

Uma vantagem do modelamento dinâmico para previsão de séries temporais é a capacidade de capturar e se adaptar a mudanças no comportamento dos dados ao longo do tempo. Isso é especialmente relevante em séries temporais que apresentam tendências, sazonalidade e outras variações não lineares. Além disso, o modelamento dinâmico permite a incorporação de informações externas, como variáveis exógenas, que podem melhorar a precisão das previsões. No entanto, o modelamento dinâmico

também apresenta desafios (PARMEZAN, 2016). A escolha adequada do modelo e dos parâmetros requer experiência e conhecimento prévio dos dados. A estimação dos parâmetros e a validação do modelo podem ser complexas, especialmente quando se lida com séries temporais de grande escala.

A complexidade de ajuste dos parâmetros dos modelos pode variar dependendo do tipo de modelo e da técnica de estimação utilizados. Alguns modelos possuem parâmetros que podem ser ajustados de forma direta e relativamente simples, enquanto outros exigem técnicas mais avançadas e computacionalmente intensivas para encontrar os valores ótimos dos parâmetros. Dentre os fatores de complexidade, podem ser citados o número de parâmetros, o tipo da função objetivo, o método de estimação, o tamanho do conjunto de dados e as restrições e condições do modelo. Quanto maior o número de parâmetros do modelo, mais complexa é sua estimação, aumentando o espaço de busca e exigindo mais iterações para encontrar os valores ótimos. É importante considerar esses fatores ao escolher e aplicar um modelo, bem como ao planejar a estimação dos parâmetros (NELLES, 2020).

A otimização dos parâmetros do modelo é uma tarefa importante, muitas vezes ilustrada pelo dilema da relação entre bias e variância. Esse dilema se baseia na decomposição do erro de modelagem em dois termos distintos: o erro de bias e o erro de variância. O erro de bias (polarização ou viés) é uma medida do viés sistemático do modelo, ou seja, a diferença entre a média das predições do modelo e o valor verdadeiro da variável de interesse. Um modelo com alto viés tende a simplificar demais a relação entre as variáveis, não capturando toda a complexidade dos dados. Por outro lado, um modelo com baixo viés é capaz de representar de forma mais precisa o fenômeno em estudo. Já o erro de variância, por sua vez, mede a sensibilidade do modelo a variações nos dados de treinamento. Um modelo com alta variância é altamente sensível às flutuações nos dados de treinamento, levando a uma grande dispersão nas predições. Isso pode ocorrer quando o modelo é excessivamente complexo e se ajusta muito bem aos dados de treinamento específicos, mas não generaliza bem para novos dados (NELLES, 2002; SCHWEDERSKY, 2017).

O dilema entre bias e variância surge porque a redução do erro de bias geralmente aumenta o erro de variância e vice-versa. Modelos mais complexos tendem a ter menor viés, pois são capazes de capturar mais detalhes e nuances dos dados. No entanto, essa complexidade adicional pode levar a uma maior variância, tornando o modelo sensível a variações nos dados de treinamento e reduzindo sua capacidade de generalização. É necessário ajustar a complexidade do modelo de acordo com a disponibilidade e qualidade dos dados, evitando tanto modelos excessivamente simples, que não conseguem capturar a complexidade do fenômeno, quanto modelos excessivamente complexos, que se ajustam demasiadamente aos dados de treinamento, perdendo capacidade de generalização (MEDEIROS, 2019).

A otimização do modelo envolve a seleção criteriosa de técnicas de modelagem, a identificação adequada de parâmetros e a utilização de técnicas de validação para avaliar o desempenho do modelo em conjuntos de dados não utilizados no treinamento. Isso permite encontrar um ponto ideal no *trade-off* entre bias e variância, resultando em um modelo que seja capaz de fazer previsões precisas e generalizáveis (MEDEIROS, 2019; SCHWEDERSKY, 2017).

Métodos meta-heurísticos podem ser poderosas ferramentas no processo de otimização, em especial quando há muitos parâmetros para serem otimizados e o problema a ser tratado é não linear. O emprego destas abordagens como alternativa de ajuste dos hiperparâmetros é bastante comum na literatura, como é o caso do algoritmo genético (AG). Esse tipo de abordagem é interessante por permitir o ajuste de todos os parâmetros ao mesmo tempo (GOLDBERG; KORB; DEB, 1989).

2.3 MODELO AUTORREGRESSIVO INTEGRADO DE MÉDIAS MÓVEIS

O modelo ARIMA é um modelo estocástico linear, uma técnica amplamente utilizada na análise e previsão de séries temporais. O mesmo pode ser utilizado para diversas atividades como planejamento, controle de sistemas e otimização de processos. O modelo permite que valores futuros de uma série sejam preditos tomando por base seus valores presentes e passados (BOX *et al.*, 2015).

Esse modelo é adequado para capturar a estrutura temporal das séries que exibem tendências, padrões sazonais (com alguma modificação no modelo) e comportamentos autorregressivos. Com a capacidade de identificar correlações sequenciais, esse modelo é construído a partir de uma sequência de impactos aleatórios independentes entre si, seguindo uma distribuição específica com média zero e variância constantes (LÚCIO *et al.*, 2010). Esses modelos são resultantes da combinação de três componentes, chamados de filtros: o componente auto-regressivo (AR, do inglês *Autoregressive*), o filtro de integração (I, do inglês *Integrated*) e o componente de médias móveis (MA, do inglês *Moving Average*) (BOX *et al.*, 2015).

O termo AR refere-se à dependência dos valores passados da série. Ele considera que o valor atual da série é linearmente relacionado com os seus próprios valores anteriores. A ordem do componente AR, denotada por p , indica quantos valores anteriores são considerados na modelagem (BOURGUIGNON; SANTOS-NETO; CASTRO, 2021). Uma forma simples de modelar uma série de dados como um processo linear é dado por:

$$x_t = c + \sum_{i=1}^p \Phi_i x_{t-i} + \varepsilon_t \quad (1)$$

onde x_t é a variável de estacionalidade, c é uma constante, Φ_i são os coeficientes de

autocorrelação na defasagem e ε_t é o resíduo.

O termo integrador refere-se à aplicação de sinais diferenciais ao invés dos absolutos na série, com o objetivo de torná-la estacionária, mesmo que tenha uma tendência de crescimento, por exemplo. A ordem de integração, denotada por d , indica quantas vezes a série foi diferenciada (BROCKWELL; DAVIS, 2002). Já o termo média móvel refere-se à dependência dos erros de predição anteriores na predição atual. Ele considera que o valor atual da série é linearmente relacionado com os erros de predição passados. A ordem do componente MA, denotada por q , indica quantos erros de predição passados são considerados na modelagem (BOX *et al.*, 2015). O tempo MA pode ser representado de forma linear por:

$$x_t = \mu + \sum_{i=1}^q \theta_i \varepsilon_{t-i} \quad (2)$$

onde μ é a expectativa de x_t , θ são os pesos dos termos aplicados aos valores atuais e anteriores de forma estocástica na série temporal e onde $\theta_0 = 1$ (RIBEIRO *et al.*, 2022).

A combinação dos 3 modelos, formando o modelo ARIMA, pode ser representada pela equação (3).

$$x_t = c + \sum_{i=1}^p \Phi_i x_{t-i} + \varepsilon_t + \mu + \sum_{i=1}^q \theta_i \varepsilon_{t-i} \quad (3)$$

Vale salientar que o componente integrado d foi incorporado usando i como um índice para representar os diferentes períodos de diferenciação. A diferença x_{t-i} refere-se à diferença entre os valores da série x no tempo t e no tempo $t-i$, representando a parte integrada do modelo ARIMA.

A principal vantagem desse tipo de modelo reside na capacidade de lidar com diversos tipos de séries, incluindo aquelas com tendências, sazonalidades e variações não lineares. Além disso, o modelo ARIMA permite a incorporação de termos sazonais por meio do modelo SARIMA, que é especialmente útil para tratar padrões sazonais mais complexos. Contudo, é necessário considerar as limitações do modelo ARIMA. O mesmo assume que a série seja linear e estacionária, o que pode restringir sua aplicabilidade em séries com comportamentos não lineares ou não estacionários. Além disso, a seleção adequada dos parâmetros do modelo, como as ordens p , d e q , exige conhecimento prévio dos dados e experiência na análise de séries temporais (BOX *et al.*, 2015).

Este modelo, baseado em ciclo iterativo dos próprios dados, é composto por quatro etapas, sendo elas: a identificação, estimação, verificação e predição. A identificação visa obter as três características do modelo que descrevem o comportamento da

série e é comum utilizar funções de autocovariância e autocorrelação para investigar as relações teóricas entre as observações (LUIZ *et al.*, 2022). Essas funções têm o objetivo de estimar o grau de interdependência numérica entre as observações de uma variável aleatória. Quando trata de diferentes séries temporais que podem estar relacionadas, é possível estabelecer uma correlação entre elas e avaliar o nível de dependência em relação a um determinado atraso ou defasagem nos dados. Na estimação, realiza a estimativa dos parâmetros auto-regressivo e média móvel de forma simples. Já na próxima etapa, é verificado se os parâmetros definidos são adequados. A estimação dos parâmetros do modelo ARIMA é geralmente realizada por meio de técnicas de otimização, como o método dos mínimos quadrados. A confiabilidade do modelo tipicamente pode ser quantificada pela raiz do erro médio quadrático (RMSE, do inglês *Root Mean Squared Error*), desejando que seu valor seja o mais próximo de zero possível, ou seja, que os valores preditos sejam mais próximos dos reais. O objetivo é encontrar os valores dos parâmetros que melhor se ajustam aos dados observados, minimizando a diferença entre as previsões do modelo e as observações reais. Dessa forma, com o modelo pronto e com dados de entrada disponíveis para determinado sistema, é possível prever os próximos dados para identificação do comportamento da série temporal (RIBEIRO *et al.*, 2022).

2.4 MODELOS DE APRENDIZADO DE MÁQUINA

O aprendizado de máquina é uma área da inteligência artificial que se concentra no desenvolvimento de algoritmos e modelos capazes de aprender a partir de dados e realizar tarefas específicas sem serem explicitamente programados. É uma abordagem que permite que as máquinas adquiram conhecimentos e façam previsões ou tomem decisões com base nesses conhecimentos. O objetivo fundamental dessa abordagem é capacitar as máquinas a aprenderem padrões e realizarem tarefas complexas com base nos dados disponíveis. Isso é alcançado por meio de um processo iterativo, no qual os modelos são treinados em conjuntos de dados de treinamento e ajustados para otimizar seu desempenho (MASINI; MEDEIROS; MENDES, 2023).

As técnicas de aprendizado de máquina podem ser divididas em três tipos principais, de acordo com o aprendizado: supervisionado, não supervisionado e por reforço. No aprendizado supervisionado, os modelos são treinados utilizando dados rotulados, ou seja, dados em que as respostas corretas são conhecidas. O modelo aprende a mapear os dados de entrada para as respostas corretas, permitindo realizar previsões ou classificações precisas em novos dados (HAYKIN, 2001). No aprendizado não supervisionado, os modelos são treinados utilizando dados não rotulados, ou seja, sem informações sobre as respostas corretas. O objetivo tipicamente é descobrir padrões, estruturas ou agrupamentos intrínsecos nos dados, possibilitando a identificação de *insights* valiosos ou segmentações relevantes. Já no aprendizado por reforço, os mo-

delos aprendem por meio de interações com um ambiente. O modelo, denominado agente, recebe um *feedback* na forma de recompensas ou penalidades baseadas em suas ações. Assim, aprende a tomar decisões que maximizem as recompensas ao longo do tempo, por exemplo, em jogos ou controle de sistemas complexos (FLECK *et al.*, 2016).

Cada tipo de aprendizado possui suas particularidades e é selecionado de acordo com o problema a ser resolvido e a disponibilidade dos dados. Com a ajuda desses algoritmos e modelos, é possível extrair informações úteis dos dados e tomar decisões inteligentes com base nesses conhecimentos adquiridos. É importante destacar que o sucesso da predição depende de vários fatores, como a qualidade dos dados utilizados para treinamento, a escolha adequada do algoritmo de aprendizado de máquina, a seleção adequada das variáveis de entrada e outros aspectos relacionados ao pré-processamento dos dados (PARENTE, 2022).

As RNAs são um subcampo do aprendizado de máquina e são amplamente utilizadas para realizar tarefas de predição, classificação, reconhecimento de padrões, processamento de linguagem natural, entre outras. Embora a aplicação de RNAs necessite da definição de uma estrutura para o modelo neural, essa técnica tem a capacidade de modelar processos mesmo com escasso ou nenhum conhecimento prévio sobre a física subjacente ao processo (SCHWEDERSKY, 2017). As RNAs tipicamente implementam aprendizado supervisionado, mas há estruturas especiais de redes para tarefas de aprendizado não supervisionado. Exemplos de aplicação de RNAs para tarefas de predição são encontradas em SANTOS *et al.* (2021), que utilizaram o modelo LSTM para predição da temperatura e irradiância solar, e Schwedersky (2017), que empregou diferentes arquiteturas de redes neurais recorrentes para identificação de sistemas dinâmicos não lineares.

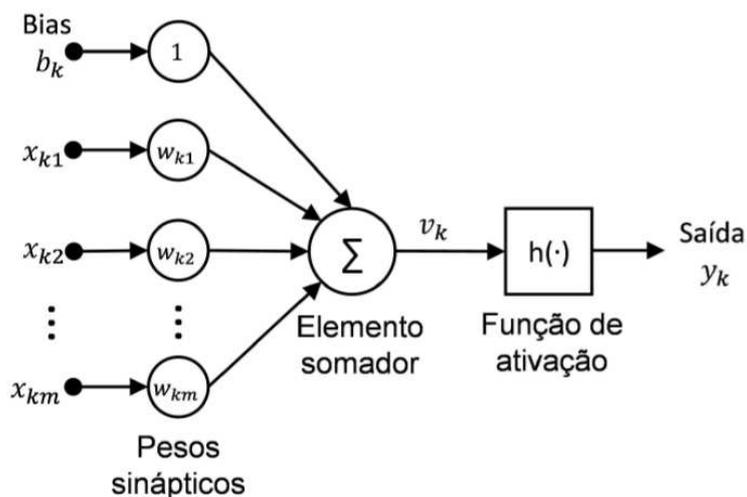
2.4.1 Visão geral sobre redes neurais

A tarefa de predição é uma das aplicações mais comuns de redes neurais artificiais. Elas são capazes de aprender padrões complexos nos dados de treinamento e, com base nesse aprendizado, podem fazer predições ou estimativas sobre dados não vistos anteriormente (HAYKIN, 2001).

Uma RNA é um modelo computacional inspirado no funcionamento do cérebro humano. Ela é composta por um conjunto interconectado de unidades de processamento chamadas de neurônios artificiais. Esses neurônios tipicamente estão organizados em camadas e cada neurônio recebe dados de entrada, realiza uma operação de processamento e gera uma saída, que é transmitida para os neurônios da camada seguinte. A estrutura básica de um neurônio de uma RNA pode ser representada pela Figura 2, onde x_{kn} são as entradas e n quantidades de entrada, b_k bias, w_{kn} os pesos sinápticos, v_k é a saída do elemento somador, $h(.)$ função de ativação e y_k a

saída e. Os neurônios artificiais, componentes essenciais das redes neurais artificiais, possuem uma estrutura fundamental que se baseia na morfologia e funcionamento dos neurônios biológicos. Eles consistem em três elementos principais: um conjunto de pesos sinápticos, um somador e uma função de ativação (SCHWEDERSKY, 2017).

Figura 2 – Estrutura típica de um neurônio artificial.



Fonte: Schwedersky (2017) e Haykin (2001)

Através de algoritmos de otimização, a rede ajusta seus parâmetros internos, também conhecidos como pesos, de modo a minimizar a diferença entre as saídas previstas e as saídas desejadas. Esse processo de aprendizado permite que a rede neural generalize a partir dos dados de treinamento e seja capaz de fazer previsões ou classificações corretas para novos exemplos que não foram vistos durante o treinamento (GONÇALVES, 2017).

Existem diferentes arquiteturas de redes neurais, como as redes neurais *feedforward* e as redes neurais recorrentes, cada qual com suas características e aplicações específicas. As redes neurais recorrentes têm a capacidade de implementar sistemas dinâmicos, algo que não é possível com redes *feedforward* convencionais. Dessa forma, essas arquiteturas são especialmente úteis para aprender e lidar com problemas envolvendo séries temporais (GOMES, 2005).

As redes neurais recorrentes apresentam algum tipo de realimentação. Assim, ao contrário das redes neurais *feedforward* convencionais, as RNNs têm a capacidade de manter uma memória interna ou estado oculto que permite que informações anteriores influenciem o processamento de informações futuras (GIGLIO, 2021). Por essa característica, as RNNs são um tipo de modelo adequado para lidar com dados sequenciais, nos quais a ordem e a dependência temporal entre os elementos são importantes. Isso se dá porque a recorrência da rede permite armazenar informações

temporais, a fim de aprender e memorizar o histórico de entrada. Assim, elas são uma ferramenta poderosa para aprendizado de trajetória, previsão de séries temporais e outras tarefas dependentes do tempo. No entanto, na prática, os métodos de treinamento empregados em redes *feedforward*, geralmente, apresentam problemas de estabilidade quando aplicados a RNNs. Além disso, mesmo quando a estabilidade é obtida, a alta complexidade numérica das recorrências faz com que sejam obtidos tempos de treinamento muito longos (PARK *et al.*, 2023).

A arquitetura tradicional de uma RNN consiste em uma sequência de células recorrentes, em que cada célula possui uma entrada, uma saída e um estado oculto. A entrada é alimentada para a célula a cada passo de tempo, juntamente com o estado oculto anterior, e a saída é gerada com base nessas duas informações de entrada. Assim, o estado oculto é atualizado a cada passo de tempo e serve como uma forma de memória interna que armazena informações contextuais das etapas anteriores. Essa conexão realimentada permite que as RNNs capturem informações contextuais de etapas anteriores e as usem para influenciar a saída atual (GRAVES, 2013). No entanto, as RNN tradicionais têm um desafio para treinamento, visto que os gradientes das informações propagadas ao longo do tempo são atenuados ou amplificados exponencialmente. Esse problema dificulta a capacidade da rede de aprender dependências de longo prazo e pode afetar o desempenho e a estabilidade do treinamento (LIPTON; BERKOWITZ; ELKAN, 2015).

Na literatura, existem alternativas para contornar os problemas apresentados por RNNs clássicas. Uma classe específica desses modelos é a computação de reservatório (RC, do inglês *reservoir computing*), que utiliza um reservatório de neurônios interconectados para gerar um modelo dinâmico. A principal arquitetura que se beneficia desse conceito é a ESN (SCHWEDERSKY, 2017). Outro método que merece destaque é a arquitetura do tipo LSTM, que foi projetada para superar o problema de desvanecimento do gradiente e capturar dependências de longo prazo em sequências de dados. Ao contrário das RNNs tradicionais, as LSTMs possuem uma arquitetura especializada com células de memória (ou *gates*) que podem armazenar informações por longos períodos de tempo. Essas células possuem mecanismos de controle de fluxo de informações, como portas de entrada, saída e esquecimento, que permitem que as LSTMs aprendam a reter informações relevantes e descartar informações irrelevantes ao longo do tempo (SANTOS *et al.*, 2021). Essas alternativas são detalhadas nas seções a seguir.

