

**UNIVERSIDADE FEDERAL DE ITAJUBÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM MEIO AMBIENTE E RECURSOS
HÍDRICOS**

**Estudo de operação otimizada de uma rede benchmark com aplicação de protótipo de
gêmeo digital adaptativo aliado a redes neurais recorrentes do tipo Long Short Term
Memory e algoritmos genéticos**

Sara Maria Marques

ITAJUBÁ (MG)

2026

**UNIVERSIDADE FEDERAL DE ITAJUBÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM MEIO AMBIENTE E RECURSOS
HÍDRICOS**

Sara Maria Marques

**Estudo de operação otimizada de uma rede benchmark com aplicação de protótipo de
gêmeo digital adaptativo aliado a redes neurais recorrentes do tipo Long Short Term
Memory e algoritmos genéticos**

Tese submetida ao Programa de Pós-Graduação em Meio Ambiente e Recursos Hídricos como parte dos requisitos para obtenção do título de Doutor em Ciências em Meio Ambiente e Recursos Hídricos.

Área de Concentração: Meio Ambiente e Recursos Hídricos

Orientador: Prof. Dr. Fernando das Graças Braga da Silva

ITAJUBÁ (MG)

2026

FOLHA DE APROVAÇÃO

AGRADECIMENTOS

A Deus, pela força, proteção e sabedoria concedidas ao longo de toda esta jornada acadêmica. Sem Sua presença, não teria sido possível chegar até aqui.

Aos meus pais e a minha irmã, pelo amor incondicional, apoio constante e incentivo diário. Vocês sempre acreditaram no meu potencial e me ensinaram o valor da dedicação e da persistência.

Aos amigos que fiz durante essa etapa, pela convivência leve, pelas discussões produtivas, pela troca de conhecimento e por todos os momentos de parceria que tornaram o caminho mais agradável e enriquecedor.

Ao meu orientador, Prof. Dr. Fernando das Graças Braga da Silva, pela orientação atenta, pela confiança no meu trabalho, pelas contribuições fundamentais e por todo aprendizado transmitido ao longo desta pesquisa.

À CAPES, pelo apoio e financiamento (nº 88887.829273/2023-00), que tornaram possível o desenvolvimento deste estudo e contribuíram de forma decisiva para minha formação acadêmica. Aos colegas do NUMMARH – Núcleo de Modelagem e Simulação em Meio Ambiente, Recursos e Sistemas Hídricos da UNIFEI, pelas contribuições técnicas realizadas ao longo do trabalho. Ao Laboratório de Hidráulica Computacional da UNIFEI (LHC), por proporcionar a infraestrutura técnica necessária ao desenvolvimento da pesquisa.

A todos vocês, meu sincero agradecimento.

" O importante é não parar de questionar; a curiosidade tem sua própria razão de existir." (Albert Einstein)

RESUMO

As mudanças climáticas e o uso indiscriminado da água têm levado à escassez de recursos hídricos, intensificando a crise hídrica. Esses fatores colocam em risco o abastecimento de água eficiente para a população. Com isso, o desenvolvimento de técnicas e metodologias para garantir a operação ótima de uma rede de distribuição de água torna-se essencial. Dessa forma, o trabalho desenvolvido elaborou um método para a previsão e otimização de perdas de uma rede de distribuição de água de um sistema benchmark. A metodologia consistiu em utilizar três ferramentas de maneira conjunta, com o objetivo de prever os piores cenários de perdas da rede de estudo para as próximas 24 horas de funcionamento. Essas ferramentas foram: o protótipo de gêmeo digital adaptativo para a troca simultânea de dados entre servidor e cliente; a rede neural recorrente LSTM para previsão dos piores cenários de perdas; e o algoritmo genético para a otimização desses cenários com base na configuração ideal dos parâmetros operacionais: bombas, válvulas e reservatórios. Para isso, foi utilizada a linguagem de programação Python e suas bibliotecas, como *WebSocket*, *Scikit-Learn* e *Keras* (*TensorFlow*). Os cenários mais adversos de perdas encontrados variaram entre 51% e 53%, representando um aumento médio de 15,44 pontos percentuais em relação à condição de referência do sistema. A partir da aplicação do algoritmo genético, foi possível obter uma redução média de 11,42 pontos percentuais nesses cenários, além da identificação das configurações operacionais responsáveis por essa melhoria, permitindo sua incorporação como estratégia de mitigação. Assim, a principal inovação do trabalho reside na integração dessas três abordagens em um único framework, possibilitando não apenas a previsão, mas também a otimização de condições críticas de operação. A utilização do gêmeo digital permitiu a realização de análises em ambiente virtual, sem riscos ao sistema real, enquanto a combinação com técnicas de inteligência artificial e otimização contribuiu para o aumento da eficiência operacional e para a antecipação e mitigação de cenários adversos de perdas.

Palavras-chave: Integração físico-virtual. Análise preditiva. Operação eficiente. Redes de distribuição de água.

ABSTRACT

Climate change and the indiscriminate use of water have led to the scarcity of water resources, intensifying the water crisis. These factors put at risk the efficient supply of water to the population. Consequently, the development of techniques and methodologies to ensure the optimal operation of a water distribution network becomes essential. Therefore, this work developed a method for forecasting and optimizing losses in a benchmark water distribution network. The proposed methodology consisted of using three tools jointly, with the objective of predicting the worst loss scenarios of the study network for the next 24 hours of operation. These tools were: an adaptive digital twin prototype for simultaneous data exchange between server and client; a recurrent LSTM neural network for forecasting the worst loss scenarios; and a genetic algorithm for optimizing these scenarios based on the ideal configuration of operational parameters: pumps, valves, and reservoirs. For this purpose, the Python programming language and its libraries were used, such as WebSocket, Scikit-Learn, and Keras (TensorFlow). The most adverse loss scenarios found ranged between 51% and 53%, representing an average increase of 15.44 percentage points compared to the system's baseline condition. By applying the genetic algorithm, it was possible to achieve an average reduction of 11.42 percentage points in these scenarios, as well as identifying the operational configurations responsible for this improvement, allowing their incorporation as a mitigation strategy. Thus, the main innovation of this work lies in the integration of these three approaches into a single framework, enabling not only prediction but also optimization of critical operating conditions. The use of the digital twin allowed for analyses to be performed in a virtual environment, without risks to the real system, while the combination with artificial intelligence and optimization techniques contributed to increased operational efficiency and the anticipation and mitigation of adverse loss scenarios.

Keywords: Physical-virtual integration. Predictive analytics. Efficient operation. Water distribution networks.

LISTA DE FIGURAS

Figura 3.1 - Evolução das perdas na distribuição no Brasil em porcentagem.....	20
Figura 3.2 - Perdas na distribuição por macrorregião do Brasil em porcentagem.	20
Figura 3.3 - Tipos de perdas aparentes.....	22
Figura 3.4 - Tipos de perdas reais.....	23
Figura 3.5 - Esquematização do funcionamento do algoritmo genético.	30
Figura 3.6 - Exemplificação do funcionamento do modelo.	35
Figura 3.7 - Exemplificação do funcionamento do modelo.	36
Figura 3.8 - Comportamento das funções de ativação mais utilizadas.....	37
Figura 3.9 – Estrutura básica da rede neural artificial recorrente.....	37
Figura 3.10 - Overfitting, Underfitting e cenário ideal nos ajustes de dados de modelos de Machine Learning.....	39
Figura 3.11 - Exemplificação da validação cruzada de dados.....	40
Figura 3.12 - Tipos de nível de integração do Gêmeo Digital.	45
Figura 4.1 - Fluxograma das etapas de metodologia do estudo.....	54
Figura 4.2 - Layout da rede de distribuição de água utilizada na etapa 1.	56
Figura 4.3 – Pressões da rede – gráfico gerado no Python.....	57
Figura 4.4 - Funcionamento básico da relação servidor/cliente, através do WebSocket.	58
Figura 4.5 - Estrutura básica da RNAR-LSTM.....	59
Figura 4.6 - Estrutura detalhada de uma célula da RNAR-LSTM.	60
Figura 4.7 - Fluxograma de funcionamento da rede neural LSTM desenvolvida.	63
Figura 4.8 - Fluxograma de funcionamento do algoritmo genético aplicado ao estudo.	65
Figura 4.9 - Fluxograma do funcionamento resumido da metodologia proposta.....	66
Figura 4.10 - Configuração da rede desenvolvida por Marques.....	68
Figura 4.11 - Configuração da rede desenvolvida por Silva.	69
Figura 4.12 – Configuração da rede Anytown.	70
Figura 4.13 – Rede de estudo.	71
Figura 4.14 – Curva de demanda da rede de estudo.....	74
Figura 5.1 – Número de elementos da rede encontrados no Python.	75
Figura 5.2 – Consoles criados no Spyder.	77
Figura 5.3 – Mensagens de comunicação iniciadas.....	77
Figura 5.4 – Histograma das pressões da rede de estudo.	80
Figura 5.5 – Boxplot das pressões da rede de estudo.	81

Figura 5.6 – Configuração das camadas do modelo de previsão.....	83
Figura 5.7 – Comparação das perdas previstas e originais.....	85
Figura 5.8 – Curva LSTM.	87
Figura 5.9 – Diagrama de dispersão.	88
Figura 5.10 – Erros absolutos por época.	89
Figura 5.11 – Geração ideal encontrada na otimização para cada hora.	90
Figura 5.12 – Histograma das pressões otimizadas da rede de estudo.....	93
Figura 5.13 – Boxplot das pressões otimizadas da rede de estudo.....	93
Figura 5.14 – Comparação das perdas originais e otimizadas.....	94
Figura 5.15 – Dados da Bomba 7, obtidos pelo algoritmo genético.	96
Figura 5.16 – Dados da Bomba 17, obtidos pelo algoritmo genético.	96
Figura 5.17 – Dados da válvula 9, obtidos pelo algoritmo genético.	97
Figura 5.18 – Dados da válvula 19, obtidos pelo algoritmo genético.	97
Figura 5.19 – Dados da válvula 23, obtidos pelo algoritmo genético.	98
Figura 5.20 – Dados do reservatório 1, obtidos pelo algoritmo genético.....	98
Figura 5.21 – Dados do reservatório 2, obtidos pelo algoritmo genético.....	99
Figura 5.22 – Boxplot das perdas originais.	99
Figura 5.23 – Boxplot das perdas previstas.....	100
Figura 5.24 – Boxplot das perdas otimizadas.....	101
Figura 5.25 – Mensagem recebida pelo servidor no final da comunicação.	102

LISTA DE TABELAS

Tabela 1 - Dados dos nós da rede de estudo.....	71
Tabela 2 - Dados dos trechos da rede de estudo.....	72
Tabela 3 – Comparação das pressões encontradas no Python e no EPANET.....	76
Tabela 4 - Pressões originais da rede de benchmark de estudo.....	78
Tabela 5 - Perdas originais da rede de benchmark de estudo.....	81
Tabela 6 – Configurações teste do LSTM de acordo com os valores das métricas.	83
Tabela 7 - Perdas previstas.	84
Tabela 8 – Comparação das perdas previstas e otimizadas.	91
Tabela 9 - Pressões da rede de estudo depois da otimização.....	92

QUADRO

Quadro 1 – Balanço hídrico proposto pela IWA.	21
--	----

LISTA DE SIGLAS

SINISA	Sistema Nacional de Informações sobre Saneamento
GD	Gêmeo Digital
ML	Machine Learning
RNA	Rede Neural Artificial
RNAR-LSTM	Rede Neural Artificial Recorrente-Long Short Term Memory
IWA	International Water Association
ETA'S	Estações de tratamento de água
MSE	Erro quadrático médio
MAE	Erro absoluto médio
R²	Coefficiente de determinação
IDE	Ambiente de desenvolvimento integrado
HTTP	Hyper Text Transfer Protocol
TCP/IP	Protocolo de Controle de Transmissão/Protocolo de Internet
PVC	Policloreto de Vinila

SUMÁRIO

1. INTRODUÇÃO	14
2. OBJETIVOS	17
3. REVISÃO BIBLIOGRÁFICA	18
3.1 SISTEMAS DE ABASTECIMENTO DE ÁGUA	18
3.1.1 Redes de distribuição de água	18
3.2 PERDAS NAS REDES DE DISTRIBUIÇÃO DE ÁGUA.....	19
3.3 OPERAÇÃO DE REDES DE DISTRIBUIÇÃO DE ÁGUA	24
3.4 ALGORITMO GENÉTICO DE OTIMIZAÇÃO	29
3.4.1 Conceituação	29
3.4.2 Aplicação de algoritmos genéticos em estudos de redes de distribuição de água	31
3.5 MACHINE LEARNING	34
3.5.1 Conceituação	34
3.5.2 Aplicação de Machine Learning em estudos de redes de distribuição de água....	40
3.6 GÊMEOS DIGITAIS.....	43
3.6.1 Conceituação	43
3.6.2 Aplicação de Gêmeos Digitais em estudo de redes de distribuição de água	49
4. MATERIAIS E MÉTODOS	54
4.1 ETAPA 1: INTEGRAÇÃO DO EPANET AO PYTHON PARA A LEITURA E SIMULAÇÃO DE ARQUIVO DO EPANET	54
4.2 ETAPA 2: ALGORITMO PARA PREVISÃO E OTIMIZAÇÃO DE PERDAS EM REDES DE BENCHMARK.....	58
4.3 ETAPA 3: DEFINIÇÃO DAS REDES DE DISTRIBUIÇÃO DE ÁGUA – BENCHMARK.....	67
4.3.1 Redes de benchmark para o treinamento do modelo de previsão.....	67
4.3.2 Rede de benchmark de estudo	70
4.4 ETAPA 4: DEFINIÇÃO DO CENÁRIO DE ESTUDO DA REDE DE BENCHMARK.....	73
4.5 ETAPA 5: ANÁLISE DOS RESULTADOS DA APLICAÇÃO DOS ALGORITMOS DESENVOLVIDOS À REDE DE ESTUDO	74
5. RESULTADOS E DISCUSSÕES.....	75
5.1 RESULTADOS DO ESTUDO DA ITERAÇÃO EPANET PYTHON.....	75
5.2 RESULTADOS DA APLICAÇÃO DO PROTÓTIPO DE GÊMEO DIGITAL ADAPTATIVO... 	76

5.3	RESULTADO DAS PERDAS PREVISTAS PELA REDE NEURAL RECORRENTE LSTM ..	82
5.4	RESULTADO DA OTIMIZAÇÃO DAS PERDAS PREVISTA PELO ALGORITMO GENÉTICO	89
6.	CONCLUSÕES.....	103
7.	RECOMENDAÇÕES.....	105
	REFERÊNCIAS.....	106
	ANEXO.....	123

1. INTRODUÇÃO

A água é um bem limitado e muito importante para o desenvolvimento humano e de suas atividades econômicas e sociais. Contudo, o uso desordenado desse recurso tem levado ao surgimento de questões ambientais, como o esgotamento de mananciais, e sociais com a impossibilidade de abastecimento (Silva Júnior, 2017).

Com isso, ainda há um fator que intensifica essa problemática: as perdas encontradas nos sistemas de abastecimento de água ao longo do seu processo. Os maiores volumes dessas perdas estão nas redes de distribuição de água (Silva Júnior, 2017; Brasil, 2021).

Segundo o levantamento realizado pelo Sistema Nacional de Informações sobre Saneamento (SINISA) em 2023 o total de perdas nas companhias nacionais foi de 37,8%, apresentando uma redução relativamente pequena de 2,9% em relação ao levantamento realizado em 2021 (Brasil, 2023).

Porém, o Brasil ainda possui um valor extremamente alto de perdas quando comparado com países de excelência no abastecimento, como Estados Unidos, Austrália e Japão, cujos índices são de 13%, 17% e 7%, respectivamente, sendo Singapura o país com maior eficiência no abastecimento com um índice de 5% de perdas (Berg, 2015; Brasil, 2021; Pathirane *et al.*, 2023). Essas perdas abordadas nos levantamentos são denominadas reais e se caracterizam por vazamentos em tubulações ou extravasamento de reservatórios (Brasil, 2014).

Devido à preocupação com a questão, estudos, metodologias e softwares vêm sendo desenvolvidos para auxiliar nessa previsão e redução das perdas reais dos sistemas de distribuição de água. Dentre os programas desenvolvidos para esse fim, tem-se o EPANET. Por ser um software de código aberto, permitindo a realização de simulações estatísticas e dinâmicas de inúmeros modelos de redes pressurizadas, mostrando o comportamento desses sistemas, possibilitando conhecer o percurso e o destino final da água, além de auxiliar na escolha das melhores configurações de funcionamento (Rossman, 2009).

Outra ferramenta que vem se destacando nos estudos de redes de água é o Python. Assim como o EPANET, possui código aberto, o que permite sua utilização de forma gratuita. Além disso, apresenta linguagem simples e ambiente integrado, bem como inúmeros pacotes que possibilitam o desenvolvimento e manipulação de dados. Também permite a geração de gráficos e tabelas, o que auxilia nos estudos hidráulicos. Uma das características mais importantes desse software é a possibilidade de interação com outros programas, como o

EPANET, possibilitando a simulação de uma rede dentro do próprio ambiente do Python, facilitando e ampliando as possibilidades de estudos sobre esses sistemas (Santos, 2020).

Com base nessas duas ferramentas acopladas, é possível realizar estudos mais profundos utilizando modelos e técnicas que permitem conhecer melhor o comportamento dos sistemas de distribuição de água, facilitando a análise de sua operação. Nesse contexto, destacam-se os gêmeos digitais (GD), machine learning (ML) e os algoritmos genéticos.

Os GD são definidos, de maneira geral, como uma representação de um sistema real em um ambiente virtual dinâmico e progressivo, onde ocorre troca de informações, em tempo real ou não, entre os ambientes físicos e virtuais (Singh et al., 2021). Porém, diferenciam-se dependendo de características como tempo de criação, nível de integração e nível de maturidade.

O tempo de criação se divide em protótipo e instância. Na prototipagem, o modelo virtual é criado antes do real, permitindo a realização de testes destrutivos sem comprometer o sistema real futuro. Na instância, o processo é inverso, pois o modelo físico já existe e o modelo virtual é criado para auxiliar na gestão do sistema (Grieves; Vickers, 2017).

Quanto ao nível de maturidade distingue-se em Pré-Digital, Gêmeo Digital e Gêmeo Digital Adaptativo. O Pré-Digital refere-se ao modelo virtual criado antes do físico para evitar problemas técnicos. O Gêmeo Digital caracteriza-se pelo ativo virtual que incorpora dados do ativo físico, com troca bidirecional. Já o Gêmeo Digital Adaptativo é semelhante ao anterior, mas inclui aplicação de aprendizado de máquina (ML) (Madni et al., 2019).

O ML é uma aplicação da inteligência artificial que possibilita que sistemas sejam capazes de aprender por si mesmos com base em suas experiências (dados), por meio de etapas como treinamento, validação e teste (Janiesch et al., 2021; Dalarmelina, 2023). Dentre seus modelos, destacam-se as Redes Neurais Recorrentes, que possuem maior capacidade de analisar dados que exigem reconhecimento de padrões, pois utilizam conexões reversas, podendo se conectar a nós de camadas anteriores e a si próprias. Com isso, esses sistemas possuem duas memórias: uma longa, proveniente das conexões entre nós, e uma curta, resultante da autoconexão, o que aumenta a capacidade desse modelo em estudos de previsão (Hochreiter; Schmidhuber, 1997; Aurélien, 2019).

O algoritmo genético é um modelo baseado em métodos evolucionários que tem como objetivo encontrar a melhor solução global para problemas de otimização. Assim, os algoritmos genéticos selecionam as melhores soluções e descartam as piores, submetendo as mais viáveis à recombinação por meio de crossover e mutação, tornando-as mais generalistas.

Essa característica os distingue dos algoritmos aleatórios, pois conseguem direcionar a busca para regiões mais promissoras do espaço de soluções, reduzindo o risco de encontrar apenas soluções ótimas locais e aumentando as chances de alcançar a solução ótima global (Srinivas; Patnaik, 1994; Lacerda; Carvalho, 1999).

O presente estudo destaca-se por sua relevância ao abordar a previsão, otimização e análise de perdas em sistemas de distribuição de água a partir das condições operacionais da rede, com foco especial na identificação de cenários adversos de funcionamento. O principal diferencial da proposta reside na utilização integrada de ferramentas avançadas, tendo como eixo central o gêmeo digital adaptativo, que permite a sincronização dinâmica entre o sistema físico e seu modelo virtual, viabilizando a representação fiel e em tempo real do comportamento hidráulico da rede.

A partir dessa estrutura, o trabalho não se limita à análise estática das perdas, mas avança no sentido de explorar cenários críticos e complexos, nos quais as perdas tendem a ser mais significativas e de difícil detecção. Nesse contexto, a integração do gêmeo digital com técnicas de inteligência artificial, como a Rede Neural Recorrente do tipo LSTM, possibilita a captura de padrões temporais e não lineares associados ao comportamento das perdas. Além disso, a aplicação de algoritmos genéticos permite a otimização do sistema, buscando soluções eficientes mesmo em condições adversas.

Dessa forma, o estudo propõe uma abordagem inovadora baseada na iteração conjunta entre modelagem hidráulica, aprendizado de máquina e otimização, o que representa um avanço em relação às metodologias tradicionais, geralmente aplicadas de forma isolada. Essa integração torna possível não apenas compreender melhor os mecanismos de perdas, mas também propor estratégias robustas para sua mitigação, especialmente em situações críticas. Dessa forma, configura-se como uma abordagem moderna e promissora para o estudo e a gestão eficiente de perdas em sistemas de abastecimento de água.

2. OBJETIVOS

O objetivo geral deste trabalho é desenvolver uma metodologia para a previsão e otimização de perdas a partir de variáveis operacionais, permitindo identificar condições de operação que minimizem perdas e evitem cenários adversos, utilizando da aplicação conjuntas de ferramentas, como gêmeo digital, redes neurais e algoritmos genéticos.

Os objetivos específicos:

- Propor uma rede de benchmark para a aplicação do protótipo de Gêmeo Digital Adaptativo, utilizando a ferramenta *WebSockets* 15.0 em Python para comunicação e atualização em tempo real dos dados da rede;
- Desenvolver um modelo de previsão de cenários adversos de perdas utilizando uma Rede Neural Recorrente do tipo Long Short-Term Memory (LSTM), empregando as bibliotecas *Scikit-learn* e *Keras (Tensorflow)*;
- Aplicar de um Algoritmo Genético para a otimização das perdas previstas, utilizando o pacote DEAP em Python, a fim de determinar os valores operacionais ótimos que minimizem perdas previstas e reduzam a ocorrência de cenários adversos;
- Avaliar a metodologia proposta, verificando o desempenho e a aplicabilidade conjunta das ferramentas desenvolvidas para o estudo de uma rede de distribuição de água de benchmark.

3. REVISÃO BIBLIOGRÁFICA

Este capítulo traz uma revisão da literatura sobre os assuntos pertinentes e conceituações teóricas necessárias ao desenvolvimento desta pesquisa.

3.1 Sistemas de abastecimento de água

Os sistemas de abastecimento de água, de maneira geral, podem ser definidos como uma união de equipamentos, serviços e obras empregados para o fornecimento de água potável à população, atendendo às diversas finalidades de uso. Esses sistemas são responsáveis por fornecer água em quantidade e qualidade adequada, atendendo às condições de pressão e vazão necessárias para o seu funcionamento (Azevedo Netto *et al.*, 1998; Porto, 2006).

Tais sistemas são compostos por etapas ou fases que consistem, basicamente, em captação, estações elevatórias, estações de tratamento de água, reservatórios de regularização e redes de distribuição (Tusutiya, 2006).

3.1.1 Redes de distribuição de água

As redes de distribuição de água são consideradas um importante elemento de infraestrutura para a sociedade, pois são responsáveis por transportar água em qualidade e quantidade suficientes, além de garantir pressões adequadas. São sistemas constituídos por diferentes elementos, como bombas, válvulas, tubulações, reservatórios, entre outros, que estão interligados e conduzem a água até os consumidores domésticos, comerciais ou industriais (Heler; Pádua, 2010; Martinho, 2020; Martinho; Melo; Sörensen, 2021).

A configuração das redes de distribuição decorre da topografia da via em que as tubulações serão instaladas e da demanda do local (Porto, 2006). De acordo com a NBR 12218/2017, independente de seu layout as redes são formadas basicamente por dois tipos de condutos: principais e secundários. Os condutos principais, também chamados de tronco ou mestres, são tubulações de maiores diâmetros que alimentam as tubulações secundárias. Estas por sua vez, são de diâmetros menores e são responsáveis por levar a água diretamente aos consumidores.

De acordo com a localização das tubulações principais e o sentido do escoamento nas tubulações secundárias, as redes podem ser classificadas em rede ramificadas, malhadas e mistas. Na rede ramificada, o abastecimento de água se dá por meio do conduto tronco para os condutores secundários, com escoamento em um único sentido. Dessa forma, esse tipo de

rede é mais utilizado em pequenas comunidades, em sistemas de irrigação e em sistemas particulares (Tsutiya, 2006).

Já a rede malhada caracteriza-se por tubulações formando anéis, permitindo o abastecimento dos pontos de consumo por mais de uma direção. Com isso, há maior flexibilidade para atender à manutenção e à demanda da rede (Porto, 2006). A reversibilidade de vazão que essa configuração proporciona possibilita a entrega de água com melhor qualidade, devido ao fluxo permanente (Heler; Pádua, 2010). Devido ao seu arranjo, é o tipo de rede mais empregado na maioria dos centros urbanos (Gomes, 2009). As redes mistas possuem um arranjo que combina os dois tipos anteriores, sendo parte malhada e parte ramificada (Porto, 2006).

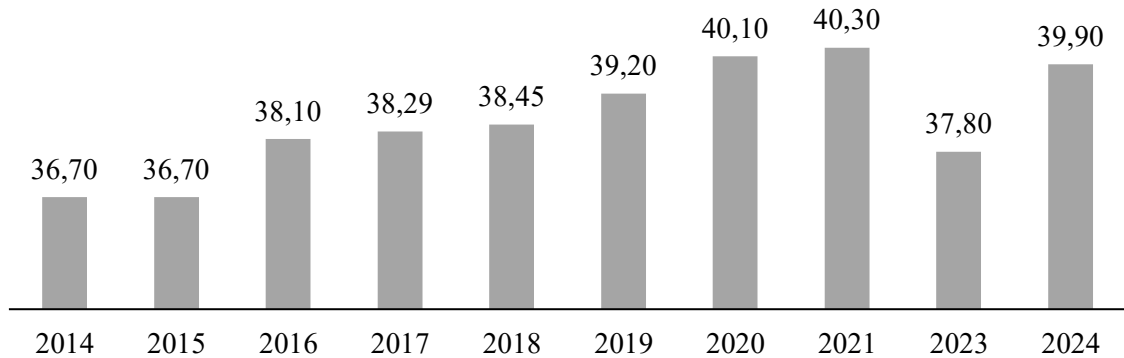
3.2 Perdas nas redes de distribuição de água

As perdas de água são um processo inerente aos sistemas de abastecimento de água, geralmente causadas por deficiência na operação e manutenção das estruturas que os compõem, além da gestão comercial ineficiente. Com isso, independente das dimensões da rede de distribuição de água, as perdas impactam não só a receita das concessionárias, mas também o desperdício de água e energia (Sobrinho; Borja, 2016; Sanchez, 2018; Brasil, 2023).

Dessa forma, a deficiência no gerenciamento, aliada à baixa manutenção efetiva das redes, tem intensificado o processo de degradação dos componentes que compõem esse sistema, o que leva ao aumento das perdas reais de água. Isso se agrava pelo fato de as estruturas estarem enterradas, sendo muitas vezes esquecidas e recebendo atenção apenas quando ocorre alguma ruptura visível (Fontana *et al.*, 2012; Mutikanga *et al.*, 2013).

O levantamento realizado pelo Sistema Nacional de Informações sobre Saneamento (SINISA) no ano de 2024 apresentou um total de perdas para as concessionárias nacionais de 39,9%, indicando uma redução de 0,4 ponto percentual em relação ao levantamento de 2021, cujo valor era de 40,3% (Brasil, 2023). Contudo, observa-se um aumento quando comparado ao ano de 2023, em que a perda levantada foi de 37,8% (Brasil, 2024). Figura 3.1 traz a evolução das perdas levantadas pelo SINISA ao longo de dez anos.

Figura 3.1 - Evolução das perdas na distribuição no Brasil em porcentagem.

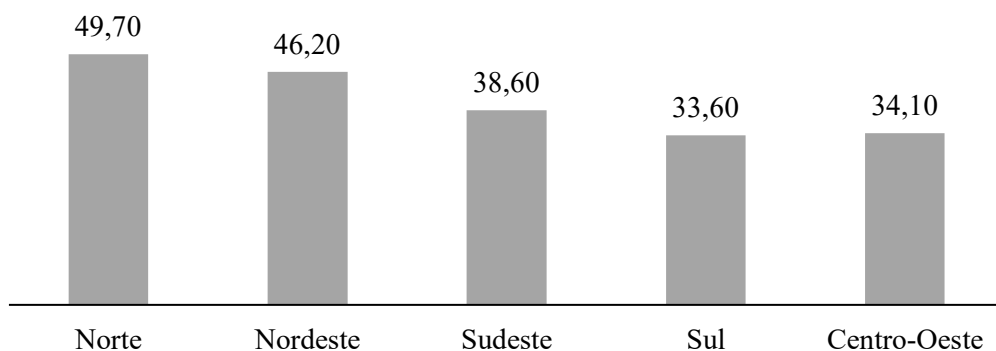


Fonte: Adaptado Brasil (2024).