2.4.2 Computação de reservatório

Em um modelo de computação de reservatório, são identificadas pelo menos duas estruturas distintas: o reservatório e a estrutura de leitura (*readout*). O reservatório é responsável por expandir os dados de entrada e as informações históricas em um

espaço de alta dimensão, por meio de uma projeção não linear. Essa projeção pode aumentar a separabilidade linear dos dados de entrada e, além disso, o reservatório retém uma transformação não linear do passado em seus estados internos, atuando como uma estrutura de memória (BOCCATO, 2013).

Por outro lado, a estrutura de leitura é projetada para ser robusta e rápida no processo de aprendizado. Uma característica distintiva desses modelos é que os pesos dos circuitos do reservatório são considerados fixos durante o treinamento. Em outras palavras, o reservatório é inicializado de forma arbitrária e permanece inalterado durante o aprendizado. O treinamento concentra-se apenas na atualização dos pesos na estrutura de leitura. Esse fato baseia-se na suposição de que, por experiência empírica, em tarefas de aprendizado, sob determinadas hipóteses algébricas, é suficiente atualizar apenas a estrutura de leitura para obter bons desempenhos (LUKOŠEVIČIUS; JAEGER, 2009).

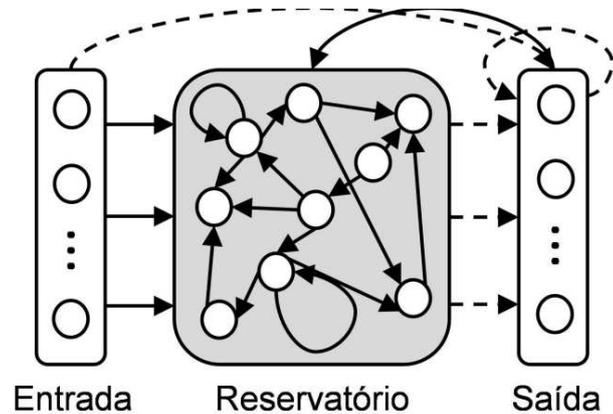
Jaeger (2002) propôs uma abordagem de compromisso para treinar redes neurais recorrentes por meio de um modelo de computação de reservatório. Essa arquitetura é caracterizada pela presença de um reservatório de dinâmicas com parâmetros fixos, que atua como uma camada intermediária recorrente, e uma camada de saída linear, que é ajustada de maneira supervisionada. Essa abordagem permite que o reservatório aprenda a capturar padrões complexos nos dados de entrada, enquanto a camada de saída é treinada apenas para mapear as representações do reservatório para as saídas desejadas. Essa separação entre o treinamento do reservatório e da camada de saída simplifica o processo de treinamento e pode levar a melhores resultados em problemas de predição e de classificação (SCHWEDERSKY, 2017).

Nesse tipo de arquitetura, o reservatório pode ser entendido como um *kernel* temporal, que projeta a entrada em um espaço dinâmico não linear. Durante a operação, os estados do reservatório formam uma trajetória que depende tanto dos estímulos externos quanto da memória dos estímulos passados. A camada de saída desempenha um papel fundamental na computação, pois lê os estados instantâneos do reservatório e realiza as operações necessárias para gerar as saídas desejadas, sendo a única parte treinável do modelo (SCHWEDERSKY, 2017; ANTONELLO, 2011).

A arquitetura de computação de reservatório empregada neste trabalho é a ESN. Na Figura 3 é ilustrada a arquitetura básica de uma rede de estados de eco. Essa arquitetura é composta por neurônios com ativação sigmoide, conectados de forma recorrente, o que confere um comportamento dinâmico à rede. As conexões entre os neurônios do reservatório são fixas e geradas de forma aleatória, sendo algumas conexões opcionais e podendo ser omitidas dependendo da aplicação. No entanto, as conexões internas ao reservatório e o mecanismo de leitura do reservatório, que liga o reservatório à camada de saída, são essenciais. A camada de leitura combina linearmente os sinais gerados pelos neurônios do reservatório para produzir as saídas

da rede (SCHWEDERSKY, 2017).

Figura 3 – Arquitetura ESN. As conexões fixas entre os neurônios são representadas por linhas sólidas, enquanto os pesos treináveis são representados por linhas tracejadas.



Fonte: Antonelo (2011)

Os estímulos de entrada u_t são combinados linearmente usando os coeficientes de W^{in} e transmitidos para o reservatório de dinâmicas. O reservatório é composto por N neurônios não lineares conectados, com ativações denotadas por k_t . As ativações dos neurônios são atualizadas de acordo com (BOCCATO, 2013):

$$k_{(t+1)} = f(W^{in}u_{(t+1)} + Wx_{(t)}), \quad (4)$$

em que a matriz $W \in R^{(N,N)}$ representa os pesos das conexões recorrentes do reservatório, $f(\cdot) = (f_1(\cdot), \dots, f_N(\cdot))$ indica as funções de ativação das unidades internas e $x_{(t)}$ é estado do reservatório.

Os valores iniciais das matrizes têm uma influência significativa no comportamento global do modelo e certas características dessas matrizes desempenham um papel relevante na dinâmica do reservatório. Os pesos do reservatório (W), que são gerados de forma aleatória e garantem a memória da rede, devem seguir certos requisitos para garantir que a dinâmica do reservatório seja estável, atingindo a propriedade de estado de eco. Os estados de eco são responsáveis por capturar informações sobre o histórico dos sinais de entrada e, assim, proporcionar uma memória dinâmica à rede. Essa propriedade é alcançada por meio da escolha adequada da matriz de pesos recorrentes do reservatório, que deve ter autovalores dentro do círculo de raio unitário no plano complexo. Essa condição garante que as informações passadas desapareçam assintoticamente, evitando que o reservatório fique sobrecarregado com informações irrelevantes e mantendo a capacidade de resposta aos sinais atuais (JAEGGER, 2002).

Além disso, para ajustar a escala de tempo do reservatório e capturar a dinâmica temporal dos sinais de entrada, as unidades integradoras com taxa de vazamento

($\alpha \in (0,1]$) desempenham um papel eficaz, alterando o nível de memória do reservatório. O parâmetro α controla a rapidez com que as informações passadas são esquecidas, sendo que $1 - \alpha$ determina a taxa de retenção das informações anteriores. Isso permite ao modelo lembrar informações relevantes do passado enquanto acompanha variações temporais nos sinais de entrada, tornando-o mais preciso na previsão de séries temporais e em problemas que exigem considerações temporais (SCHWEDERSKY, 2017). Vale destacar que não há uma técnica singular e precisa para a escolha do valor ideal desse parâmetro. Apesar da falta de uma abordagem determinística, é possível recorrer a métodos meta-heurísticos, como os algoritmos genéticos, para explorar eficientemente o espaço de soluções em busca de valores mais apropriados para o parâmetro α .

A atualização do estado do modelo considerando taxa de vazamento é dada por:

$$x_{(t+1)} = (1 - \alpha)x_{(t)} + \alpha f(W^{in}u_{(t)} + W^{res}x_{(t)} + W^{bias} + W^{out,res}y_{(t)}). \quad (5)$$

A equação (5) é comumente denominada equação de estados, a qual determina o estado futuro do modelo. Nesse contexto, os valores atuais dos neurônios de entrada são representados por $u_{(t)}$ e W são os pesos, sendo os índices "res" referente ao reservatório, "in" aos neurônios de entrada e "bias" indicando o viés, respectivamente. A função $f(\cdot)$, também conhecida como função base na teoria de identificação de sistemas (NELLES, 2002), é geralmente a função de ativação utilizada, frequentemente expressa como $\tanh(\cdot)$.

O reservatório de dinâmicas tem como objetivo gerar um conjunto diversificado de comportamentos dinâmicos. No entanto, ele é projetado sem informações específicas do sinal que se deseja aproximar, o que resulta em uma perda de capacidade de representação em comparação com uma estrutura ajustada idealmente com base em um critério de mínimo erro quadrático médio. Nesse contexto, torna-se importante uma definição adequada dos parâmetros da camada recorrente de antemão, como o número de neurônios, a taxa de vazamento e o raio espectral (BOCCATO, 2013).

O objetivo do aprendizado é encontrar a solução ótima para as matrizes W^{out} e $W^{bias,out}$, que minimiza o erro quadrático entre a saída predita $y_{(t)}$ e a saída desejada $\hat{y}_{(t)}$ para cada instante de tempo t . A saída predita é obtida diretamente a partir do estado do reservatório como:

$$y_{(t)} = W^{out}x_{(t)} + W^{bias,out} \quad (6)$$

O problema se resume em encontrar a solução para um sistema de equações lineares que garanta a melhor correspondência entre as previsões e as saídas desejadas (ANTONELO, 2011; SCHWEDERSKY, 2017). É importante reforçar que esse ajuste é realizado apenas nos coeficientes da camada de saída, dado que os pesos do reservatório são inicializados a priori e não mudam durante o treinamento. Sendo

assim, os coeficientes ótimos podem ser determinados usando métodos de regressão linear de mínimos quadrados (SHI; HAN, 2007), devido à natureza linear do problema resultante, o que simplifica significativamente o processo de treinamento em comparação com outras arquiteturas de redes neurais recorrentes (BOCCATO, 2013; YAO; WANG; ZHANG, 2019).

Ao ajustar apenas os pesos do mecanismo de leitura do reservatório, o treinamento dessa arquitetura de rede evita várias desvantagens, como o longo tempo de treinamento e o problema do desaparecimento do gradiente, que afeta redes profundas treinadas com métodos baseados no cálculo do gradiente. Para o processo de criação e treinamento do modelo ESN, podem ser otimizados alguns hiperparâmetros, relacionados à arquitetura do reservatório e ao procedimento de treinamento, sendo eles: raio espectral, número de neurônios, taxa de vazamento e parâmetro de regularização (LUKOŠEVIČIUS; JAEGER, 2009).

2.4.3 Redes neurais LSTM

A rede neural LSTM é uma família de arquiteturas de redes compostas por uma ou mais camadas de unidades de memória especializadas, denominadas memórias LSTM. Cada unidade LSTM apresenta uma estrutura que foi desenvolvida para superar uma das principais limitações das redes neurais recorrentes tradicionais, que é o aprendizado de dependências de longo prazo, devido ao processo de treinamento apresentar problemas como o desaparecimento do gradiente, o que não ocorre nas redes LSTM (HOCHREITER; SCHMIDHUBER, 1997).

As unidades de memória LSTM são compostas por uma estrutura com três componentes fundamentais: os portões de entrada (*input gate*), de esquecimento (*forget gate*) e o de saída (*output gate*). Esses portões são estruturas de controle que regulam o fluxo de informação dentro da célula LSTM, conforme representado na Figura 4, onde uma memória LSTM é apresentada em três instantes de tempo distintos, apresentando de forma visual o fluxo de informações entre os instantes de tempo $t - 1$, t e $t + 1$ (HOCHREITER; SCHMIDHUBER, 1997; SANTOS *et al.*, 2021).

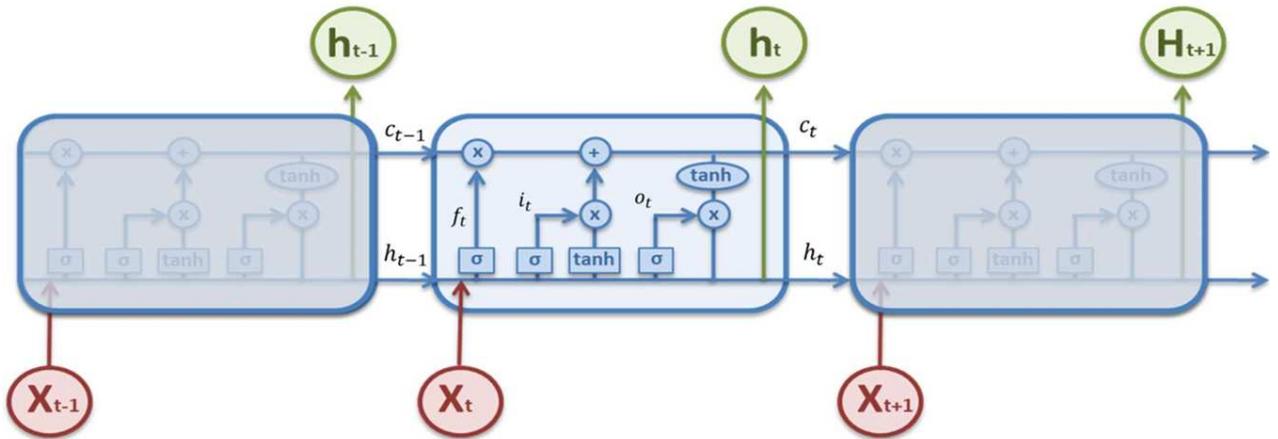
O portão de entrada (i_t) é configurado para determinar quais informações novas devem ser armazenadas no estado atual, c_t , da memória LSTM, conforme:

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i), \quad (7)$$

onde W_{xi} e W_{hi} representam as matrizes de pesos associados à entrada atual e à saída passada da memória, b_i representa o viés associados ao portão de entrada, x_t é o vetor de entrada no instante de tempo t , h_t é a saída da memória LSTM no instante de tempo t e σ é a função de ativação sigmoide.

O estado atual da memória será composto por informações selecionadas pelo portão de entrada e informações que estavam armazenadas na memória no instante

Figura 4 – Arquitetura LSTM. Fluxo de informação de um bloco de memória LSTM no tempo t



Fonte: Sá *et al.* (2020)

passado, o que é representado por c_{t-1} . A quantidade de informação que será mantida é determinada pelo portão de esquecimento, o qual determina quais porção da informação armazenada em c_{t-1} será mantida, por meio de:

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f), \quad (8)$$

onde as matrizes W_{xf} e W_{hf} representam os pesos do portão de esquecimento, associados à entrada atual e estado passado, respectivamente; e b_f representa o viés associado ao portão de esquecimento.

O estado atual da memória LSTM (c_t) é a soma da porção que será mantida do estado passado e de uma porção que será adicionada, a qual é determinada pelo portão de entrada i_t , o qual filtra informações de x_t e h_{t-1} conforme:

$$\tilde{c}_t = \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c), \quad (9)$$

fazendo com que o estado atual da memória seja definido como:

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t, \quad (10)$$

onde $\tanh(\cdot)$ é a função tangente hiperbólica; c_t é o estado da célula de memória no tempo t ; W_{xc} e W_{hc} são as matrizes de pesos associados à entrada e saída passada, respectivamente; b_c é um viés e \odot refere-se ao produto de Hadamard, que é uma multiplicação elemento a elemento.

Após realizar a atualização do estado interno da memória, o portão de saída (o_t), decide quais informações da memória devem ser retornados na saída. Para escolher a porção que será retornada, o portão utiliza informações da entrada e da saída no

instante de tempo passado, conforme:

$$o_t = \sigma(W_{x0}x_t + W_{h0}h_{t-1} + b_o). \quad (11)$$

onde W_{x0} e W_{h0} representam as matrizes de pesos associadas à entrada e saída passada da memória e b_o representa o viés associado ao portão de saída (GREFF *et al.*, 2016). Por fim, o valor de saída é determinado por:

$$h_t = o_t \odot \tanh(c_t). \quad (12)$$

O estado interno de cada bloco é responsável pelo armazenamento de informação, o que possibilita a atividade de mapeamento das sequências de observações anteriores de entrada, para determinação da saída predita (CHEN *et al.*, 2021). Esse modelo é bastante utilizado na literatura para problemas em que o modelo deve ser capaz de armazenar informações do passado, como em problemas de predição de séries temporais, como em Brownlee (2019).

Para utilização desse modelo é necessária a configuração inicial de suas camadas ocultas, definida pelo número de entradas, quantidade de unidades LSTM e função de ativação das unidades (XIA *et al.*, 2022). Outras configurações consistem de taxa de aprendizado (TA) inicial, fator de queda da taxa de aprendizado (QTA), período de QTA e número de épocas. A TA inicial em um modelo LSTM é um hiperparâmetro que define o tamanho do passo utilizado durante o processo de atualização dos pesos do modelo. Uma TA maior pode acelerar o processo de aprendizado, mas também pode resultar em instabilidade. O QTA determina a taxa com a qual a TA é reduzida ao longo do treinamento, geralmente aplicando-se uma multiplicação pelo QTA a cada período de QTA. O período de QTA indica a frequência em que a TA é reduzida. O número de épocas define o número total de vezes que o modelo percorre todo o conjunto de treinamento durante o treinamento. Um número maior de épocas permite um ajuste mais refinado do modelo, mas também pode levar a um maior tempo de treinamento e risco de *overfitting* (CHEN *et al.*, 2021). Mais detalhes sobre esses parâmetros são apresentados em Hochreiter e Schmidhuber (1997).

2.5 ALGORITMOS GENÉTICOS

AGs são um método de otimização inspirado na teoria da evolução. O método utiliza uma abordagem baseada em populações de soluções candidatas e emprega operadores genéticos, como seleção, cruzamento e mutação, para explorar um espaço de busca e encontrar soluções melhores que as anteriores (GOLDBERG; KORB; DEB, 1989).

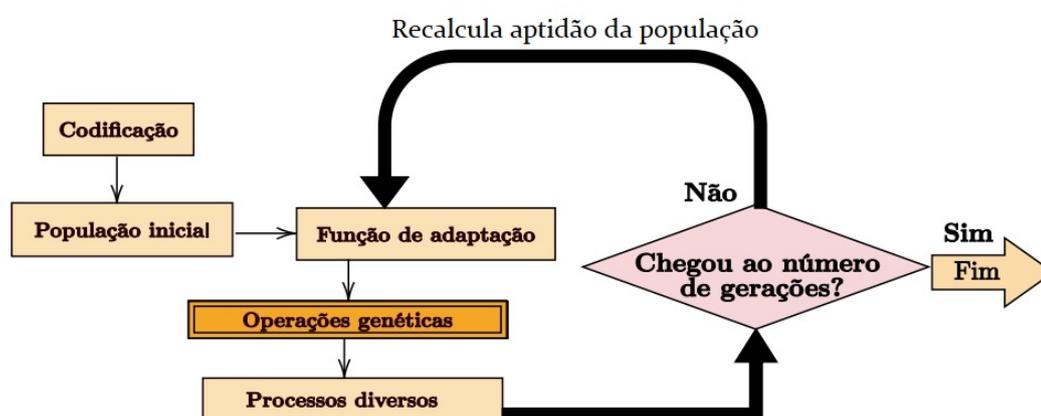
Segundo Dréo *et al.* (2006a), no contexto do algoritmo evolutivo, os participantes passam por um processo de transformação no qual se apresentam como possíveis

soluções para um determinado problema. Cada solução possui um nível variável de eficiência dentro do espaço de busca da otimização. Essas soluções, em conjunto, formam uma população, tratada simultaneamente pelo algoritmo evolutivo. O processo evolutivo ocorre ao longo de uma série de iterações, persistindo até que um critério de conclusão seja satisfeito, considerando-se a qualidade prévia das soluções obtidas.

A ideia de utilizar esse método meta-heurístico para identificação automática de parâmetros de uma rede neural já é usado e apresenta resultados significativamente positivos, como em Santos *et al.* (2019), que utilizaram um AG para regulação dos hiperparâmetros de uma RNA.

A base do algoritmo genético consiste da seleção de três componentes aplicadas ao problema, sendo elas: população inicial, função de adaptação e operadores genéticos (ZHAN; LI; ZHANG, 2022). Com a população inicial gerada dentro das margens pré-definidas, a mesma é submetida à evolução, obtendo soluções mais ou menos eficientes. Essas soluções pertencem ao subespaço de pesquisa do problema de otimização, como representado na Figura 5. Essa metodologia pode ser trabalhada com valores binários e absolutos.

Figura 5 – Representação algoritmo genético



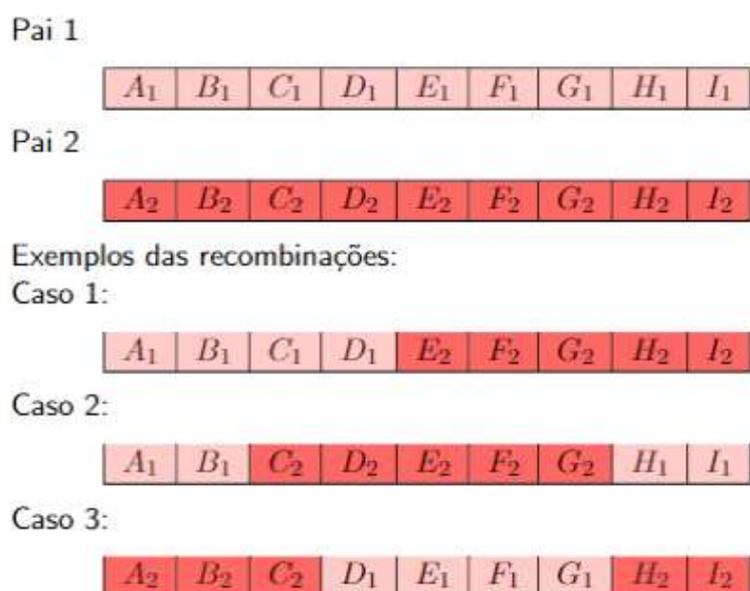
Fonte: Elaborado pela autora baseado na explicação do algoritmo de Dréo *et al.* (2006a)

Em cada iteração do algoritmo, nomeado por geração, os indivíduos são avaliados com base em uma função de aptidão, que mede o quão bem cada solução resolve o problema. Com base nessa avaliação, os indivíduos mais aptos têm uma maior probabilidade de serem selecionados para reprodução. Ou seja, a evolução durante uma sucessão de iterações é executada até um critério de término, que leva em consideração a melhoria da qualidade das soluções obtidas por meio da função de aptidão (DRÉO *et al.*, 2006a).

Durante o processo de reprodução, ocorre a recombinação genética, na qual os genes dos indivíduos selecionados são combinados para criar novos indivíduos, chamados de descendentes. Essa recombinação pode ser através de técnicas como cruzamento de pontos de corte ou *crossover* uniforme (DRÉO *et al.*, 2006a; GOLDBERG; KORB; DEB, 1989).

Os operadores genéticos são responsáveis por gerar novos indivíduos para a nova população e são definidos pela recombinação, mutação e elitismo. A recombinação gera um novo indivíduo a partir das informações que foram atribuídas para dois indivíduos que exercem a função de reprodução (pais). A Figura 6 traz três exemplos de recombinações possíveis para gerar novos filhos a partir da recombinação dos pais. Cada caso conta com a determinação de 1 ou 2 posições de corte dos genes. As posições de cortes podem ser fixadas ou selecionadas de forma aleatória entre os intervalos determinados e a recombinação é feita a partir dos genes dos pais, com a contribuição de cada um determinada pelo corte.

Figura 6 – Representação do processo de recombinação



Fonte: Elaborado pela autora

A mutação gera um novo indivíduo a partir de apenas um pai (esse operador gera um equivalente a uma busca local). Com uma dada probabilidade, é escolhido o melhor indivíduo da geração para sofrer ao menos uma mutação. Caso o sorteio defina que ele não deve ser escolhido, são escolhidos indivíduos aleatoriamente para sofrerem mutação. O elitismo, por sua vez, garante que a melhor solução da geração se propague para a geração seguinte (GOLDBERG; KORB; DEB, 1989). Dessa forma através de um processo de iterativo, a população é melhorada após passar por uma avaliação de aptidão.

A utilização de algoritmos genéticos como alternativa de ajuste dos hiperparâmetros é bastante comum na literatura, sendo interessante por permitir o ajuste de todos os parâmetros ao mesmo tempo. Esse tipo de abordagem se baseia em um processo que emula um processo de sobrevivência e evolução das soluções que tendem a ser mais aptas. Sua capacidade de explorar eficientemente espaços de busca complexos e encontrar soluções de alta qualidade faz deles uma ferramenta valiosa na área de otimização. A vantagem desse método em comparação com outros é a resistência à convergência em mínimos locais, devido a sua variação de informações durante o processo de seletividade da melhor solução (BAUTISTA, 2019).

2.6 COMENTÁRIOS FINAIS

Neste capítulo de revisão da literatura, são abordadas as características e modelamento de séries temporais utilizando métodos tradicionais, como o modelo ARIMA, e soluções baseadas em redes neurais recorrentes, como LSTM e ESN, para a predição dessas séries. Também são apresentados os fundamentos para a avaliação desses métodos. As arquiteturas derivadas do modelo ARIMA são as mais comuns e representam uma escolha padrão. Por outro lado, as arquiteturas que empregam o conceito de computação de reservatório e LSTM oferecem uma alternativa às redes neurais recorrentes clássicas, com um grande potencial para a aplicação em problemas de predição. Uma das limitações desses métodos reside na configuração dos hiperparâmetros, visto que a configuração adequada da rede é importante para que se obtenham bons resultados. Uma solução meta-heurística explorada é o algoritmo genético, que pode ser utilizado para buscar a melhor solução de acordo com a função de aptidão.

No Capítulo 3, a combinação dos conceitos de predição e RNA é utilizada para estabelecer uma metodologia de identificação de sistemas empregando RNAs recorrentes e modelos ARIMA. Já no Capítulo 4, esses modelos são submetidos a avaliações experimentais para a predição de dados reais, proporcionando uma análise da aplicabilidade dessas arquiteturas.

3 MÉTODO E ESTUDO DE CASO

Este capítulo apresenta o método empregado e o estudo de caso considerado, com medições reais de uma área contaminada. Ele está dividido em três seções. A Seção 3.1, inicia o capítulo apresentando uma discussão sobre a depleção natural na zona da fonte e sua conexão com a temperatura do subsolo. Em seguida, na Seção 3.2 concentra-se em um estudo de caso específico que envolve a previsão. São abordados tópicos como o sistema de aquisição de dados, o método empregado para a previsão de séries temporais de temperaturas no subsolo e a descrição das séries de dados utilizadas para o estudo. Além disso, são detalhadas as métricas de avaliação da qualidade das previsões. Por fim, a Seção 3.3 apresenta as considerações finais que sintetizam os principais resultados.

3.1 DEPLEÇÃO NATURAL NA ZONA DA FONTE E SUA RELAÇÃO COM A TEMPERATURA DO SUBSOLO

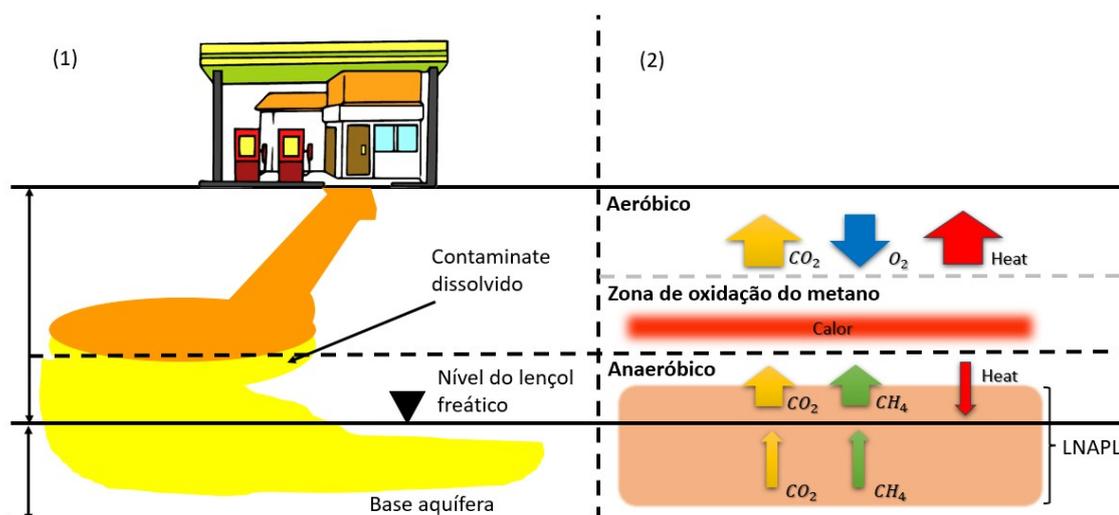
O derramamento de derivados de petróleo em solo resulta na formação do LNAPL, um líquido não aquoso menos denso que a água. No início, o LNAPL se infiltra no solo por forças gravitacionais e pode ser retido por barreiras físicas ou forças de empuxo. Ele pode se mover verticalmente devido à oscilação do lençol freático ou horizontalmente devido ao gradiente hidráulico. O espalhamento do material oleoso depende do volume derramado e das propriedades do solo. Durante o transporte, parte do LNAPL permanece retida no solo, atuando como uma fonte contínua de contaminação do solo, água subterrânea e ar atmosférico (STOCKWELL, 2015). Com a mobilidade do contaminante para a base aquífera é formada uma pluma de contaminante e, ao longo do tempo, esses compostos se quebram em compostos menores por efeito de biodegradação, que pode ocorrer de forma natural ou estimulada (PERES, 2019).

A NSZD está se tornando importante para o gerenciamento de áreas contaminadas por LNAPL, visto que ela se refere aos mecanismos naturais presentes no solo que atuam na atenuação do LNAPL, como biodegradação direta, volatilização, dissolução na água subterrânea e precipitação química. Para aplicar a NSZD como método de remediação com sucesso, é necessário compreender e medir os processos naturais que depletam o LNAPL (ASKARANI; SALE, 2020; PERES, 2019; SWEENEY; RIRIE, 2014).

Stockwell (2015) mostra o processo de depleção dividido em estágios que ocorrem em diferentes camadas do subsolo. Tal processo é ilustrado na Figura 7. A camada onde fica depositada a pluma de contaminante na base aquífera é identificada como anaeróbica, ou zona de formação de metano. Nela ocorre o processo de metanogênese dos hidrocarbonetos, gerando calor, CO_2 e CH_4 . Tais gases sobem para a camada superior, onde ocorre o encontro com o oxigênio advindo da atmosfera, gerando a

oxidação do metano, reação que também libera calor. Os gases passam, por último, pela zona de transporte aeróbico e são liberados para a atmosfera.

Figura 7 – Modelo conceitual de fluxo de carbono associado com NSZD em locais com a presença de LNAPL: (1) dinâmica de distribuição do LNAPL; (2) processo de depleção natural das fontes de contaminação.



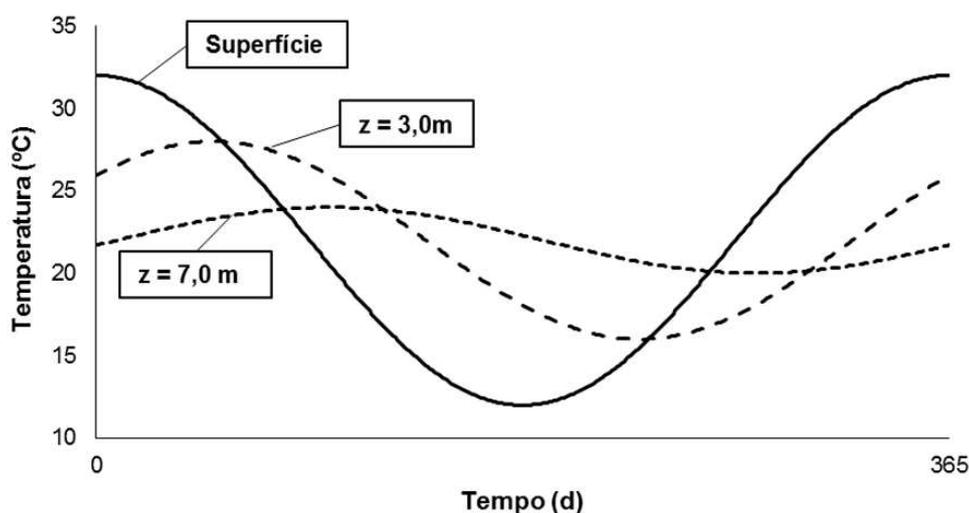
Fonte: Elaborado pela autora, baseado em Stockwell (2015) e Peres (2019).

A compreensão aprofundada da NSZD revelou que a degradação do LNAPL é amplamente influenciada pela metanogênese, especialmente em casos de derramamentos antigos. O metano, por sua vez, é frequentemente oxidado e transformado em gás carbônico na zona de oxidação de metano, liberando uma quantidade considerável de calor durante a reação. Pesquisas recentes destacaram que tanto a formação do metano quanto sua oxidação têm um efeito mensurável na temperatura da zona da fonte. Esse efeito levou ao desenvolvimento de um método de cálculo da biodegradação do LNAPL baseado na temperatura. Essa abordagem utiliza o gradiente de temperatura para estimar o fluxo de calor, que está relacionado ao volume de produtos de degradação liberados, por meio de relações termodinâmicas (SWEENEY; RIRIE, 2014). Esse método oferece uma maneira contínua e de baixo custo para estimar a taxa de NSZD (PERES, 2019)

De acordo com Sweeney e Ririe (2014), as interações térmicas entre o solo e a atmosfera são comumente investigadas utilizando modelos unidimensionais, que permitem analisar como a temperatura varia ao longo da profundidade do solo. Em situações nas quais a influência da fonte de calor geotérmica é insignificante, espera-

se que as curvas de temperatura no solo se assemelhem à curva de temperatura atmosférica, porém com atenuação na amplitude e atrasos na fase, que se tornam mais pronunciados à medida que a profundidade aumenta. A Figura 8 apresenta essa relação para diferentes profundidades de medição, denotadas por “z”.

Figura 8 – Influência da variação da temperatura atmosférica na temperatura do solo.



Fonte: Peres (2019).

A determinação da taxa de NSZD tradicionalmente é realizada por meio da medição da liberação de gás carbônico ou metano na superfície da fonte, podendo ser obtido pelo método do gradiente de concentração (LUNDEGARD; JOHNSON, 2006) ou pelo método dinâmico de câmara fechada (SIHOTA; SINGURINDY; MAYER, 2011). Os métodos, no entanto, apresentam desvantagem pela complexidade de instalação e manutenção (SWEENEY; RIRIE, 2014).

As previsões da taxa de biodegradação utilizando um modelo fenomenológico simplificado são afetadas por várias grandezas atmosféricas, além do perfil de temperatura. Essas grandezas incluem umidade do ar e do solo, precipitação atmosférica, temperatura ambiente, além de nível e temperatura do lençol freático, que acabam influenciando também propriedades do solo. O fato de considerar as heterogeneidades e incertezas na matriz do solo acaba limitando a eficácia das previsões feitas com base em modelos simplificados, porém modelos mais complexos podem demorar muito para serem resolvidos (PERES, 2019).

Crutcher, Rovers e McBean (1982) utilizaram a temperatura da subsuperfície para monitorar os processos de biodegradação em aterros sanitários. A partir dos resultados desse estudo pioneiro, abordagens baseadas em relações termodinâmicas passaram a ser avaliadas para a quantificação da taxa de NSZD. Pode-se citar, por exemplo, a documentação dos dados de temperatura resultantes da biodegradação

comparados com dados de um local semelhante, porém sem contaminação, por Davis *et al.* (2005). Em seguida, Sweeney e Ririe (2014) apontaram as oportunidades, limitações e potencial aplicação para essa técnica. Com essas informações, surgiram métodos que produzem estimativas de taxas de NSZD usando dados de temperatura de subsuperfície (STOCKWELL, 2015). Tais métodos foram aperfeiçoados para técnicas que não necessitam de medições em regiões sem contaminação para referência, com um conjunto de transdutores de temperatura medindo diversas profundidades do solo apenas na região contaminada para gerar uma estimativa on-line da taxa de NSZD (ASKARANI; SALE, 2020).

3.2 ESTUDO DE CASO: PREDIÇÃO DE TEMPERATURA DO STMF

A medição contínua das temperaturas de subsuperfície de regiões contaminadas permite a utilização de técnicas de modelamento dinâmico capazes de realizar a predição das dinâmicas em futuro próximo. Por meio da relação entre as temperaturas e a taxa de NSZD, tais predições podem, então, ser utilizadas para estimar o comportamento da taxa de NSZD em futuro próximo. Esta seção apresenta um estudo de caso com base na medição de temperaturas em uma área contaminada real.

3.2.1 Sistema de aquisição de dados

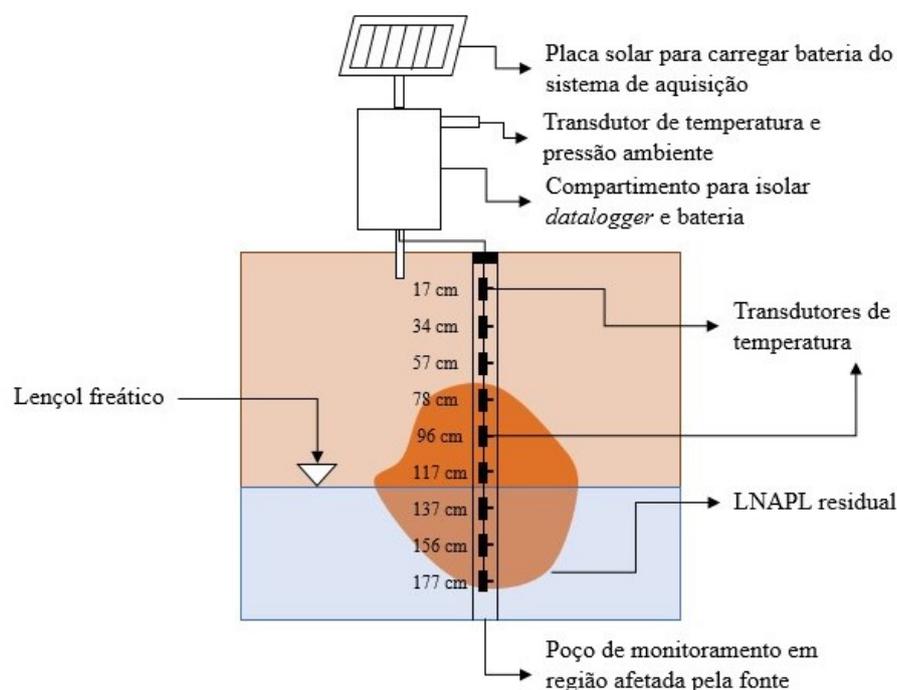
Este trabalho utilizou medições reais do perfil de temperatura de subsuperfície em uma área contaminada com gasolina no Núcleo Ressacada de Pesquisas em Meio Ambiente (REMA), da Universidade Federal de Santa Catarina (UFSC).