Apesar de apresentar uma redução nos valores de perdas de água no sistema nacional, o Brasil ainda possui um índice elevado quando comparado a países referência na gestão de sistemas de abastecimento de água, como Estados Unidos, Austrália, Japão e Singapura, cujos índices são de 13%, 17%, 7% e 5%, respectivamente (Berg, 2015; Brasil, 2021).

Nos dados levantados por macrorregião no Brasil, observa-se uma heterogeneidade dos valores, em que as regiões Norte e Nordeste apresentam os maiores índices de perdas. Esse fato ocorre por serem regiões com baixa infraestrutura e menores índices de abastecimento. Já as regiões Sul, Sudeste e Centro-Oeste apresentam valores de perdas abaixo da média nacional, devido aos maiores níveis de infraestrutura (Brasil, 2024). A Figura 3.2 apresenta os valores das perdas para cada uma das macrorregiões do Brasil.

Figura 3.2 - Perdas na distribuição por macrorregião do Brasil em porcentagem.



Fonte: Adaptado Brasil (2023).

Para determinar as perdas nos sistemas de distribuição, é habitualmente utilizado uma metodologia técnica proposta pela *International Water Association* (IWA), que consiste em uma matriz na qual os processos de produção da água tratada são especificados desde a sua entrada no sistema até o consumo final, possibilitando quantificar o volume consumido e o

perdido (Oliveira *et al.*, 2020). O Quadro 1 apresenta a matriz do balanço hídrico proposto pela IWA.

Quadro 1 – Balanço hídrico proposto pela IWA.

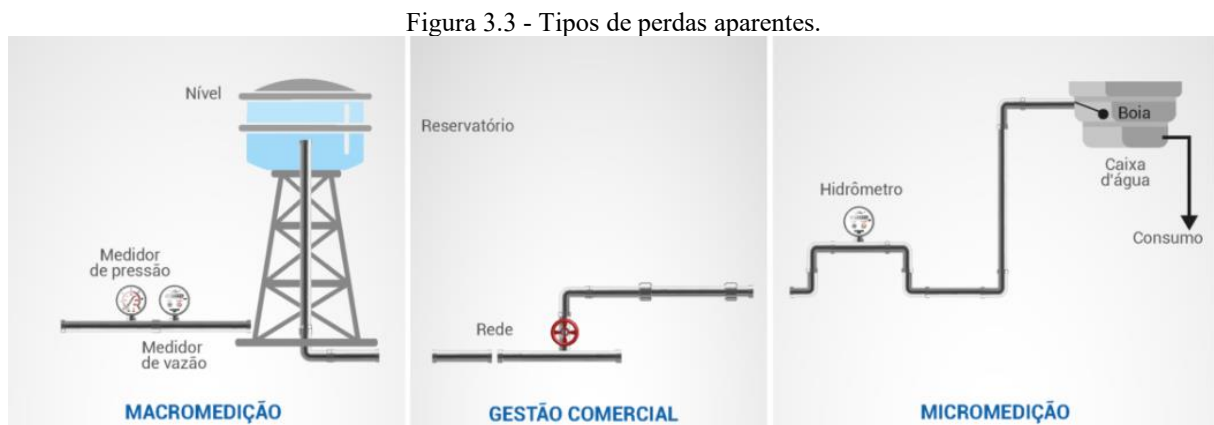
Água que entra no sistema (inclui água importada)	Consumo autorizado	Consumo autorizado faturado	Consumo faturado medido (inclui água exportada)	Água faturada
			Consumo faturado não medido (estimado)	
		Consumo autorizado não faturado	Consumo não faturado medido (uso próprio, caminhão pipa, entre outros)	Água não faturada
			Consumo não faturado não medido	
	Perdas de água	Perdas aparentes (comerciais)	Uso não autorizado (fraudes e falhas de cadastro)	
			Erros de medição (macro e micromedição)	
		Perdas reais (físicas)	Vazamentos e extravasamento nos reservatórios (de adução e/ou distribuição)	
			Vazamento nas adutoras e/ou redes (de distribuição)	
Vazamento nos ramais até o ponto de medição do cliente				

Fonte: Oliveira *et al.* (2020).

O Balanço Hídrico considera como ponto inicial a água que entra no sistema e se divide no processo de distribuição em consumo autorizado, que corresponde ao volume de água que foi fornecido ao cliente autorizado (sendo faturado ou não), e perdas de água, que se dão pela diferença entre o volume que entra no sistema pelo volume contabilizado e podem ser classificadas de acordo com a sua natureza em dois tipos: reais e aparentes (Hirner; Lambert, 2000; Oliveira, 2020).

As perdas aparentes, também denominadas de perdas não físicas, se caracterizam pelo consumo da água tratada, que não foi contabilizada pela companhia. Sua maior ocorrência se dá em pontos próximos do consumidor final e se caracteriza por problemas de medição em hidrômetros, ligações clandestinas, aferição inadequada do consumo de água, entre outras utilizações não permitidas da água (Kanakoudis; Muhammetoglu, 2014; Vieira, 2019; Oliveira *et al.*, 2023).

Ela pode ocorrer em três tipos diferentes: a macromedicação, que se caracteriza por erros de medição da água que entra nas redes, causados principalmente por erros nos equipamentos de medição. A gestão comercial, que se refere à gestão ineficiente do sistema, caracteriza-se por ligações clandestinas nas redes, cadastros desatualizados de clientes e fraudes. E a micromedicação, definida por erros de medição dos hidrômetros dos consumidores finais (Chini; Stillwell, 2018; Mutikanga; Sharma; Vairavamoorthy, 2013). A Figura 3.3 exemplifica esses três tipos de perdas aparentes.



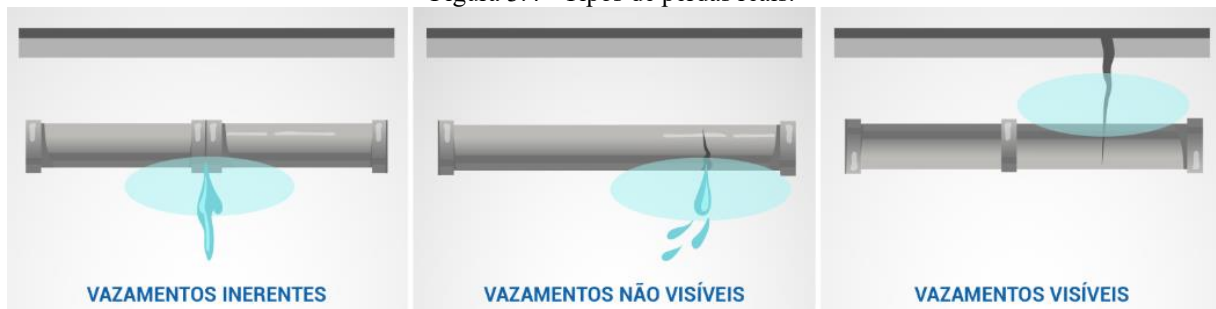
Fonte: DAE Jundiá (2023).

Para Kanakoudis e Muhammetoglu (2014), esse tipo de perda pode ser evitado por meio do gerenciamento efetivo do sistema com ações como fiscalização da rede, atualização de cadastro, automatização do sistema e aferição dos equipamentos de medição.

Já as perdas reais, também chamadas de físicas, são determinadas pelo volume de água que se perde ao longo da operação do sistema de abastecimento de água: captação, tratamento, armazenamento e distribuição (Oliveira *et al.*, 2023). Sua ocorrência ao longo do processo se dá devido a alguns fatores, como qualidade dos materiais, vida útil, manutenção deficiente, obras externas próximas às redes e principalmente às mudanças de pressão, que podem levar a vazamentos nas tubulações (Andrade Sobrinho; Borja, 2016).

Essa perda também se difere em três tipos diferentes, inerentes, não visíveis e visíveis. As perdas inerentes se caracterizam por serem não visíveis e de difícil detecção por equipamentos de acústica. Já as perdas não visíveis são determinadas por não afloram na superfície, porém são detectadas por equipamentos de acústica. E por fim as perdas visíveis, como o próprio nome especifica, afloram à superfície, sendo assim fácil de serem detectadas (Brasil, 2014). A Figura 3.4 apresenta os três tipos principais de perdas reais encontradas nas redes de distribuição de água.

Figura 3.4 - Tipos de perdas reais.



Fonte: DAE Jundiá (2023).

Com isso, observa-se que as perdas físicas de água são preocupantes, pois causam impactos ambientais com o desperdício de água, aumento da pressão sobre as fontes hídricas, aumento dos custos de mitigação de impactos negativos dessa atividade, decaimento de qualidade, mas também impactos na produção com a elevação dos custos de operação e manutenção o que acarreta na diminuição de receita (Al-Washali, 2020). Além disso, trazem outra questão relevante, a redução da eficiência hidráulica, o que leva ao aumento do consumo elétrico, pois a ocorrência de vazamentos nos condutos exige aumento de pressão e vazão para continuar atendendo a demanda o que conduz para a ineficiência energética da rede (Sousa; Soares, 2020).

No Brasil, esse consumo de energia nos sistemas de abastecimento de água é de aproximadamente 11,84 TWh/ano, o que representa um gasto de R\$6,35 bilhões para as companhias de abastecimento (Brasil, 2021). No trabalho realizado por Silqueira *et al.* (2024) apresentou em seu estudo a correlação entre as perdas e o consumo energético para um sistema de abastecimento de água de pequeno porte da região sudeste de Minas Gerais. De acordo com os autores a perda de água levou a uma perda de energia durante os últimos cinco anos de funcionamento foi de 192.539,13 kWh resultando em uma perda monetária de R\$249.949,06. Isso demonstra que as perdas de água em redes de distribuição de água levam a um desperdício energético e consequentemente de perda monetária. Assim, num sistema eficiente, não há somente a redução da perda de água, mas também de energia, uma vez que estão interligados (Guanais; Cohim; Medeiros, 2016).

Portanto, a questão abordada exige medidas para a sua mitigação com uma operação e gestão eficiente, em que as características dos materiais aplicados sejam adequadas para a prevenção de vazamentos, bem como o funcionamento adequado dos elementos ativos do sistema (bombas e válvulas redutoras de pressão), como a troca de bombas ineficazes, a utilização de bombas adequadas às características de vazão e pressão do sistema, evitando o

superdimensionamento, o balanceamento do rotor e o emprego de inversores de frequência no lugar de válvulas de controle (Makisha; Kazimirova, 2018).

3.3 Operação de redes de distribuição de água

As redes de distribuição de água são um sistema constituído por elementos interligados entre si com o objetivo de conduzir a água em quantidade e qualidade das ETA's para o consumidor (Kurian *et al.*, 2018). A alimentação das redes vai depender da topografia do local onde foram instaladas e do seu desenho. Assim, ocorrem através de reservatórios, localizados a jusante ou montante da rede, podendo ser elevados, apoiados ou semienterrados ou diretamente por estações elevatórias (Tsutiya, 2006).

Quanto ao fluxo de água no sistema, ele pode ser contínuo, que se caracteriza por um sistema em que as tubulações são pressurizadas durante todo o dia e com isso, a água é fornecida a todo momento ao consumidor final, não havendo a necessidade de um armazenamento pelo consumidor. E o fluxo intermitente em que a água é fornecida durante um período de tempo, podendo variar entre uma ou duas vezes ao dia, sendo necessário que o consumidor realize o seu armazenamento. Esse último tipo ocorre pela necessidade de limitar a quantidade de água que é fornecida devido a uma menor disponibilidade da água (Kurian *et al.*, 2018)

Por serem sistema que ficam enterrados nos logradouros públicos e terem um acesso difícil, sofrem com uma operação inadequada e a falta de manutenção (Tsutiya, 2006). Assim, observa-se a degradação das tubulações, o que leva ao seu rompimento e, conseqüentemente, a vazamentos, que são os maiores problemas das redes. Ademais, existem mais problemas que essa questão pode trazer como a falta de água devido a uma oferta insuficiente, subdimensionamento e entupimento dos condutos, contaminação da água e principalmente variação de pressão e vazão (Tardelli Filho, 2015).

Para a literatura a operação de uma rede não consiste somente na variação de válvulas ou acionamento de bombas, mas também no conjunto de ações que irão permitir o funcionamento de maneira ideal. Essas ações segundo Tardelli Filho (2015) envolve: dispor de um cadastro técnico e comercial atualizado e confiáveis; o monitoramento das vazões na entrada dos reservatórios, das estações elevatórias e de áreas específicas da rede, além dos pontos de pressão importantes, verificação se as variáveis operacionais estão de acordo com as normas; utilização de indicadores de performance operacional e gerencial; e o

desenvolvimento de relatórios das variáveis da operação do sistema aferido em todos os setores da rede e zonas de pressão.

Já para Ougui (2003), a operação de um sistema de distribuição de água envolve fatores, como o custo de funcionamento das bombas, que estão instaladas na rede. Para ele, a questão do bombeamento tem relevância, pois a operação eficiente depende da interconexão entre a sua adequação à distribuição de água, à realidade atual dos custos com energia.

Com isso, diversos estudos têm sendo desenvolvidos com o objetivo de melhorar a eficiência do processo de operação das redes de distribuição de água.

Sousa e Soares (2020) trouxeram em seu estudo um modelo de operação eficiente para uma rede de distribuição hipotética desenvolvida em 1985 na “Batalha dos Modelos de Rede”, conhecida como Anytown, por meio da junção do simulador hidráulico EPANET e algoritmos genéticos. O objetivo do trabalho era de otimizar o funcionamento das bombas, reduzindo o consumo de energia elétrica e as perdas de água da rede, aumentando a confiabilidade de operação do sistema com a operação e vazões ideais. Segundo os autores o estudo apresentado se mostrou favorável, pois um sistema com funcionamento adequado reduz os custos e permite o aumento da eficiência hidroenergética, além de auxiliar na tomada de decisão por parte das companhias de água.

Egito (2022) apresentou um modelo operacional de rede de distribuição de água por meio da combinação do uso eficiente de reservatório de armazenamento com o melhor acionamento de bombas, com o objetivo de reduzir os custos operacionais do sistema de abastecimento. A ferramenta aplicada para o desenvolvimento do modelo foi o toolkit WNTR que possibilita a junção do EPANET ao um algoritmo genético de otimização. O resultado obtido foi uma redução de 30% nos custos da rede hipotética estudada, o que mostra a importância de um estudo que busca um melhor arranjo no funcionamento de elementos essenciais nesse tipo de sistema como reservatórios e bombas.

Móller (2021) trouxe um modelo para a operação eficiente do bombeamento de uma rede de distribuição de água. O trabalho consistiu em encontrar o melhor ponto de operação de bombas e o número das mesmas operando em paralelo com o uso da técnica Particle Swarm Optimization (PSO), com o objetivo de minimizar os custos de energia do sistema e de perdas de água. Além disso, também utilizaram da análise da rede tronco por meio de grafos para a definição da melhor localização das estações elevatórias. Para isso, utilizaram de três redes amplamente estudadas na literatura, a Anytown, NET3 e a D-Town. O resultado mostrou que um sistema mais operacional mais flexível quanto o acionamento das bombas e

sua localização mais otimizada permite um sistema com operação otimizada quanto aos custos.

Em Kurian *et al.* (2018), foi apresentada uma metodologia para a operação ideal das bombas e válvulas das redes, com a finalidade de minimizar a energia consumida quando há variação da demanda. A rede abordada pelos autores é um sistema real da Índia que se caracteriza por possuir uma única bomba em um sistema com oito reservatórios de alimentação. De acordo com os autores, o resultado da pesquisa mostrou que a operação ideal se dá com a instalação de reservatórios de armazenamento intermediários no sistema. Também é necessário a aplicação de um cronograma de acionamento de bombas e válvulas para a redução do consumo energético. Com isso, o sistema torna-se mais eficiente com um menor custo da atividade e maior produção.

Zhang *et al.* (2020) trouxeram uma pesquisa sobre o funcionamento ideal de uma rede de distribuição de água com o foco na pressão, em que a pressão a ser entregue aos consumidores é otimizada no ajuste da pressão de saída da bomba. A rede estudada fica na cidade de Xangai na China que possui uma área total de 250 km². Nessa metodologia, a pressão do sistema é retratada pela pressão dos nós-chaves e das bombas principais, embasada em cálculos dos métodos de transferência e do método de simplificação de rede, em que os inúmeros modelos lineares são utilizados para a relação de transferência de pressão. Ou seja, os dados reais coletados na rede de Xangai são comparados com métodos heurísticos existentes. O resultado mostrou eficácia do método utilizado na otimização das pressões, mantendo o atendimento de forma eficiente da rede e diminuindo a probabilidade de vazamento e consumo de energia elétrica. Ademais, é de fácil aplicação em redes existentes.

Já Chen *et al.* (2021) apresentaram um estudo sobre a operação das estações de bombeamento aplicada a uma rede genérica. O modelo consiste em integrar a operação de bombeamento da rede em dois estágios (captação e abastecimento) com um algoritmo genético (NSGA-II), com o objetivo de garantir confiabilidade, eficiência e durabilidade da operação e a minimização de custos de funcionamento, baseado no controle de nível dinâmico dos reservatórios de água tratada. Os resultados apresentados expõem que o método utilizado reduz significativamente os custos operacionais de energia do sistema, atendendo aos objetivos expostos, além de manter o funcionamento adequado das bombas e impedir o transbordamento ou esvaziamento dos reservatórios de água.

Souza (2021) trouxe em seu estudo a avaliação da operação eficiente hidroenergética de redes de distribuição de água sobre dois tipos de operação intermitente e contínua. As

redes utilizadas foram as benchmarking ZJ e OBCL-1 e as variáveis consideradas para a melhor solução operacional foram o consumo de energia e o volume de vazamentos. Foram gerados cenários no abastecimento intermitente variando o tempo de operação e acionamento dos elementos ativos como as válvulas e bombas, e assim, comparando os resultados com o abastecimento contínuo, com o uso do algoritmo de otimização Particle Swarm Optimization. O resultado apresentado foi que o uso da operação intermitente foi mais viável do que o contínuo, obtendo uma eficiência no custo de operação de 60,9 % na rede ZJ e 47,4 % na rede OBCL-1.

Balekelayi *et al.* (2021) propuseram uma abordagem de otimização da operação de sistemas de distribuição de água com múltiplos reservatórios de armazenamento e estações de bombeamento, visando à redução dos custos energéticos e ao aumento da eficiência operacional. Dessa forma, foi empregado um algoritmo genético aliado a simulações hidráulicas offline realizadas no EPANET. A metodologia visou manter os níveis de água e as pressões da rede dentro de limites operacionais aceitáveis, ao mesmo tempo em que minimiza o consumo de energia nos períodos de tarifa elevada, sendo aplicado ao sistema de distribuição de água da cidade de Asmara, na Eritreia. A aplicação do modelo ao sistema de abastecimento de água demonstrou a viabilidade da estratégia proposta, evidenciando a necessidade de restrições adicionais para reservatórios não controlados e destacando contribuições relevantes para a gestão de sistemas hídricos urbanos em países em desenvolvimento.

Evangelista (2022) analisou a eficiência operacional de redes de distribuição de água ao longo de sua vida útil, considerando a evolução temporal de características físicas e hidráulicas dessas infraestruturas. O estudo propõe a modelagem matemática de parâmetros como rugosidade das tubulações, índices de vazamentos, crescimento da demanda e deterioração de bombas, utilizando o ambiente MATLAB integrado ao simulador hidráulico EPANET, aplicados em duas redes da literatura Anytown e D-town. Os resultados indicam que os vazamentos exercem maior influência sobre os custos operacionais, enquanto o aumento da rugosidade impacta principalmente as pressões da rede.

Marques *et al.* (2022) investigaram o comportamento hidráulico de uma rede de distribuição de água por meio de simulações que consideram variações nos níveis dos reservatórios, na rugosidade das tubulações e nos diâmetros dos elementos da rede. A metodologia consistiu em modelar cenários hidráulicos, com o objetivo de identificar configurações ideais de operação que possibilitem a gestão de pressão e a redução das perdas

físicas no sistema de abastecimento. A aplicação de simulações de Monte Carlo permitiu a análise de múltiplas condições operacionais, culminando na identificação de um cenário que reduziu em 11,6 % as perdas físicas em comparação ao caso base, demonstrando a eficácia da modelagem computacional para otimização de redes de abastecimento de água e com isso, a eficiência operacional dos sistemas.

Silva (2023) trouxe um estudo sobre a operação ideal de rede de distribuição de água com a utilização de redes neurais artificiais do tipo perceptron com uma camada oculta para o reconhecimento de parâmetros operacionais do sistema e uma com sete camadas para realizar a otimização das pressões da rede teórica estudada. Os parâmetros operacionais utilizados foram o funcionamento das válvulas redutoras de pressão e acionamento das bombas. Segundo o autor, o modelo desenvolvido foi eficiente pois foi capaz de prever o funcionamento da rede e com isso prever as perdas por meio do controle das pressões nos nós.

Li *et al.* (2023) propuseram a aplicação de um esquema de layout de sensores de pressão em um sistema de abastecimento de água, utilizando o algoritmo de agrupamento EGAE. O esquema desenvolvido consiste em integrar as características estruturais da rede como topologia com os dados de pressão coletados, e com isso, selecionar os nós mais representativos do sistema como o local para os sensores, utilizando para isso do coeficiente de correlação. O sistema utilizado no estudo foi uma rede experimental, desenvolvida pelos próprios autores. O resultado encontrado foi que a metodologia aplicada foi eficiente no posicionamento ideal de sensores e que a relação pressão e topologia são pontos essenciais para a identificação do posicionamento ideal de sensores para o monitoramento da rede de distribuição de água.

Gobbo *et al.* (2024) desenvolveram um estudo de redes de distribuição de água por meio da otimização do funcionamento de válvulas redutoras de pressão visando a redução de perdas e a estabilidade mecânica das tubulações. Para isso foi utilizado o sensoriamento remoto de um sistema real onde foi aplicado a otimização, e o software utilizado foi o *EclipseE3* que possui uma plataforma HMI/SCADA. O resultado apresentado pelos autores mostrou que um sistema automatizado de VRP é capaz de aumentar o controle de pressão, pois esse acionamento pode ser realizado de maneira remota o que torna a resposta mais rápida e com isso as perdas de água também foram reduzidas.

Ayyash *et al.* (2024) analisaram os desafios operacionais de um sistema de benchmark (D-town) de abastecimento de água, considerando limitações na disponibilidade de água, os custos operacionais elevados e a desigualdades no acesso a água. Assim, a metodologia foi

desenvolvida em duas etapas, sendo composta primeiramente pela setorização heurística da rede, baseada na conectividade hidráulica e no acesso às fontes de abastecimento, e na sequência pela aplicação de um método de otimização multiobjetivo (NSGA-II), com o objetivo de reduzir custos operacionais, promover maior equilíbrio de pressão entre os usuários e aumentar a equidade no fornecimento de água. Os resultados obtidos em uma rede de referência indicam que a combinação entre setorização e operação otimizada contribui para a melhoria da uniformidade de pressão, a ampliação da equidade no acesso e a redução do consumo energético ao longo do período de operação, configurando-se como uma estratégia eficaz para a gestão de sistemas de abastecimento de água sob condições de escassez hídrica.

Cheng *et al.* (2024) trouxeram em seu trabalho um estudo sobre a localização ótima de sensores para a identificar as perdas do sistema, considerando variáveis como topologia da rede e a ligação entre os nós da rede. O método desenvolvido utilizou matriz de covariância sensíveis à pressão do sistema, para que a variável topologia tivesse mais peso, e assim, com o uso do Graph Laplace Regularization (GLR) e Gershgorin Disc Alignment (GDA) pudessem ser identificados os pontos onde seriam localizados os sensores. Essa metodologia foi aplicada a três tipos de redes de proporções e comportamentos diferentes a Hanoi, Net 3 e PA2. O resultado apresentado pelos autores foi que o método aplicado foi eficiente para a localização das perdas nos três tipos de rede estudadas, mostrando a aplicabilidade do método.

Godase *et al.* (2025) propuseram uma abordagem para a gestão eficiente e otimizada de um sistema de distribuição de água em escala laboratorial, buscando a detecção e o isolamento rápido de vazamentos na rede. A metodologia baseou-se no uso de controladores lógicos programáveis (PLC) para o monitoramento em tempo real de dados de fluxo e pressão, possibilitando a identificação de anomalias operacionais e o acionamento automático de válvulas para o isolamento de trechos afetados. Os resultados indicam que o protótipo é capaz de detectar e responder a vazamentos em curto intervalo de tempo, reduzindo perdas de água em relação a métodos convencionais, o que evidencia contribuições relevantes para a modernização e a eficiência operacional de sistemas de abastecimento de água em contextos urbanos.

3.4 Algoritmo Genético de otimização

3.4.1 Conceituação

O algoritmo genético é um modelo de otimização desenvolvido por John Henry Holland em 1975, com base na teoria evolutiva de Charles Darwin, em que o indivíduo mais

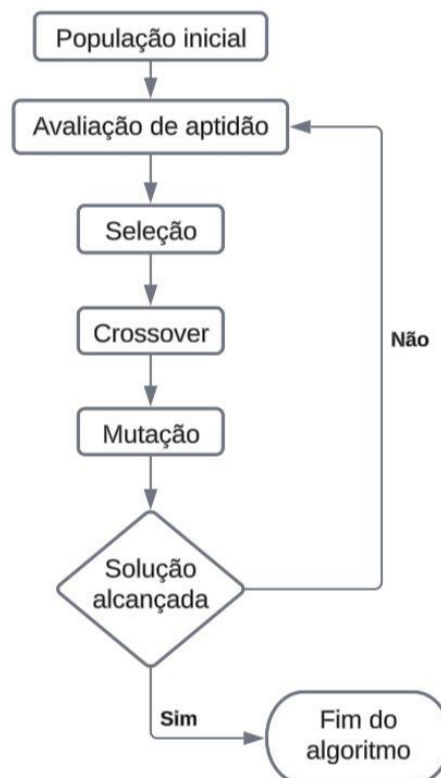
adaptado a seu ambiente tem maiores chances de sobreviver e produzir descendentes (Massago, 2013).

Assim, o funcionamento do algoritmo genético é baseado no funcionamento dos genomas humanos, em que o cruzamento de materiais genéticos dos pais resulta em uma nova combinação, gerando os filhos que herdam as características de seus pais. Além disso, também ocorre a mutação, que são alterações aleatórias que o cromossomo pode sofrer (Lacerda; Carvalho, 1999).

Dessa forma, o algoritmo genético inicia seu funcionamento por meio do cromossomo (população) gerados aleatoriamente ou com base nos dados obtidos. Esse cromossomo passa por seleção para selecionar o mais apto. Na sequência o indivíduo passa por cruzamento e mutação para que a variabilidade da próxima geração seja assegurada. Essa etapa é repetida até que a solução ótima seja alcançada (Galvão et al.; 1999).

O número de iterações, chamada de gerações, que o algoritmo genético irá realizar pode ser determinada previamente ou essa repetição pode ser atendida quando os requisitos são alcançados (Lacerda; Carvalho, 1999). A Figura 3.5 apresenta esse esquema de funcionamento do algoritmo genético.

Figura 3.5 - Esquematização do funcionamento do algoritmo genético.



Fonte: Adaptado de Chagas; Rodrigues; Tavares (2014).

A utilização dos operadores de cruzamento e mutação são importantes no melhoramento dos cromossomos e no conhecimento de outras regiões no espaço de busca. Contudo, durante a aplicação do cruzamento e mutação nas gerações podem ocorrer a perda dos melhores indivíduos, sendo necessário a utilização da técnica de elitismo. Essa abordagem consiste em identificar e manter os melhores indivíduos para as próximas gerações, sendo um processo iterativo em que um indivíduo considerado elite em uma geração pode ser substituído por um novo caso o algoritmo identifique como melhor (Formiga, 2005; Ferreira, 2007).

A utilização do algoritmo genético em estudos apresenta algumas vantagens como a capacidade de otimizar grandes números de variáveis; facilidade em aplicar parâmetros contínuos e discretos ou a combinação de ambos; a utilização de populações, o que permite a busca simultânea em várias regiões; fácil aplicação computacional; capacidade de otimizar múltiplas funções que possuem objetivos conflitantes; podem ser utilizados com outras técnicas e heurísticas, entre outras (Lacerda; Carvalho, 1999).

Dentre as vantagens que o algoritmo possui, se destacam para a presente pesquisa a capacidade de otimização utilizando de um grande número variáveis e a facilidade de aplicação computacional, sendo a primeira característica importante para o trabalho, uma vez que serão utilizadas mais de uma variável operacional da rede de distribuição de água para a otimização do sistema.

3.4.2 Aplicação de algoritmos genéticos em estudos de redes de distribuição de água

O emprego de algoritmos genéticos nos estudos da área de recursos hídricos já vem sendo utilizado, pois apresenta-se como bom solucionador quando comparado com métodos convencionais.

Como pode ser observado em Neto (2021), que apresentou um estudo de calibração otimizada de redes de distribuição de água por meio da variação da demanda. Para isso, foram realizadas várias simulações com 100, 200, 400, 500, 800 e 1000 gerações e quanto maior o número de gerações menor era o erro médio relativo em ambas as representações, porém o tempo de simulação foi maior. Mesmo com o número de erros pequenos nas duas abordagens, segundo o autor, a representação real obteve coeficiente de rugosidade melhores que o binário.

Já em Romero-Ben *et al.* (2022) foi realizado um estudo sobre uma rede de benchmark L-Town, com objetivo de definir os locais ideais para a colocação de sensores de pressão para

o controle de pressões da rede para prever as perdas, de acordo com a distância entre os nós. O resultado mostrou que o método foi viável, encontrando os pontos mais importantes para a implementação dos sensores, aumentando o controle das pressões e assim reduzindo as perdas.

Sangroula *et al.* (2022) desenvolveram um algoritmo denominado Smart Optimization Program for Water Distribution Networks, que aplica um algoritmo genético à otimização do projeto de três redes de distribuição de água, Hanoi, GoYan e Two-loop, com o objetivo de minimizar os custos e garantir a eficiência do desempenho hidráulico. A pesquisa enfatiza a importância da otimização dos diâmetros dos tubos em redes predeterminadas, uma vez que esses sistemas representam uma parcela substancial dos custos de construção, operação e manutenção em projetos de abastecimento. Os resultados foram consistentes para as três redes utilizadas, contribuindo como uma ferramenta promissora na projeção de redes e na reabilitação de sistemas de distribuição de água.

O estudo de Islam *et al.* (2022) apresentou uma revisão sistemática das tecnologias atuais e direções futuras para a detecção de vazamentos em redes de distribuição de água, destacando a importância crítica dessa temática devido às perdas econômicas significativas causadas por vazamentos em sistemas de abastecimento subterrâneos. A pesquisa analisou 47 artigos publicados, identificando tendências e lacunas nas abordagens tecnológicas empregadas, especialmente no uso de sensores de vibração, acústicos e de fluxo, bem como a crescente aplicação de algoritmos de aprendizado de máquina e tecnologias de Internet das Coisas (IoT) para melhorar a detecção e localização de vazamentos. Além disso, o estudo discutiu o papel das tecnologias de comunicação que dominam cerca de 80 % dos métodos analisados, e aponta para futuras direções de pesquisa que envolvem integração de métodos avançados e automação no monitoramento de redes hidráulicas.

Santos-Ruiz *et al.* (2022) apresentaram uma metodologia para a seleção ótima de locais de instalação de sensores de pressão em redes de distribuição de água, empregando princípios da teoria da informação para maximizar a utilidade dos dados coletados na localização de vazamentos. A proposta foi baseada na análise de mudanças de pressão em diferentes cenários de vazamentos gerados sinteticamente por meio de simulações no software EPANET, a fim de identificar subconjuntos de nós que forneçam maior informação relevante e menor redundância entre medições correlacionadas. A implementação foi realizada sobre a rede Hanoi e os resultados mostraram que o modelo permitiu determinar os nós que ofereceram a melhor combinação para detectar e localizar vazamentos.

Marques (2023), trouxe em sua pesquisa uma metodologia para a otimização de uma rede de distribuição de benchmark utilizando de duas ferramentas: a busca aleatória e o algoritmo genético, com o objetivo da redução de perdas. Os parâmetros utilizados para a realização da otimização foram a rugosidade, diâmetro, nível do reservatório bombas e as válvulas redutoras de pressão. O resultado mostrou que o algoritmo genético foi melhor em encontrar a menor perda para o sistema, sendo essa redução de 15,75%.

Silva e Rodrigues (2023) apresentou um modelo computacional para a minimização dos custos em projetos de redes de distribuição de água, utilizando de algoritmo genético com penalização adaptativa para o controle da diversidade populacional do algoritmo. O estudo foi aplicado a três redes de benchmark de distribuição para a validação da metodologia. Os resultados mostraram que o algoritmo desenvolvido foi eficiente em minimizar o custo obtendo a melhor configuração em projeto com o menor custo.

Ortiz *et al.* (2023) trouxeram um estudo sobre a otimização da calibração de duas redes de distribuição de água, uma da cidade de Zacatecas no México e uma rede de benchmark. O algoritmo desenvolvido consistiu na aplicação do algoritmo genético - FEMH. O modelo apresentado foi eficiente para calibração das redes, sendo assim, segundo os autores, um modelo confiável porque representou as condições reais das redes.

Parvaze *et al.* (2023) apresentaram uma revisão abrangente sobre a aplicação de algoritmos genéticos na otimização de redes de distribuição de água, destacando sua importância na minimização do consumo de água e na maximização da eficiência operacional. A revisão realizada mostrou que os algoritmos genéticos representam uma técnica eficaz para enfrentar problemas de otimização não lineares típicos desses sistemas, além de contribuir com perspectivas sobre expansão, reabilitação e reforço de redes existentes. Dessa forma, o estudo oferece uma visão consolidada das aplicações e desafios associados à otimização de redes de distribuição de água, servindo como referência para futuras pesquisas e atividades práticas voltadas à sustentabilidade e tomada de decisão baseada em métodos computacionais avançados.

Lara *et al.* (2024) desenvolveram uma metodologia para a comparação de dois algoritmos, genético evolucionário e de busca aleatória para a operação eficiente, ou seja, funcionamento sobre pressões adequadas de duas redes de benchmark de distribuição de água. Segundo os autores foram desenvolvidos nove cenários com a variação dos parâmetros de recombinação e mutação do algoritmo. Os resultados mostraram que ambos os algoritmos

conseguiram adequar as pressões tornando o sistema mais eficiente, porém aqueles que possuem taxa de mutação menor que 5% apresentaram os melhores resultados.

Em Shiddiqi *et al.* (2024) foi desenvolvido uma pesquisa sobre a otimização de uma rede de distribuição de água utilizando de algoritmos genéticos, para encontrar os pontos excelentes para a aplicação de sensores aumentando a confiabilidade do sistema na detecção de perdas. As variáveis utilizadas para o modelo foram a intensidade dos vazamentos, locais de ocorrência e frequência dos eventos. Para encontrar o cenário ideal foram realizadas inúmeras simulações de vazamento, além da avaliação do método.

A pesquisa desenvolvida por Shiu *et al.* (2024), trouxe um estudo de calibração de uma rede de distribuição de água na cidade de Taipei na China. O algoritmo genético realiza a otimização dos coeficientes de rugosidade através da comparação dos dados obtidos com os dados simulados, adquiridos por meio da automatização da variação das rugosidades. O modelo possibilitou encontrar a rugosidade ideal do sistema, além disso, auxiliou na gestão do sistema de água.

Pavansimha e Javeed (2025) investigaram a aplicação de um algoritmo genético integrado ao EPANET e o Python para otimização de redes de distribuição de água, com foco na redução dos custos totais das tubulações sem comprometer a pressão nos nós da rede. A pesquisa foi realizada sobre duas redes: a Hanoi e a Koodlahalli. Os resultados demonstraram que a combinação entre o algoritmo evolutivo e a simulação hidráulica possibilitou uma redução significativa no custo da rede ao mesmo tempo em que manteve a pressão mínima exigida durante a operação.

3.5 Machine Learning

3.5.1 Conceituação

A técnica de Machine Learning é definida como uma subárea da inteligência artificial, que tem como tradução Aprendizado de Máquina, que permite que os sistemas submetidos a essa ferramenta sejam capazes de tomar decisões de acordo com suas próprias experiências (Janiesch *et al.*, 2021).

As etapas de aplicação dessa técnica se dividem em partes básicas: treinamento, execução ou teste. Na primeira etapa, o treinamento, o sistema passam por uma contextualização do problema, onde dados são fornecidos para que o padrão seja analisado. Esses dados disponibilizados podem ter passado por uma classificação prévia, onde as entradas e saídas são conhecidas, sendo denominada aprendizado supervisionado. Já os dados

que não passam por essa classificação prévia são chamados de aprendizado não supervisionado. Com o modelo treinado é possível realizar a fase de teste (execução), que é caracterizada pela determinação dos resultados esperados, com base nos padrões definidos anteriormente (Dalarmelina, 2023; Janiesch *et al.*, 2021).

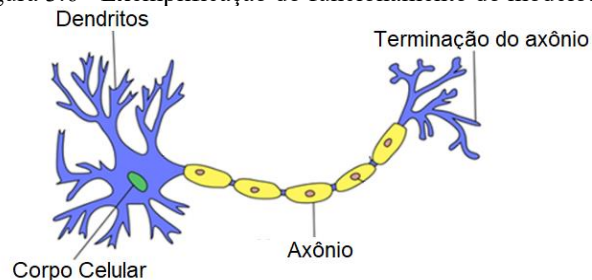
A aprendizagem de máquina também pode se diferenciar em tipos de classificação, podendo ser de Regressão, Classificação e Clusterização. Sendo que os modelos do tipo Regressão e Classificação são técnicas de aprendizado supervisionado e a Clusterização é não supervisionado (Campesato, 2020; Marsland, 2011).

Segundo Marsland (2011), ainda existem dois tipos de aprendizado, o de reforço e o evolutivo. No reforço o algoritmo é informado do erro, porém, não é dito como resolver, sendo necessário que o algoritmo encontre a resposta por si mesmo, testando as diversas possibilidades. Já o aprendizado evolutivo se define como um processo de aprendizado em que os organismos se adaptam para que possam ter maior chance de sobrevivência e de reproduzir, se assemelhando nesses aspectos ao algoritmo genético.

Dentre os modelos existentes de Machine Learning as Redes Neurais Artificiais (RNA) se destacam pela capacidade de aprendizado e com isso, sendo mais eficiente na resolução de problemas das diversas áreas do conhecimento.

Com isso, a RNA é uma técnica de aprendizado de máquina que foi desenvolvido com base no funcionamento do cérebro humano, em que há relação entre o conjunto de sinais de entrada e saída, assim como o funcionamento dos neurônios. Os neurônios biológicos recebem sinais de entrada nos dendritos, localizados no corpo celular, que se acumulam e a célula dispara o sinal de saída pelo axônio até sua terminação, que irá transmitir para os próximos neurônios (Dangeti, 2017). A estrutura do neurônio biológico está exemplificada na Figura 3.6.

Figura 3.6 - Exemplificação do funcionamento do modelo.

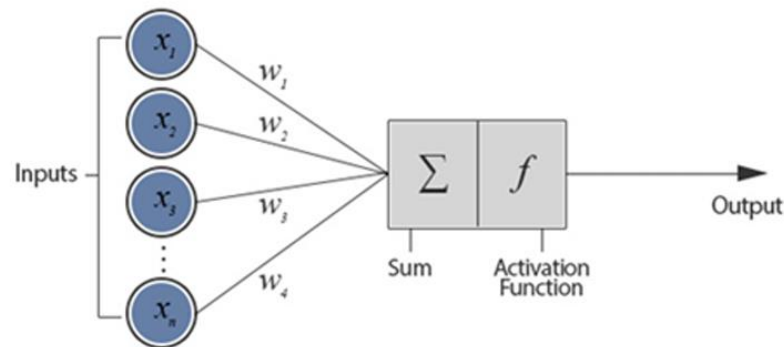


Fonte: Adaptado de Dangeti (2017).

Com base no funcionamento do neurônio biológico a RNA utiliza de neurônios artificiais interconectados para a resolução de problemas. Assim, seu funcionamento se dá

basicamente com cada neurônio calculando uma função de acordo com as entradas ponderadas recebidas, que são formadas por um conjunto de entradas, em que cada uma recebe um peso específico. Um neurônio linear realiza a combinação dessas entradas ponderadas e aplica uma função de ativação da soma agregada. E por fim é retornada a saída de acordo com o resultado da função (Dangeti, 2017). A Figura 3.7 apresenta de maneira generalista o funcionamento de um neurônio em RNA.

Figura 3.7 - Exemplificação do funcionamento do modelo.



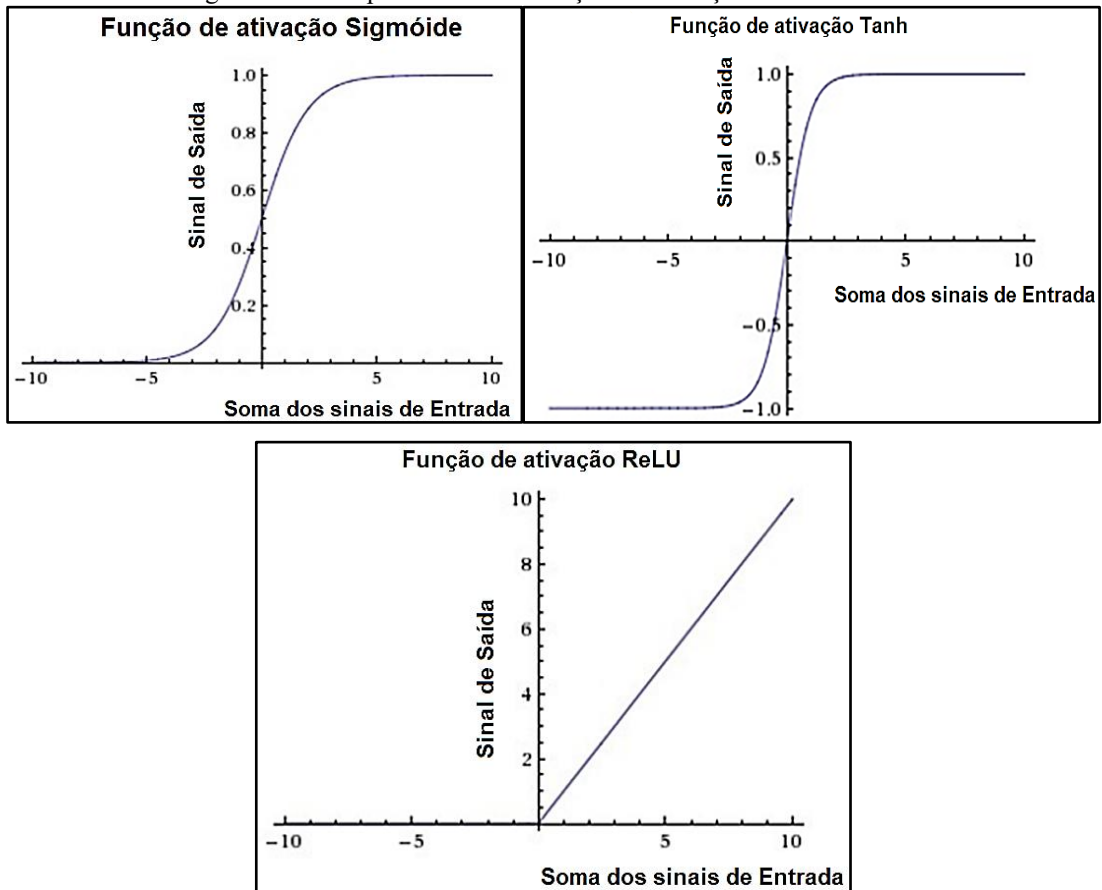
Fonte: Adaptado de Dangeti (2017).

O neurônio artificial pode ser representado pela Equação 1. Onde os pesos (W) possibilitam que cada uma das n entradas de x contribua com uma quantidade maior ou menor para a soma dos sinais de entrada. O $f(x)$ corresponde a função de ativação e o axônio de saída está representado pelo $y(x)$ (Dangeti, 2017).

$$y(x) = f\left(\sum_{i=1}^n w_i x_i\right) \quad (1)$$

Nas redes neurais artificiais podem existir diferentes tipos de função de ativação dependendo do nível de controle do neurônio, sendo as mais comuns as Sigmoid, Tanh e Relu. A função Sigmoid transforma os valores reais em intervalos de 0 a 1. E tem a forma matemática $\sigma(x) = 1 / (1 + e^{-x})$. Já a função Tanh transforma os valores reais no intervalo -1 e 1 e sua equação é dada por $\tanh(x) = (e^x - e^{-x}) / (e^x + e^{-x})$. E a função de Unidade Linear Retificada-ReLU sua ativação se dá em limiares de zero, tendo sua fórmula em $f(x) = \max(0, x)$ (Dangeti, 2017). O comportamento dessas funções está exemplificado na Figura 3.8.

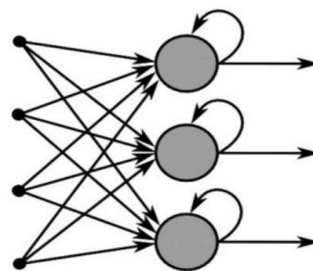
Figura 3.8 - Comportamento das funções de ativação mais utilizadas.



Fonte: Adaptado de Dangeti (2017).

Existem alguns tipos de redes neurais artificiais, se destacando para o presente estudo as RNA Recorrentes (RNAR), que se destacam por ter um tempo de treinamento e a complexidade de implementação minimizados. Esse tipo de RNA se caracteriza pela realimentação do modelo, além da conexão entre os neurônios, assemelhando na estrutura com as redes neurais de feedforward, mas se diferenciam no treinamento do modelo, pois as RNAR necessitam de poucas iterações e parâmetros para seu treinamento (Unadkat; Ciocoiu; Medsker, 2001; Toscano; Cabral, 2008). A Figura 3.9 mostra como é a estrutura básica de uma rede neural recorrente.

Figura 3.9 – Estrutura básica da rede neural artificial recorrente.



Fonte: Data Science Academy (2025).

Assim, destaca-se a Rede Neural Artificial Recorrente - Long Short-Term Memory (RNAR-LSTM), que foi proposta por Hochreiter e Schmidhuber (1997). Ela foi originalmente avaliada em problemas artificiais de dependência temporal de longo prazo. Esse experimento demonstrou a capacidade da rede em armazenar e recuperar informações ao longo de grandes intervalos de tempo, superando limitações das redes neurais recorrentes tradicionais. Assim, a finalidade do desenvolvimento dessa ferramenta foi de superar as limitações que as primeiras redes neurais recorrentes possuíam em relação ao seu treinamento (Correia *et al.*, 2021).

Essas redes neurais usam conexões reversas, em que os nós podem se conectar com os nós de camadas anteriores e consigo mesmo, construindo o ciclo direcionado. Com isso, essa rede tem capacidade de memorização, o neurônio de entrada que possui o valor passado do neurônio anterior junto com a entrada da camada anterior (Aurélien, 2019). Assim, a saída atual é influenciada pelos valores de saída passados para o neurônio e pelas entradas anteriores.

Esse tipo de rede é usado em problemas que exigem o reconhecimento de padrões em que o resultado anterior influencia os resultados atuais, como é o caso de dados de séries temporais. Contudo, esse tipo de rede sofre com a questão de se lembrar das entradas recentes e esquecer de maneira rápida das entradas de longo prazo (Hochreiter e Schmidhuber, 1997).

As Redes Neurais Recorrentes Long Short Term Memory (RNAR-LSTM) possuem uma estrutura que se diferencia das redes neurais recorrentes tradicionais, principalmente por apresentarem duas memórias, uma curta e uma longa, determinados pelos portões internos de cada célula e pelo seu estado atual. Esses portões geralmente são formados por funções sigmóides que tem a funcionalidade de controlar o quanto de dados devem ser mantidos em cada um dos módulos (Aurélien, 2019).

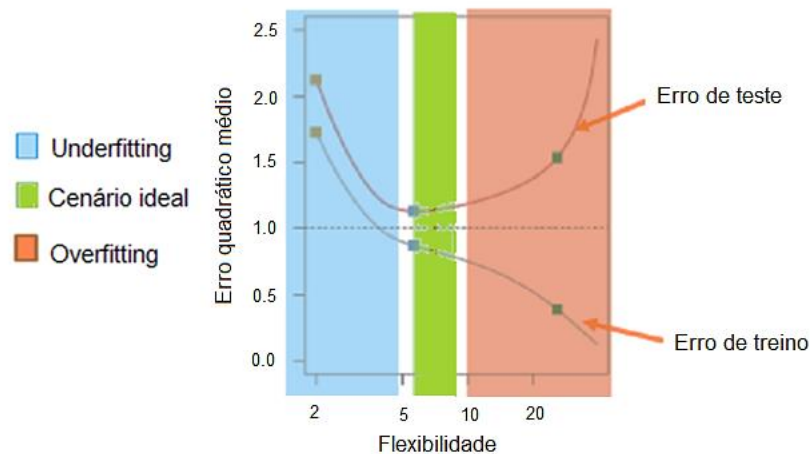
A aplicação do modelo de aprendizado de máquina em estudo traz um problema que está no tratamento dos dados a serem utilizados, pois, as bases de dados constantemente não estão limpas e nem consistentes para serem aplicadas, podendo com isso, prejudicar os processos e análises, levando o erro de interpretação e diminuindo a precisão do modelo. Assim, a correção de dados errados e faltantes, a retirada de dados duplicados, determinação sobre outlier são tarefas que envolvem a limpeza dos dados, que melhoram a consistência dos mesmos a serem utilizados, melhorando também a resposta entregue pelo modelo (Unpingo, 2016).

Os modelos de aprendizado de máquina de maneira geral podem passar por dois problemas devido ao ajuste de seus dados, que pode levar ao erro e à imprecisão do modelo,

sendo eles o overfitting e underfitting. Overfitting trata do ajuste em excesso dos parâmetros tornando o problema específico para um certo conjunto de treinamento, ao invés de ser generalista (Chaves, 2012; Mohri *et al.*, 2012).

De acordo com Dangeti (2017) o overfitting pode ser representado por um baixo treino, altos erros de teste e baixa precisão de teste. Já o underfitting os dados de treinamento e teste não geram taxas de erro baixas, sendo necessário o reajuste do algoritmo para melhorar o seu desempenho (Chaves, 2012). É representado por um alto treino e erro de teste e baixa precisão (Dangeti, 2017). A Figura 3.10 apresenta o overfitting, underfitting e o cenário ideal para os modelos de aprendizado de máquina, em que o treino e o teste são baixos e a precisão é alta.

Figura 3.10 - Overfitting, Underfitting e cenário ideal nos ajustes de dados de modelos de Machine Learning.



Fonte: Adaptado de Dangeti (2017).

Devido a isso, são utilizadas ferramentas de validação como a hold-out e validação cruzada. A hold-out consiste na divisão simples dos dados em treino e teste, sendo que o desempenho do modelo é realizado pela porção destinada ao teste, e a divisão geralmente é realizada em 80% para treino e 20% para teste. Porém, quando são utilizados modelos diferentes, uma parte dos dados pode ser dívida para o cálculo dos erros e com isso ser possível realizar a modificação dos parâmetros do modelo para reduzir os erros (Scaccia, 2020).

A outra técnica é a validação cruzada em que os dados se dividem em dois grupos com n partes iguais, e cada uma dessas partes será subdividida ocasionando duas amostras, uma para o treino e validação e outra para teste. Essa etapa pode ser repetida de acordo com a necessidade do modelo e seu erro geral é determinado frequentemente pela média das estimativas dos erros (Unpingo, 2016; Dangeti, 2017; Mohri *et al.*, 2018).

Dangeti (2017) apresentou em seu estudo um exemplo de como a validação cruzada dos dados é realizada. Ele considerou uma validação cruzada quádrupla dos dados, onde os dados foram divididos em cinco partes iguais e o treinamento foi realizado em quatro partes (80%) e o teste em uma parte (20%) desses dados. Esse processo foi repetido cinco vezes para que todo o conjunto de dados fosse coberto. E o erro final foi determinado pela média dos erros. A Figura 3.11, traz essa exemplificação de como foi realizada a validação cruzada.

Figura 3.11 - Exemplificação da validação cruzada de dados.



Fonte: Adaptado de Dangeti (2017) e Mohri *et al.* (2012).

3.5.2 Aplicação de Machine Learning em estudos de redes de distribuição de água

A seguir, são apresentados os estudos que se utilizam da técnica de Machine Learning para o estudo de redes de distribuição de água.

Em Alves (2022) apresenta uma metodologia para a previsão de anomalias e consumo de uma rede de distribuição de água da região de Algarve em Portugal, na zona de medição da Quinta do Lago. Para isso, foram utilizados dados de uma série temporal (2015 a 2019) do consumo doméstico da rede de estudos. A metodologia desenvolvida utilizou inicialmente a técnica de clustering para que os nós semelhantes fossem agrupados, e a esse agrupamento foi aplicado um modelo de Machine Learning baseado em três modelos de aprendizado, o ARRIMA, RNA e Random Forest. O resultado encontrado pelo autor é de que as técnicas utilizadas para a previsão de consumo foram eficientes, porém, para a previsão de anomalias o resultado não foi eficiente, pois a falta de dados no período analisado, de alguma anomalia que já havia ocorrido, prejudicou a definição da veracidade da ferramenta, mostrando assim, a necessidade de aprimoramento da mesma em trabalhos futuros.

Gouveia (2022) trouxe em seu trabalho a avaliação do potencial de utilização dos classificadores de Machine Learning para a gestão adequada de ativos e controle de perdas, com o uso de dados físicos, comerciais e ambientais de uma rede da CAESB - DF. Foram utilizados 12 modelos de ML, na linguagem Python com auxílio de bibliotecas do mesmo de

ML e ciência de dados. Segundo o autor, os modelos do tipo Ensemble Learning (junção de mais de um modelo de Machine) apresentaram melhor resultado na acurácia, sendo que o modelo do tipo *Adaboost* se destacou, com um valor de 59,70% na acurácia do modelo.

Roller *et al.* (2022) desenvolveram um estudo para o monitoramento de rede de distribuição de água com foco na detecção de vazamentos de água nesses sistemas. Para isso, foi aplicada uma técnica de aprendizado de máquina que se baseia em Redes Neurais de Grafo, que levam em consideração variáveis medidas pelos sensores presentes no sistema, como a pressão, além da localização dos nós, para que fosse possível identificar as perdas reais. A rede de distribuição utilizada no estudo foi desenvolvida pelos próprios autores, sendo assim um sistema hipotético de distribuição. Os resultados encontrados mostraram-se promissores para a detecção de perdas físicas em redes de distribuição com o uso de RN de Grafos, porém, segundo os autores os resultados ainda são preliminares, necessitando de mais tempo de simulação para os resultados definitivos dessa aplicação.

Cariço e Ramos (2023) propuseram um modelo de gestão de pressão para auxiliar na gestão de sistemas de distribuição de água com o objetivo da redução de perdas físicas, visando a eficiência hidroenergética. Para a elaboração do modelo foi utilizado o EPANET para a simulação hidráulica da rede e a técnica de Redes Neurais Artificiais para a redução das perdas no sistema, tendo em vista a instalação de micro turbinas do tipo PAT. A instalação desse elemento ativo no sistema foi favorável para o sistema de estudo, pois foi capaz de reduzir as perdas durante sua operação, e de gerar energia elétrica tornando o sistema mais sustentável.

Oliveira *et al.* (2021) trouxeram um estudo para auxiliar na gestão de redes de distribuição de água quanto a previsão de perdas físicas de água desses sistemas. Para isso, foi utilizada a ferramenta de geoprocessamento SIG, para o levantamento exato da rede e o algoritmo de otimização Random Forest Regressor. O sistema de estudo foi uma rede do município de Vila Nova de Gaia em Portugal, onde dados históricos da rede do período de 2017 a 2019, foram utilizados para o aprendizado do modelo desenvolvido. O resultado encontrado com a aplicação do modelo desenvolvido foi eficiente pois conseguiu prever as áreas mais suscetíveis a perdas e com isso, possibilitar a implementação de medidas para evitar que ocorram.

Sousa *et al.* (2022) mostraram um estudo de calibração da rugosidade de um sistema de distribuição de água com a aplicação de um modelo híbrido que associa o método iterativo do gradiente hidráulico - Migha e o Rede Neurais Artificiais do tipo Perceptron, com o

objetivo de controle de pressão da rede e de perdas. A rede empregada no estudo é hipotética, elaborada por Sousa (2020) em que todos os parâmetros da rede eram conhecidos. O modelo híbrido utilizado foi capaz de encontrar uma calibração ideal para o sistema, levando em consideração as pressões em que o erro médio foi inferior de 0,016 mm.

Farghadan *et al.* (2023) propuseram um conceito de injeção e propagação de falhas em um sistema de rede de distribuição de água digital para a detecção das características das perdas. A metodologia consistiu em gerar um vazamento no sistema com características conhecida e com isso propagar os efeitos do vazamento original, com um classificador de Ensemble do tipo Bayesiano que processa os dados conjuntos e determina a localização do vazamento e de futuros. As redes utilizadas no estudo foram a Hanói e Modena, que são redes utilizadas em inúmeros estudos da literatura. Segundo os autores, o resultado encontrado se mostrou preciso para a previsão das perdas, com uma precisão de 94,6% e 93,76% para as redes Hanói e Modena, respectivamente.

Fu *et al.* (2024) apresentaram um modelo para a eficiência da operação de redes de distribuição de água por meio de um modelo de previsão e classificação de perdas baseado em Deep Learning (GA-LSTM) e Reconhecimento dinâmico de padrões multilíneares. A rede aplicada no estudo é uma rede real da cidade de Gwangju, na Coreia do Sul. Para o desenvolvimento da técnica aplicada foi utilizado dados de operação da rede, como pressão, para a verificação e treinamento dos modelos. Também foi realizada a análise da influência de diferentes arranjos de funcionamento dos elementos ativos para auxiliar na estabilização do modelo. A aplicação do modelo GA-LSTM-DMTPR (junção de duas técnicas de Machine Learning) foi favorável para o estudo, pois apresentou uma alta precisão na previsão de demanda e das perdas, bem como na classificação dessas perdas.

Silva *et al.* (2024) utilizaram dois modelos, ARIMA e Redes Neurais Artificiais multicamadas do tipo perceptron, para a previsão de pressão e vazão em um sistema de distribuição de água real, em uma cidade do sul de Minas Gerais, Brasil, com base em dados da rede. O objetivo foi aumentar a confiabilidade do sistema em atender as demandas da rede, além de auxiliar na redução das perdas de água. O resultado apresentado pelos autores mostrou que o modelo ARIMA possui maior capacidade de previsão que as RNA do tipo perceptron, além disso, são mais fáceis de serem aplicadas para o tipo de estudo que foi desenvolvido.