O Sistema Termodinâmico de Monitoramento de Fontes (STMF), é um conjunto de hardware e software para monitoramento remoto, *in situ*, de áreas contaminadas, visando a redução de custos de processo de monitoramento e remediação. De acordo com Peres (2019), seu objetivo principal é obter dados de biodegradação da fonte que sirvam como evidências da eficácia dos processos de NSZD. O STMF calcula a taxa de NSZD por meio da análise do calor liberado, que é estimado a partir das medições de temperatura no subsolo.

O STMF é composto por uma série de transdutores de temperatura distribuídos verticalmente para monitorar calor gerado pela depleção de massa de hidrocarbonetos de petróleo, conforme indicado na Figura 9.

A quantidade de transdutores de temperatura e o espaçamento entre os mesmos são fixos, com espaçamento de 20 cm. Adicionalmente, um sensor de concentração de metano, localizado próximo à superfície do solo, pode ser utilizado como uma segunda forma de estimar a depleção de massa de hidrocarbonetos. Essa estação apresenta comunicação sem fio, como aponta a Figura 10, fazendo com que o método de avaliação possa ser realizado de maneira remota e, conseqüentemente, diminuindo

Figura 9 – Esquemático de instalação para monitoramento de temperatura.



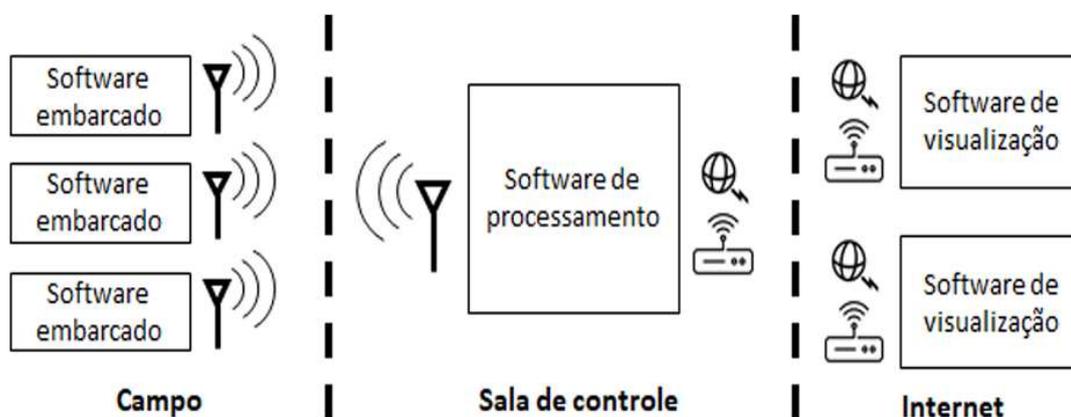
Fonte: Elaborado pela autora.

contato humano com a área contaminada (LIAE, 2019). As grandezas medidas com o STMF e consideradas neste trabalho são:

- pressão atmosférica;
- temperatura ambiente;
- gradiente de temperatura de subsolo em diferentes profundidades, a cada 20 cm;
- umidade relativa do ar e do solo;
- temperatura e nível do lençol freático;
- precipitação;
- radiação solar.

O sistema realiza medições do perfil de temperatura na região afetada pela contaminação e em uma região não afetada (*background*) de forma a permitir em análise posterior a separação do efeito de aumento de temperatura causado pela contaminação do efeito causado por mudanças na temperatura ambiente. Vale ressaltar que existem abordagens que trabalham só com uma haste de medição e não utilizam os dados do nível do lençol freático (ASKARANI; SALE, 2020). A umidade do solo também contribui para as variações no perfil de temperatura medido.

Figura 10 – Sistema responsável pela coleta e envio dos dados adquiridos pelo STMF.



Fonte: LIAE (2019)

3.2.2 Método para predição de séries temporais de temperaturas de subsolo

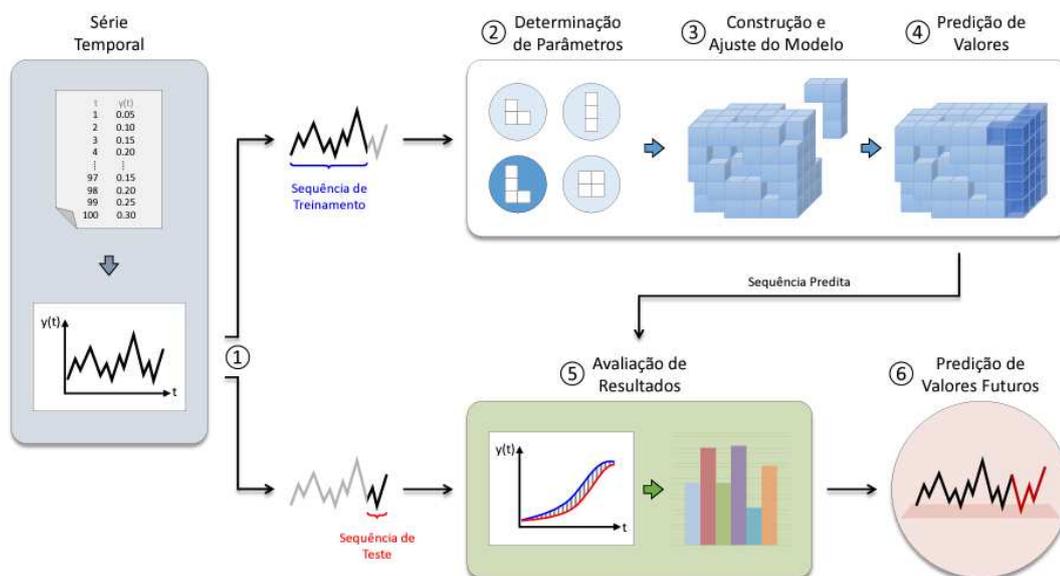
Os dados coletados no STMF, depois de agrupados, são classificados como STs, visto que atendem as características descritas na Seção 2.1. Esses dados são adequados para a aplicação de técnicas de predição, permitindo a construção de modelos explicativos que projetam informações históricas para o futuro, considerando diferentes horizontes temporais.

Neste trabalho, a predição do comportamento futuro de sistemas e séries temporais é associada com a utilização de redes neurais e do modelo ARIMA. Nesta seção, são exploradas abordagens gerais discutidas por Nelles (2002), Schwedersky (2017) e Parmezan (2016) para a predição de dados desse tipo, com adaptações para o problema específico tratado neste trabalho.

O processo de predição de uma ST genérica é representado na Figura 11, que esquematiza os passos fundamentais para modelar sistemas dinâmicos usando séries temporais. Esse esquema descreve o processo de modelagem de dados que é aplicável a qualquer um dos métodos de identificação usados neste trabalho. A diferença é a maneira de abordagem de cada método empregado para realizar a construção e o ajuste do modelo.

O método utilizado segue um procedimento de predição que inclui várias etapas. Primeiramente, antes de utilizar os dados para o treinamento, aplicou-se um filtro de média móvel com o objetivo de atenuar o efeito de ruídos de medição e de pontos discrepantes, como no caso de falha de leitura dos transdutores. Após essa preparação dos dados, eles foram divididos em duas partes distintas: uma destinada ao treinamento e ao ajuste dos modelos, conhecida como sequência de treinamento (85% dos dados), e a outra para avaliar a qualidade do modelo, chamada de sequência de teste

Figura 11 – Processo de predição de uma série temporal



Fonte: Parmezan (2016)

(15% dos dados). A determinação da divisão dos dados de treinamento foi estabelecida após uma análise detalhada da dinâmica dos dados, levando em consideração que tanto a parcela destinada ao treinamento quanto a parcela de teste refletem o padrão subjacente dos dados.

Além disso, a sequência de treinamento foi subdividida em duas partes adicionais: uma para o treinamento efetivo do modelo e outra para a validação dos principais hiperparâmetros. A sequência de teste desempenha um papel relevante na avaliação da capacidade preditiva do modelo, pois consiste em dados que não foram previamente observados durante o processo de treinamento. É importante destacar que essa divisão de dados foi mantida a mesma para os três modelos empregados neste estudo.

A estrutura do modelo de predição é adaptada às características dos dados, ou seja, os hiperparâmetros de cada modelo são ajustados de acordo com a dinâmica da ST. Para realizar essa adaptação, é necessário conduzir um estudo que avalie o impacto dos hiperparâmetros no desempenho dos modelos. Neste trabalho, foi proposta a aplicação de uma técnica heurística de busca, o AG, que possibilita a otimização dos hiperparâmetros para cada modelo específico. No caso do modelo ARIMA, foram ajustados os valores de p , d e q , conforme detalhado na Seção 2.3. Para o modelo baseado na rede LSTM, foram otimizados os hiperparâmetros relacionados à taxa de aprendizado inicial, ao fator e ao período de queda da taxa de aprendizado. Quanto ao modelo que utiliza a rede ESN, foram ajustados o raio espectral, a taxa de vazamento e o parâmetro de regularização. A utilização do AG como técnica de busca se mostra

mais eficiente em termos de velocidade em comparação à busca em grade, permitindo alcançar uma boa configuração para o modelo dentro de um tempo de treinamento razoável.

Com os hiperparâmetros definidos para cada modelo, passou-se à etapa de treinamento propriamente dita. Esse processo é tipicamente iterativo e procura encontrar os melhores valores para os pesos ou coeficientes dos modelos, de modo que métricas de desempenho sejam minimizadas. Neste trabalho, foi usada a raiz do erro médio quadrático (RMSE, do inglês *Root Mean Squared Error*) como métrica de desempenho para avaliação dos modelos durante a fase de treinamento.

Finalmente, após os modelos terem sido treinados e terem sido excitados com os dados de entrada das séries temporais de teste, foi avaliada a qualidade das previsões realizadas. Isso envolve a avaliação das métricas, que quantificam o grau de divergência entre os valores preditos e os valores reais da sequência de teste. Essa análise é importante para comparar diferentes modelos e determinar qual deles possui a melhor qualidade em dados não observados durante o ajuste do modelo.

3.2.3 Séries de dados utilizadas

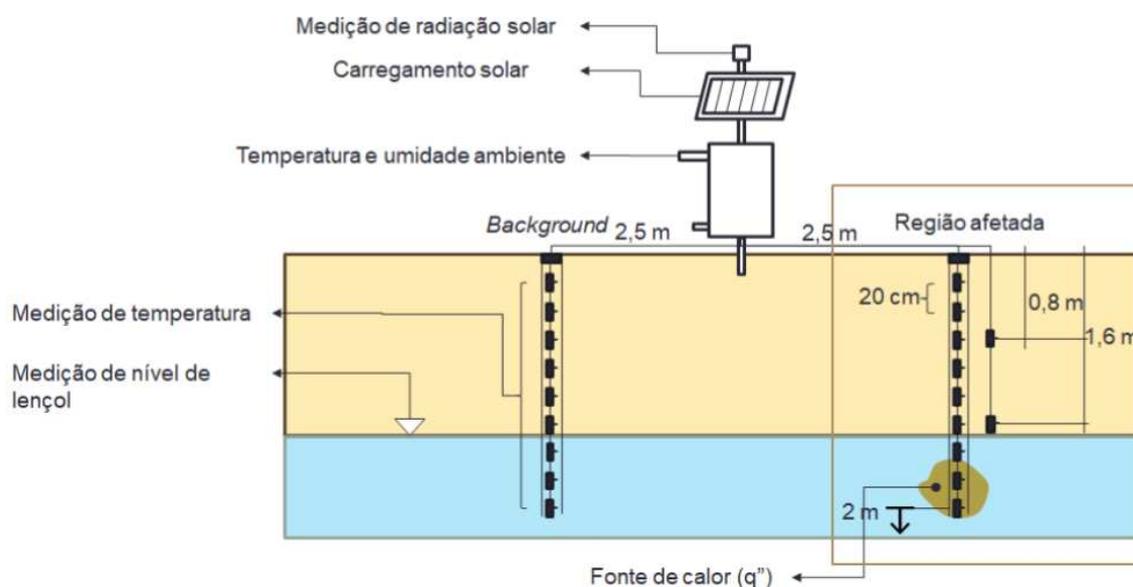
Para o modelamento de um sistema através de redes neurais recorrentes e pelo modelo ARIMA, é necessário obter séries de dados, que são adquiridos no STMF. Para este trabalho, foram empregados dados do perfil de temperatura do subsolo da área contaminada adquiridos durante o mês de setembro de 2022 em uma região de contaminação real no REMA.

Em posse dos dados obtidos no STMF, é possível realizar a implementação proposta neste projeto. Porém, como forma de minimizar o efeito de ruídos, interferências e pontos espúrios obtidos por erros de medição, foi realizada uma etapa de pré-processamento. Para isso, foram implementados um filtro do tipo média móvel e um procedimento que identifica valores que estejam fora de uma faixa razoável de temperatura para as medições realizadas. O filtro de média móvel realiza o cálculo da média de 100 valores consecutivos, buscando suavizar interferências e evidenciar tendências de longo prazo. Isso proporciona uma visão mais nítida do perfil de temperatura ao longo do tempo. Por sua vez, o procedimento para identificar valores fora de determinada faixa de amplitudes de temperatura foi empregado para identificar pontos espúrios, advindos de erros de medição.

O sistema realiza medições do perfil de temperatura na região afetada pela contaminação e em uma região não afetada (*background*), de forma a permitir em análise posterior a separação do efeito de aumento de temperatura causado pela contaminação do efeito causado por mudanças na temperatura ambiente. A Figura 12 apresenta um diagrama simplificado de como foi instalada a estação no primeiro momento de monitoramento. Nela é possível observar como estão alocados os instrumentos de

medição de temperatura em dois locais diferentes, empregados para adquirir os dados empregados neste projeto.

Figura 12 – Esquemático de instalação para monitoramento de solo com o STMF.



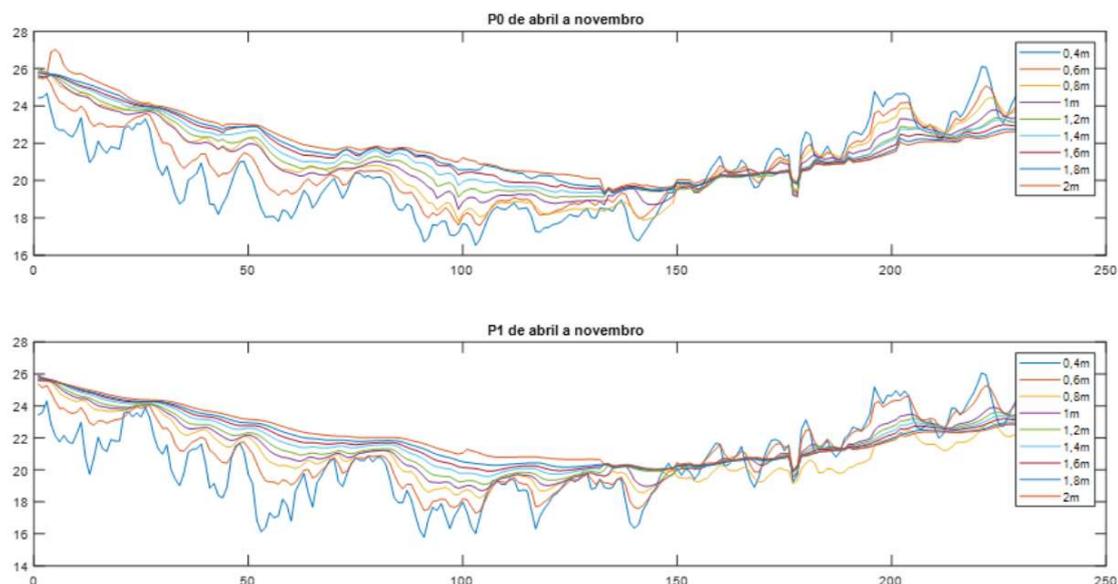
Fonte: LIAE (2019)

Os dados de temperatura adquiridos são apresentados na Figura 13. As principais informações medidas são os perfis de temperatura na região P0, em área contaminada e os perfis de temperatura P1, da região não afetada. Foram identificados problemas de padronização da taxa de aquisição nos dados utilizados. Além disso, ocorreram erros de leitura nos transdutores em determinados momentos, o que exigiu a aplicação dos filtros previamente descritos para corrigir essas inconsistências. Adicionalmente, a ocorrência de eventos no sítio de medição reduziu a quantidade de dados que podem ser utilizados nos estudos.

Para este trabalho, são utilizados apenas os dados dos transdutores localizados nas profundidades de 17 cm, 96 cm e 117 cm, como forma de avaliar medições com diferentes características, sendo uma mais próxima da superfície, uma mais profunda mas não em contato direto com o lençol freático e outra que pode sofrer influência direta do lençol. Assim, espera-se que as medições possuam dinâmicas variadas, o que permite uma comparação mais justa entre as técnicas de modelamento dinâmico que são avaliadas para estimar o comportamento futuro. Os perfis de temperatura desses dados são apresentados na Figura 14.

O objetivo no modelamento dinâmico das séries temporais das temperaturas de subsuperfície nas posições selecionadas consiste em uma estimativa a curto prazo. Os dados originais foram adquiridos com um período de amostragem de 5 minutos. No

Figura 13 – Séries temporais resultante das medições



Fonte: Dados da pesquisa.

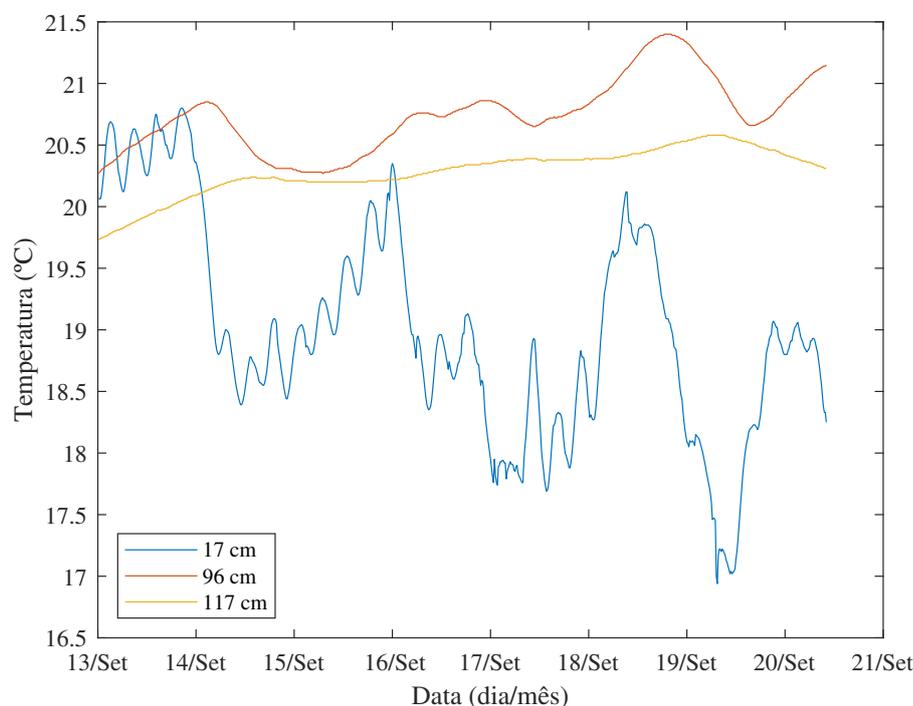
entanto, para o modelamento dinâmico foi feita uma subamostragem para uma amostra a cada 1 hora, uma vez que as dinâmicas não apresentam variações significativas em curtos períodos de tempo. Cada série temporal é composta por dados referentes a 1 mês de medições de temperatura dos transdutores localizados em cada uma das três posições mencionadas. Cada série possui 742 pontos, sendo que os dados referentes aos últimos 42 pontos foram reservados para o conjunto de teste dos modelos. O restante (700 pontos) foi utilizado para o treinamento dos modelos.