Pires (2024) apresentou um estudo sobre a comparação entre modelos de previsão de séries temporais, Holt-Winters, ARIMA, LSTM e Prophet, para dados caudais de água

distribuídos em quatro setores de abastecimento em zonas rurais localizadas em Portugal. Para determinar o melhor modelo para a previsão foram utilizadas métricas de avaliação como erro médio absoluto, raiz quadrática média e coeficiente de determinação. A previsão foi realizada para dois períodos: um de curto e um de longo prazo. Os resultados mostram que o modelo de previsão LSTM foi mais eficiente para previsões de longo prazo.

Gómez-Coronel *et al.* (2024) desenvolveram um estudo para localização de vazamento em uma rede de distribuição de água utilizando a separação de setores para a definição de nós próximos e a rede neural artificial recorrente LSTM para encontrar cenários de vazamento de perdas que podem ocorrer no sistema estudado. Para a alimentação do modelo RNAR-STM foram utilizados dados de pressão e vazão do sistema estudado. O resultado mostrou que o modelo foi eficiente em encontrar vazamento no sistema.

Gómez-Coronel *et al.* (2025) propuseram uma metodologia para a determinação de vazamentos em sistemas de distribuição de água de benchmark utilizando a rede neural LSTM sendo aplicada ao software MATLAB. Para o desenvolvimento do modelo foram utilizados dados de topologia da rede e as pressões de alguns nós. Os resultados mostram o bom desempenho da metodologia na identificação dos nós em que ocorrem os vazamentos.

Mahdi *et al.* (2024) apresentaram um estudo sobre identificação e localização de perdas em um sistema teórico de escala laboratorial sendo uma rede com layout ramificado. Para o desenvolvimento da metodologia foram utilizados dados históricos coletados de sensores presentes no sistema, como de acelerômetros, pressão dinâmica, coeficiente de autocorrelação e energia do sinal. Esses parâmetros foram aplicados a Rede Neural Artificial com neurônios e camadas ocultas com o objetivo de alcançar eficiência na localização e detecção das perdas. O resultado apresentado pelos autores mostrou que o modelo desenvolvido foi eficiente em alcançar o objetivo proposto de detectar perdas da rede de distribuição de estudo.

3.6 Gêmeos Digitais

3.6.1 Conceituação

Gêmeos Digitais foi um termo desenvolvido em 2002 na Universidade de Michigan nos EUA pelo Dr. Michael Grieves para o desenvolvimento de Product Lifecycle Management (Mykoniatis; Harris, 2021). Este conceito consiste em um modelo virtual dinâmico e progressivo de um objeto do ambiente real, podendo ser uma máquina, um processo, peças entre outros, que caracteriza o gêmeo físico de maneira precisa em tempo

real, por meio da troca de informações. Além de realizar a manutenção de dados históricos (Singh *et al.*, 2021).

De acordo com Nascimento *et al.* (2022), o Gêmeo Digital pode ser dividido em três partes: o produto físico, correspondendo ao objeto do ambiente físico; o produto virtual, é o objeto físico definido no ambiente virtual; e a conexão dos dados entre os dois produtos, ou seja, a transcrição em tempo real das informações do ambiente real para o virtual. Englobando dados do tipo quantitativos, qualitativos, dados históricos e dados do ambiente (Singh *et al.*, 2021). Além disso, esse modelo permite que as informações do gêmeo virtual possam ser aperfeiçoadas em relação à quantidade disponível, sendo adicionadas características e comportamentos para que possibilite testar a capacidade de desempenho, bem como, a visualização do produto (Grieves, 2014).

Devido à sua aplicabilidade ser ampla e multidisciplinar, Singh *et al.* (2021) afirmam que os Gêmeos Digital podem ser classificados em diferentes tipos, que levam em consideração características e peculiaridades do modelo, tendo como destaque o tempo de criação, nível de integração e nível de maturidade.

O tempo de criação se subdivide em protótipo e instância. O protótipo corresponde ao modelo virtual que contém todos os dados necessários para a criação do modelo real. Sua utilização se dá para que testes possam ser realizados antes da criação do modelo físico, permitindo a identificação dos cenários mais adversos que muitas vezes são difíceis de serem identificados no sistema real (Singh *et al.*, 2021). Dessa forma, quanto mais precisas forem as simulações melhor será o gêmeo físico.

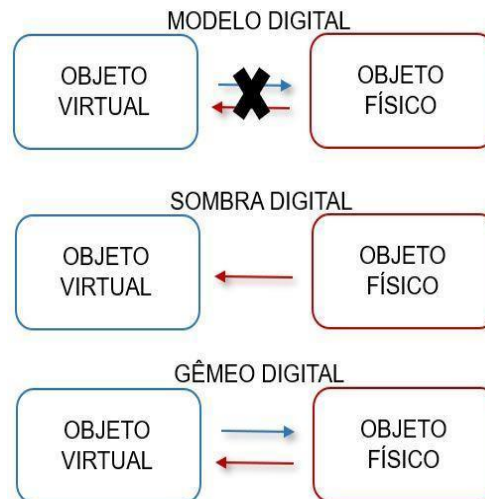
A Instância se caracteriza pelo processo inverso, o gêmeo físico já existe e é desenvolvido o virtual para a previsão e monitoramento do sistema, ocorrendo uma comunicação direta do físico para o virtual e do virtual para o físico. Sua aplicação permite verificar se o comportamento do sistema físico é ideal e também analisar os cenários adversos que podem ser evitados (Grieves; Vickers, 2017).

Ainda segundo Grieves e Vickers (2017) o uso da prototipagem tem o menor custo de utilização, porque possibilita a realização de testes destrutivos, a recriação do sistema, otimizar a operação e prolongar a vida útil de equipamentos sem nenhum custo adicional, pois essas alterações ocorrem somente no ambiente virtual.

Quanto ao nível de integração, os GD se caracterizam em Modelo, Sombra e Gêmeo Digital. O Modelo é uma representação de um objeto físico, porém não ocorre troca dinâmica de dados, sendo estes inseridos manualmente. Já na Sombra há uma representação do objeto

físico, porém a interação de informações do físico para o virtual. E por fim o GD é um modelo mais evoluído e a troca de dados é bidirecional, ou seja, há troca entre o virtual e real e vice-versa (Kritzinger *et al.*, 2018). De acordo com Singh *et al.* (2021) a sombra é considerada como um modelo estático em que os dados são adicionados com o tempo, sendo melhor para o planejamento estratégico de sistemas e o GD é um modelo dinâmico, com os dados sendo transmitidos em tempo real. A Figura 3.12 exemplifica o funcionamento dos níveis de integração dos gêmeos digitais.

Figura 3.12 - Tipos de nível de integração do Gêmeo Digital.



Fonte: Adaptado de Nascimento (2022).

Já o nível de maturidade dos gêmeos pode ser de três tipos: Pré-Digital, Gêmeo Digital e Gêmeo Digital Adaptativo. O Pré-Digital (nível 1), é um modelo virtual criado antes do físico, com o objetivo de evitar os problemas e riscos técnicos que podem aparecer em um sistema real. No Gêmeo Digital (nível 2), o ativo virtual incorpora informações do ativo físico, podendo o virtual controlar o físico, caracterizando assim, uma troca mútua entre os dois ativos. O GD adaptativo (nível 3), utiliza o aprendizado de máquina para entender as preferências e prioridades dos operadores, tornando possível a tomada de decisão e o planejamento em tempo real (Madni *et al.*, 2019).

Assim, os Gêmeos Digitais podem ser aplicados em diferentes necessidades de produção, como no apoio na gestão, monitoria, melhoria e flexibilização dos processos de produção, auxílio no ciclo de vida de máquinas, permite o aumento da segurança do sistema e da interação homem máquina (Cimino; Negri; Fumagali, 2019). Em sistemas de distribuição de água eles podem ser aplicados para o gerenciamento de pressão, simulação, operação e otimização de válvulas e bombas, na previsão de demanda da rede, previsão e avaliação de perdas de água e na análise de contaminação da água da rede (Singh *et al.*, 2021).

Sua utilização também pode ser observada em estudo na área de recursos hídricos como a observada no estudo apresentado por Li *et al.* (2022), que investigou a aplicação da tecnologia de gêmeos digitais na gestão integrada de água e energia no contexto da construção de cidades sustentáveis, com foco no Sistema de Drenagem Urbana. Os autores detalham a construção do modelo de gêmeo digital, identificando os elementos essenciais do sistema de drenagem e os dados necessários para sua modelagem, e propõem um algoritmo de escalonamento dinâmico de prioridades e importância, para melhorar o desempenho de estações de bombeamento em cenários de chuva intensa. Os resultados experimentais demonstraram que o algoritmo apresentou melhorias substanciais na operação do sistema, o que sugere que a adoção de gêmeos digitais pode reduzir perdas por transbordamento e aumentar a eficiência operacional de sistemas de drenagem urbana, contribuindo para o desenvolvimento sustentável das cidades.

Zekri, Jabeur e Gharrad (2022) propuseram um modelo de gestão inteligente da água fundamentado em gêmeos digitais, com o objetivo de otimizar o uso e a distribuição de água potável em grandes comunidades diante da escassez de recursos hídricos. Os autores destacam que há deficiência de autonomia ao longo de todo o ciclo de aquisição, processamento de dados e controle de ativos, o que limita a eficácia das estratégias de gerenciamento. Nesse sentido, o modelo apresentado pelos autores combinou análise de dados para avaliação do consumo, mecanismos de feedback dos usuários e simulação das operações de infraestrutura sob restrições específicas de trabalho e detecção de falhas ou defeitos. A proposta evidenciou o potencial dos gêmeos digitais, como uma abordagem promissora para aprimorar a eficiência, a equidade e a sustentabilidade no gerenciamento de recursos hídricos.

O trabalho de Wu, Lu e Wu (2023) apresentou uma revisão abrangente sobre a utilização de tecnologias de sensoriamento remoto na construção de bacias digitais por meio de gêmeos digitais, utilizando como variáveis hidrológicas essenciais a precipitação, temperatura da superfície, evapotranspiração, níveis de água, descarga dos rios, umidade do solo e vegetação. Neste estudo, os autores ressaltaram que o sensoriamento remoto com a aplicação dos gêmeos digitais possibilita uma cobertura espacial ampla e potencializa funções como monitoramento de secas, previsão de precipitação e gestão de recursos hídricos, embora apontem a necessidade de melhorias na precisão e latência dos dados. Assim, a pesquisa contribuiu para aprofundar a compreensão das aplicações atuais e futuras dos gêmeos digitais no contexto da hidrologia moderna, fornecendo um fundamento teórico e técnico para

avanços futuros na integração de sensoriamento remoto das bacias com a aplicação da digitalização das bacias.

Araujo e Silva (2023) abordaram a aplicação de uma plataforma de gêmeos digitais à modelagem e simulação computacional de um sistema de microdrenagem de uma região da cidade de Salvador com o objetivo de mitigar os picos de vazão durante os eventos críticos de chuva, aplicado em lagoas de retenção. Os resultados evidenciaram que o modelo conseguiu reduzir significativamente os picos de vazão, embora apresente sobrecarga em situações de precipitações superiores a 10 mm. Adicionalmente, a adoção da ferramenta demonstrou capacidade de previsão em tempo real, contribuindo para a integração de ações preventivas e corretivas, fortalecendo a resiliência urbana frente aos riscos de inundação. Assim, o trabalho reforça a importância da incorporação de gêmeos digitais na gestão de sistemas urbanos de água como instrumento de apoio à tomada de decisões por órgãos gestores e à promoção de cidades mais seguras e sustentáveis.

Daneshgar *et al.* (2024) apresentaram o desenvolvimento e a implementação de um gêmeo digital operacional em escala real para a recuperação de recursos hídricos de Eindhoven, nos Países Baixos, como um exemplo robusto de aplicação de tecnologia digital ao setor de tratamento de águas residuais. O estudo descreveu a integração de um modelo detalhado de toda a planta e uma interface de usuário desenvolvida com os operadores, permitindo simulações em tempo quase real a cada duas horas e a previsão dinâmica de configurações operacionais com até 48 horas de antecedência. A ferramenta demonstrou capacidade de suportar análises de cenários operacionais, mitigação de falhas, otimização de processos e suporte à decisão de longo prazo, destacando o papel dos gêmeos digitais como instrumentos avançados para melhorar a eficiência, resiliência e confiabilidade de sistemas complexos de recuperação de água.

Syed *et al.* (2024) trouxeram uma revisão abrangente da literatura sobre a aplicação integrada de gêmeos digitais e técnicas avançadas de inteligência artificial no contexto da gestão da demanda de água, identificando lacunas nos métodos tradicionais de previsão e controle. Os autores argumentam que a integração de um gêmeo digital com modelos de IA permite superar limitações das abordagens isoladas, proporcionando maior precisão na previsão da demanda de água, melhor eficiência operacional e potencial contribuição para a sustentabilidade hídrica global. Além disso, os autores destacam a importância de utilizar dados em tempo real e modelos preditivos robustos para aprimorar a tomada de decisão em

sistemas complexos de gestão de água, estabelecendo uma base conceitual para futuras implementações práticas desse modelo.

Joaquim, Piqueiro e Kai (2025) abordaram o desenvolvimento de sistemas de gestão de águas pluviais em áreas urbanas por meio da utilização de gêmeos digitais, destacando os desafios e as soluções associadas ao manejo de escoamentos. O trabalho foi realizado sobre a região de Luanda na Angola, utilizando análise de dados cartográficos e pluviométricos para identificar causas de eventos de inundação em um trecho urbano específico dessa região, propondo intervenções como a integração de sistemas de drenagem funcional e dispositivos complementares que aumentem a eficiência do fluxo de água e reduzam a ocorrência de acúmulo superficial, tornando o sistema mais eficiente, mostrando que os gêmeos foram eficaz para o que se proporam.

Ma *et al.* (2025) apresentaram um modelo integrado de gêmeos digitais com machine learning para apoiar a transformação digital de instalações de tratamento de água. Foi utilizado um estudo de caso aplicado a plantas de tratamento em escala real da região de Pequim na China. A infraestrutura de dados foi baseada na Internet das coisas-IoT e computação em nuvem conectando equipamentos de tratamento, sistemas intermediários e análise preditiva para permitir controle e manutenção proativa dos processos e na prevenção de fenômenos no sistema, além da otimização multiobjetivo de consumo energético. Os resultados demonstraram que o modelo proposto encontrou reduções no consumo de energia e proporcionou insights práticos para a adoção de sistemas inteligentes na gestão do processo, além de reduzir os desafios técnicos e organizacionais que limitam a ampla implementação dos gêmeos digitais no setor de saneamento.

Gümüş, Alver e Akbal (2026) desenvolveram um gêmeo digital orientado por dados para enfrentar desafios operacionais dos sistemas de tratamento de água potável, com foco especial na remoção de matéria orgânica natural. A estrutura proposta integra modelos de aprendizado de máquina — em particular o algoritmo XGBoost para prever indicadores críticos de desempenho, como carbono orgânico dissolvido, e condições hidráulicas e químicas do sistema. Os modelos demonstraram alta acurácia preditiva, tanto para sistemas somente de ultrafiltração quanto para configurações híbridas de adsorção com carvão ativado, possibilitando a simulação de cenários de operação em tempo real e oferecendo um suporte à decisão robusto para operadores, incluindo limites de ciclo de filtração e avaliação da eficácia do pré-tratamento de água.

3.6.2 Aplicação de Gêmeos Digitais em estudo de redes de distribuição de água

Observa-se que essa ferramenta permite uma ampla aplicação em diversas áreas, devido às suas inúmeras características e níveis. Porém, não são muitos os trabalhos encontrados durante o levantamento bibliográfico, em que os gêmeos digitais foram utilizados para o estudo de redes de distribuição de água. Os estudos identificados são apresentados a seguir.

Nascimento (2022) trouxe em seu trabalho o desenvolvimento de Gêmeo Digital aplicado a um sistema de bombeamento de uma rede de distribuição de água, com o objetivo de evitar que o sistema opere com pressões de risco para o seu funcionamento. Para isso, foi utilizado um sistema experimental, desenvolvido no laboratório de eficiência energética e hidráulica em saneamento da Universidade Federal da Paraíba. O modelo foi desenvolvido em etapas, que consistiram primeiramente no modelo digital, seguido da sombra digital e por fim o GD. Também foi empregado redes neurais artificiais para avaliar as funções de ativação do sistema. Segundo o autor os resultados apresentados pelo autor se mostraram favoráveis, pois, o GD desenvolvido foi capaz de proteger o sistema de bombeamento quando ao funcionamento inadequado, por meio de simulações de condições extremas e sem a comunicação com o ambiente físico, impedindo assim, danos no sistema.

Em Oliveira (2022) foi apresentado uma metodologia para a calibração de uma rede de distribuição de água teórica e experimental, desenvolvida na Universidade Federal da Paraíba. A metodologia consistiu no desenvolvimento de um GD e um algoritmo genético para a calibração da rugosidade, usando como variáveis de otimização a rugosidade das tubulações, perdas de cargas e pressão. A otimização do sistema visa variar a rugosidade por meio da redução dos erros entre os valores de pressão e vazão real e virtual. Os resultados mostraram que a calibração utilizando gêmeos digitais foi eficiente pois o erro entre o calculado e o simulado foi menor que 5%, comprovando a eficiência da ferramenta.

Pompeu e Nepomuceno (2023) aplicaram um modelo linear e um Gêmeo Digital para o apoio na tomada de decisão quanto a otimização energética de um sistema de distribuição de água. Para isso, foi utilizado um sistema SCADA (Supervisory Control and Data Acquisition) para a obtenção de dados da rede. O modelo para a tomada de decisão e previsão de demandas da rede estudada se deu por meio de bibliotecas do Python, como WNTR, Water Networks Tool for Resilience e Pyomo. Com isso, foram gerados vários cenários de simulação até que o sistema apresentasse uma otimização ideal. O resultado encontrado com a aplicação dessa

metodologia foi de uma redução 19,5% de economia energética, mostrando a eficiência do modelo.

Fuertes *et al.* (2020) apresentaram o desenvolvimento de um modelo de um Gêmeo Digital para o estudo de uma rede da cidade de Valência, na Espanha, que atende cerca de 1,6 milhão de habitantes. O modelo utilizado possui alguns objetivos como: planejamento ideal para a expansão; encontrar um design ótimo para a rede a ser expandida; detecção de perdas e previsão da qualidade da água. A ferramenta desenvolvida foi capaz de realizar a construção do modelo virtual da rede real, bem como a simulação em tempo real e de cenários passados. Porém, o modelo não foi capaz de realizar a simulação da qualidade da água, sendo necessário segundo os autores mais estudos.

Ramos *et al.* (2022) trouxeram uma metodologia para a detecção de perdas em um sistema real com a aplicação do Gêmeo Digital. A rede utilizada foi um distrito de medição 320 da cidade de Lisboa-Portugal. O modelo desenvolvido considerou alguns parâmetros e funcionamento do sistema para que a determinação das perdas fosse realizada. Esses dados foram: a evolução da rede ao longo do dia, a variação do consumo, os horários de picos de consumo, vazão e pressão ao longo do dia. A simulação com a utilização do GD foi eficiente, pois foi capaz de encontrar a perda do sistema e reduzi-la, sendo uma redução de 15% do volume perdido.

Em Persantez *et al.* (2022), foi apresentado um estudo para avaliar os impactos causados sobre um sistema de distribuição de água, devido ao aumento do consumo que ocorreu devido a pandemia da Covid-19 em uma rede de distribuição de água da cidade de Lakewood na Califórnia. Para isso, foi desenvolvido um modelo de Gêmeo Digital para acoplar os dados de medição da rede e um modelo hidráulico, para auxiliar nas simulações do sistema. Vários cenários foram simulados para que fossem avaliadas as mudanças de consumo e de qualidade da água durante o período da pandemia. Segundo os autores, o modelo desenvolvido foi capaz de identificar as mudanças de qualidade e da demanda e permitiu também a tomada de decisão para a otimização desse sistema quanto a esses dois parâmetros.

Brahmbhatt (2023) desenvolveu um estudo sobre a calibração e operação de uma rede de distribuição de água, denominada C-Town, muito utilizada na literatura para tal estudo. Para a calibração do modelo foi utilizado parâmetros hidráulicos e de qualidade do sistema estudado. Para o estudo foi desenvolvido um modelo de Gêmeo Digital para a simulação em tempo real da rede, com o objetivo de detectar vazamentos e o nível de desinfetante na água, e com isso, pode definir o funcionamento ótimo do sistema. O resultado mostrou que a

metodologia aplicada foi capaz de encontrar com precisão as perdas do sistema, bem como o nível de qualidade da água através da definição do nível ideal de desinfetante.

Mücke *et al.* (2023) propuseram um gêmeo digital aplicado à localização de vazamentos em redes de distribuição de água. A metodologia utilizou um modelo generativo, integrando técnicas de aprendizado profundo e bayesiana, incorporando incertezas inerentes ao sistema e possibilitando a combinação entre os valores observados de sensores e a distribuição estimada dos locais de vazamento. Os resultados da validação do modelo aplicado em três redes de teste (Hanoi, Net3 e Modena) com complexidade crescente mostraram que a abordagem produz estimativas rápidas, precisas e confiáveis da localização de vazamentos, com desempenho expressivo em termos de distância topológica média entre as previsões e os vazamentos reais e acurácia de localização mesmo em face de medições ruidosas.

Singh *et al.* (2024) apresentaram um estudo da aplicação do Gêmeo Digital para a calibração e detecção de vazamentos de um sistema de distribuição de água do campus do Instituto de Tecnologia indiana em Jodhpur no Karwar, com base em dados coletados desse sistema. O modelo desenvolvido apresentou ser eficiente segundo os autores, pois apresentou uma precisão na calibração e identificação de vazamentos, porém ressaltam que o modelo apresentado precisa de aprimoramento quanto a aplicação de dados em tempo real.

Xu *et al.* (2024) trouxeram um estudo sobre a identificação de falhas em sistemas de bombeamento de redes de distribuição de água com a aplicação de Gêmeo Digital alimentado por modelos de aprendizado de grafos. A metodologia apresentada consistiu em desenvolver o GD para a obtenção de dados de vibração e falhas causada pelo fluxo do impulsor sob diferentes estados e a elaboração de uma rede neural convolucional de grafos para determinar as falhas do sistema. O sistema em que o estudo foi aplicado foi uma bomba centrífuga do laboratório da Universidade de Xangai na China. Segundo os autores, os resultados foram favoráveis para o estudo, pois o modelo desenvolvido foi capaz de identificar as falhas com banco de dados reduzido. Com isso, foi possível reduzir os custos de manutenção além de aumentar a confiabilidade operacional do sistema.

Silva *et al.* (2025) exploraram a aplicação de gêmeos digitais como ferramenta para otimizar os custos de distribuição de água, destacando sua importância em resolver os desafios energéticos e operacionais enfrentados pelos sistemas de distribuição de água. A pesquisa apresentou o desenvolvimento e a implementação de um sistema de apoio à decisão baseado em gêmeos digitais, capaz de prever a demanda de água, simular o comportamento

da rede e ajustar operações de bombeamento em tempo real, conciliando o funcionamento das bombas com a redução do consumo. O estudo foi realizado em uma rede de abastecimento de água de Cantanhede, em Portugal. O resultado demonstrou que o modelo encontrou reduções significativas nos custos energéticos e operacionais, reforçando o potencial dessa tecnologia para melhorar a eficiência e a sustentabilidade das operações de distribuição de água.

Dui, Cao e Wang (2025) desenvolveram uma arquitetura de gêmeo digital para redes de distribuição de água, com foco na avaliação de resiliência e no gerenciamento emergencial dessas infraestruturas. O estudo foi realizado sobre uma rede de distribuição real da cidade de Gyeongsang, na Coreia do Sul. Os autores propuseram um modelo de avaliação de desempenho baseado na satisfação dos usuários, capaz de refletir de forma mais intuitiva a eficácia dos serviços de abastecimento urbano, além de um método para quantificar a resiliência residual das tubulações considerando o tempo de vida, o que permite otimizar a sequência de recuperação de falhas e estabelecer estratégias preventivas de manutenção. A aplicação da metodologia demonstrou que a abordagem pode melhorar significativamente a resiliência global do sistema, reforçando a utilidade dos gêmeos digitais como ferramenta de monitoramento em tempo real e controle inteligente.

Park *et al.* (2026) utilizaram de uma ferramenta de gêmeo digital integrada com realidade aumentada (AR) e o EPANET para o monitoramento de distribuição de água, visando superar as limitações das representações bidimensionais tradicionais que não capturam a complexidade espacial dessas infraestruturas. O estudo apresenta uma interface 3D que sobrepõe, mediante AR e coordenadas GPS, modelos de tubulações e dados de simulação em tempo real, permitindo aos usuários visualizar variáveis operacionais como taxas de fluxo e dispersão de poluentes de forma espacialmente contextualizada, o que favorece a análise mais intuitiva e precisa das redes de distribuição. A ferramenta foi aplicada a um sistema real de distribuição de água da cidade de Daegu na Coreia do Sul, que foi modelado inicialmente no EPANET. O resultado demonstrou que a sua utilidade possibilitou um suporte maior à tomada de decisão quanto à operação e a gestão do sistema

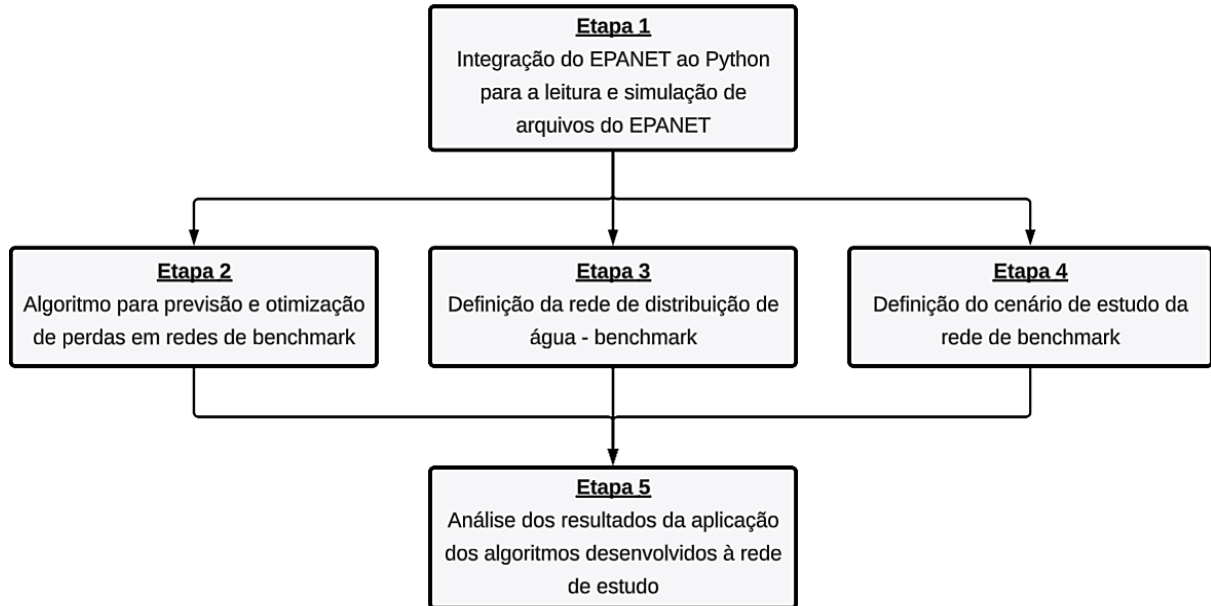
Diante do exposto, observa-se que os estudos existentes têm contribuído para a compreensão dos sistemas de abastecimento de água, principalmente em relação as perdas. No entanto, ainda não se verificou na literatura uma aplicação como a apresentada pela pesquisa, em que foram utilizadas três ferramentas de maneira conjunta para a o estudo de perdas na operação dinâmica desses sistemas.

Neste contexto, o trabalho propõe a utilização conjunta do protótipo de gêmeos digitais, rede neural recorrente long short term memory e algoritmos genéticos para a previsão e otimização de cenários adversos de perdas em uma rede de distribuição de água de benchmark, desenvolvida para a pesquisa.

4. MATERIAIS E MÉTODOS

O trabalho desenvolvido foi dividido em etapas, como está exemplificado na Figura 4.1.

Figura 4.1 - Fluxograma das etapas de metodologia do estudo.



Fonte: Próprio autor (2024).

4.1 Etapa 1: Integração do EPANET ao Python para a leitura e simulação de arquivo do EPANET

Para o desenvolvimento da pesquisa foi preciso estudar as ferramentas que podem auxiliar na sua elaboração. Assim, nessa primeira etapa foi estudado o pacote WNTR (Water Network Tool for Resilience) do Python, que permite a utilização do EPANET dentro do Python (Klise *et al.*, 2020). A aplicação de bibliotecas de código aberto possibilita a otimização do tempo de estudo, uma vez que elimina a necessidade de desenvolver um algoritmo próprio do zero para a simulação hidráulica. Além disso, confere maior confiabilidade ao trabalho, pois a ferramenta já foi testada e amplamente utilizada em diferentes tipos de redes, oferecendo também maior flexibilidade de aplicação.