Além da filtragem, os dados foram submetidos a um pré-processamento para evitar problemas no aprendizado dos modelos. Antes de serem utilizados nos modelos do tipo RNA, tanto o conjunto de entrada quanto o conjunto de saída foram normalizados, levando-os para uma escala entre -1 e $+1$. Essa normalização é importante para melhorar a estabilidade dos algoritmos de treinamento, evitando problemas causados por valores extremamente altos nos dados de entrada que poderiam saturar os neurônios da camada inicial. No contexto deste trabalho, a normalização do conjunto de saída é particularmente importante, pois alguns modelos utilizam conexões recorrentes a partir da camada de saída, tratando a própria saída como uma entrada adicional que também requer o processo de normalização (SCHWEDERSKY, 2017).

3.2.4 Avaliação da qualidade da predição

A seleção de algoritmos para a predição de séries temporais dinâmicas envolve a combinação de conhecimento especializado com métodos de avaliação empírica.

Figura 14 – Perfis de temperatura empregados para realizar o modelamento



Fonte: Dados da pesquisa.

A questão-chave nesse processo é determinar quais índices ou métricas devem ser considerados para avaliar e comparar o desempenho dos diferentes métodos.

Pensando nisso, o cálculo de medidas de erro desempenha um papel relevante ao quantificar a discrepância entre os valores reais observados e aqueles preditos pelo modelo ajustado. No entanto, quando o horizonte de predição é suficientemente longo, é importante utilizar índices de desempenho mais abrangentes, que permitam uma avaliação mais completa da qualidade dos dados projetados. Esses índices fornecem uma estimativa mais precisa do verdadeiro erro de um determinado algoritmo quando aplicado a uma série temporal (PARMEZAN, 2016).

Existem várias métricas para avaliação dessa qualidade, dentre elas estão: erro médio percentual absoluto (MAPE, do inglês *Mean Absolute Percentage Error*), erro quadrático médio (MSE, do inglês *Mean Squared Error*), RMSE, erro médio quadrático ponderado (WMSE, do inglês *Weighted Mean Squared Error*), erro médio absoluto (MAE, do inglês *Mean Absolute Error*) e o coeficiente de determinação (R^2). Cada uma delas possui suas características específicas, e a escolha da métrica adequada depende do objetivo da análise e da interpretação desejada dos resultados.

3.2.4.1 Erro médio percentual absoluto

O MAPE mede a porcentagem média do erro absoluto dos valores reais com os valores preditos, fornecendo uma medida de quão próximas as predições estão dos valores reais. No entanto, o MAPE pode ser sensível a valores extremos ou zeros nos dados, o que pode afetar sua interpretação. O MAPE é calculado pela seguinte equação:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100, \quad (13)$$

onde:

- n é o número de observações;
- y_i é o valor real observado;
- \hat{y}_i é o valor predito pelo modelo.

3.2.4.2 Erro médio absoluto

O MAE calcula a média das diferenças absolutas entre os valores preditos e os valores reais, fornecendo uma medida direta do desvio médio do modelo em relação aos dados observados. Quanto menor o valor do MAE, mais preciso é o modelo em suas predições ou estimativas. O MAE é definido como:

$$MAE = \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|. \quad (14)$$

3.2.4.3 Erro quadrático médio

O MSE calcula a média dos erros quadrados entre os valores reais e os valores preditos pelo modelo e fornece uma medida de dispersão dos erros. O MSE é uma métrica que atribui maior peso a erros maiores devido ao uso do quadrado dos desvios, o que pode ser útil em casos nos quais é importante penalizar erros maiores de forma mais significativa. O MSE é calculado pela seguinte equação:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2. \quad (15)$$

3.2.4.4 Raiz do erro quadrático médio

O RMSE é uma versão modificada do MSE que calcula a raiz quadrada da média dos erros quadrados entre os valores reais e os valores preditos pelo modelo. O RMSE fornece uma medida da dispersão dos erros em relação à escala dos dados. O

RMSE é calculado pela seguinte equação:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}. \quad (16)$$

3.2.4.5 Erro médio quadrático ponderado

O WMSE é uma métrica utilizada para avaliar a qualidade de uma predição levando em consideração diferentes pesos para cada observação. Essa métrica é útil quando há a necessidade de atribuir maior importância a determinadas observações específicas no cálculo do erro, o que pode ser útil em situações em que algumas observações são consideradas mais relevantes para a análise. O WMSE é calculado pela equação:

$$WMSE = \frac{1}{\sum_{i=1}^n w_i} \sum_{i=1}^n w_i \cdot (y_i - \hat{y}_i)^2, \quad (17)$$

onde w_i é o peso atribuído a cada observação.

3.2.4.6 Coeficiente de determinação

O coeficiente de determinação, denotado como R^2 , é uma métrica estatística que mede a proporção da variância dos dados dependentes que é explicada pelo modelo de regressão. Ele fornece uma medida da qualidade do ajuste do modelo aos dados observados. O coeficiente de determinação varia entre -1 e $+1$, onde 0 indica que o modelo não explica nenhuma variabilidade e $+1$ indica que o modelo explica toda a variabilidade dos dados. O coeficiente é definido como:

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}, \quad (18)$$

onde:

- SS_{res} é a soma dos quadrados dos resíduos (a diferença entre os valores reais e os valores predito pelo modelo);
- SS_{tot} é a soma total dos quadrados (a variabilidade total dos valores reais em relação à média).

O coeficiente de determinação é uma métrica amplamente utilizada na análise de regressão para avaliar o desempenho e a capacidade de explicação de um modelo. No entanto, é importante considerar outras métricas e análises para uma avaliação completa do modelo, como a significância dos coeficientes, a análise de resíduos e a validação cruzada.

3.3 CONSIDERAÇÕES FINAIS

Neste capítulo foi apresentado o método utilizado para avaliar as arquiteturas de redes neurais recorrentes e o modelo ARIMA. Além do método, o capítulo apresentou

os dados empregados para o estudo de caso que serviu como base para a avaliação experimental das arquiteturas de redes neurais. Para isso, foi detalhado o processo de aquisição de dados que foi realizado para obter as séries de dados usadas nas etapas de predição.

Enquanto este capítulo aborda o primeiro passo do método, que é comum a todos os modelos, o Capítulo 4 descreve os passos subsequentes, que são apresentados para cada tipo de modelo considerado neste trabalho. A análise dos resultados apresentada no próximo capítulo considera os dados do conjunto de teste.

4 RESULTADOS DA PREDIÇÃO DE TEMPERATURAS DE DEPLEÇÃO DE HIDRO-CARBONETOS

Neste capítulo, é apresentado o desenvolvimento dos modelos para predição de temperaturas de depleção de hidrocarbonetos. Para cada modelo, é apresentado o processo realizado para criação e ajuste do mesmo, seguindo a metodologia única para cada modelo, a qual foi apresentada na Subseção 3.2.2. Três arquiteturas de modelos são apresentadas neste trabalho: a primeira consistiu em um modelo do tipo ARIMA, o qual é usado como base de comparação para os modelos mais complexos, baseados em redes neurais recorrentes; a segunda consiste em uma rede neural recorrente com blocos de memória do tipo LSTM; e a terceira é uma rede de estado de eco.

Na Seção 4.1, são apresentados o desenvolvimento e resultados obtidos para a predição com modelo ARIMA. Os resultados para a rede com memórias LSTM é apresentado na Seção 4.2. Em seguida, na Seção 4.3, são apresentados a criação, a parametrização e os resultados para o modelo baseado na ESN. Por fim, na Seção 4.4, é apresentada uma comparação das diferentes abordagens de predição desenvolvidas neste trabalho, levando em conta não apenas os resultados para a tarefa de predição, mas também aspectos relacionados à aplicabilidade na indústria de tais técnicas, como os tempos demandados para obtenção de cada modelo.

4.1 PREDIÇÃO COM MODELO ARIMA

Nesta seção, é apresentado o desenvolvimento do modelo de predição de temperaturas de depleção utilizando a estrutura de modelo ARIMA. Na Subseção 4.1.1, é apresentado o processo de criação e seleção dos principais hiperparâmetros associados ao modelo ARIMA, baseado no uso de algoritmo genético (AG). Os resultados obtidos com as configurações selecionadas para cada modelo ARIMA são detalhados na Subseção 4.1.2.

4.1.1 Otimização do modelo ARIMA com algoritmo genético

A criação de um modelo de predição utilizando a estrutura ARIMA requer uma etapa de seleção da configuração do modelo. Nesse tipo de modelo, podem ser utilizadas três parcelas, cujas formulações e utilidade são detalhadas na Seção 2.3. Tipicamente, a escolha da estrutura do modelo pode ser realizada usando um método, como o apresentado em Box *et al.* (2015), no qual os modelos são determinados a partir da definição das ordens p , d e q do modelo ARIMA. Já na etapa de identificação da ordem, são utilizadas ferramentas estatísticas, como gráficos da função de autocorrelação (ACF) e autocorrelação parcial (PACF), a fim de identificar a ordem apropriada do modelo ARIMA. Depois disso, seus coeficientes são ajustados e os resultados do modelo passam por um diagnóstico para avaliar a qualidade do ajuste.

De forma a realizar a seleção automática da estrutura do modelo ARIMA, nesta pesquisa, foi desenvolvida uma ferramenta baseada em AG para otimização da estrutura, com base no exposto na Seção 2.5. Essa ferramenta é capaz de escolher os valores p , d e q , os quais definem o funcionamento do modelo, utilizando o procedimento iterativo do AG. Esse método usa dados da depleção para treinar modelos ARIMA com determinada configuração. A faixa de valores considerados para cada parâmetro, faixa de 0 e 2, e foi definida com base na análise de ACF e PACF. Para esse modelo simples, poderia ter sido realizada uma busca em grade para determinar as melhores configurações, mas optou-se por empregar o AG para manter o padrão do método proposto.

A utilização da ferramenta de otimização da estrutura do modelo ARIMA requer a pré-definição de alguns parâmetros que definem o funcionamento do AG. É necessário ajustar aspectos relacionados à população a ser gerada, além dos critérios de parada para garantir que o algoritmo tenha convergência para a melhor resposta. Inicialmente, é selecionada a população inicial do modelo. Devido ao fato de os modelos considerados neste estudo apresentarem uma quantidade variada de parâmetros a serem otimizados, suas matrizes referentes às populações iniciais são diferentes. Dessa forma, a população inicial para o modelo ARIMA é definida por uma matriz cuja dimensão é de 3, referente a p , d e q pelo número de indivíduos na população (n_{pop}).

A função de adaptação é utilizada para avaliar a aptidão de cada indivíduo em relação aos demais e, neste trabalho, ela é determinada pelo cálculo do RMSE entre a saída medida e a saída predita. Ao final de cada iteração, calcula-se a RMSE obtida, que representa a aptidão de cada indivíduo na população. As aptidões são usadas, então, para fins de comparação de aplicação dos operadores genéticos. A Tabela 1 apresenta os parâmetros de configuração do algoritmo para cada um dos modelos de predição considerados neste estudo.

Tabela 1 – Configuração da ferramenta de otimização baseada em AG para o modelo ARIMA

Parâmetro	Valor
Número de variáveis	3
Número de indivíduos na população	10
Número de genes	5
Número de gerações/iterações	5
Número de recombinações	1
Número de indivíduos no duelo	2
Número de mutações	1
Número de elite	2

Fonte: Dados da pesquisa.

A ferramenta de ajuste é aplicável a qualquer série temporal dos dados de depleção que tenha o mesmo comportamento que o perfil utilizado como referência.

As configurações ótimas obtidas por meio do AG são apresentadas na Tabela 2 e essas configurações são utilizadas para realizar o treinamento do modelo, empregando os dados experimentais.

Tabela 2 – Configurações ótimas para o modelo ARIMA

Medição	p	d	q
17 cm	0	1	2
96 cm	2	1	2
117 cm	2	1	1

Fonte: Dados da pesquisa.

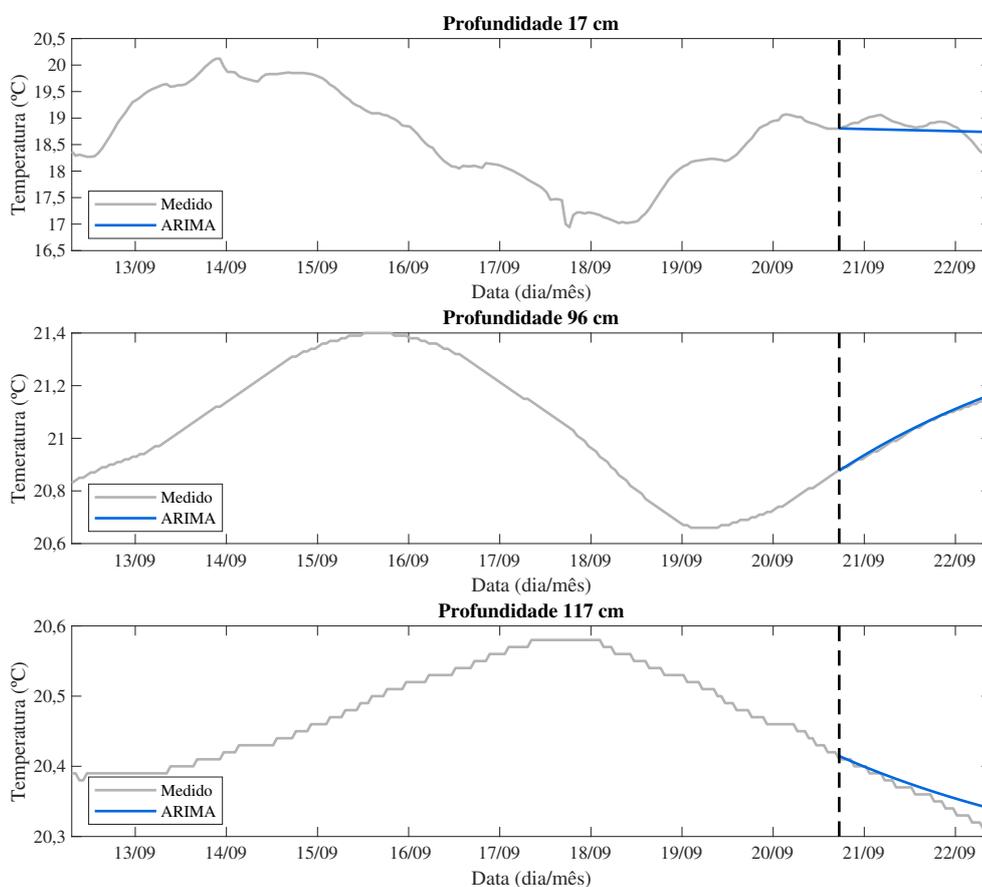
4.1.2 Resultados para o modelo ARIMA

Nesta seção, são apresentados os resultados obtidos pelo modelo ARIMA para três diferentes profundidades das sondas de medição de temperatura: 17 cm, 96 cm e 117 cm. Esses resultados estão representados na Figura 15 e foram obtidos utilizando as configurações otimizadas pelo AG. É importante ressaltar que, para cada perfil de temperatura, os resultados correspondem a uma porção do ensaio que não fez parte do conjunto de dados usado para determinar as configurações do modelo, garantindo, assim, a validade das análises subsequentes. Após a apresentação dos resultados, é realizada uma análise gráfica das predições das temperaturas, sustentada por uma avaliação com emprego das métricas de desempenho descritas na Subseção 3.2.4.

Uma etapa importante antes da realização das predições de temperatura é a inicialização do modelo. Para garantir que o modelo esteja em um estado inicial adequado, ele é alimentado com uma parte inicial da série de dados correspondentes (treinamento ou validação) antes de realizar as primeiras predições. Após essa etapa inicial, os modelos são executados em uma configuração recursiva, em que cada predição é usada como entrada para a próxima, emulando assim o comportamento de uma tarefa de predição do futuro, na qual as saídas reais não estão disponíveis. O objetivo é predizer N amostras à frente, sendo que os resultados dessa predição são empregados para a análise subsequente do desempenho de cada modelo.

A análise dos resultados referentes à profundidade de 17 cm, tanto pela inspeção da Figura 15 quanto pelas métricas apresentados na Tabela 3, indica que o modelo ARIMA apresentou um desempenho insatisfatório, já que o modelo não foi capaz de acompanhar as leves variações de temperatura que acontecem nas medições mais próximas à superfície do solo. A análise das métricas suporta essa conclusão, já que valores elevados para as métricas RMSE (0,191 °C) e WMSE (3,646 °C) foram verificados, o que indica que as predições estão distantes dos valores ideais. Por sua vez, o coeficiente de determinação (R^2) foi baixo (0,075), sugerindo que o modelo explica apenas uma pequena parte da variabilidade dos dados observados.

Figura 15 – Resultados para modelos ARIMA avaliados na tarefa da predição a temperatura em três profundidades distintas.



Fonte: Dados da pesquisa.

Tabela 3 – Métricas de desempenho para o modelo ARIMA

Métrica	17 cm	96 cm	117 cm
RMSE (°C)	0,191	0,008	0,015
WMSE (°C ²)	3,646	0,183	0,163
MAE (°C)	0,155	0,007	0,011
MAPE (%)	0,371	0,0018	0,027
MSE (°C)	0,036	10 ⁻⁴	2 × 10 ⁻⁴
R ²	0,075	0,988	0,767

Fonte: Dados da pesquisa.

Ao analisar os resultados referentes à profundidade de 96 cm, é observado que o modelo foi capaz de modelar o comportamento da temperatura com maior fidelidade que para a profundidade de 17 cm, apresentando um desempenho mais próximo do ideal. Os valores de RMSE (0,008 °C) e WMSE (0,183 °C) são consideravelmente menores em comparação aos observados para a profundidade de 17 cm, indicando que as predições estão mais próximas dos valores reais. A métrica MAPE (0,018%)

também apresentou valores melhores que os verificados para a profundidade de 17 cm, com uma porcentagem de erro médio absoluto mais baixa, o que indica uma melhor precisão das predições. Para a métrica coeficiente de determinação (R^2), foi verificado um valor de 0,988, demonstrando que o modelo explica uma parte mais substancial da variação nos dados.

Por fim, a análise dos resultados referentes à profundidade de 117 cm mostra um desempenho semelhante ao da profundidade de 96 cm. O modelo ARIMA obteve valores de RMSE (0,015 °C) e WMSE (0,163 °C) mais baixos em comparação à profundidade de 17 cm, indicando predições mais precisas. A métrica MAPE (0,027%) também apresentou uma porcentagem de erro médio absoluto relativamente baixa. Para o coeficiente de determinação (R^2) foi verificado um valor mais próximo da unidade (0,767), indicando que o modelo explica uma porção significativa da variabilidade dos dados observados.

Portanto, com base nas métricas de avaliação de desempenho, é possível concluir que o modelo ARIMA apresentou um desempenho que variou de acordo com a profundidade de medição. Em profundidades com dinâmicas mais complexas, o modelo teve dificuldades em capturar a natureza não linear e oscilatória dos dados, resultando em predições menos precisas. No entanto, para profundidades com dinâmicas mais simples, o modelo mostrou um desempenho relativamente melhor, com predições mais próximas dos valores reais.

4.2 PREDIÇÃO COM A REDE NEURAL LSTM

Nesta seção, é apresentada a elaboração de modelos baseados em uma arquitetura de RNN com memórias LSTM, os quais são usados para modelar o comportamento das temperaturas associadas à depleção de hidrocarbonetos de petróleo. Por simplicidade, esses modelos são denominados apenas como modelos LSTM no decorrer do trabalho. Os procedimentos para criação, configuração e otimização dos hiperparâmetros do modelo LSTM são apresentados na Subseção 4.2.1. Os resultados obtidos pelos modelos LSTM são discutidos na Subseção 4.2.2.

4.2.1 Otimização da rede LSTM com Algoritmo Genético

Os modelos baseados na rede neural com camadas LSTM podem ser usados para modelar séries temporais utilizando a formulação apresentada subseção 2.4.3.

A configuração da rede possui uma estrutura simples com um único recurso de entrada, uma camada LSTM com 10 unidades ocultas, uma camada totalmente conectada para regressão com uma única saída e uma camada de regressão para calcular o erro. As opções de treinamento incluem o uso do otimizador do algoritmo de estimativa de momento adaptativo (ADAM, do inglês *Adaptive Moment Estimation*),

combinando momento e adaptação de taxa de aprendizado para acelerar o treinamento e lidar com gradientes ruidosos. É eficaz para problemas não convexos e oferece robustez na escolha de hiperparâmetros. Além disso, é definido o limite de gradiente de 1, 500 épocas máximas de treinamento, uma taxa de aprendizado inicial de 0,01 com programação de taxa de aprendizado *piecewise*, que ajusta a taxa de aprendizado em diferentes estágios do treinamento, normalmente reduzindo-a em etapas predefinidas para melhorar a convergência e evitar oscilações. A taxa de aprendizado é reduzida pela metade a cada 100 épocas. O diferencial do modelo é a existência de erro no código, pois o modelo está sendo atualizado com dados de treinamento em vez de dados de teste em suas iterações de previsão para a etapa de teste.

Além desses parâmetros fixos, existem 3 hiperparâmetros que definem a dinâmica do aprendizado, sendo eles a taxa de aprendizado (TA) inicial, período e fator de queda da taxa de aprendizado (QTA). Para otimização do processo, assim como realizado para o modelo ARIMA, foi implementado um AG desses hiperparâmetros. Para isso, foram definidos os limites máximos e mínimos de cada sendo: TA inicial 10^{-5} a 0,5; fator de QTA 10^{-2} a 300; e período QTA 10^{-5} a 10^3 . Os parâmetros considerados para configuração do AG são apresentados na Tabela 4. A população inicial foi definida com uma matriz de dimensão $4 \times n_{pop}$.

Tabela 4 – Parâmetros considerados para o AG para o modelo LSTM

Parâmetro	Valor
Número de variáveis	4
Número de indivíduos na população	50
Número de genes	10
Número de gerações/iterações	20
Número de recombinações	10
Número de indivíduos no duelo	5
Número de mutações	10
Número de elite	3

Fonte: Dados da pesquisa.

Após realizar o processo de otimização dos hiperparâmetros, o AG retorna as melhores combinações de parâmetros para cada uma das profundidades. Os valores finais, obtidos para o modelo relacionado a cada temperatura são iguais, sendo TA inicial 0,01, fator de QTA 0,05 e período de QTA 100.

A validação do modelo desempenha um papel fundamental na avaliação da sua capacidade de generalização. Para validação, foi realizado um processo de divisão dos dados em conjuntos de treinamento e validação, seguindo a mesma estratégia adotada para o modelo ARIMA. No caso do AG, essa divisão inclui uma porção de treinamento e validação das séries temporais, definidas previamente e mantidas consistentes para todos os modelos avaliados. Durante a validação, o AG treina modelos usando o conjunto de treinamento e avalia o desempenho da configuração que está sendo

testada utilizando o conjunto de validação, permitindo ajustar os hiperparâmetros e prevenir problemas de sobreajuste. É importante destacar que o conjunto de testes não é utilizado nesse processo, sendo reservado exclusivamente para a etapa subsequente, que avalia os modelos finais, fazendo uso de métricas relevantes detalhadas na Seção 3.2.4.

Utilizando os parâmetros ótimos obtidos por meio do AG, prosseguiu-se com o treinamento dos modelos finais para cada temperatura especificada. Com essas configurações aprimoradas, os modelos foram capazes de capturar com maior precisão os padrões intrínsecos às séries temporais, o que se refletiu em um desempenho superior durante a avaliação no conjunto de validação.

Uma vez concluídas as etapas de configuração e validação do modelo LSTM, torna-se viável realizar testes finais no conjunto de teste, com o intuito de avaliar o desempenho do modelo em dados não utilizados durante o treinamento. Essa etapa final é essencial para obter uma estimativa precisa do desempenho do modelo em cenários do mundo real e para fornecer uma validação robusta das capacidades do modelo na tarefa de predição de séries temporais.

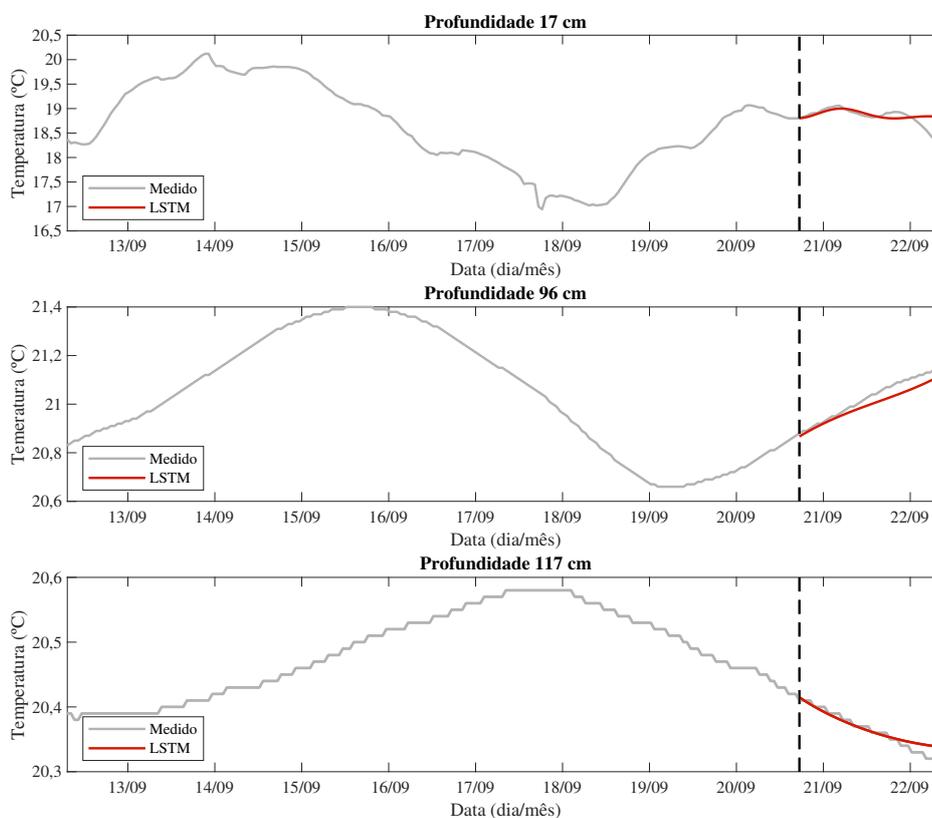
4.2.2 Resultados para o modelo LSTM

A análise dos resultados alcançados pelo modelo LSTM na predição de séries temporais de temperatura foi realizada de forma semelhante ao apresentado para o modelo ARIMA, com o modelo sendo testado na tarefa de predição. A análise da Figura 16 revela que o modelo desenvolvido é capaz de acompanhar de forma eficaz o comportamento geral do fenômeno modelado. Em particular, destaca-se o desempenho notável para a profundidade de 17 cm, na qual o modelo demonstrou uma capacidade superior em capturar as oscilações da temperatura, um aspecto que não foi adequadamente refletido pelo modelo ARIMA. Esse resultado sugere que a abordagem otimizada pelo AG oferece uma maior adaptação aos padrões específicos das séries temporais, proporcionando uma predição mais precisa e detalhada em comparação com métodos tradicionais.

A análise dos resultados das métricas, os quais são apresentados na Tabela 5. Para a profundidade de 117 cm, o LSTM alcança o menor RMSE de 0,015 °C e o maior coeficiente de determinação (R^2) de 0,734, indicando um ajuste muito bom aos dados. À medida que a profundidade diminui, o RMSE aumenta gradualmente, mas o modelo ainda mantém um desempenho sólido, com RMSE de 0,031 °C para 96 cm e 0,187 °C para 17 cm. Essa análise sugere que o LSTM é particularmente eficaz em previsões de temperatura em profundidades mais profundas, mas mantém um desempenho aceitável em profundidades menores.

Esses resultados destacam a eficácia do modelo LSTM na predição de séries temporais de temperatura em diferentes profundidades de medição. A utilização do al-

Figura 16 – Resultados obtidos com o modelo LSTM avaliado no conjunto de testes para a tarefa de predição de múltiplos instantes à frente.



Fonte: Dados da pesquisa.

Tabela 5 – Métricas de desempenho para o modelo LSTM

Métrica	17 cm	96 cm	117 cm
RMSE (°C)	0,187	0,031	0,015
WMSE (°C ²)	1,286	0,500	0,320
MAE (°C)	0,114	0,027	0,014
MAPE (%)	0,272	0,066	0,035
MSE (°C)	0,0351	0,001	0,0003
R ²	0,136	0,858	0,734

Fonte: Dados da pesquisa.

goritmo genético para otimização dos parâmetros de ajuste do modelo LSTM contribuiu para a obtenção de modelos adequados, resultando em um desempenho satisfatório para a tarefa de predição, com o modelo sendo capaz de se ajustar aos dados medidos. No entanto, é importante destacar uma limitação significativa do modelo LSTM, que é a demanda por um alto poder computacional durante o processo de treinamento. Esse requisito resulta em tempos de treinamento prolongados, o que pode ser problemático

em aplicações que exigem adaptação rápida do modelo conforme novas medições são recebidas. Esse desafio é particularmente relevante para a aplicação do AG, pois envolve a avaliação de várias configurações de parâmetros. Para contextualizar essa limitação, é útil considerar que o tempo médio de treinamento da LSTM em comparação com métodos alternativos, como ESN e ARIMA, pode ser substancialmente maior, prejudicando a eficiência do AG ao explorar um espaço de parâmetros extenso. Portanto, embora a LSTM demonstre um excelente desempenho de predição, suas demandas computacionais podem ser uma restrição em cenários nos quais o tempo é um fator crítico.

4.3 PREDIÇÃO COM REDE ESN

Nesta seção, são apresentadas a construção e a análise dos resultados obtidos com redes neurais de estado de eco para a tarefa de predição de temperaturas. O funcionamento da ESN, bem como sua formulação e treinamento, foram discutidos na Seção 2.4. Nesta seção, são apresentados os passos relativos à criação de modelos ESN para a tarefa de predição, sendo detalhados os passos restantes da metodologia de predição da série temporal, que incluem a definição da estrutura do modelo e seus parâmetros, apresentados na Seção 4.3.1, e a etapa de validação dos resultados, detalhada na Seção 4.3.2.

4.3.1 Otimização da rede neural ESN com Algoritmo Genético

A construção do modelo ESN envolve a definição da estrutura da rede e o ajuste dos principais hiperparâmetros que ajustam o funcionamento do modelo. A estrutura da ESN é composta por três camadas, a camada de entrada, a camada de reservatório e a camada de saída. A camada de entrada recebe os dados de entrada, que podem ser as séries temporais ou outros tipos de informações relevantes para a predição. A camada de reservatório é composta por uma quantidade grande de neurônios, os quais são interconectados de forma recorrente, por meio de conexões definidas aleatoriamente. Essa camada possui uma dinâmica não linear e é responsável por processar e representar os padrões presentes nos dados de entrada. Por fim, a camada de saída mapeia a saída do reservatório para a variável de saída do modelo, que é a temperatura a ser predita.

Conforme apresentado a arquitetura na Seção 2.4.2, foi fixado que o modelo possui uma camada de entrada e uma camada de saída, sendo o número de entradas e saídas igual a 1. As conexões da camada de entrada para a camada de saída estão desativadas, mas as conexões da camada de entrada para o reservatório estão ativas. O tamanho do reservatório é definido como 1000 neurônios.

O método de treinamento é regressão com regularização e lasso (do inglês *train*

Ridge regression) para treinar a camada de leitura e inclui parâmetros de regularização. Esse método busca encontrar pesos que minimizem o erro entre as predições do modelo e os valores reais, enquanto adiciona uma penalização para controlar a complexidade do modelo e evitar *overfitting*. A regularização Ridge é aplicada através da introdução de um termo de penalização L2 nos coeficientes dos pesos.

Existem diferentes abordagens para otimizar os hiperparâmetros da ESN, sendo uma delas o uso de AG, assim como para os demais modelos. Os parâmetros que podem ser otimizados incluem a taxa de vazamento, parâmetro regularização e o raio espectral. O objetivo é encontrar uma combinação ótima desses parâmetros que resulte em um desempenho superior do modelo na tarefa de predição.

Com base no método descrito na Seção 2.5, é necessário atribuir a margem de busca dos parâmetros da população inicial com uma matriz de dimensão $3 \times n_{pop}$ e para seus operadores genéticos. Essa margem é de variação para o raio espectral varia entre 0,8 a 0,9999999, a taxa de vazamento entre 0 e 1 e os parâmetros de regularização entre 10^{-5} e 10^3 .

A avaliação do desempenho dos indivíduos em um algoritmo genético é fundamental para determinar sua aptidão. Neste estudo, a aptidão dos indivíduos é medida pelo RMSE entre a saída medida e a saída predita. Ao final de cada iteração, o RMSE é calculado e utilizado para comparar a aptidão dos indivíduos na população. Os parâmetros de configuração do algoritmo para cada modelo de predição considerado neste estudo são apresentados na Tabela 6.

Tabela 6 – Parâmetros considerados para o AG para o modelo ESN

Parâmetro	Valor
Número de variáveis	3
Número de indivíduos na população	50
Número de genes	10
Número de gerações/iterações	20
Número de recombinações	10
Número de indivíduos no duelo	5
Número de mutações	10
Número de elite	3

Fonte: Dados da pesquisa.

Os resultados finais das combinações otimizadas dos parâmetros, incluindo raio espectral, taxa de vazamento e parâmetro de regularização, para o modelo podem ser encontrados na Tabela 7. Essa tabela apresenta as configurações finais que foram identificadas pelo algoritmo genético como as mais adequadas para obter o melhor desempenho do modelo. Ao analisar os resultados apresentados, é possível observar as combinações de parâmetros que proporcionam um menor erro médio quadrático e, conseqüentemente, um melhor ajuste e predição das séries temporais para as profundidades consideradas neste estudo.

Tabela 7 – Valores dos parâmetros para o modelo ESN

Medição	Raio espectral	Taxa de vazamento	Parâmetro regularização
17 cm	0,99	0,4	0,01
96 cm	0,98	0,04	0,01
117 cm	0,99	0,16	845

Fonte: Dados da pesquisa.

Assim como foi feito para os demais tipos de modelo, as redes ESN treinadas para cada uma das profundidades foram avaliadas com dados do conjunto de teste, que não foi utilizado durante o treinamento.

4.3.2 Resultados para rede ESN

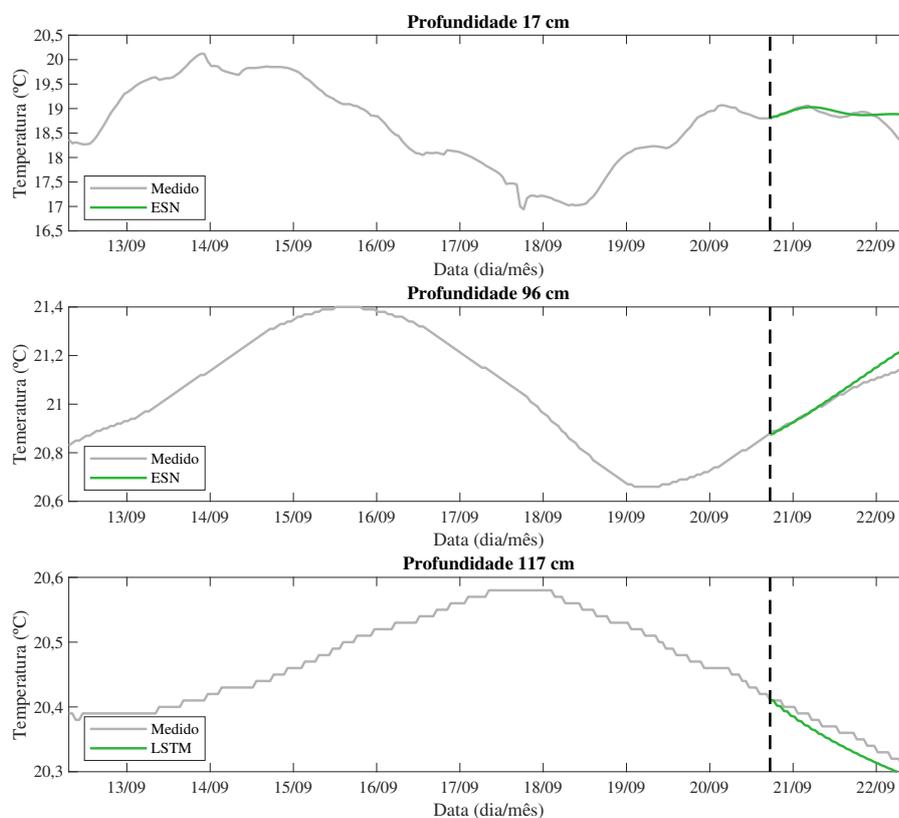
A seção de resultados apresenta os principais resultados obtidos através da aplicação do AG para ajuste dos hiperparâmetros do modelo de rede neural ESN. Os resultados obtidos a partir desta abordagem podem ser observados na Figura 17, que apresenta as predições geradas pelo modelo ajustado em comparação com os valores reais da série temporal. Essa análise visual dos resultados permite avaliar o desempenho do modelo ESN na tarefa de predição de séries temporais.