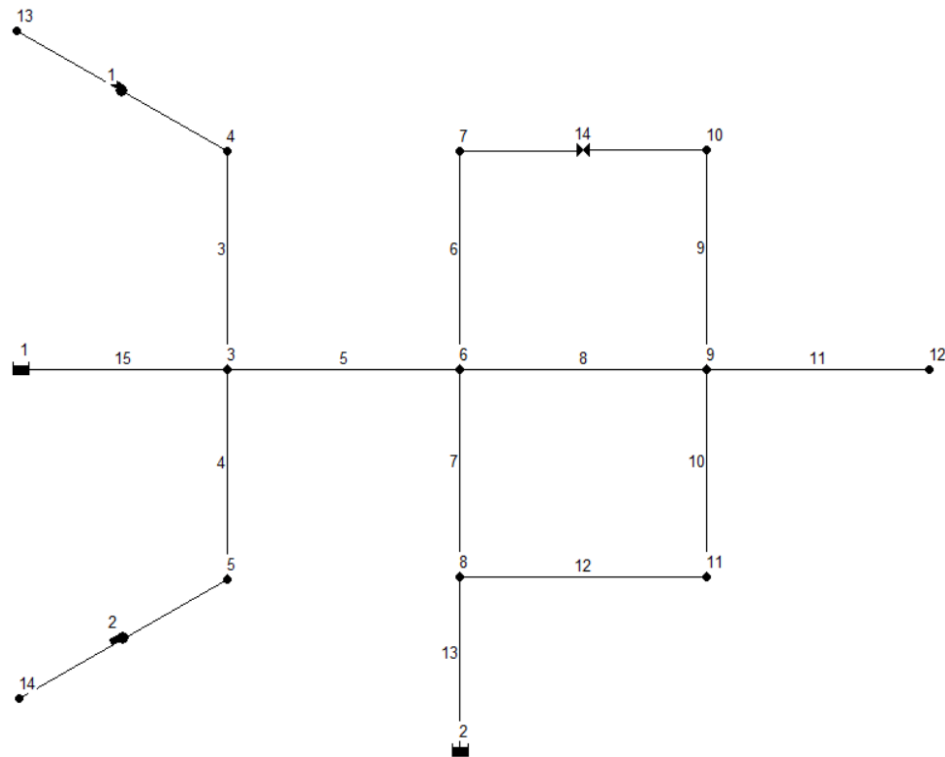
A interação entre o Python e EPANET ocorre devido a extensão *Toolkit* do EPANET, que é uma biblioteca de vínculo dinâmico de função (DLL). Ou seja, essa ferramenta permite a customização e o aprimoramento do motor computacional do EPANET de acordo com as necessidades de simulação de quem está programando. Também possibilita a associação com outros programas escritos em linguagem C/C++, Delphi Pascal, Visual Basic entre outros que

possam chamar as funções do arquivo DLL (Costa; Frota, 2018). O mecanismo de funcionamento do DLL para a incorporação com outros softwares é de fácil utilização segundo Rossman (2000), pois seu funcionamento se inicia com o salvamento de um arquivo em formato “.inp”. Para a abertura desse arquivo em qualquer linguagem de programação deve ser solicitado dentro da IDE do Python que está sendo utilizada, junto ao nome do arquivo “.inp”, anteriormente salvo, para que esse possa ser aberto para a realização de simulações necessárias.

Ainda segundo Rossman (2000), existem outras funções fornecidas pela *Toolkit* do EPANET que podem ser utilizadas na manipulação dessa ferramenta. Sendo as mais básicas o “*ENset _*” que permitem modificar as características da rede; o “*ENsolveH*” que realiza a simulação hidráulica completa da rede e salva o resultado em um arquivo; “*ENsolveQ*” que efetua uma simulação completa de qualidade da água; “*Enget _*” que possibilita saber os resultados ao longo da simulação; e “*ENclose*” que encerra o motor de simulação com todos os arquivos e libera a memória do sistema. Assim, o nome das funções da biblioteca do WNTR é desenvolvido de acordo com os nomes da extensão do EPANET.

Com isso, foi desenvolvido uma rotina computacional para a simulação de uma rede de distribuição de água. O IDE utilizado foi o Spyder, obtido através do Anaconda, para Windows 10-64 bits. A escolha desta interface se deu por ser um ambiente leve e de fácil utilização devido à sua interface intuitiva que possibilita analisar as linhas de comando, o console e os valores das variáveis em uma única janela (Kyriakou et al., 2023). A rede utilizada nesta etapa foi inicialmente desenvolvida no software EPANET e é composta por doze nós, quinze trechos, dois reservatórios de nível fixo, duas bombas e uma válvula redutora de pressão. Sua configuração foi definida de modo a garantir um escoamento predominante no sentido montante–jusante, isto é, dos reservatórios em direção aos nós mais a jusante. A topografia adotada apresenta características irregulares, com o objetivo de representar de forma mais realista as condições observadas em sistemas de distribuição de água, o que também se reflete na variabilidade dos comprimentos dos trechos. A Figura 4.2 ilustra o layout da rede desenvolvida para esta etapa do estudo.

Figura 4.2 - Layout da rede de distribuição de água utilizada na etapa 1.



Fonte: Próprio autor (2024).

Assim, a primeira simulação realizada no Python, foi de contagem dos elementos da rede, verificando se o código foi capaz de reconhecer a rede e seus elementos. O algoritmo desenvolvido foi eficiente, pois ele reconheceu a rede e os elementos que a compõem, retornando na tela a quantidade de cada elemento.

Também foi verificado o que foi instalado em cada nó e trecho da rede bem como os índices de cada elemento. A identificação dessa parte é importante, pois para realizar qualquer alteração na rede é preciso conhecer em que id o elemento, objeto da alteração se encontra, para que somente ele seja modificado e não a rede como um todo.

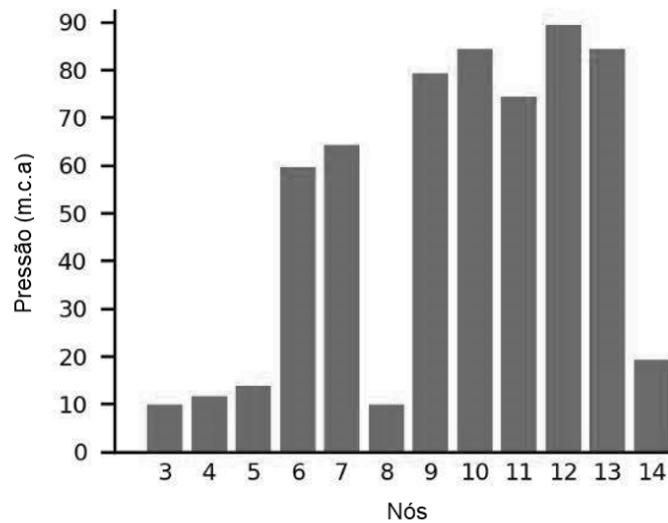
Em seguida, realizou-se uma simulação hidráulica da rede no Python com o código desenvolvido para analisar se a simulação hidráulica da rede estava funcionando corretamente e se os valores de pressão eram os mesmos encontrados no EPANET.

Assim, o código desenvolvido mostrou ser eficiente, pois identificou corretamente os id's da rede e também conseguiu encontrar as pressões correspondentes a cada um dos nós, que foram as mesmas encontradas na simulação hidráulica no EPANET.

Também foi testada a ferramenta de geração de gráfico disponível no Python, que é uma linguagem que permite a apresentação de resultados de inúmeras formas. Para isso, foi utilizado o pacote *Matplotlib*, que é uma biblioteca do Python que permite a criação de

gráficos. A Figura 4.3 apresenta os valores de pressão em cada nó da rede, por meio de gráfico de barras.

Figura 4.3 – Pressões da rede – gráfico gerado no Python.



Fonte: Próprio autor (2024).

Com a simulação hidráulica da rede foi possível determinar a função das perdas de água da rede, utilizando como base a equação apresentada por Silva (2003) (Equação 2). Trata-se de uma adaptação do estudo desenvolvido por Germanopoulos e Jowitt (1989), no qual a perda é correlacionada às pressões. Nessa equação são considerados pequenos vazamentos em torno de cada nó e a pressão da água nos mesmos. Seu valor é dado em porcentagem e a pressão média é determinada pela média aritmética das pressões de cada um dos nós da rede. As constantes presentes na equação foram ajustadas por Silva (2003) com base na faixa média das perdas em redes de distribuição de água.

$$\sum_{i=1}^{nodes} = P_{mean}^{0,5} * 7,27 \quad (2)$$

Esta equação de perdas também foi utilizada nos trabalhos desenvolvidos por Silva et al. (2020), Marques et al. (2022), Lara et al. (2024) e Barbedo et al. (2025).

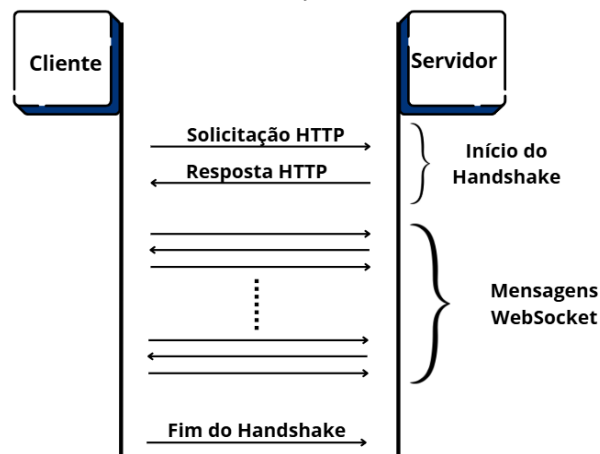
4.2 Etapa 2: Algoritmo para previsão e otimização de perdas em redes de benchmark

Para o desenvolvimento do protótipo de Gêmeo Digital Adaptativo foi utilizado como base o pacote *WebSocket* que é um pacote do Python que possibilita a interação entre um cliente (quem controla os dados) e um servidor (que disponibiliza os dados), permitindo a comunicação e alterações de dados (Augustin *et al.*, 2013).

O *WebSocket* é um protocolo de comunicação em tempo real TCP/IP, entre um servidor e o cliente. Esse protocolo utiliza a tecnologia de HTTP (*Hyper Text Transfer Protocol*) para o transporte de informações que possui as infraestruturas proxy (conexão servidor/cliente), filtragem (controle do tráfego de dados) e a autenticação que verifica a identidade do servidor antes de iniciar a comunicação (Fette; Melnikov, 2011).

O funcionamento desse protocolo ocorre em duas etapas: o handshake e a transferência de dados. Dessa forma, o cliente solicita ao servidor a conexão por meio do HTTP (utilizado somente na primeira comunicação para a abertura do canal) e se o servidor suportar a conexão e se não houver erro na chamada o servidor envia uma mensagem ao cliente e o handshake será iniciado. Nesse processo inicial é necessário que o endereço do servidor (host) seja especificado pelo cliente. Uma vez iniciado a comunicação, ambos os ativos podem enviar dados de um para o outro de maneira simultânea, necessitando de um único TCP/IP para essa comunicação, o que diferencia esse protocolo de outros (Fette; Melnikov, 2011; Srinivasan *et al.*, 2013; Bayılmış *et al.*, 2022). A Figura 4.4 traz o funcionamento básico dessa comunicação realizada pelo *WebSocket*.

Figura 4.4 - Funcionamento básico da relação servidor/cliente, através do WebSocket.



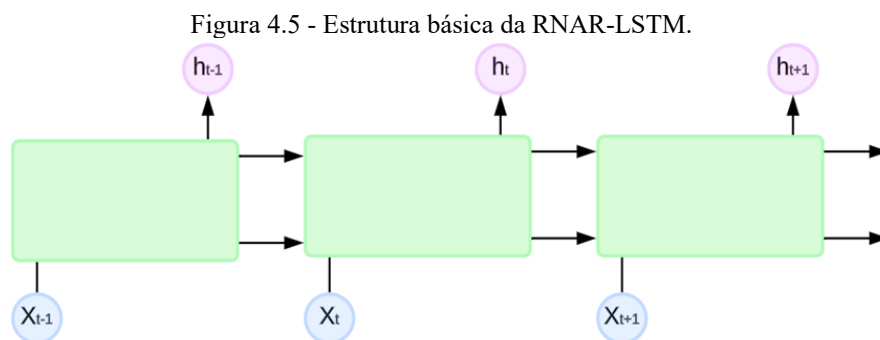
Fonte: Adaptado de Bayılmış *et al.* (2022).

A utilização deste protocolo se dá por meio de comando, se destacando como principais o “onopen”, “onclose”, “onmessage”, “onerror”, “send” e “enclose”. O comando

“onopen” é utilizado quando a comunicação servidor/cliente é iniciada e o “onclose” finaliza a comunicação. A transferência de mensagens entre o servidor e o cliente é realizada pelo “onmessage” e o comando “send” faz a troca do cliente para o servidor. Em ambas as comunicações é necessário utilizar na estrutura do código o comando “onerror” para que algum erro que venha ocorrer na iteração seja identificado. Porém, esses dados aplicados na troca precisam estar no mesmo padrão (Srinivasan *et al.*, 2013; Bayılmış *et al.*, 2022).

Para a aplicação da ferramenta de Machine Learning foram utilizadas as bibliotecas do Python *Scikit-learn* e *Keras (Tensorflow)*, que são ferramentas específicas para ML, mas especificamente para as redes neurais. Elas dispõem de recursos que permitem a análise preditiva de dados, podem ser aplicadas a diferentes estudos, e possui código aberto, permitindo seu acesso sem taxas ou licenças (Cournapeau *et al.*, 2007; Chollet *et al.*, 2015; Abadi *et al.*, 2016).

Assim, o modelo definido para a previsão das perdas foi a Rede Neural Artificial Recorrente de Long Short Term Memory, que é uma rede neural que possui capacidade de incluir ou não informações ao estado da célula, se caracterizando por ter duas memórias (curta e longa). A Figura 4.5 traz a configuração do LSTM, apresentando de maneira geral como são as células da rede neural.



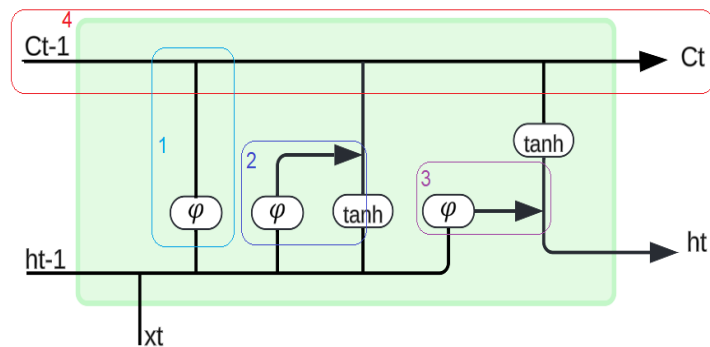
Fonte: Adaptado de Teixeira *et al.* (2020).

Cada uma das células da rede neural LSTM é formada por portões (representados pelos números 1 a 4 na Figura 4.6): um portão de esquecimento, um portão de entrada, um portão de memória e um portão de saída, além do estado atual da célula. Como a estrutura desse tipo de rede neural possui duas memórias, as funções sigmóide são utilizadas para definir o quanto de informações será mantida em cada uma delas. Já as funções tanh são aplicadas para impedir a explosão de gradiente (Zhang *et al.*, 2018; Hochreiter e Schmidhuber, 1997).

Para o melhor entendimento do funcionamento interno de cada célula a Figura 4.6 traz de maneira mais detalhada sua operação por meio do detalhamento de uma célula apresentada na Figura 4.5.

Na Figura 4.6 O x_t corresponde a entrada de dados no tempo t e o h_{t-1} refere-se a saída no tempo $t-1$, sendo considerada a memória de curto prazo da célula. A região 1 corresponde ao local onde é definido o quanto da memória longa (c_{t-1}) será mantido ou descartado. Na região 2 é definido o quanto da memória curta (h_{t-1}) e da entrada (x_t) será incorporado ao estado atual da célula, sendo chamada de portão de entrada. Em 3, tem se o portão de saída, onde é realizado a combinação do estado atual da célula com a entrada para formar a saída (h_t), porém sem mudar o estado atual (c_t) e isso é levado para a próxima célula em paralelo a saída (h_t). E por fim a região 4, onde é definido o estado atual da célula com a entrada (c_{t-1}) que corresponde a memória longa (Hochreiter and Schmidhuber 1997; Aurélien, 2019).

Figura 4.6 - Estrutura detalhada de uma célula da RNAR-LSTM.



Fonte: Adaptado de Correia *et al.* (2021) e Contarini (2020).

Para as funções de ativação presentes nos portões, são utilizados as Equações 3 a 8, em que as W_i , W_f , W_g e W_o , correspondem à matriz de pesos que podem ser podados entre as entradas da rede LSTM e os portões. U_i , U_f , U_g e U_o são as matrizes de peso entre os valores dos propagadores internos de dados (h_t). E os b_i , b_f , b_g e b_o são os biases que estão em cada portão da rede neural. Já o x_t corresponde ao conjunto de dados de entrada na célula no tempo t e c_t o dissipador de informações no tempo t (Zhong *et al.*, 2021).

$$i_t = \sigma (W_i x_t + U_i h_{t-1} + b_i) \quad (3)$$

$$f_t = \sigma (W_f x_t + U_f h_{t-1} + b_f) \quad (4)$$

$$g_t = \sigma (W_g x_t + U_g h_{t-1} + b_g) \quad (5)$$

$$o_t = \sigma (W_o x_t + U_o h_{t-1} + b_o) \quad (6)$$

$$h_t = o_t * \tanh (c_t) \quad (7)$$

$$C_t = f_t * c_{t-1} + i_t + g_t \quad (8)$$

Para a utilização de redes neurais LSTM, é preciso estabelecer outros parâmetros para o bom funcionamento da rede neural LSTM, como a quantidade de camadas, neurônios em cada camada e o número de épocas. Porém, não é possível encontrar uma padronização que relacione a quantidade de dados disponibilizados com os valores que esses parâmetros podem ter, assim a escolha da quantidade de cada parâmetro depende de processos mais empíricos do problema que está sendo estudado com o uso desse tipo de rede neural (Zhang et al., 2018; Hochreiter e Schmidhuber, 1997).

Assim, para a definição desses parâmetros para o desenvolvimento do estudo, utilizou-se o método empírico, no qual os valores foram ajustados iterativamente até encontrar a combinação que melhor se adequasse à previsão proposta, considerando a quantidade de dados disponíveis para o modelo. Além disso, foram considerados alguns trabalhos levantados que aplicaram essa técnica para o estudo de recursos hídricos. Como Lee e Yo (2021) que utilizaram de 3 camadas para a previsão de vazamento em uma rede de distribuição de água, sendo que a quantidade de neurônios para cada uma das camadas foi de 128, 64 e 48. Em relação ao número de épocas empregadas foi de 120 pelo fato da quantidade de dados disponibilizados no estudo elaborado pelo autor. Já Yu et al. (2025) aplicaram duas camadas, ambas com 128 neurônios, no estudo de previsão de demanda de uma rede de distribuição de água. E para o número de épocas o valor foi de 64 épocas.

Desta forma, a melhor configuração encontrada para a previsão proposta foi composta por duas camadas, com 128 neurônios na primeira camada e 48 na segunda e para a época o valor foi de 300. Observa-se que em estudos que se utilizam de aprendizado profundo, como redes neurais do tipo perceptron, geralmente são utilizados números maiores de iterações, porém para o presente estudo o número de 300 épocas foi o suficiente para encontrar o objetivo proposto (Shi *et al.*, 2024). Além disso, os estudos levantados também aplicaram um número pequeno de iterações utilizando do mesmo tipo de rede do estudo a RNAR-LSTM.

Na simulação do algoritmo de previsão também foi estabelecido a quantidade de dados que seria utilizada para o treinamento, teste e validação do modelo, sendo estabelecido em 70% para o treinamento, 15% para o teste e 15% para a validação (Philip *et al.*, 2024).

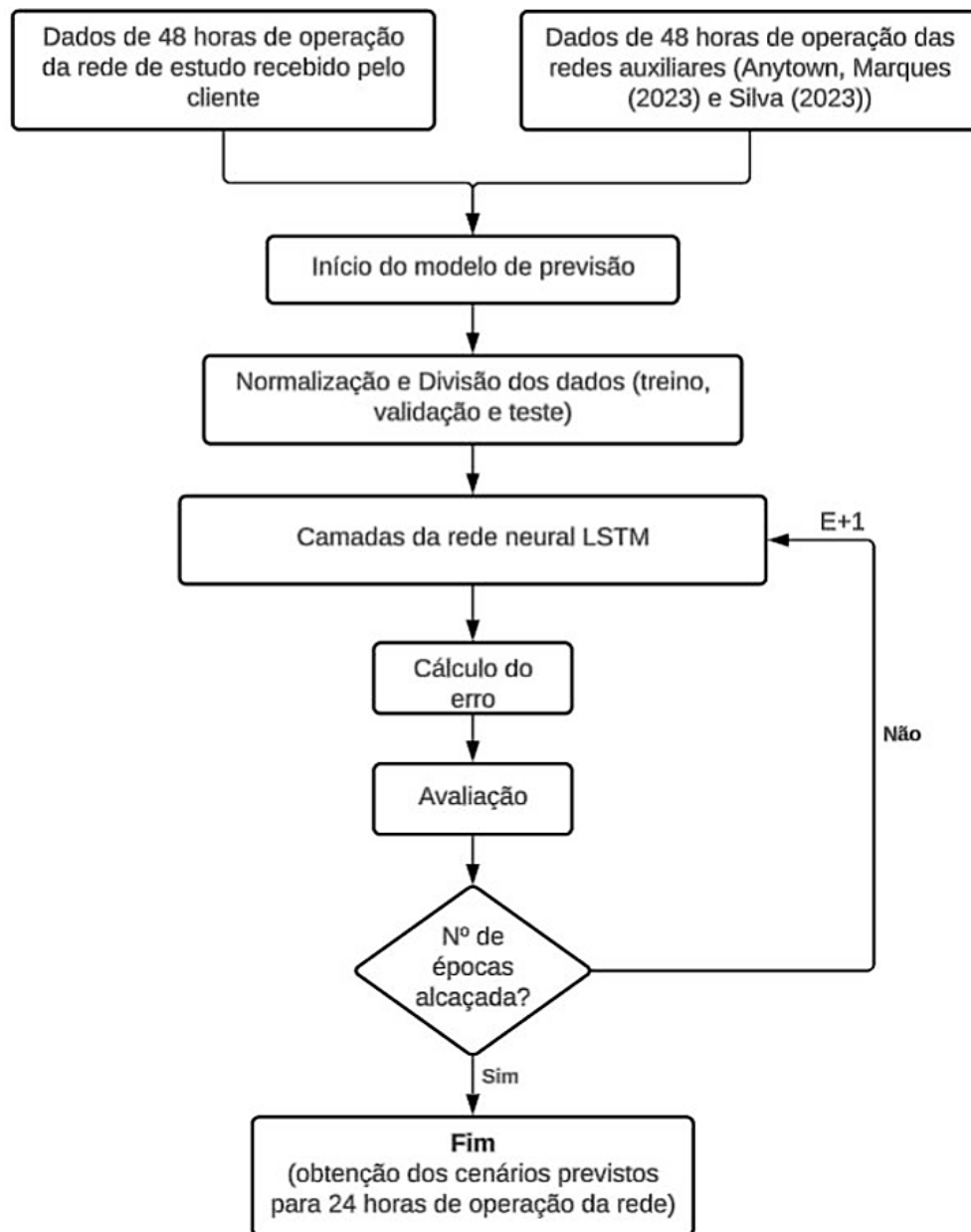
Para o estudo, utilizou-se o algoritmo de otimização Adam (*Adaptive Moment Estimation*), que realiza o ajuste automático dos pesos da rede neural durante o treinamento para a minimização dos erros de previsão, no caso do estudo, o erro quadrático médio. Assim, o algoritmo faz o ajuste das taxas de aprendizado automaticamente a cada peso, sendo aplicado o valor inicial padrão do algoritmo de 0,001 (Kingma; Ba, 2015; Le *et al.*, 2019; Pires, 2024).

A Figura 4.7 apresenta um fluxograma do funcionamento da rede neural recorrente LSTM desenvolvida para a pesquisa. Em que o início do funcionamento se dá com a entrada dos dados da rede de estudo que foi comunicada ao cliente por meio do gêmeo digital, e os dados das redes auxiliares que são chamadas diretamente no cliente por meio do motor do EPANET dentro do Python pelo pacote WNTR. Com todos os dados de operação das redes dentro do cliente a rede neural de previsão LSTM é iniciada. Para que os dados pudessem ser utilizados na previsão foi preciso realizar a normalização das variáveis para evitar a discrepância gerada pela ordem de grandeza das mesmas. Assim, foi aplicada a técnica de MinMax (Equação 9), em que a normalização se baseia na amplitude dos valores, ou seja, os valores das variáveis são reescalados para o intervalo de 0 e 1, considerando seus valores mínimos e máximos.

$$x_{normalizado} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (9)$$

Além disso, os dados foram divididos em treino (70%), validação (15%) e teste (15%), e em seguida se iniciou as camadas do modelo de previsão e seus neurônios respectivamente. A cada época, o modelo realiza o cálculo do erro, por meio do erro quadrático médio já mencionado. Sendo o melhor valor aquele em que o erro quadrático for o menor. Com isso, é possível conhecer o comportamento do modelo, se sofreu overfitting ou underfitting, ou se conseguiu realizar bem a previsão. O modelo é encerrado assim que o número de épocas determinados a ele for alcançado.

Figura 4.7 - Fluxograma de funcionamento da rede neural LSTM desenvolvida.



Fonte: Próprio autor (2025).

Com o modelo de previsão por meio do RNAR-LSTM, definido e estruturado, será utilizado a técnica de otimização através do algoritmo genético para otimizar esses dados previstos. Para isso, foi utilizada a biblioteca *DEAP* do Python que é um pacote utilizado para a elaboração de algoritmos genéticos (Fortin *et al.*, 2012). O algoritmo foi elaborado para otimizar as pressões médias previstas através do objetivo de alcançar os pontos ótimos dos parâmetros ativos da rede, como bomba, reservatórios e válvulas, para que as pressões sejam mantidas de acordo com a norma NBR 12218:2017 (10 a 50 m.c.a) e com isso, minimizar as perdas previstas. Assim, foi definido a função objetiva de minimização que está descrita na

Equação 10, bem como as restrições. Onde “P” corresponde a pressão em cada nó e “n” o número de nós da rede.

$$F.O = \min[f], f = \frac{\sum_i^n P_{(i)}}{n}, \text{sujeito a } 10 \leq P_{(i)} \leq 50 \quad (10)$$

A escolha por um algoritmo genético de objetivo simples se deu para evitar uma maior complexidade na rotina computacional e, com isso, evitar que o tempo de execução pudesse ser maior, o que impactaria o desenvolvimento do trabalho, já que a otimização é realizada para um período de 24 horas de operação da rede de distribuição água.

Assim, o algoritmo desenvolvido visa minimizar as perdas previstas, através da otimização das pressões médias. Para isso, foram definidos intervalos de variação para as variáveis operacionais para que o algoritmo pudesse utilizar para a otimização, pois eles são responsáveis pela variação das pressões durante a operação da rede. Esse processo de escolha do indivíduo se dá através do torneio, que consiste na escolha dos indivíduos de maneira aleatória (Ferreira, 2007). Assim, para os reservatórios foram definidos valores de variação de 900 a 940 m, para as bombas uma variação da rotação de 1 a 3 e para as válvulas redutoras de pressão de 0 a 100, considerando sua abertura e fechamento. Para evitar discrepância entre as variáveis aplicadas devido à diferença entre suas ordens de grandeza foi utilizado a normalização dos valores. A definição dos valores para cada um dos parâmetros foi realizada com base na pesquisa elaborada por Silva (2023) e Marques (2023).