O modelo ESN demonstrou um desempenho notável na modelagem das séries temporais consideradas neste estudo, especialmente na profundidade de 117 cm. Os resultados obtidos nas diferentes posições dos transdutores foram avaliados por meio das métricas apresentadas na Tabela 8. Essas métricas refletem o desempenho dos modelos na fase de teste, na qual foram utilizados dados que não haviam sido apresentados durante o treinamento ou ajuste dos modelos.

Ao analisar os resultados na posição de 17 cm, onde a temperatura medida apresenta uma oscilação maior, o ESN se mostrou capaz de capturar de forma mais satisfatória a dinâmica não linear dos dados. Os resultados obtidos para o ESN nessa profundidade estão resumidos na Tabela 8 e na Figura 17, onde são apresentadas as métricas de desempenho. Observa-se que para previsões a 17 cm, o RMSE é de 0,206 °C, indicando um erro moderado, enquanto para previsões a 96 cm e 117 cm, é significativamente menor, 0,038 °C e 0,018 °C, respectivamente, indicando maior precisão nas previsões. O coeficiente de determinação (R^2) também reflete essa variação, com valores negativos para previsões a 17 cm, sugerindo um ajuste inadequado, e valores positivos para previsões a 96 cm e 117 cm, indicando um bom ajuste aos dados. Na posição de 96 cm, onde a dinâmica da temperatura é menos oscilatória, o modelo ESN conseguiu estimar satisfatoriamente o comportamento da porção inicial do conjunto de treinamento, porém suas predições acabaram se desviando ligeiramente do comportamento real posteriormente.

O uso do ESN, juntamente com o ajuste dos hiperparâmetros por meio de

Figura 17 – Resultados obtidos com o modelo ESN avaliado no conjunto de testes para a tarefa de predição de múltiplos instantes à frente



Fonte: Dados da pesquisa.

Tabela 8 – Métricas de desempenho para o modelo ESN

Métrica	17 cm	96 cm	117 cm
RMSE (°C)	0,206	0,038	0,018
WMSE (°C ²)	1,377	0,300	0,366
MAE (°C)	0,122	0,026	0,017
MAPE (%)	0,290	0,062	0,041
MSE (°C)	0,042	0,001	0,0003
R ²	-0,045	0,794	0,630

Fonte: Dados da pesquisa.

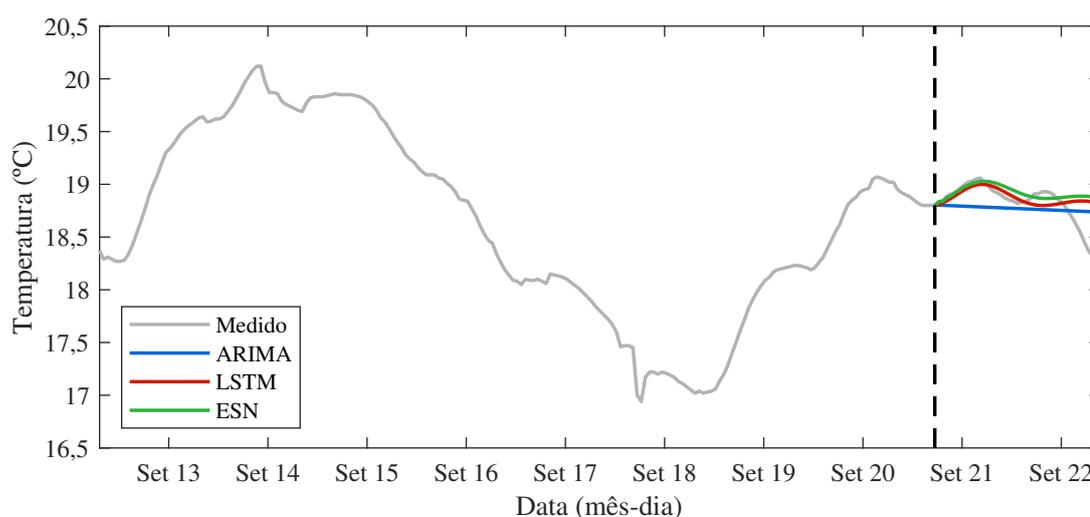
algoritmos genéticos, contribuiu para a obtenção de resultados mais confiáveis e representativos.

4.4 COMPARAÇÃO DOS RESULTADOS

A análise comparativa dos modelos de predição (ARIMA, LSTM e ESN) com base nas tabelas apresentadas nas Subseções 4.1, 4.2 e 4.3 revela as diferenças significativas em seu desempenho na modelagem de séries temporais em diferentes profundidades. Essas diferenças podem ser atribuídas à dinâmica das temperaturas em cada posição, fornecendo informações valiosas para a compreensão do comportamento desses modelos e para a escolha do modelo com características mais adequadas para modelar o comportamento da série de temperaturas em cada profundidade.

Ao observar a posição de 17 cm, Tabela 9 e Figura 18, identificou-se uma maior oscilação na temperatura medida, o que sugere uma maior influência da temperatura ambiente nessa localização próxima à superfície. Essa dinâmica oscilatória representou um desafio para o modelo ARIMA, que se baseia em uma abordagem linear de modelagem. Conseqüentemente, o modelo ARIMA enfrentou dificuldades em aprender a dinâmica do conjunto de treinamento, resultando em predições que se distanciaram consideravelmente das medições reais. Em contrapartida, os modelos LSTM e ESN, com sua capacidade de lidar com relações não lineares e complexas, apresentaram um desempenho mais satisfatório, com predições mais próximas das medições reais. O modelo ESN demonstrou ser o mais eficiente, seguido pelo LSTM, conforme indicado pelo WMSE.

Figura 18 – Resultados dos diferentes modelos avaliados no conjunto de testes para a tarefa de predição de múltiplos instantes à frente na profundidade de 17 cm



Fonte: Dados da pesquisa.

Na profundidade de 96 cm, Tabela 10 e Figura 19, observou-se uma dinâmica de temperatura menos oscilatória em comparação com a posição de 17 cm, tornando

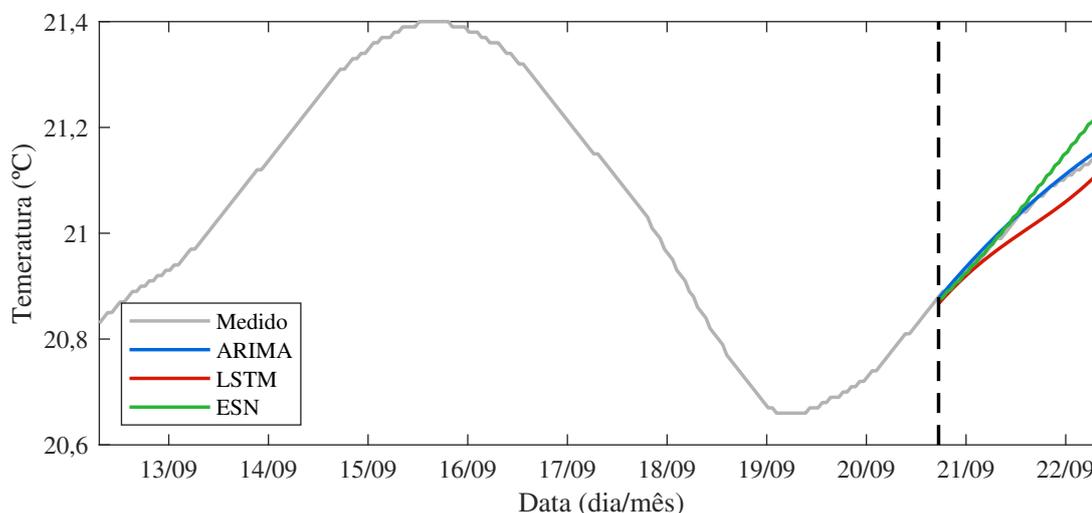
Tabela 9 – Métricas de desempenho para todos os modelos para profundidade de 17 cm

Métrica	ARIMA	LSTM	ESN
RMSE (°C)	0,191	0,187	0,206
WMSE (°C ²)	3,464	1,286	1,377
MAE (°C)	0,155	0,114	0,122
MAPE (%)	0,371	0,272	0,290
MSE (°C)	0,036	0,035	0,042
R ²	0,075	0,136	-0,045

Fonte: Dados da pesquisa.

a modelagem mais viável para modelos mais simples. Nesse caso, o modelo ARIMA obteve o menor RMSE ponderado, demonstrando uma melhor capacidade de prever o comportamento do conjunto de teste em comparação aos outros modelos. Os modelos LSTM e ESN, embora mais complexos que o ARIMA, conseguiram estimar razoavelmente o comportamento inicial do conjunto de teste, embora suas predições tenham se afastado ligeiramente do comportamento real posteriormente.

Figura 19 – Resultados dos diferentes modelos avaliados no conjunto de testes para a tarefa de predição de múltiplos instantes à frente na profundidade de 96 cm



Fonte: Dados da pesquisa.

Quanto à posição de 117 cm, Tabela 11 e Figura 20, a dinâmica foi semelhante à observada na posição de 96 cm, porém com uma amplitude de variação de temperatura menor. Nessa profundidade, o modelo LSTM, embora tenha um RMSE ligeiramente menor de 0,011(°C), possui um R² um pouco superior de 0,869, indicando que ele também se ajusta muito bem aos dados, sendo ligeiramente superior em termos de precisão, indicando um desempenho superior em relação aos outros modelos. Os

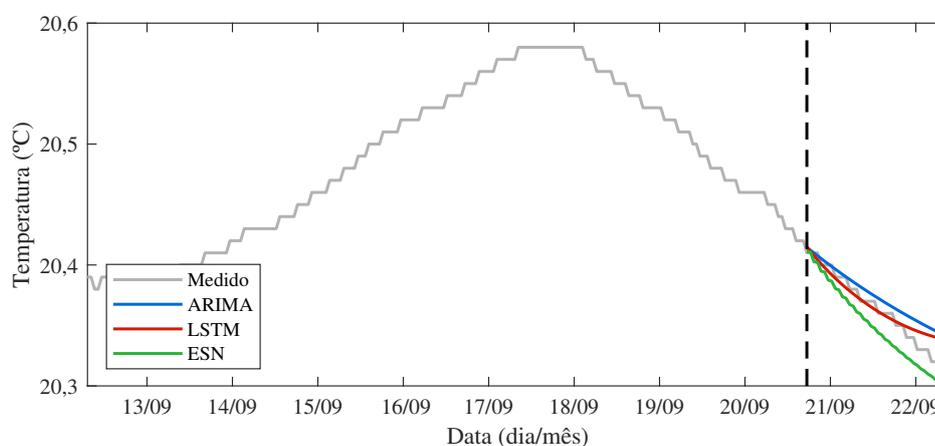
Tabela 10 – Métricas de desempenho para todos os modelos para profundidade de 96 cm

Métrica	ARIMA	LSTM	ESN
RMSE (°C)	0,008	0,031	0,038
WMSE (°C ²)	0,183	0,500	0,300
MAE (°C)	0,007	0,027	0,026
MAPE (%)	0,018	0,066	0,062
MSE (°C)	0,0001	0,001	0,001
R ²	0,988	0,858	0,794

Fonte: Dados da pesquisa.

modelos ARIMA e ESN foram capazes de estimar razoavelmente o comportamento inicial do conjunto de treinamento, com destaque para o ESN, que se aproximou mais da tendência correta no final do período de teste. De maneira geral, nesse caso, os três modelos mostram um desempenho excepcionalmente bom para a profundidade de 117 cm, com o LSTM tendo uma vantagem ligeira em termos de precisão RMSE e R².

Figura 20 – Resultados dos diferentes modelos avaliados no conjunto de testes para a tarefa de predição de múltiplos instantes à frente na profundidade de 117 cm



Fonte: Dados da pesquisa.

Esses resultados destacam a importância de considerar a dinâmica das séries temporais ao escolher um modelo de predição adequado. Isso pode ser claramente observado por meio das análises gráficas apresentadas nas figuras desta seção. Enquanto o modelo ARIMA, por sua natureza linear, enfrenta limitações na modelagem de dados mais oscilatórios, os modelos LSTM e ESN, com sua flexibilidade para lidar com relações não lineares, são mais adequados para capturar essas dinâmicas complexas, resultando em predições mais precisas. A escolha do modelo apropriado depende da

Tabela 11 – Métricas de desempenho para todos os modelos para profundidade de 117 cm

Métrica	ARIMA	LSTM	ESN
RMSE (°C)	0,015	0,011	0,015
WMSE (°C ²)	0,118	0,403	0,320
MAE (°C)	0,011	0,078	0,014
MAPE (%)	0,027	0,018	0,035
MSE (°C)	0,0002	0,001	0,0003
R^2	0,767	0,869	0,734

Fonte: Dados da pesquisa.

natureza específica da dinâmica dos dados e da precisão desejada para as predições.

Essa análise comparativa dos modelos fornece um embasamento para a compreensão de suas capacidades e limitações na modelagem de séries temporais em diferentes profundidades. Os resultados obtidos contribuem para aprimorar a compreensão das dinâmicas térmicas nesses cenários específicos e servem de base para o desenvolvimento de modelos mais complexos e baseados em mais dados para efetiva aplicação em campo.

5 CONSIDERAÇÕES FINAIS

Neste capítulo final, são apresentadas as conclusões gerais e sugestões para trabalhos futuros. A Seção 5.1 enumera as contribuições do trabalho, juntamente com as conclusões gerais em relação ao problema abordado. Também são apresentadas as conclusões específicas relacionadas aos modelos de redes neurais artificiais e ao estudo de caso utilizado. Por sua vez, a Seção 5.2 detalha as sugestões para o prosseguimento da pesquisa e a realização de trabalhos futuros.

5.1 CONCLUSÕES

Em um escopo geral, o trabalho trouxe contribuições relevantes. O estudo da literatura sobre o emprego de redes neurais artificiais recorrentes e modelos ARIMA na predição de temperatura em áreas contaminadas com hidrocarboneto demonstrou o sucesso dessas técnicas em diversos exemplos. Os resultados experimentais, obtidos por meio de um estudo de caso com base nos dados coletados pela STMF, contribuem para a literatura ao empregar diferentes arquiteturas de RNA e modelos tradicionais, além de estabelecer um caso real para comparação dessas abordagens. Todas as arquiteturas utilizadas apresentaram resultados satisfatórios, com algumas se adequando melhor a determinadas dinâmicas de temperatura. Conclusões específicas sobre a metodologia e as técnicas aplicadas também podem ser destacadas.

A metodologia empregada para a predição de temperatura do solo em áreas contaminadas com hidrocarboneto mostrou-se efetiva, com boas adequações ao utilizar redes neurais artificiais recorrentes e o modelo ARIMA. Essa metodologia serviu como base para todo o processo de construção dos modelos, permitindo uma adequada comparação de desempenho entre eles.

A análise comparativa dos modelos de previsão, como ARIMA, LSTM e ESN, revelou diferenças significativas em seu desempenho na modelagem de séries temporais em diferentes profundidades. Ao observar os resultados obtidos, podem ser inferidas algumas conclusões importantes.

Primeiramente, verificou-se que o modelo ARIMA apresentou dificuldades em lidar com dinâmicas mais oscilatórias, como observado na posição de 17 cm. Essa limitação é atribuída à abordagem linear do ARIMA, que não é capaz de capturar relações não lineares presentes nessas séries temporais. Como resultado, as predições geradas pelo ARIMA se distanciaram consideravelmente das medições reais. Por outro lado, os modelos LSTM e ESN, que possuem maior flexibilidade para lidar com essas relações não lineares, apresentaram um desempenho mais satisfatório, com predições mais próximas das medições reais. O modelo ESN demonstrou ser o mais eficiente, seguido pelo LSTM, conforme indicado pelo WMSE.

Na posição de 96 cm, na qual a dinâmica da temperatura apresentou menor

oscilação em comparação com a posição de 17 cm, o modelo ARIMA obteve o menor RMSE ponderado. Isso indica que o ARIMA teve uma melhor capacidade de prever o comportamento do conjunto de teste em comparação aos outros modelos. Os modelos LSTM e ESN, apesar de mais complexos, não apresentaram comportamento tão bom para essa profundidade de medição, porém também conseguiram estimar razoavelmente o comportamento inicial do conjunto de teste.

Já na posição de 117 cm, que apresentou uma dinâmica semelhante à observada na posição de 96 cm, porém com uma amplitude de variação de temperatura menor, o modelo LSTM mostrou melhor desempenho que os demais modelos, com menor WMSE. No entanto, é importante ressaltar que, assim como o LSTM, o modelo ESN também foi capaz de estimar razoavelmente o comportamento inicial do conjunto de teste, aproximando-se mais da tendência correta no final do período de teste.

Esses resultados ressaltam a importância de considerar a dinâmica específica das séries temporais ao escolher um modelo de predição adequado. Enquanto o ARIMA pode ser mais apropriado para dados com menor oscilação, os modelos LSTM e ESN, devido à sua flexibilidade para lidar com relações não lineares, são mais adequados para capturar dinâmicas complexas. A escolha do modelo correto depende, portanto, da natureza específica dos dados e da precisão desejada para as predições.

Essa análise comparativa dos modelos contribui para um embasamento na compreensão de suas capacidades e limitações na modelagem de séries temporais de temperatura de subsolo em diferentes profundidades. Além disso, os resultados obtidos podem ter implicações práticas relevantes, como o desenvolvimento de estratégias mais eficientes de gerenciamento térmico em aplicações específicas. Para o caso que serviu de motivação para este trabalho, o conhecimento da tendência dessas temperaturas pode ser explorado para a tomada de decisões informadas acerca do processo de remediação de solo.

5.2 SUGESTÕES PARA TRABALHOS FUTUROS

Esta seção lista as principais sugestões vislumbradas para estudos futuros. No âmbito geral, seria proveitoso investigar a implementação de métodos de treinamento específicos destinados à tarefa de predição. No presente trabalho, todos os modelos foram treinados utilizando abordagens convencionais, as quais são mais adequadas para simulações. Assim, é recomendável, para futuras pesquisas, a exploração de alternativas na definição da função custo que permitam a avaliação dos erros para uma janela de horizontes à frente.

Apesar dos resultados satisfatórios obtidos com o emprego dos modelos ARIMA, LSTM e ESN, sugere-se a avaliação de outros modelos da literatura para a tarefa de predição do comportamento da temperatura. Existem diversas abordagens e algoritmos disponíveis, como redes neurais convolucionais, redes neurais de atenção

(*Transformer*) e modelos de aprendizado de máquina não lineares, que podem ser investigados e comparados em termos de desempenho para a predição de séries temporais. Esse estudo pode incluir diferentes profundidades e localizações, como forma de conter ainda mais cenários de estudo.

Além disso, é relevante considerar outras variáveis além da temperatura do subsolo. Fatores ambientais, como umidade do solo, radiação solar, velocidade do vento e temperatura ambiente, podem influenciar as séries temporais e afetar as predições. Recomenda-se avaliar o impacto dessas variáveis adicionais e avaliar se a consideração das medições dessas variáveis como entradas dos modelos podem aprimorar os resultados de predição. Na mesma linha, o uso de modelos multivariáveis, que considere temperaturas medidas em pontos próximos do solo como entradas, pode trazer melhora para a estimativa realizada.

A realização de estudos de longo prazo também é uma sugestão valiosa. Este estudo abrangeu um período de tempo específico, dada a natural limitação de tempo para desenvolvimento das estações e coleta dos dados dentro do período de realização do mestrado. Estender a análise para incluir dados de séries temporais ao longo de vários anos pode permitir a identificação de padrões sazonais, tendências de longo prazo e mudanças no comportamento das séries, o que é importante para aplicação prática dos métodos avaliados. Esta sugestão apenas será possível depois de o sistema estar operando em campo por mais tempo, porém ele continua operacional e adquirindo dados em uma região contaminada real, o que deverá viabilizá-la com o passar do tempo. Essas observações, aliadas a um acompanhamento dos contaminantes por métodos tradicionais, também permitirão uma avaliação mais clara da capacidade do método proposto de realizar predições não apenas das temperaturas, mas do uso das temperaturas preditas como entrada para modelos que estimam a taxa de NSZD.

Um fato observado em campo e para o qual o presente trabalho pode contribuir é que alguns transdutores de temperatura apresentam uma deriva ao longo do tempo, possivelmente por efeito da umidade ou da contaminação do solo. Isso ocorre apenas em alguns transdutores, porém atrapalha a análise dos resultados e a apresentação de uma estimativa realista da taxa de biodegradação dos contaminantes. Uma sugestão que merece ser estudada em trabalho futuro é a avaliação da aplicabilidade do método proposto para identificar esses problemas nos transdutores de temperatura. Por meio da comparação entre as predições feitas e as medições reais, possivelmente envolvendo também dados obtidos com transdutores em posições próximas, deve ser possível identificar que algum transdutor está com problema.

REFERÊNCIAS

- ADIPAH, Sylvia. Introduction of petroleum hydrocarbons contaminants and its human effects. **Journal of Environmental Science and Public Health**, Fortune Journals, v. 3, n. 1, p. 1–9, 2019.
- ANTONELO, Eric. **Reservoir computing architectures for modeling robot navigation systems**. 2011. Tese (Doutorado) – Ghent University.
- ASKARANI, Kayvan Karimi; SALE, Thomas Clay. Thermal estimation of natural source zone depletion rates without background correction. **Water research**, Elsevier, v. 169, p. 115245, 2020.
- BAROT, Tomas; BURGSTEINER, Harald; KOLLERITSCH, Wolfgang. Comparison of Discrete Autocorrelation Functions with Regards to Statistical Significance. *In*: SILHAVY, Radek (Ed.). **Applied Informatics and Cybernetics in Intelligent Systems**. Cham: Springer International Publishing, 2020. P. 257–266.
- BAUTISTA, Luis Gustavo Cordero. **Aplicação do algoritmo genético adaptativo com hipermutação no ajuste dos parâmetros dos controladores suplementares e dispositivo FACTS IPFC**. 2019. Diss. (Mestrado) – Universidade Estadual Paulista (Unesp).
- BOCCATO, Levy. **Novas propostas e aplicações de redes neurais com estados de eco**. 2013. Tese (Doutorado) – Universidade Estadual de Campinas.
- BOURGUIGNON, Marcelo; SANTOS-NETO, Manoel; CASTRO, Mário de. A new regression model for positive random variables with skewed and long tail. **METRON**, v. 79, p. 33–55, 2021.
- BOX, George E.P.; JENKINS, Gwilym M; REINSEL, Gregory C; LJUNG, Greta M. **Time series analysis: forecasting and control**. [S.l.]: John Wiley & Sons, 2015. Fifth edition.
- BROCKWELL, Peter J.; DAVIS, Richard A. (Ed.). State-Space Models. *In*: **Introduction to Time Series and Forecasting**. New York, NY: Springer New York, 2002. P. 259–316. ISBN 978-0-387-21657-7.

BROWNLEE, Jason. Comparing classical and machine learning algorithms for time series forecasting. **Machine Learning Mastery**, 2019.

CARLOS, Jonas Hendler. **Modelos dinâmicos para séries temporais positivas**. 2022. Diss. (Mestrado) – Universidade Federal do Rio Grande do Sul.

CHATFIELD, Chris; XING, Haipeng. **The analysis of time series: an introduction with R**. [S.l.]: CRC press, 2019.

CHEN, Jinyin *et al.* E-LSTM-D: A deep learning framework for dynamic network link prediction. **IEEE Transactions on Systems, Man, and Cybernetics: Systems**, v. 51, n. 6, p. 3699–3712, 2021.

CHOUDHURY, Amartya; MIDDYA, Asif Iqbal; ROY, Sarbani. Attention enhanced hybrid model for spatiotemporal short-term forecasting of particulate matter concentrations. **Sustainable Cities and Society**, Elsevier, v. 86, p. 104112, 2022.

CRUTCHER, AJ; ROVERS, FA; MCBEAN, EA. Temperature as an indicator of landfill behavior. **Water, Air, and Soil Pollution**, Springer, v. 17, p. 213–223, 1982.

DAVIS, G B; RAYNER, J L; TREFRY, M G; FISHER, S J; PATTERSON, B M. Measurement and modeling of temporal variations in hydrocarbon vapor behavior in a layered soil profile. **Vadose Zone Journal**, Soil Science Society of America, v. 4, n. 2, p. 225–239, 2005.

DRÉO, Johann; PÉTROWSKI, Alain; SIARRY, Patrick; TAILLARD, Eric. **Metaheuristics for hard optimization: methods and case studies**. [S.l.]: Springer Science & Business Media, 2006.

DRÉO, Johann; SIARRY, Patrick; PÉTROWSKI, Alain; TAILLARD, Eric. Evolutionary Algorithms. **Metaheuristics for Hard Optimization: Simulated Annealing, Tabu Search, Evolutionary and Genetic Algorithms, Ant Colonies, ... Methods and Case Studies**, Springer, p. 75–122, 2006.

EMPRESA DE PESQUISA ENERGÉTICA, EPE. **Balço energético nacional: ano base 2020**. Rio de Janeiro, 2021. Disponível em: <https://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-601/topico-596/BEN2021.pdf>. Acesso em: 1 nov. 2021.

EMPRESA DE PESQUISA ENERGÉTICA, EPE. **Cenários de Demanda para o Plano Nacional de Energia 2050**. Rio de Janeiro, 2018. Disponível em: <https://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-227/topico-202>. Acesso em: 21 out. 2021.

FLECK, Leandro; TAVARES, Maria Herminia Ferreira; EYNG, Eduardo; HELMANN, Andrieli Cristina; ANDRADE, Minéia Aparecida de Moares. Redes neurais artificiais: Principios básicos. **Revista Eletrônica Científica Inovação e Tecnologia**, v. 1, n. 13, p. 47–57, 2016.

GAUTAM, Krishna *et al.* A review on control and abatement of soil pollution by heavy metals: Emphasis on artificial intelligence in recovery of contaminated soil. **Environmental Research**, Elsevier, p. 115592, 2023.

GIGLIO, José Euclides Fernandes. **Delineamento de experimentos no treinamento de redes neurais artificiais para o problema de previsão de séries temporais não lineares**. 2021. Diss. (Mestrado) – Universidade Federal de Itajubá.

GOLDBERG, David E; KORB, Bradley; DEB, Kalyanmoy. Messy genetic algorithms: Motivation, analysis, and first results. **Complex systems**, Complex Systems Publications, Champaign, IL, USA, v. 3, n. 5, p. 493–530, 1989.

GOMES, Daniel Takata. Redes neurais recorrentes para previsão de séries temporais de memórias curta e longa. **Master's thesis, Department of Statistics, Campinas State University, Campinas, Brazil**, p. 153, 2005.

GONÇALVES, Victor Henrique. **Previsão de séries temporais econômicas usando redes neurais caóticas**. 2017. Diss. (Mestrado) – Universidade de São Paulo.

GRAVES, Alex. Generating sequences with recurrent neural networks. **arXiv preprint arXiv:1308.0850**, 2013.

GREFF, Klaus; SRIVASTAVA, Rupesh K; KOUTNIK, Jan; STEUNEBRINK, Bas R; SCHMIDHUBER, Jürgen. LSTM: A search space odyssey. **IEEE transactions on neural networks and learning systems**, IEEE, v. 28, n. 10, p. 2222–2232, 2016.

HASTIE, Trevor; TIBSHIRANI, Robert; FRIEDMAN, Jerome H; FRIEDMAN, Jerome H. **The elements of statistical learning: data mining, inference, and prediction**. [S.l.]: Springer, 2009. v. 2.

HAYKIN, Simon. **Redes Neurais: Princípios e Práticas**. 2. ed. Porto Alegre: Bookman, 2001. P. 900.

HOCHREITER, Sepp; SCHMIDHUBER, Jürgen. Long short-term memory. **Neural computation**, MIT press, v. 9, n. 8, p. 1735–1780, 1997.

JAEGER, Herbert. Adaptive nonlinear system identification with echo state networks. **Advances in neural information processing systems**, v. 15, 2002.

LEVINE, William S. **The Control Handbook (three volume set)**. [S.l.]: CRC press, 2018.

LIAE, REMA; **Ferramentas de monitoramento e quantificação da depleção da massa de hidrocarbonetos de petróleo em fontes de contaminação para o apoio no encerramento de casos de contaminação**. [S.l.: s.n.], 2019.

LIPTON, Zachary C; BERKOWITZ, John; ELKAN, Charles. A critical review of recurrent neural networks for sequence learning. **arXiv preprint arXiv:1506.00019**, 2015.

LÚCIO, Paulo Sérgio *et al.* Um modelo estocástico combinado de previsão sazonal para a precipitação no Brasil. **Revista Brasileira de Meteorologia**, SciELO Brasil, v. 25, p. 70–87, 2010.

LUIZ, Thiago Boeno Patricio *et al.* **Previsão de dados de níveis de água subterrânea utilizando modelos baseados em aprendizado de máquina**. 2022. Tese (Doutorado) – Universidade Federal de Santa Maria.

LUKOŠEVIČIUS, Mantas; JAEGER, Herbert. Reservoir computing approaches to recurrent neural network training. **Computer science review**, Elsevier, v. 3, n. 3, p. 127–149, 2009.

LUNDEGARD, Paul D; JOHNSON, Paul C. Source zone natural attenuation at petroleum hydrocarbon spill sites—II: application to a former oil field. **Groundwater Monitoring & Remediation**, Wiley Online Library, v. 26, n. 4, p. 93–106, 2006.

AL-MAJED, Abdul Aziz; ADEBAYO, Abdulrauf Rasheed; HOSSAIN, M Enamul. A sustainable approach to controlling oil spills. **Journal of environmental management**, Elsevier, v. 113, p. 213–227, 2012.

MARIANO, Adriano Pinto. **Avaliação do potencial de biorremediação de solos e de águas subterrâneas contaminados com óleo diesel**. 2006. Tese (Doutorado) – Universidade Estadual Paulista.

MASINI, Ricardo P; MEDEIROS, Marcelo C; MENDES, Eduardo F. Machine learning advances for time series forecasting. **Journal of economic surveys**, Wiley Online Library, v. 37, n. 1, p. 76–111, 2023.

MEDEIROS, Talles Henrique de. **Estratégias de decisão em aprendizado de máquina multi-objetivo**. 2019. Tese de Doutorado – Universidade Federal de Minas Gerais, Minas Gerais, Brasil. Programa de Pós-Graduação em Engenharia Elétrica.

MENDES, Marcel M; RESENDE FILHO, Levi W de; VARGAS, Lucas V; PEIXOTO, Alessandro Jacoud. Visão computacional e redes neurais convolucionais aplicadas à detecção de vazamentos de Óleo. *In*: 1. CONGRESSO Brasileiro de Automática-CBA. [S.l.: s.n.], 2020. v. 2.

MUEEN, Abdullah; KEOGH, Eamonn; ZHU, Qiang; CASH, Sydney; WESTOVER, Brandon. Exact discovery of time series motifs. *In*: SIAM. PROCEEDINGS of the 2009 SIAM international conference on data mining. [S.l.: s.n.], 2009. P. 473–484.

NELLES, Oliver. Nonlinear system identification. **Measurement Science and Technology**, v. 13, n. 4, p. 646–646, 2002.

NELLES, Oliver. **Nonlinear system identification: from classical approaches to neural networks, fuzzy models, and gaussian processes**. [S.l.]: Springer Nature, 2020.

NIKBAKHT, Saeid; ANITESCU, Cosmin; RABCZUK, Timon. Optimizing the neural network hyperparameters utilizing genetic algorithm. **Journal of Zhejiang University-Science A**, Springer, v. 22, n. 6, p. 407–426, 2021.

PARENTE, Regina Rosa. **Classificação de uma classe para seleção de conjuntos de dados sintéticos em meta-aprendizado**. 2022. Doutorado em Ciência da Computação – Universidade Federal de Pernambuco, Recife.

PARK, Soyoun; JUNG, Solyoun; LEE, Jaegul; HUR, Jin. A Short-Term Forecasting of Wind Power Outputs Based on Gradient Boosting Regression Tree Algorithms. **Energies**, Multidisciplinary Digital Publishing Institute, v. 16, n. 3, p. 1132, 2023.

PARMEZAN, Antonio Rafael Sabino. **Predição de séries temporais por similaridade**. 2016. Diss. (Mestrado) – Universidade de São Paulo, São Carlos, Brazil.

PERES, Caio Brito. **Algoritmo para estimativa da depleção natural de hidrocarbonetos de petróleo a partir de perfis verticais de temperatura**. 2019. Diss. (Mestrado) – Universidade Federal de Santa Catarina, Florianópolis.

RIBEIRO, Iran F; KROHLING, Breno; COMARELA, Giovanni; MOTA, Vinicius FS. Geração de Séries Temporais Utilizando Redes Generativas Adversárias: da Teoria à Prática. **Sociedade Brasileira de Computação**, 2022.

SÁ, Luisa Castello Branco de *et al.* **Desenvolvimento de um modelo de predição de temperatura e umidade de uma estufa baseado em aprendizado de máquina**. 2020. Diss. (Mestrado) – Universidade Federal de Minas Gerais.

SALLES, Andre Assis de; LONDE, Mariana Alves. Determinação de expectativas dos preços do petróleo no mercado internacional através de modelos clássicos de previsão. **Brazilian Journal of Business**, v. 2, n. 1, p. 228–241, 2020.

SANTOS, Edson Simões dos; MATIAS, Italo; PONTES, Márcio Oliveira; SCHWARTZ, Gustavo Lemos. Configuração automática de arquitetura de rede neural artificial por algoritmo genético. *In*: XIX Simpósio de Pesquisa Operacional e Logística da Marinha. [S.l.: s.n.], 2019.

SANTOS, Paulo Guilherme Silva dos *et al.* **Previsão de variáveis ambientais na Amazônia com uso de redes neurais artificiais do tipo Long Short-Term Memory**. 2021. Diss. (Mestrado) – Universidade Federal do Oeste do Pará.

SCHWEDERSKY, Bernardo Barancelli. **Avaliação experimental de redes neurais artificiais recorrentes para identificação de sistemas dinâmicos: um estudo de**

caso. 2017. Diss. (Mestrado) – Programa de Pós-Graduação em Engenharia de Automação e Sistemas. Universidade Federal de Santa Catarina, Florianópolis.

SHI, Zhiwei; HAN, Min. Ridge regression learning in ESN for chaotic time series prediction. **Control and Decision**, v. 22, n. 3, p. 258, 2007.

SHUMWAY, Robert H; STOFFER, David S; SHUMWAY, Robert H; STOFFER, David S. ARIMA models. **Time Series Analysis and Its Applications: With R Examples**, Springer, p. 75–163, 2017.

SIHOTA, Natasha J; SINGURINDY, Olga; MAYER, K Ulrich. CO₂-efflux measurements for evaluating source zone natural attenuation rates in a petroleum hydrocarbon contaminated aquifer. **Environmental Science & Technology**, ACS Publications, v. 45, n. 2, p. 482–488, 2011.

SILVA, Anderson P da; LUDERMIR, Teresa B; ALMEIDA, Leandro M. Um Metodo para Busca Automatica de Redes Neurais Artificiais. **arXiv preprint arXiv:2107.04702**, 2021.

SMITH, Jonathon J.; BENEDE, Enrique; BEUTHE, Brigitta; MARTI, Manuel *et al.* A comparison of three methods to assess natural source zone depletion at paved fuel retail sites. **Quarterly Journal of Engineering Geology and Hydrogeology**, Geological Society of London, v. 54, n. 4, 2021.

STOCKWELL, Emily Beth. **Continuous NAPL loss rates using subsurface temperatures**. 2015. Diss. (Mestrado) – Colorado State University, Fort Collins.

SWEENEY, Robert E; RIRIE, G Todd. Temperature as a tool to evaluate aerobic biodegradation in hydrocarbon contaminated soil. **Groundwater Monitoring & Remediation**, Wiley Online Library, v. 34, n. 3, p. 41–50, 2014.

TIAN, Yuandong; CHEN, Xinlei; GANGULI, Surya. Understanding self-supervised learning dynamics without contrastive pairs. *In*: PMLR. INTERNATIONAL Conference on Machine Learning. [S.l.: s.n.], 2021. P. 10268–10278.

VAN DE VEN, Cole JC; SCULLY, Keelin H; FRAME, Mikaela A; SIHOTA, Natasha J; MAYER, K Ulrich. Impacts of water table fluctuations on actual and perceived natural source zone depletion rates. **Journal of Contaminant Hydrology**, Elsevier, v. 238, p. 103771, 2021.

- XIA, Dawen; YANG, Nan; JIAN, Shunying; HU, Yang; LI, Huaqing. SW-BiLSTM: a Spark-based weighted BiLSTM model for traffic flow forecasting. **Multimedia Tools and Applications**, Springer, v. 81, n. 17, p. 23589–23614, 2022.
- YAMAK, Peter T; YUJIAN, Li; GADOSEY, Pius K. A comparison between arima, lstm, and gru for time series forecasting. *In: PROCEEDINGS of the 2019 2nd International Conference on Algorithms, Computing and Artificial Intelligence*. [S.l.: s.n.], 2019. P. 49–55.
- YAO, Xianshuang; WANG, Zhanshan; ZHANG, Huaguang. A novel photovoltaic power forecasting model based on echo state network. **Neurocomputing**, Elsevier, v. 325, p. 182–189, 2019.
- ZHAN, Zhi-Hui; LI, Jian-Yu; ZHANG, Jun. Evolutionary deep learning: A survey. **Neurocomputing**, Elsevier, v. 483, p. 42–58, 2022.