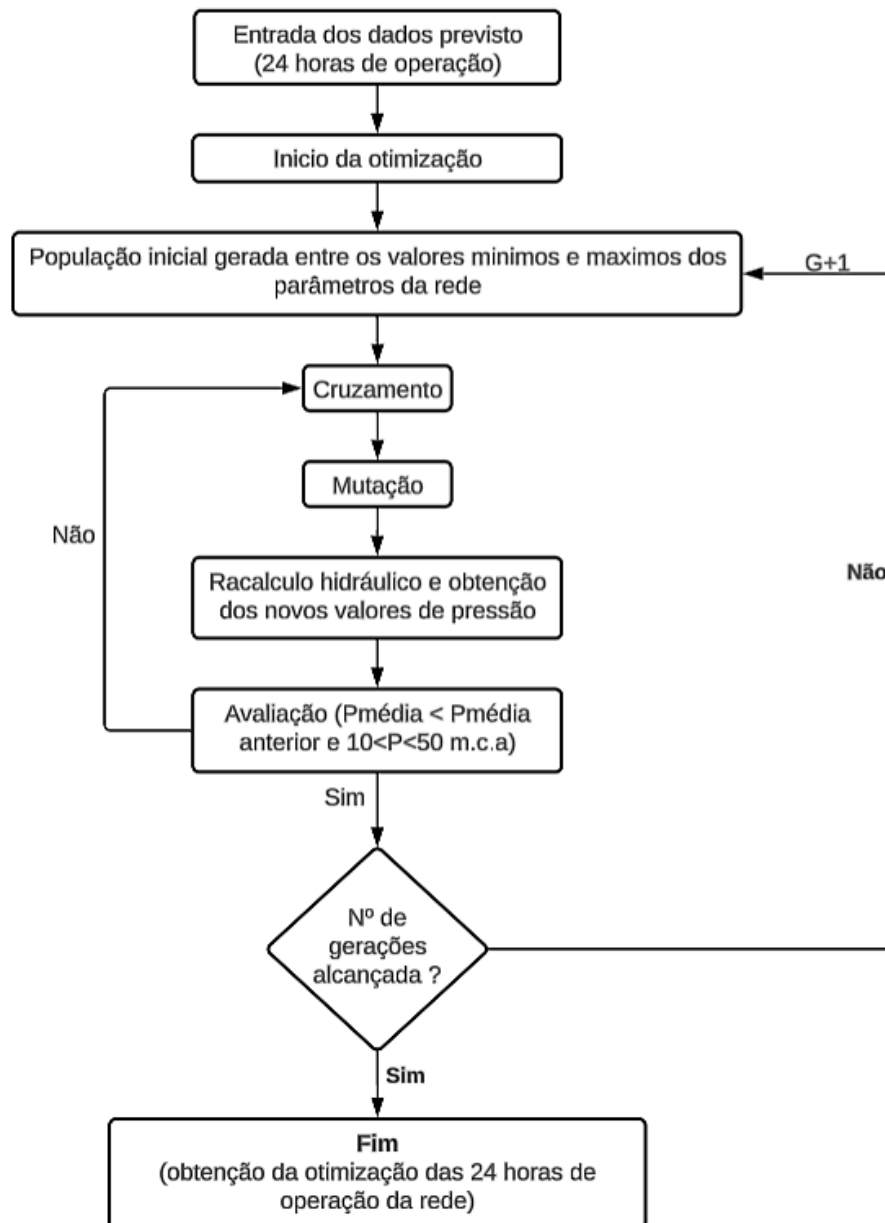
Outros parâmetros importantes para o funcionamento do algoritmo genético foram estabelecidos, como o número de gerações, que foi definido com base em Scrucca (2013) que determina que o número de geração em 100 é o suficiente para que o algoritmo encontre os cenários ideais de otimização. Porém, outros estudos como Lara et al. (2024) e Marques (2023) mostraram a utilização do algoritmo genético em estudos de rede de distribuição de água aplicando números elevados de gerações. Assim, para a presente pesquisa foi definido em 300 gerações devido ao tempo computacional levado para a execução da otimização das 24 horas de operação dos cenários previstos.

Para a mutação e cruzamento estabeleceu-se valores de 5% e 80%, respectivamente (Simpson *et al.*, 1994). Também foi determinado que o tipo de algoritmo utilizado seria o geracional elitista, para que as melhores soluções pudessem ser preservadas para as gerações seguintes (Fogel, 1994).

Para exemplificar o funcionamento do algoritmo genético desenvolvido para a otimização, foi elaborado o fluxograma da Figura 4.8, evidenciando os pontos de inicialização

e decisão que o algoritmo realiza para encontrar os resultados desejados. Ressaltando que o algoritmo encerra seu funcionamento quando o número de gerações determinada é atingida.

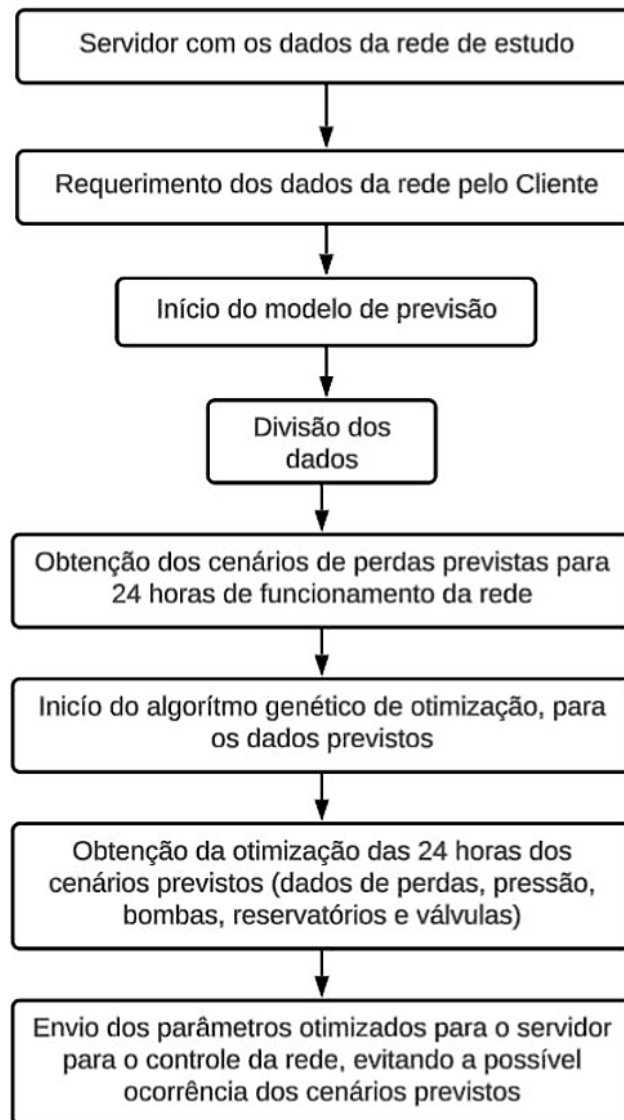
Figura 4.8 - Fluxograma de funcionamento do algoritmo genético aplicado ao estudo.



Fonte: Próprio autor (2025).

De acordo com as ferramentas descritas, foi desenvolvida a metodologia para o estudo de previsão e otimização de perdas da operação da rede de distribuição de água escolhida. Assim, para melhor explicar e resumir o funcionamento delas de forma conjunta foi elaborada a Figura 4.9.

Figura 4.9 - Fluxograma do funcionamento resumido da metodologia proposta.



Fonte: Próprio autor (2025).

O procedimento da metodologia desenvolvida se inicia com a comunicação entre o servidor e cliente. O servidor possui os dados operacionais da rede, a serem utilizados para a previsão das perdas e, por meio do *WebSocket* (biblioteca do Python), esses dados serão passados para o cliente, que será responsável por realizar as simulações com o algoritmo de rede neural recorrente LSTM (por meio da biblioteca *Scikit-learn*), para prever as perdas do sistema de distribuição a ser estudado. Na simulação com o modelo de previsão, será considerado a divisão descrita anteriormente. Porém, se essa porcentagem não se mostrar eficiente para o estudo proposto ela poderá ser alterada passando para 80% o treinamento e 20% para teste e validação (Scaccia, 2020).

Com a finalização da simulação do algoritmo de previsão, será realizada a otimização das perdas previstas utilizando o algoritmo genético, por meio da biblioteca *DEAP* do Python, aplicando os parâmetros já definidos anteriormente. Por fim, o cliente irá enviar para o servidor os parâmetros das bombas, válvulas redutoras de pressão e dos reservatórios encontrados na otimização de perdas previstas, para que o servidor verifique se estes valores estão ocorrendo ou irá ocorrer na rede e com isso, ser possível evitar esse cenário adverso.

Para a análise de desempenho do modelo desenvolvido foram utilizadas algumas métricas, que estão apresentadas pelas Equações 11 a 13, que são o erro quadrático médio, erro absoluto médio e coeficiente de determinação, respectivamente. Sendo que a função objetivo utilizada pela rede neural foi a minimização do erro quadrático médio. Além disso, essas métricas também auxiliaram na escolha da melhor configuração dos parâmetros do modelo desenvolvido, para a rede neural, sendo realizada a comparação dos valores encontrados em cada métrica para cada uma das configurações testadas. Uma vez que são geralmente aplicados para a comparação de modelos (Bai *et al.*, 2021; Bonilla *et al.*, 2025).

$$MSE = \frac{1}{z} \sum_{i=1}^z (y_i - \hat{y}_i)^2 \quad (11)$$

$$MAE = \frac{1}{z} \sum_{i=1}^z |y_i - \hat{y}_i| \quad (12)$$

$$R^2 = 1 - \frac{\sum_{i=1}^z (y_i - \hat{y}_i)^2}{\sum_{i=1}^z (y_i - \bar{y})^2} \quad (13)$$

Nas Equações o z representa o número de valores previstos, y_i corresponde aos valores observados e \hat{y}_i são as médias dos valores previstos, no intervalo de tempo de i hora. Para o MSE e MAE para determinar o bom desempenho do modelo de previsão precisam apresentar valores baixos, enquanto o R^2 necessita apresentar valores próximo a 1 para estabelecer o bom desempenho do modelo (Bai *et al.*, 2021; Zanfei *et al.*, 2022; Bonilla *et al.*, 2025).

4.3 Etapa 3: Definição das redes de distribuição de água – benchmark

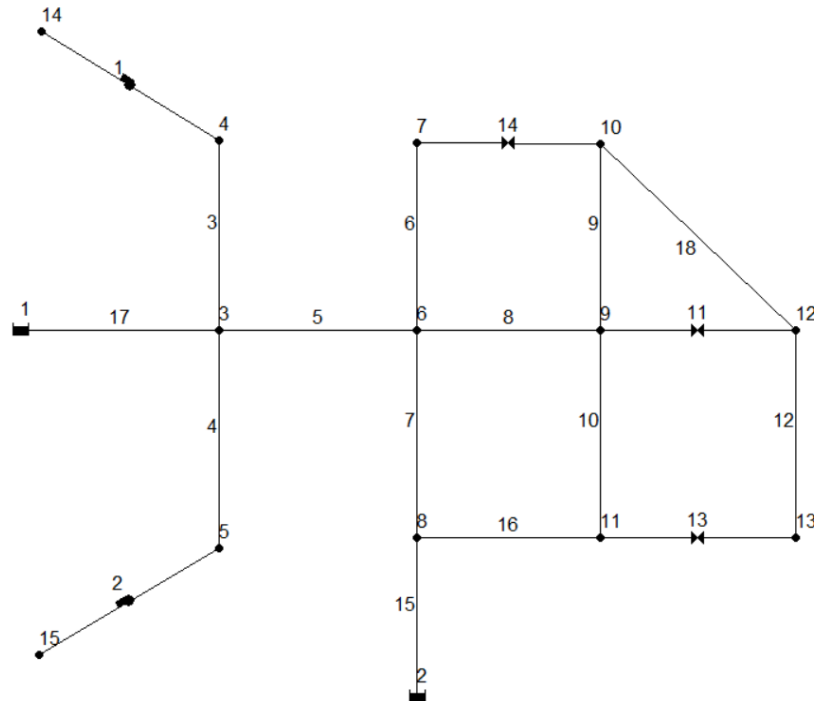
4.3.1 Redes de benchmark para o treinamento do modelo de previsão

Para o desenvolvimento da metodologia de aprendizado de máquina, é necessário o fornecimento de dados para que este possa aprender e, com isso, ser possível encontrar os melhores resultados. Assim, foram utilizadas algumas redes com configuração e

comportamentos diversos, para o melhor aprendizado do modelo a ser desenvolvido. Essas redes são redes já aplicadas em estudos presentes na literatura, comprovando assim a aplicabilidade das mesmas.

A primeira rede escolhida foi uma rede desenvolvida por Marques (2023), que está apresentada na Figura 4.10.

Figura 4.10 - Configuração da rede desenvolvida por Marques.



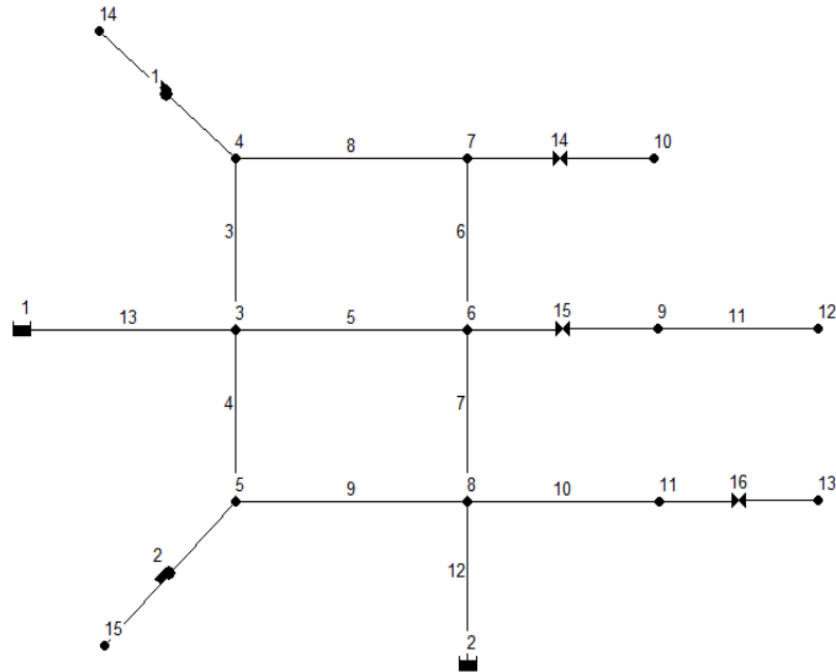
Fonte: Marques (2023).

A rede é composta por dois reservatórios de nível fixo, treze nós, treze trechos, duas bombas e três válvulas redutoras de pressão. Os dados de entrada dos nós da rede, como as cotas, representam uma região montanhosa onde o desnível é grande e as demandas dos nós variaram de 1 a 3 l/s. Esses valores foram estabelecidos para que o escoamento ocorresse da esquerda para a direita. Os valores de pressão foram delimitados com base em uma variação de pressão de 20 a 90 m.c.a., que são pressões próximas das apresentadas em redes reais, representando uma operação desfavorável, com uma topografia irregular (Marques, 2023).

Os valores de entrada dos trechos, segundo Marques (2023), foram delimitados para que representassem um sistema real tanto no comprimento das tubulações, que variam de 100 a 500 m. O tipo de material aplicado, foi o PVC. O diâmetro varia de 50 a 200, sendo os maiores localizados próximo aos reservatórios e os menores mais a jusante da rede.

A segunda rede foi desenvolvida por Silva (2023) e é constituída por dois reservatórios de nível fixo, treze nós, onze trechos, duas bombas e três válvulas reguladoras de pressão. A Figura 4.11, traz o desenho da rede desenvolvida por Silva (2023).

Figura 4.11 - Configuração da rede desenvolvida por Silva.

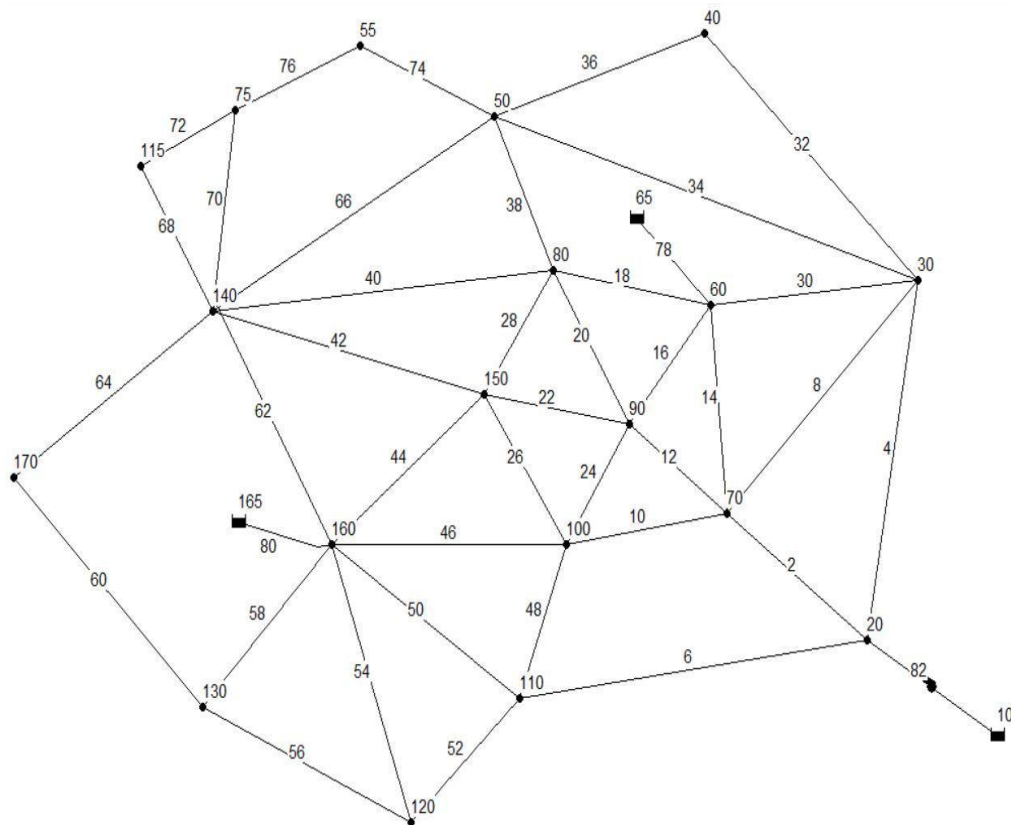


Fonte: Silva (2023).

Assim como a rede 1, também foi possível obter os dados de entrada dos nós e dos trechos. Seus valores também foram escolhidos para que representassem um sistema real, com topografia irregular e valores de pressão desfavoráveis ao seu funcionamento. As demandas também variaram de 1 a 3 l/s. Os valores de comprimento variam de 100 a 500 m e o diâmetro das tubulações foi de 50 a 200, e a rugosidade das tubulações são 0,06 mm, utilizados para tubulações de PVC.

A terceira rede é a Anytown, é uma rede que foi publicada pela primeira vez por Walski *et al.* (1987). Essa rede possui três reservatórios de nível fixo responsáveis pela alimentação do sistema e uma bomba localizada após o reservatório de alimentação ao sudeste da rede. Para o material dos trechos, foi utilizado o PVC, que apresenta uma rugosidade de 0,06 mm. O desenho da rede está mostrado na Figura 4.12, constituída de 42 trechos e 19 nós.

Figura 4.12 – Configuração da rede Anytown.



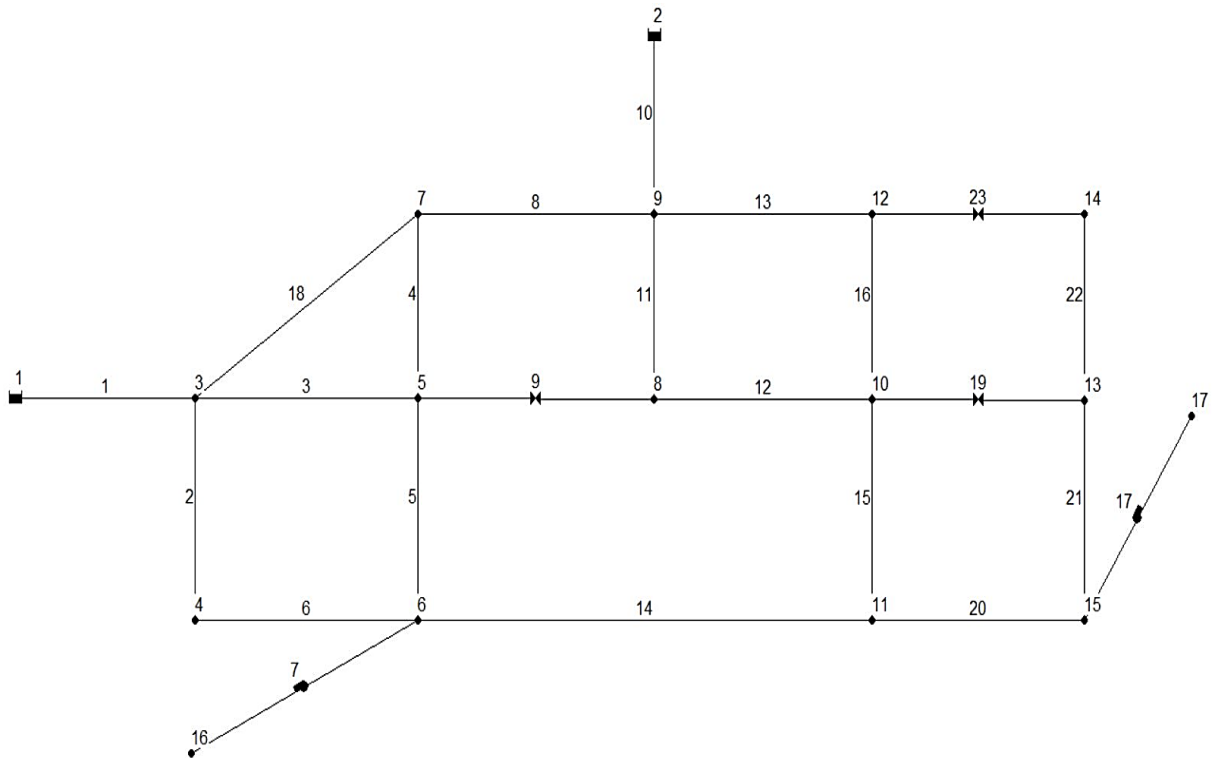
Fonte: Walski *et al.* (1987).

Para o melhor treinamento do modelo de previsão, estas redes foram utilizadas no modelo proposto por apresentarem uma configuração que se assemelha à rede de estudo apresentada a seguir e para aumentar a variação nos dados de sua composição, o que pode permitir um aprendizado maior do modelo, possibilitando que seja mais generalista.

4.3.2 Rede de benchmark de estudo

Para o estudo, sobre a qual foi aplicado o protótipo de Gêmeo Digital Adaptativo, o Aprendizado de Máquina e o Algoritmo Genético, com o objetivo de prever e otimizar os cenários de perdas adversas do sistema com base nas variáveis operacionais, foi elaborada uma rede com base em redes reais da região do sul de Minas Gerais e seu layout está exemplificado na Figura 4.13. Ela é composta por 15 nós, 21 trechos, 2 reservatórios de nível fixo, 2 bombas e 3 válvulas redutoras de pressão. O fluxo da água se dá da esquerda para a direita, para que a água escoe dos reservatórios para nós mais a jusante. Também apresenta uma topografia irregular, representando o sistema real brasileiro e, com isso, possui uma variação de pressão maior, o que permitirá uma visualização melhor das ferramentas aplicadas para o estudo. Os parâmetros dos nós estão exemplificados na Tabela 1.

Figura 4.13 – Rede de estudo.



Fonte: Próprio autor (2024).

Tabela 1 - Dados dos nós da rede de estudo.

Id dos nós	Cota (m)	Demanda (l/s)
Nó 3	888	2
Nó 4	860	1
Nó 5	835	3
Nó 6	890	1
Nó 7	850	1
Nó 8	825	2
Nó 9	887	3
Nó 10	820	2
Nó 11	855	2
Nó 12	860	1
Nó 13	815	1
Nó 14	850	1
Nó 15	810	1
Nó 16	925	1
Nó 17	930	1
Reserv. 1	900	–
Reserv. 2	900	–

Fonte: Próprio autor (2024).

Para os trechos também foram escolhidos os valores para que a rede fosse semelhante às redes real. Assim, para os trechos, foram escolhidos comprimentos entre 100 e 150, que são

valores usuais das quadras dos loteamentos. O tipo de tubulação foi escolhido por utilizar PVC, cuja rugosidade, segundo Porto (2006), é de 0,06 mm. Definiu-se por colocar bombas em pontos extremos onde não era possível que a água chegasse por gravidade. Os dados de todos os parâmetros dos trechos das redes estão na Tabela 2.

Tabela 2 - Dados dos trechos da rede de estudo.

Id dos trechos	Comprimento (m)	Diâmetro (mm)	Rugosidade (mm)
Trecho 1	400	200	0,06
Trecho 2	100	100	0,06
Trecho 3	100	100	0,06
Trecho 4	100	100	0,06
Trecho 5	100	100	0,06
Trecho 6	100	100	0,06
Trecho 8	150	100	0,06
Válvula 9	-	100	-
Trecho 10	300	200	0,06
Trecho 11	100	100	0,06
Trecho 12	150	75	0,06
Trecho 13	150	75	0,06
Trecho 14	300	100	0,06
Trecho 15	100	75	0,06
Trecho 16	100	75	0,06
Trecho 18	100	100	0,06
Válvula 19	-	75	-
Trecho 20	100	75	0,06
Trecho 21	100	50	0,06
Trecho 22	100	50	0,06
Válvula 23	-	75	-
Bomba 7	-	-	-
Bomba 17	-	-	-

Fonte: Próprio autor (2024).

4.4 Etapa 4: Definição do cenário de estudo da rede de benchmark

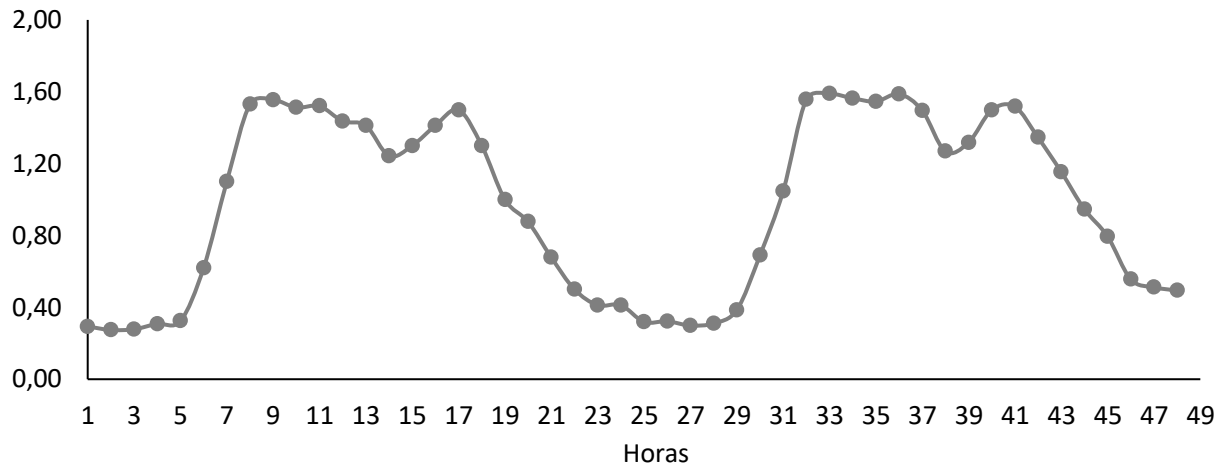
Para o desenvolvimento da metodologia de estudo da operação de uma rede de benchmark, com o objetivo de prever e otimizar as perdas desse sistema, serão utilizadas variáveis operacionais como as válvulas redutoras de pressão, bombas e reservatórios de nível fixo.

Nas válvulas redutoras de pressão, será considerada a pressão a jusante, causada pela variação de abertura e fechamento. Para as bombas, será utilizada a rotação, considerando os estados de parada ou em funcionamento nominal, pois a rotação altera fundamentalmente seu comportamento, em que a variação da rotação desloca o ponto de operação ao longo da curva do sistema permitindo adequar a bomba à demanda da rede, em horários de pico ou baixa (Hieninger et al., 2021). Já os reservatórios de nível fixo serão entendidos como variações provocadas pelo bombeamento em suas saídas.

Os dados a serem utilizados nas simulações serão elaborados no EPANET e posteriormente utilizados no Python, e será considerada a simulação dinâmica da rede para que as variações de funcionamento da mesma sejam observadas e armazenadas para realização do estudo. Para que fossem geradas as horas de funcionamento da rede, para 48 horas, foram utilizados os fatores de demanda para que a rede apresentasse a variação de operação durante as horas do dia e com isso o sistema fosse representado mais próximo da realidade e para que o estudo pudesse ser desenvolvido.

Para a obtenção dos dados de funcionamento da rede no período mencionado, foi necessário definir a curva de variação de demanda para as horas de operação, de modo a aproximar a rede de um sistema real, no qual ocorrem variações de pressão ao longo do dia (Silva, 2003; Silva, 2023). As demandas maiores foram definidas para o período das 7 às 20 horas e as menores demandas foram das 21 às 6 horas (Hossain *et al.*, 2021; Hernandez-Samaniego, *et al.*, 2023). Com isso, as pressões têm maior probabilidade de serem maiores nos períodos de menores demandas e, conseqüentemente, é maior a probabilidade de perdas (Tsuboiya, 2006). A Figura 4.14 traz a curva de variação de demanda utilizada na rede de estudo, onde a variação é causada pelo multiplicador (fator), que é adimensional.

Figura 4.14 – Curva de demanda da rede de estudo.



Fonte: Próprio autor (2025).

As curvas de demanda das redes usadas para o treinamento do modelo de previsão bem como os dados de pressão e perdas das 48 horas de operação estão em anexo.

Além disso, também será avaliado o desempenho em relação ao tempo de simulação e sua eficiência em encontrar o resultado esperado, podendo realizar os possíveis ajustes, caso necessário.

4.5 Etapa 5: Análise dos resultados da aplicação dos algoritmos desenvolvidos à rede de estudo

Com a metodologia desenvolvida, a análise dos resultados foi realizada por partes: primeiramente pela verificação dos piores cenários de perdas encontrados pelo algoritmo desenvolvido utilizando as redes neurais artificiais recorrentes LSTM. Em seguida foi examinado se esses valores foram otimizados por meio do algoritmo genético desenvolvido para essa otimização dos piores cenários, por meio da variação dos parâmetros operacionais da rede, das válvulas redutoras de pressão, das bombas e dos reservatórios de nível fixo. Essas duas ferramentas aplicadas no estudo foram realizadas dentro de um ambiente de comunicação simultânea, o *WebSocket* que possibilitou a aplicação da prototipagem do gêmeo digital adaptativo e também sendo eficiente na transmissão dos dados da rede original e dos otimizados.

5. RESULTADOS E DISCUSSÕES

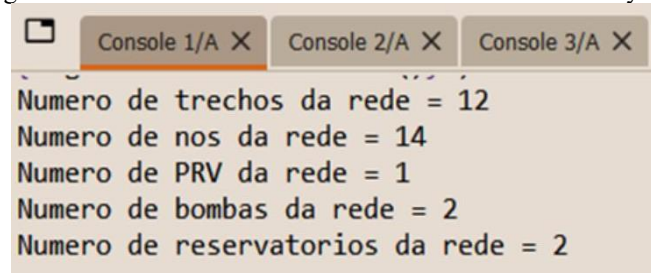
5.1 Resultados do estudo da iteração EPANET Python

Os primeiros resultados obtidos foram provenientes da aplicação do EPANET em Python, por meio do pacote WNTR. Essa simulação, conforme descrito na metodologia (item 4.1), teve como objetivo avaliar a interação entre os dois softwares, verificar o comportamento da simulação da rede e determinar se seria possível estimar a porcentagem de perda do sistema. Essa etapa foi realizada sobre a rede desenvolvida especificamente para esta fase, também detalhada no item mencionado da metodologia.

O algoritmo desenvolvido para essa etapa se mostrou eficiente, pois conseguiu reconhecer a rede, seus elementos, as pressões e a perda da rede de distribuição de água usada nessa fase.

Assim, na Figura 5.1, está apresentado o resultado encontrado do reconhecimento do algoritmo para os elementos da rede de estudo utilizada nessa fase da pesquisa, em que o número de válvulas e bombas presentes no sistema foram identificadas facilmente, não sendo consideradas como trecho.

Figura 5.1 – Número de elementos da rede encontrados no Python.

A screenshot of a Python console window with three tabs labeled 'Console 1/A X', 'Console 2/A X', and 'Console 3/A X'. The first tab is active and displays the following text:

```
Numero de trechos da rede = 12  
Numero de nos da rede = 14  
Numero de PRV da rede = 1  
Numero de bombas da rede = 2  
Numero de reservatorios da rede = 2
```

Fonte: Próprio autor (2024).

Com o reconhecimento da rede e de seus elementos, foi realizada uma simulação hidráulica dentro do Python e no EPANET para verificar se o algoritmo foi capaz de encontrar o valor igual do EPANET, para com isso validar essa parte do estudo.

A Tabela 3 traz os valores de pressão encontrados com o algoritmo desenvolvido no Python e no EPANET. É possível observar que os valores encontrados pelo algoritmo desenvolvido foram os mesmos apresentados pelo EPANET, com uma pequena variação nos decimais. Isso mostra que o algoritmo desenvolvido é viável para o estudo, uma vez que a correta identificação das pressões é fundamental para o avanço da pesquisa.

Tabela 3 – Comparação das pressões encontradas no Python e no EPANET.

Identificador dos nós	Pressões (m.c.a.)	
	Python	EPANET
3	44,696	44,70
4	49,516	49,52
5	45,114	45,11
6	59,616	59,62
7	64,447	64,45
8	59,856	59,86
9	79,447	79,45
10	84,447	84,45
11	74,707	74,71
12	88,722	88,72
13	15,498	15,50
14	06,070	06,07

Fonte: Próprio autor (2024).

Com esses valores de pressão encontrados na simulação hidráulica da rede no Python e utilizando a Equação 2, o valor da perda encontrada nesta rede criada para esse estudo inicial foi de 54,41%. Dessa forma, a utilização dessas ferramentas se mostrou viável para a continuação do estudo sobre operação otimizada de redes de distribuição de água.

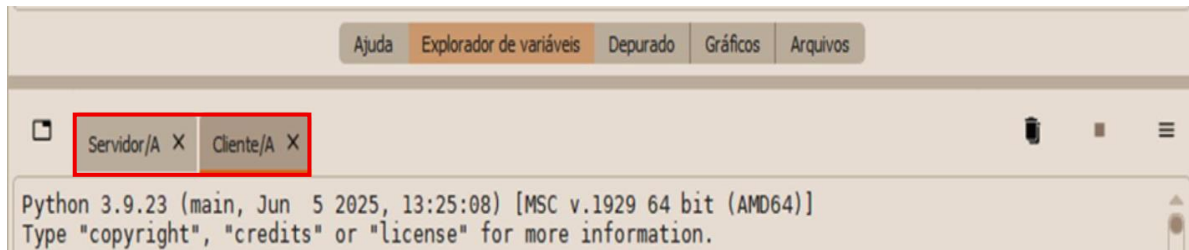
5.2 Resultados da aplicação do protótipo de gêmeo digital adaptativo

Com esses resultados iniciais de teste do pacote WNTR no Python, no reconhecimento de redes de distribuição de maneira correta, possibilitou o desenvolvimento da pesquisa e a obtenção dos resultados desejados.

Como a pesquisa trata da utilização de gêmeos digitais, mais especificamente de protótipo de gêmeo digital adaptativo, foi então realizada primeiramente a comunicação simultânea entre o servidor e o cliente com a utilização do *WebSockets*. Para isso, foi definida uma porta do computador para que a comunicação fosse implementada, uma vez que a conexão se dá pelo TCP/IP, ou seja, ela precisa de uma porta e um IP para a execução (Fette; Melnikov, 2011; Bayılmış *et al.*, 2022). Assim, a porta utilizada foi a de número 8765, que é uma das portas locais do computador, e o IP também foi local, ou seja, o mesmo da máquina utilizada no estudo.

Além da definição dos parâmetros, foi necessário abrir dois consoles para que a comunicação fosse iniciada. Como foi utilizada a IDE-Spyder do Anaconda para a execução das simulações, não foi necessário abrir dois ambientes de trabalho como seria no VScode, somente criar dois consoles em um mesmo ambiente do Spyder. A Figura 5.2 exemplifica essa criação dos dois consoles em uma única IDE, que está evidenciado em vermelho.

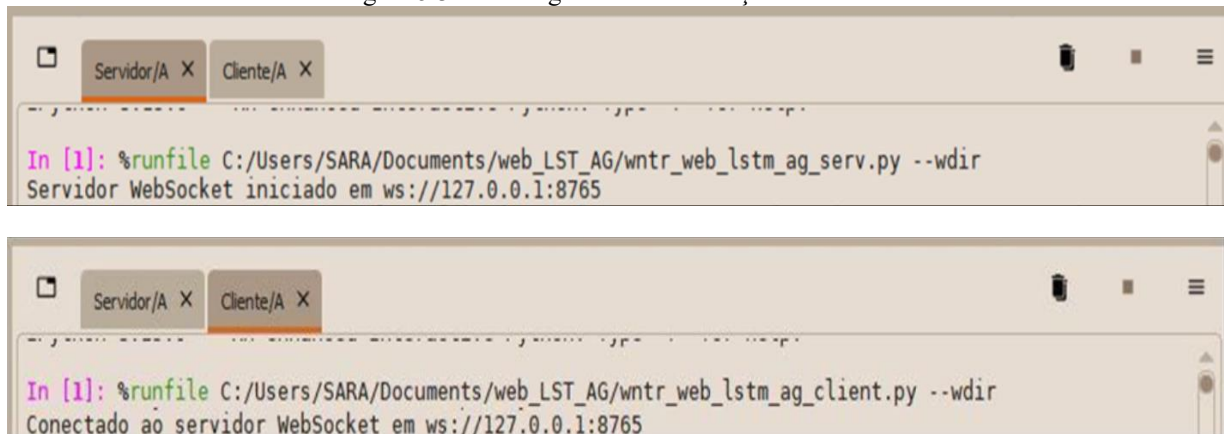
Figura 5.2 – Consoles criados no Spyder.



Fonte: Próprio autor (2025).

Com os consoles necessários para a comunicação do gêmeo digital definidos, estabeleceu-se a conexão cliente/servidor, sendo o servidor o responsável por transmitir os dados da rede para o cliente e este de realizar os estudos de previsão e otimização. Assim, a comunicação foi iniciada e uma mensagem é emitida em cada um dos consoles. O servidor mostra que a comunicação foi iniciada e em qual porta (8765) e host (127.0.0.1) está localizada, e o cliente apresenta que a conexão com o servidor foi realizada pela mesma porta e host. A Figura 5.3 traz as mensagens exibidas por cada um.

Figura 5.3 – Mensagens de comunicação iniciadas.



Fonte: Próprio autor (2025).

Com essa comunicação iniciada, o handshake estabelecido, o cliente pede os dados da rede para o servidor por meio de mensagem, e este retorna os dados solicitados. Esses dados transmitidos são os valores de pressão, os parâmetros operacionais (bombas, VRP e RNF) e a perda de cada hora, das 48 horas de funcionamento da rede de estudo. O servidor operou como se a rede de benchmark estudada estivesse funcionando em tempo real, pois a rede foi chamada para dentro do servidor por meio do WNTR utilizando o arquivo “.inp” do EPANET e em seguida realizou a simulação hidráulica da rede para que os dados de operação das 48 horas fossem extraídos.

Após a transferência dos dados da rede, uma mensagem de “dados recebidos” é emitida pelo cliente e o servidor emite a mensagem de “dados enviados”. Mesmo com essas mensagens emitidas, a conexão não é encerrada até que o cliente finalize essa comunicação com os comandos “*await websocket.send(fim)*”, que envia a mensagem de fim da conexão para o servidor e “*await websocket.close()*” que efetivamente finaliza a comunicação.

Como o *WebSockets* utiliza a porta do computador para a aplicação de gêmeo digital a cada simulação completa realizada (transmissão dos dados, simulação da previsão e otimização), a porta precisa ser liberada a cada simulação completa. Para isso, é preciso abrir o prompt de comando (CMD) da máquina e aplicar o comando “*netstat -ano | findstr 8765*”, que retorna o valor do PID (identificador do processo) que está sendo utilizado pela porta 8765. Para conhecer todas as portas que estão em utilização, o comando é o mesmo, porém sem a parte que especifica uma porta, sendo ele “*netstat -ano*”. Conhecendo o valor do PID correspondente à porta “8765”, é executado o comando “*taskkill /PID nº do PID encontrado /F*”, que vai liberar a porta correspondente (no caso do estudo a porta 8765) para que uma nova simulação completa possa ser realizada (Microsoft, 2016).

Assim, os valores de pressão da rede de estudo recebidos pelo cliente estão expressos na Tabela 4, sendo a unidade de pressão utilizada de m.c.a.

Tabela 4 - Pressões originais da rede de benchmark de estudo.

(continuação)

Horas	Identificador dos nós														
	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	11,97	39,94	69,94	9,92	49,96	50,94	12,97	69,84	44,84	56,84	51,84	39,84	42,91	10,92	10,42
2	11,97	39,93	69,93	9,91	49,95	50,93	12,97	69,87	44,82	56,88	51,82	39,83	42,85	10,91	10,40
3	11,68	39,38	69,39	9,2	49,59	50,39	12,74	69,82	43,38	53,45	50,38	38,45	42,79	10,20	8,74
4	11,40	38,81	68,84	8,49	49,22	49,84	12,5	66,91	41,91	54,04	48,91	37,04	39,45	9,48	7,05
5	11,43	38,87	68,9	8,57	49,26	49,9	12,52	67,08	42,08	52,19	49,08	37,19	39,64	9,56	7,24
6	11,51	39,02	69,05	8,76	49,36	50,05	12,59	67,46	42,46	52,56	49,46	37,56	40,08	9,75	7,68
7	11,57	39,15	69,16	8,91	49,44	50,16	12,64	67,78	42,78	52,87	49,78	37,87	40,45	9,90	8,05
8	11,44	38,89	68,91	8,58	49,27	49,91	12,53	67,11	42,11	52,23	49,11	37,23	39,67	9,57	7,27
9	11,73	39,49	69,49	9,33	49,65	50,49	12,78	68,64	43,64	53,70	50,64	38,7	41,43	10,33	9,04
10	11,87	39,74	69,75	9,67	49,83	50,75	12,89	69,33	44,33	54,36	51,33	39,36	42,23	10,67	9,83
11	11,95	39,89	69,9	9,86	49,93	50,9	12,96	69,73	44,73	54,74	51,73	39,74	42,68	10,86	10,29
12	11,97	39,94	69,94	9,92	49,96	50,94	12,97	69,84	44,84	54,84	51,84	39,84	42,81	10,92	10,42
13	11,97	39,95	69,95	9,94	49,97	50,95	12,98	69,87	44,87	54,88	51,87	39,88	42,85	10,94	10,46
14	11,97	39,93	69,93	9,91	49,95	50,93	12,97	69,82	44,82	54,83	51,82	39,83	42,79	10,91	10,4
15	11,68	39,38	69,39	9,20	49,59	50,39	12,74	68,38	43,38	53,45	50,38	38,45	41,13	10,2	8,74
16	11,40	38,81	68,84	8,49	49,22	49,84	12,5	66,91	41,91	52,04	48,91	37,04	39,45	9,48	7,05
17	11,43	38,87	68,9	8,57	49,26	49,90	12,52	67,08	42,08	51,19	49,08	37,19	39,64	9,56	7,24
18	11,51	39,02	69,05	8,76	49,36	50,05	12,59	67,46	42,46	52,56	49,46	37,56	40,08	9,75	7,68
19	11,57	39,15	69,16	8,91	49,44	50,16	12,64	67,78	42,78	52,87	49,78	37,87	40,45	9,90	8,05
20	11,44	38,89	68,91	8,58	49,27	49,91	12,53	67,11	42,11	52,23	49,11	37,23	39,67	9,57	7,27

Fonte: Próprio autor (2024).

Tabela 4 - Pressões originais da rede de benchmark de estudo.

(conclusão)

Horas	Identificador dos nós														
	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
21	11,73	39,48	69,49	9,33	49,65	50,49	12,78	68,64	43,64	53,7	50,64	38,7	41,43	10,33	9,04
22	11,87	39,75	69,75	9,67	49,83	50,75	12,89	69,33	44,33	54,36	51,33	39,36	42,23	10,67	9,83
23	11,95	39,89	69,9	9,86	49,93	50,9	12,96	69,73	44,73	54,74	51,73	39,74	42,68	10,86	10,29
24	11,97	39,94	69,94	9,92	49,96	50,94	12,97	69,84	44,84	54,84	51,84	39,84	42,81	10,92	10,42
25	11,99	39,95	69,95	9,94	49,97	50,95	12,98	69,87	44,87	54,88	51,87	39,88	42,85	10,94	10,46
26	11,98	39,94	69,94	9,90	49,96	50,92	12,98	69,82	44,83	54,83	51,83	39,84	42,79	10,92	10,41
27	11,66	39,39	69,40	9,23	49,6	50,37	12,73	68,83	43,39	53,46	50,39	38,47	41,13	10,21	8,75
28	11,42	38,82	68,87	8,47	49,24	49,86	12,51	66,93	41,92	52,03	48,93	37,02	39,46	9,49	7,06
29	11,44	38,86	68,92	8,59	49,25	49,92	12,53	67,07	42,09	52,2	49,07	37,2	40,56	9,58	7,23
30	11,52	39,03	69,07	8,78	49,38	50,06	12,6	67,48	42,48	52,57	49,47	37,57	40,08	9,76	8,90
31	11,59	39,16	69,18	9,45	49,46	50,17	12,63	67,79	42,77	52,86	49,79	37,86	40,46	9,91	8,06
32	11,46	38,90	68,90	8,56	49,26	49,90	12,55	67,14	42,13	52,24	49,13	37,25	39,69	9,58	8,60
33	11,90	39,48	69,50	9,35	49,66	50,47	13,30	68,63	41,23	53,73	50,65	38,72	41,45	10,34	9,70
34	11,86	39,76	69,77	9,66	49,82	50,76	12,40	69,35	44,34	54,37	50,9	39,37	42,24	10,60	9,84
35	11,97	39,87	69,93	9,84	49,91	50,91	12,98	68,74	44,72	54,73	51,74	39,73	42,69	10,87	10,28
36	11,96	39,95	68,90	9,93	49,98	50,95	12,96	69,86	44,86	54,85	51,54	39,85	42,91	10,91	10,43
37	11,98	39,96	69,96	9,95	49,96	50,97	12,56	69,86	44,85	54,89	51,88	39,89	42,86	10,93	10,32
38	11,98	39,96	69,95	9,42	49,97	50,96	12,98	69,83	44,83	54,84	51,83	39,84	42,78	10,92	10,41
39	11,67	39,37	69,37	9,91	49,57	50,38	12,76	68,39	43,39	53,46	50,39	38,46	41,14	10,22	8,75
40	11,41	38,83	68,85	8,50	49,23	49,85	12,51	66,92	41,92	52,05	48,92	37,05	39,46	9,45	10,30
41	11,44	38,88	68,91	8,56	49,28	49,93	12,53	67,07	42,07	52,18	49,07	37,18	39,67	9,57	8,69
42	11,53	39,01	69,09	8,78	49,35	50,04	12,58	67,48	42,47	52,57	49,49	37,57	40,09	10,48	8,34
43	11,99	39,17	69,17	10,1	49,46	50,18	12,66	67,79	42,79	52,89	49,76	37,88	40,46	9,92	8,06
44	11,45	38,87	68,93	8,59	49,29	49,93	12,52	67,13	42,44	52,25	49,13	37,24	39,68	9,58	7,28
45	11,74	39,46	69,46	9,89	49,66	50,50	12,79	68,66	43,66	53,72	50,66	38,73	41,45	10,34	9,05
46	11,86	39,76	69,77	9,98	49,85	50,77	12,90	69,35	44,34	54,38	51,43	39,37	42,27	10,68	10,00
47	11,29	39,90	69,93	9,87	49,96	50,92	12,97	69,76	44,46	54,75	51,74	39,86	42,69	10,89	10,28
48	11,96	39,96	69,93	9,13	49,94	50,95	12,98	69,88	44,86	54,86	51,85	39,82	42,8	10,94	10,46

Fonte: Próprio autor (2024).

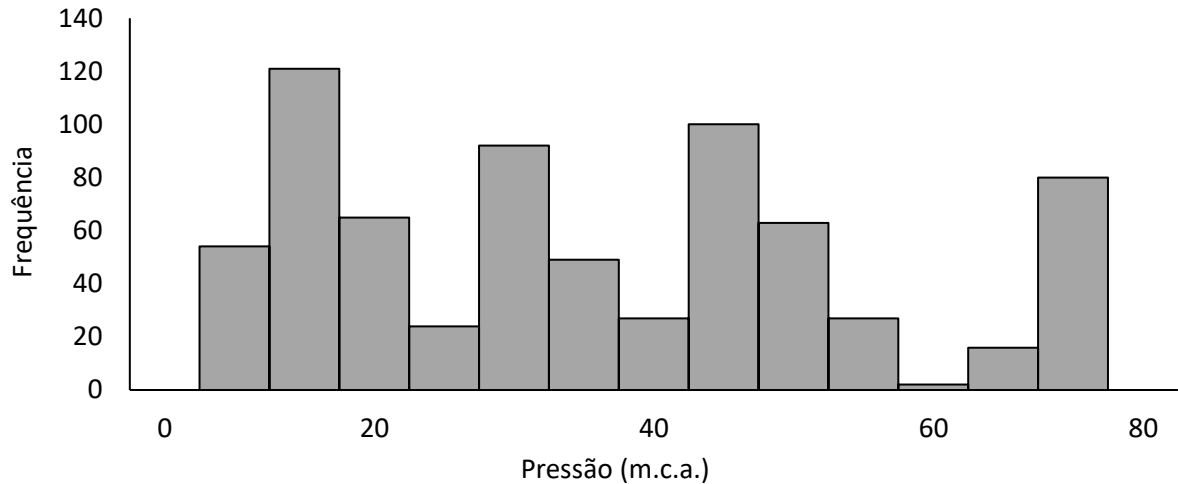
Os dados de pressões observados na Tabela 3 sofreram variações ao longo das horas devido à oscilação da demanda utilizada na rede; contudo a variação dos valores de pressões ao longo dos dias para cada um dos nós não sofreu grandes oscilações pois a demanda não apresentou uma variação grande ao longo das horas observando os dias simulados. Além disso, o sistema de estudo possui uma configuração malhada, o que pode causar interferência nas pressões já que o fluxo ocorre em mais de uma direção.

A rede em estudo apresenta uma distribuição heterogênea de pressões, com variações significativas entre valores mínimos e máximos. As maiores frequências encontradas de dados de pressões são as que possuem valores baixos, variando entre 10 e 20 m.c.a. Porém, o sistema também apresenta uma frequência alta de valores de pressão mais elevados, sendo que esses valores estão acima de 50 m.c.a., estando dessa forma acima do que especificado pela norma NBR 12218:2017.

Essa diferença entre os valores de pressão que a rede de estudo possui, aumenta a probabilidade de a mesma sofrer com perdas, o que justifica a sua utilização no estudo

desenvolvido pela pesquisa. A Figura 5.4 traz essas distribuições de pressão através do gráfico de histograma.

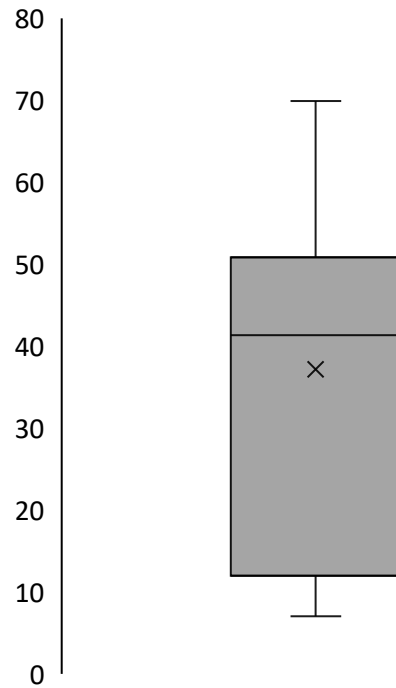
Figura 5.4 – Histograma das pressões da rede de estudo.



Fonte: Próprio autor (2024).

Quando se observa esses dados de pressão por meio do gráfico de boxplot, tem-se que a mediana desses valores está em torno de 41 m.c.a., o que significa que metade dos valores de pressão está abaixo desse valor e o outro está acima. Os demais quartis do gráfico Q1 e Q3 possuem valores de 12 m.c.a. e 50,9 m.c.a., respectivamente. Isso representa que 25% dos valores de pressão são menores que o apresentado por Q1 e 75% dos valores estão abaixo do valor de Q3. Além disso, também é possível observar os valores extremos de pressão da rede, sendo o mínimo de 7,05 m.c.a. e o máximo de 70 m.c.a., afirmando assim que a rede possui uma amplitude grande dos dados, confirmando assim o que foi dito anteriormente sobre a maior probabilidade de sofrer com perdas. A Figura 5.5, traz o gráfico de boxplot desenvolvido para a análise feita.

Figura 5.5 – Boxplot das pressões da rede de estudo.



Fonte: Próprio autor (2024).

Com os valores de pressões obtidos de cada nó para cada uma das horas de operação da rede, foi utilizada a Equação 2 para encontrar a perda que a rede apresentava originalmente. Assim, a Tabela 5 apresenta as perdas de cada uma das 48 horas de funcionamento da rede de benchmark estudada.

Tabela 5 - Perdas originais da rede de benchmark de estudo.

(continuação)

Horas	Perdas (%)	Horas	Perdas (%)
1	44,94	25	44,87
2	44,93	26	44,85
3	44,40	27	44,30
4	43,78	28	43,70
5	43,76	29	43,80
6	43,92	30	43,97
7	44,04	31	44,07
8	43,77	32	43,83
9	44,38	33	44,35
10	44,66	34	44,62
11	44,81	35	44,77

Fonte: Próprio autor (2024).

Tabela 5 - Perdas originais da rede de benchmark de estudo.
(conclusão)

Horas	Perdas (%)	Horas	Perdas (%)
12	44,85	36	44,81
13	44,87	37	44,85
14	44,85	38	44,83
15	44,28	39	44,31
16	43,70	40	43,83
17	43,72	41	43,82
18	43,92	42	43,98
19	44,04	43	44,11
20	43,77	44	43,79
21	44,38	45	44,41
22	44,66	46	44,69
23	44,81	47	44,79
24	44,85	48	44,83

Fonte: Próprio autor (2024).

Apesar de os valores de perdas da rede, antes do início das simulações de previsão e otimização, não apresentarem valores extremamente altos, os valores estão acima de 43%, o que representa um dado significativo de perdas para um sistema de distribuição de água, estando próximo ao valor apresentado pelas concessionárias brasileiras (Brasil, 2024). O que torna a rede aplicável no estudo.

5.3 Resultado das perdas previstas pela rede neural recorrente LSTM

Com os dados da rede já dentro do cliente, foi possível iniciar o modelo de previsão dos piores cenários de perdas. Para isso, foi utilizada a Rede Neural Artificial Recorrente Long Short Term Memory (RNAR-LSTM), por ser uma rede neural que melhor realiza previsões para dados de séries temporais, além de ser capaz de identificar desvios significativos do comportamento, como anomalias (Hochreiter; Schmidhuber, 1997; Aurélien, 2019; Lee; Yoo, 2021).

Assim, a arquitetura estabelecida para o modelo de previsão, foi na camada de entrada, os valores de pressões e perdas da série temporal da rede de estudo e das redes de distribuição de benchmarks Marques (2023), Silva (2023) e Anytown (Walski *et al.*, 1987), cujos dados de pressão e perdas estão em anexo. O número de épocas estabelecido foi de 300 e as camadas internas do modelo LSTM foram definidas em duas camadas, sendo a primeira com 128 nós e a segunda com 48 nós e a camada “dense”. A primeira camada tem a função de aprender padrões temporais de baixo nível, ou seja, padrões locais na sequência, e a segunda camada aprende os padrões de alto nível de acordo com as representações realizadas na primeira camada. Por fim a camada “dense” responsável por receber o estado oculto final produzido

pela LSTM e a transformá-lo nos dados previsões desejadas (Goodfellow; Bengio; Courville, 2016).

Outro fator definido nesta etapa foi o dropout, que tem a função de reduzir o sobreajuste durante o treinamento do modelo, tornando-o mais generalista, por meio do desligamento de frações de neurônios da camada durante o treinamento, assim o modelo não fica dependente de neurônios específicos (Srivastava et al., 2014; Goodfellow; Bengio; Courville, 2016; Keras, 2021). A Figura 5.6 traz esses parâmetros definidos na criação do modelo de previsão.

Figura 5.6 – Configuração das camadas do modelo de previsão.

```
def create_lstm_pressure_model(input_shape, n_outputs):
    model = Sequential()
    model.add(Input(shape=input_shape))
    model.add(LSTM(128, return_sequences=True))
    model.add(LSTM(48))
    model.add(Dropout(0.2))
    model.add(Dense(n_outputs))
    model.compile(optimizer='adam', loss='mse')
    return model
```

Fonte: Próprio autor (2025).

Como mencionado na metodologia, esta configuração foi definida de maneira empírica e com base em alguns trabalhos levantados, com isso, algumas configurações foram testadas para que a melhor para o estudo fosse encontrada. Para isso, foi utilizada a comparação dos valores encontrados nas métricas, em que valores de R^2 devem estar próximos de 1 e os demais devem ter os menores valores para que o modelo seja considerado viável. A Tabela 6 apresenta os valores encontrados para o MAE, R^2 e o MSE de cada configuração, sendo que a melhor está destacada em vermelho.

Tabela 6 – Configurações teste do LSTM de acordo com os valores das métricas.

(continuação)

Teste	Nº Camadas	Nº Neurônios	Nº Épocas	MAE (%)	R^2	MSE (%)
1	2	64	100	0,08	0,897	0,015
		48				
2	2	128	100	0,07	0,899	0,014
		48				
3	2	128	100	0,09	0,885	0,016
		64				
4	2	64	200	0,1	0,857	0,023
		48				
5	2	128	200	0,07	0,898	0,015
		48				

Fonte: Próprio autor (2024).

Tabela 6 – Configurações teste do LSTM de acordo com os valores das métricas.

(conclusão)

Teste	Nº Camadas	Nº Neurônios	Nº Épocas	MAE (%)	R ²	MSE (%)
6	2	128	200	0,08	0,891	0,017
		64				
7	2	64	300	0,1	0,804	0,032
		48				
8	2	128	300	0,05	0,918	0,01
		48				
9	2	128	300	0,07	0,901	0,031
		64				

Fonte: Próprio autor (2024).

Com a configuração da rede neural estabelecida, foram encontrados os valores das perdas previstas pelo modelo desenvolvido, que estão apresentados na Tabela 7. Os piores cenários de perdas para as 24 horas de funcionamento variaram entre 51 e 53%, mostrando que o modelo foi capaz de realizar a previsão dos cenários adversos de perdas para a rede de estudo. Isso confirma que a aplicação da prototipagem em gêmeos digitais é importante nos estudos desse tipo de sistema, pois possibilita estudar, verificar e otimizar seu funcionamento antes que venha a ocorrer.

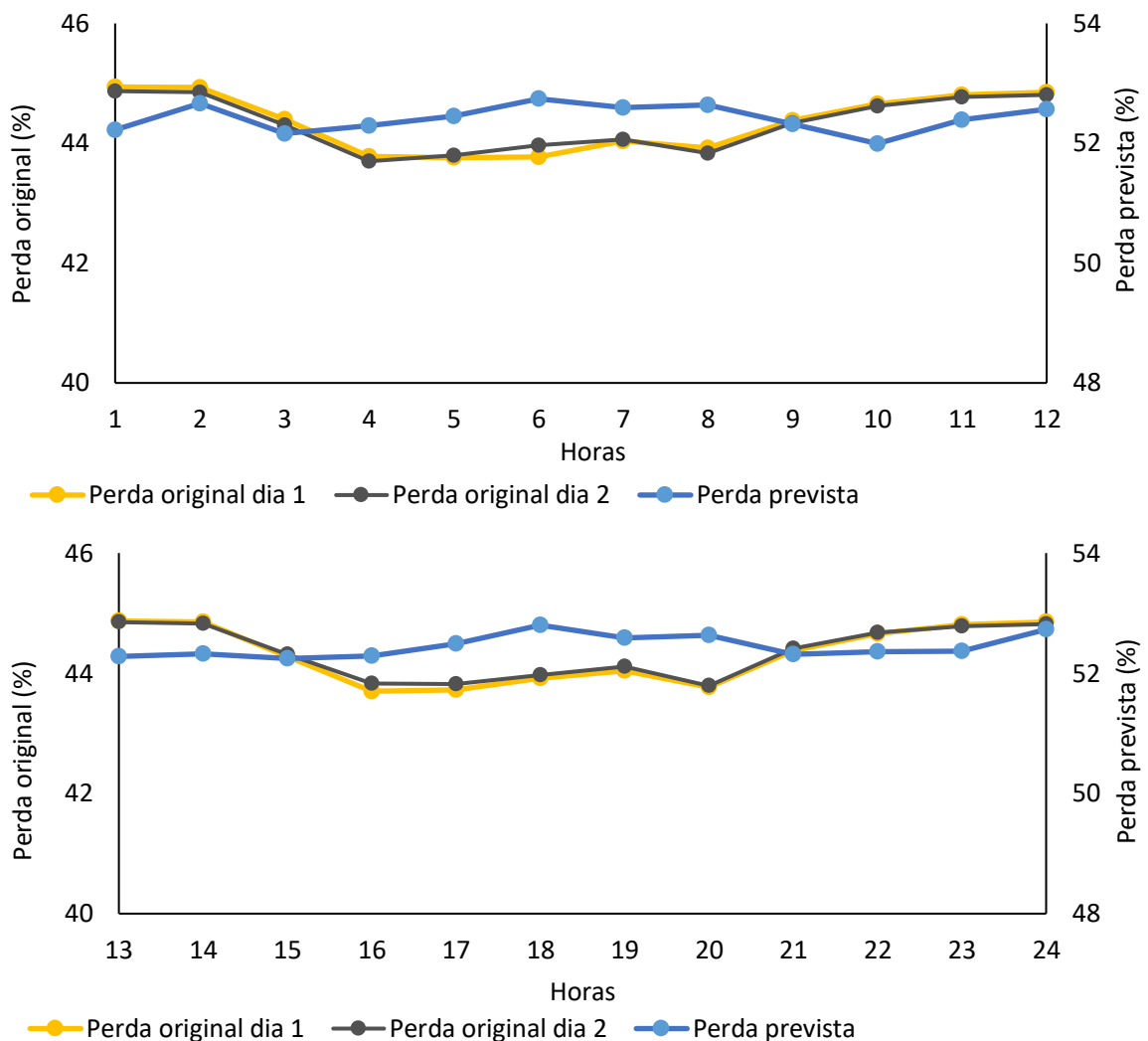
Tabela 7 - Perdas previstas.

Hora	Perda prevista (%)
1	52,22
2	52,67
3	52,16
4	52,29
5	52,45
6	52,74
7	52,60
8	52,64
9	52,32
10	51,99
11	52,39
12	52,56
13	52,28
14	52,33
15	52,24
16	52,28
17	52,49
18	52,80
19	52,59
20	52,63
21	52,32
22	52,36
23	52,36
24	52,73

Fonte: Próprio autor (2024).

Comparando os dados previstos com o original, observa-se que os valores de previsão encontrados pelo modelo apresentam um comportamento mais uniforme que os dados da rede original; contudo, as perdas previstas apresentaram uma diferença das perdas originais entre as horas 4 a 10 e das 16 a 20, mostrando uma elevação para os dois intervalos de perdas previstas enquanto a perda original possui uma diminuição. Mas nas demais horas, apresentaram uma semelhança no desenho das curvas. Como foram utilizados mais de uma rede para a alimentação do modelo de previsão, isso pode justificar a diferença apresentada pelas curvas para as horas mencionadas. Contudo, essa prática torna o modelo mais abrangente e generalista (Waqas *et al.*, 2024). A Figura 5.7 apresenta a variação das perdas originais da rede e as previstas pelo modelo de previsão da rede neural LSTM para o período de 24 horas. Para a melhor visualização dos dados o gráfico foi dividido em duas partes.

Figura 5.7 – Comparação das perdas previstas e originais.



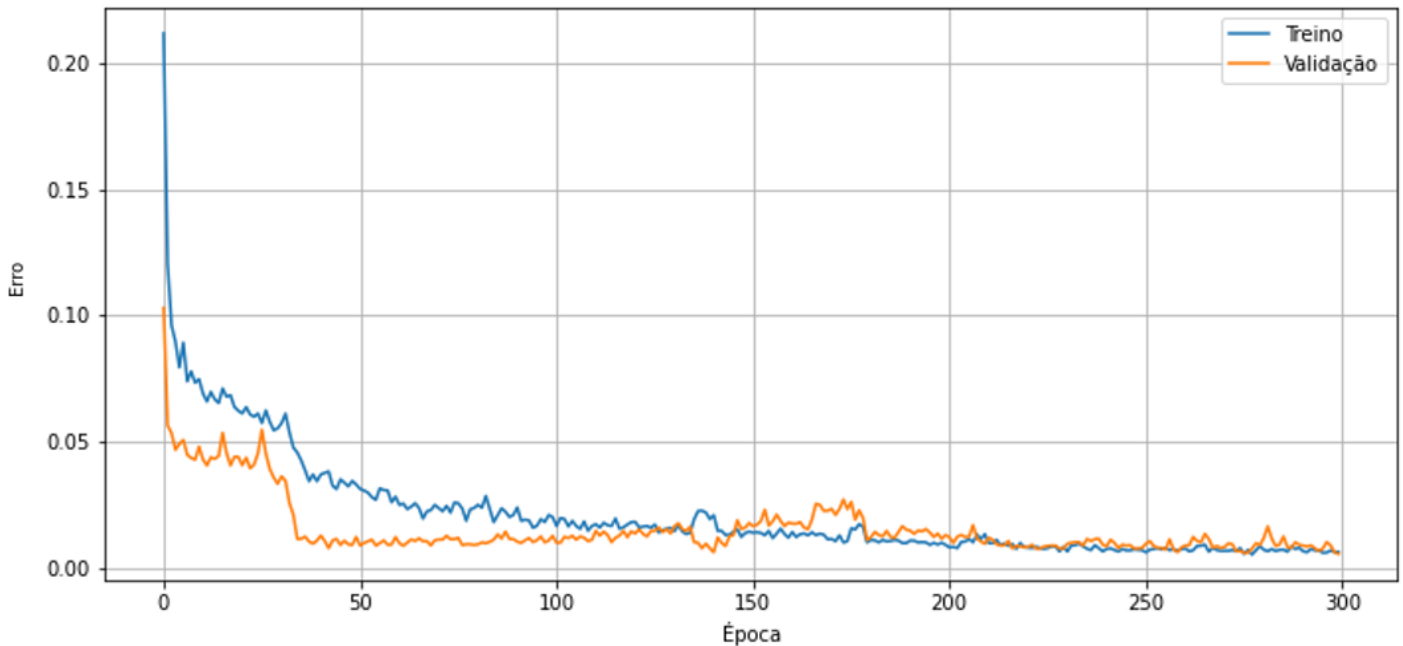
Fonte: Próprio autor (2025).

Embora o treinamento tenha sido executado por 300 épocas, o melhor desempenho de validação ocorreu na época 278, cujos pesos foram mantidos como modelo final, o que confirma que o número de 300 épocas foi o suficiente para o modelo encontrar seu objetivo.

Por se tratar de um modelo de previsão, é preciso realizar a análise dos dados previstos por meio de métrica, como apresentada na metodologia. Com isso, foi realizada a curva de perda ao longo das épocas com os dados de treinamento e validação usados no modelo de previsão. A determinação das perdas por épocas foi realizada pela função MSE (Erro Quadrático Médio), que está exemplificada na Equação 11. Essa curva de aprendizado permite avaliar o comportamento do modelo durante o treinamento. Sendo possível observar se o modelo sofreu overfitting ou undefitting. Pois, em um cenário de sobreajuste se caracteriza por um erro de treinamento baixo e o erro de validação significativamente mais alto. Já o subajuste ocorre quando ambos os erros permanecem elevados (Shirdel *et al.*, 2021; Stearns e Ruslan, 2024; Marc e Zain, 2024).

Dessa forma, observa-se no gráfico, um comportamento semelhante entre as curvas de treinamento e validação, indicando que o modelo está se ajustando aos dados durante o treinamento e não está sofrendo com sub ou superestimação dos dados, apesar da curva de validação apresentar um pequeno aumento durante as épocas 150 e 200 e uma redução antes da época 150. Contudo, ambas as curvas estão em declínio até estabilizarem no erro mínimo, e essa proximidade entre elas e seus baixos valores indicam que o modelo atingiu um ajuste adequado, sem evidências de alta variância ou alto viés (Dragicevic *et al.*, 2016; Stearns e Ruslan, 2024; Marc e Zain, 2024). O valor do erro apresentado no final da validação foi de 0,01. A Figura 5.8 traz a curva encontrada no treinamento e validação do modelo de previsão proposto.

Figura 5.8 – Curva LSTM.

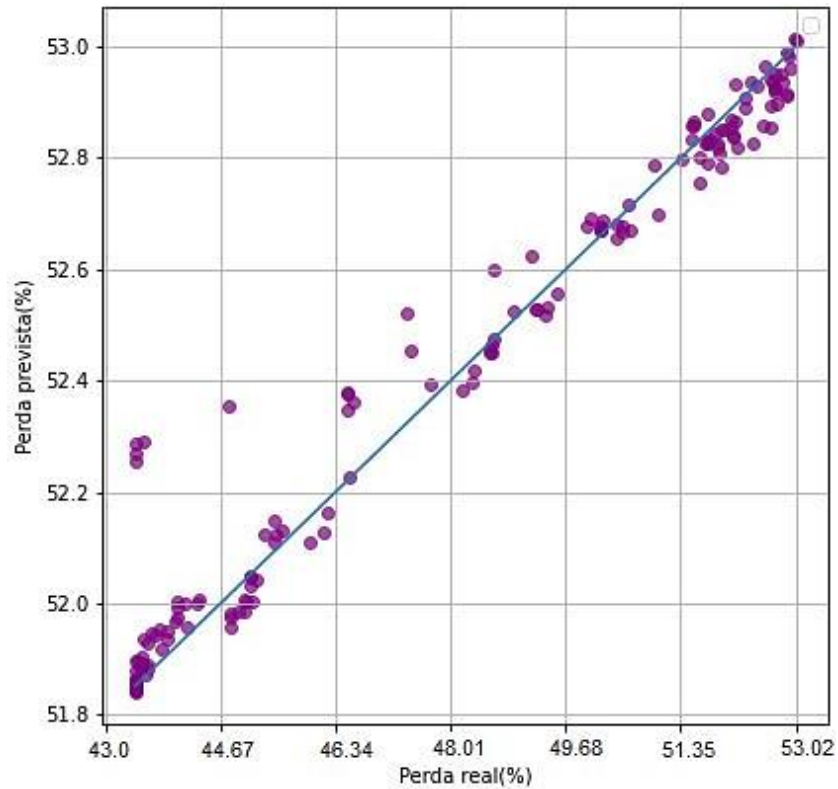


Fonte: Próprio autor (2025).

Apesar de a aplicação das métricas a seguir ser realizada para a comparação de modelos de previsão, no presente trabalho foram utilizadas para auxiliar na escolha da melhor configuração de camadas e épocas para o modelo desenvolvido (Bai *et al.*, 2021; Zanfei *et al.*, 2022; Bonilla *et al.*, 2025). Assim, serão apresentados a seguir os valores encontrados por essas outras métricas (R^2 , MAE) e como elas podem ser viáveis para a confirmação do modelo desenvolvido quanto à sua viabilidade para o estudo.

Com isso, para a verificação dos dados que foram previstos pelo modelo, foi realizada a aplicação de técnicas de análise sobre os dados selecionados para teste, que corresponde a 15% dos dados de entrada, ou seja, dos dados totais. Dessa forma, foi elaborado o gráfico de dispersão entre os dados de entrada (real) e os de saída (previstos) e verificou-se que a maioria dos dados previstos está relativamente simétricos em relação a sua distribuída em torno da reta de referência sem superestimativas ou subestimativas significativas, como mostra a Figura 5.9. Dessa forma, apresenta uma correlação positiva dos dados. Outro fator importante é o valor encontrado pelo R^2 que precisa estar próximo de 1 para ser considerado ideal, ou seja, que o modelo conseguiu se ajustar ao problema (Yu *et al.*; 2025). Assim, o modelo apresentou um valor de 0,918, o que significa que o modelo apresentou um bom desempenho de previsão, pois seu valor está próximo do valor considerado ideal.

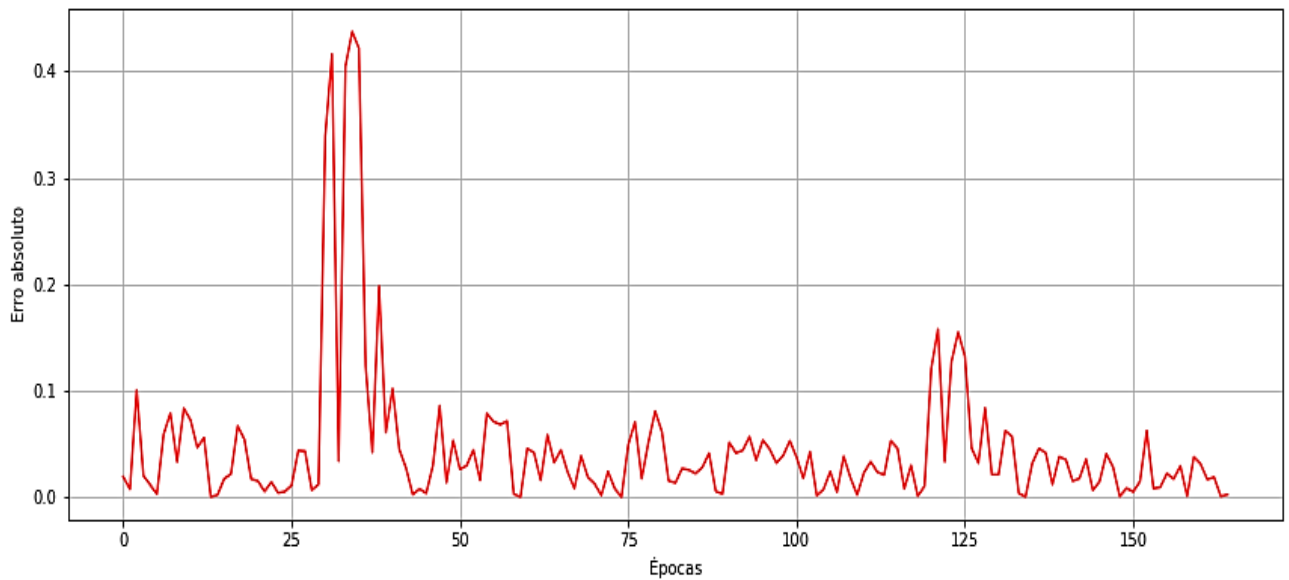
Figura 5.9 – Diagrama de dispersão.



Fonte: Próprio autor (2025).

Também foi realizada a análise do erro absoluto encontrado pelo modelo de previsão a cada época com os dados de teste. Esses valores estão mostrados no gráfico da Figura 5.10 e observa-se que os maiores erros ocorreram no início da simulação entre as épocas 25 e 50. Outro ponto que apresentou erro mais elevado foi próximo à época 125. Contudo, o valor do erro é pequeno ao longo do tempo, levando a um erro médio absoluto (MAE) de 0,05, que é considerado bom para o modelo, por ser um valor muito pequeno (Bai *et al.*, 2021; Zanfei *et al.*, 2022; Bonilla *et al.*, 2025).

Figura 5.10 – Erros absolutos por época.



Fonte: Próprio autor (2025).

5.4 Resultado da otimização das perdas prevista pelo algoritmo genético

Com a finalização da rede neural LSTM e com os valores dos piores cenários de pressão e perdas da rede para as 24 horas de operação, foi realizada a implementação do algoritmo genético para otimização das perdas previstas.

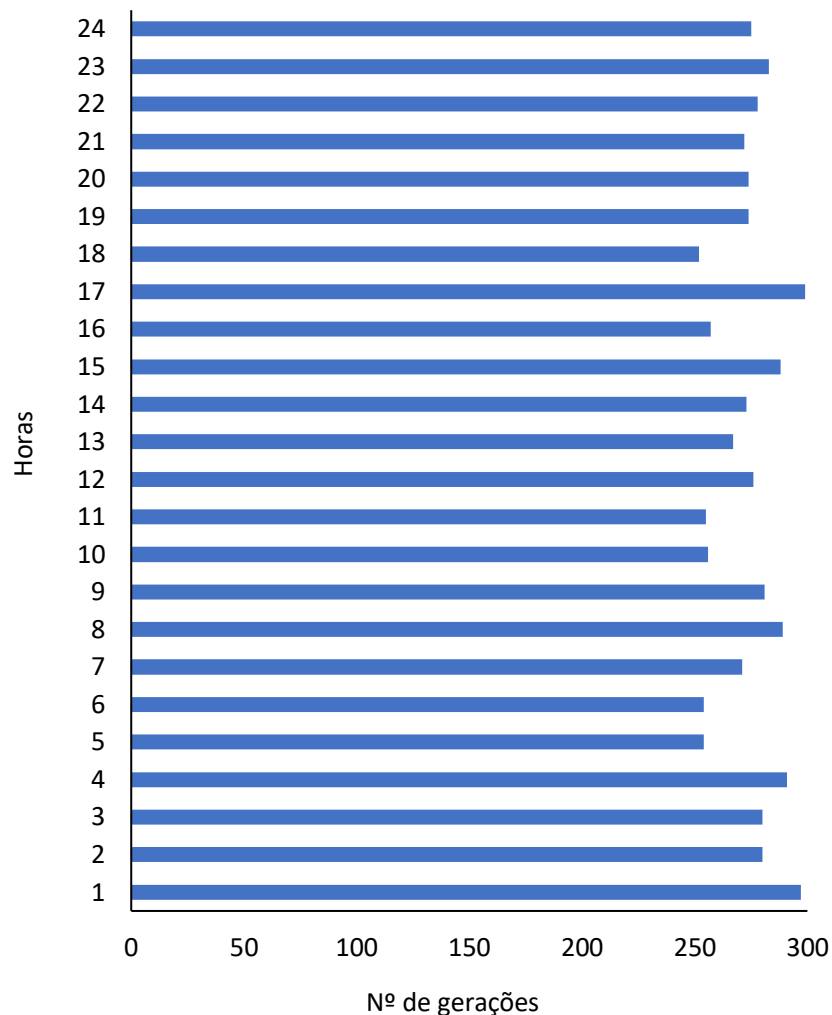
O algoritmo desenvolvido utiliza semente aleatória para a inicialização das simulações e, com isso, o algoritmo busca os valores de cada parâmetro operacional na população dentro do intervalo determinado para cada um: bombas, válvulas redutoras de pressão e reservatórios. Este intervalo de variação está especificado na seção metodologia. Além disso, é importante determinar a quantidade de população que será necessária para o algoritmo alcançar o resultado esperado, sendo então definido em 200 (Reca et al., 2017; Marques, 2023; Silva, 2023).

Além disso, também foi definido o número de gerações necessárias para o funcionamento do algoritmo, estabelecida em 300. Alguns trabalhos levantados na literatura, como Lara et al. (2024), Marques (2023), utilizaram valores maiores de gerações em seus estudos de otimização de redes de distribuição de água, com valores de 10000, 20000 e 1000. Porém, nos trabalhos trazidos por Reca et al. (2017) e Martínez-Bahena et al. (2018), o número de gerações apresentados é menor, de 200, 300 e 100. Assim, a escolha de 300 gerações para o modelo está embasada na literatura, mesmo apresentando um valor menor que alguns estudos. Outro fator que auxiliou na determinação desse valor foi o tempo de simulação, pois a utilização de 300 gerações levou 6 horas para realizar a otimização

completa para as 24 horas de operação futura da rede. Com isso, se o número de gerações fosse grande como utilizado por Lara et al. (2024) e Marques (2023), o tempo gasto seria de dias, o que inviabilizaria o estudo.

Com o algoritmo estruturado, foi realizada a otimização das piores perdas previstas anteriormente pelo modelo de previsão e, para cada hora, foram encontradas a geração ideal para essa otimização. A Figura 5.11 traz as gerações em que o algoritmo convergiu a cada hora de otimização.

Figura 5.11 – Geração ideal encontrada na otimização para cada hora.



Fonte: Próprio autor (2025).

De acordo com os resultados alcançados, observa-se que para todas as horas as melhores gerações se deram acima da geração de número 250, contudo todas encontraram o melhor resultado antes do limite determinado. Isso possibilita inferir preliminarmente que o algoritmo foi capaz de encontrar solução antes do limite determinado, que foi de 300 gerações.

Com a otimização dos cenários adversos previstos anteriormente, o algoritmo genético aplicado retornou às pressões, perdas otimizadas e as configurações dos parâmetros dos elementos operacionais que levaram à redução das perdas.

Os valores das perdas otimizadas pelo algoritmo apresentaram uma redução significativa quando comparadas às perdas previstas, com diferença variando entre 10,71 e 12,02 pontos percentuais, sendo que a hora com maior redução foi a hora 18 e a menor redução ocorreu na hora 1. Com isso, o algoritmo alcançou uma redução média de 11,42 p.p., que é um valor significativo para o sistema de distribuição de água estudado. A Tabela 8 apresenta os valores das perdas otimizadas para cada hora, ao lado dos valores das perdas prevista, para melhor visualização das reduções encontradas pelo algoritmo.

Tabela 8 – Comparação das perdas previstas e otimizadas.

Horas	Perda prevista (%)	Perda otimizada (%)
1	52,22	41,51
2	52,67	41,46
3	52,16	41,29
4	52,29	40,87
5	52,45	40,58
6	52,74	40,77
7	52,60	40,86
8	52,64	41,03
9	52,32	40,63
10	51,99	41,10
11	52,39	41,55
12	52,56	41,23
13	52,28	41,27
14	52,33	41,30
15	52,24	41,20
16	52,28	40,90
17	52,49	40,61
18	52,80	40,78
19	52,59	40,84
20	52,63	40,90
21	52,32	40,64
22	52,36	40,86
23	52,36	41,05
24	52,73	41,17

Fonte: Próprio autor (2025).

Como as perdas são definidas pelo controle das pressões nas redes, é importante analisar as pressões alcançadas pelo algoritmo genético que levaram às perdas otimizadas

apresentadas. A Tabela 9 traz os dados das pressões de cada um dos nós da rede referente às horas de otimização em m.c.a.

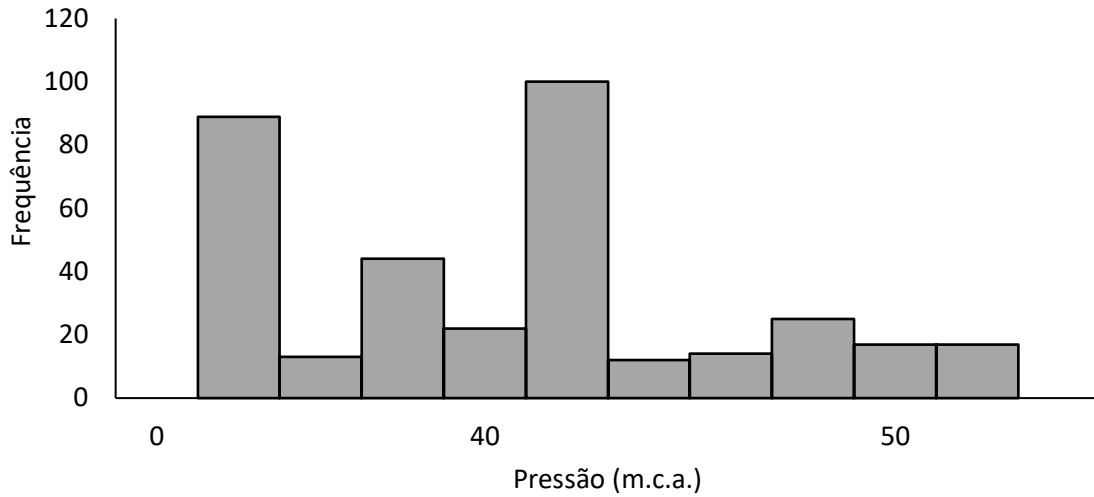
Tabela 9 - Pressões da rede de estudo depois da otimização.

Horas	Identificador dos nós														
	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	12,93	40,17	48,96	12,93	39,11	36,08	15,06	51,11	40,93	46,96	37,07	39,96	42,94	12,28	12,55
2	12,48	39,98	48,92	12,95	38,16	35,99	15,81	51,21	40,70	46,92	37,18	39,93	42,90	12,00	12,61
3	12,03	40,16	48,96	12,03	38,05	36,03	13,42	52,11	40,02	46,94	36,08	40,41	42,92	12,01	12,58
4	12,05	39,45	47,49	11,73	39,68	35,49	12,93	50,49	39,68	45,49	35,63	38,56	41,24	11,73	12,33
5	12,46	38,97	46,03	11,58	39,34	34,96	13,51	50,03	39,87	44,03	35,88	37,14	39,57	11,58	12,50
6	12,46	39,23	46,35	11,84	39,55	35,19	13,81	51,35	39,95	44,35	36,03	37,45	39,91	11,84	12,53
7	12,31	39,57	46,82	11,71	39,74	35,45	13,92	52,82	39,78	44,82	35,78	37,90	40,44	11,48	11,33
8	13,09	39,85	47,26	12,02	39,88	35,67	13,25	50,26	40,44	46,27	36,17	38,32	40,93	11,71	12,55
9	12,14	39,22	46,40	11,63	39,56	35,21	13,18	50,40	39,58	44,40	35,60	37,49	39,96	11,27	12,41
10	12,23	39,80	47,91	11,97	39,93	35,77	14,05	51,91	40,07	45,91	36,19	38,95	41,70	11,97	11,01
11	12,12	39,79	48,40	11,90	39,89	35,81	13,00	51,40	39,95	46,40	45,90	39,44	42,30	11,90	11,67
12	12,01	39,92	48,77	11,96	39,96	35,93	13,44	51,77	40,00	46,76	36,05	39,78	42,72	11,96	11,49
13	12,29	39,98	48,92	12,09	38,90	36,89	13,53	51,92	40,18	46,92	36,13	39,94	42,89	11,09	11,63
14	13,10	40,03	49,04	12,63	38,11	36,09	13,56	50,04	40,83	47,02	36,63	40,09	43,41	11,63	11,92
15	12,20	39,93	48,82	12,03	39,95	35,93	13,34	50,82	40,10	46,82	36,05	39,83	42,79	11,03	11,99
16	12,07	39,49	47,50	11,73	39,71	35,52	13,01	51,50	39,70	45,50	35,63	38,56	41,25	11,48	12,12
17	12,48	39,06	46,17	11,87	39,47	35,09	13,01	50,17	39,77	44,17	35,63	37,28	39,71	11,63	12,61
18	12,63	39,09	46,24	11,95	39,44	35,08	13,40	52,24	40,01	44,24	35,95	37,35	39,80	11,67	12,89
19	12,28	39,27	46,62	11,29	39,54	35,23	14,10	52,62	39,87	44,62	36,02	37,71	40,23	11,80	12,12
20	12,89	39,28	46,94	11,05	39,58	35,31	12,92	52,94	40,20	44,94	35,87	38,01	40,60	11,69	12,43
21	12,30	38,90	46,14	11,55	39,28	34,93	12,96	52,14	39,66	44,14	35,57	37,26	39,71	11,45	12,67
22	12,07	39,48	47,64	11,61	39,65	35,49	12,87	50,64	39,76	45,64	35,70	38,70	41,43	11,73	11,37
23	12,43	39,74	48,33	11,53	39,83	35,75	12,99	50,33	40,16	46,33	36,29	39,36	42,23	11,87	11,08
24	12,22	39,89	48,73	12,01	39,93	35,90	13,10	50,73	40,10	46,73	36,02	39,74	42,68	11,95	11,29

Fonte: Próprio autor (2024).

Para analisar os dados de pressão otimizados encontrados pelo algoritmo de otimização, foi elaborado um gráfico de histograma para verificar como os valores de pressão estão distribuídos, que está exemplificado na Figura 5.12. Os valores de pressão que apresentam maior frequência estão entre o intervalo de 40 a 50 m.c.a., com o pico de maior frequência em torno de 41 m.c.a. A rede também apresenta uma distribuição elevada entre as faixas de 13 a 15 m.c.a. Além disso, é possível observar uma frequência pequena de valores de pressão que estão acima do valor de 50 m.c.a., que é um valor que está acima do estabelecido pela norma NBR 12218. Esse valor é causado pelo nó 10 da rede, como exemplifica a Tabela 9, em que os valores estão entre 50,03 e 52,94 m.c.a.

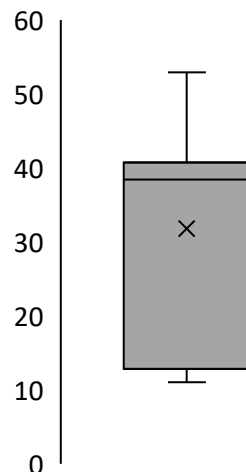
Figura 5.12 – Histograma das pressões otimizadas da rede de estudo.



Fonte: Próprio autor (2024).

Com a utilização do gráfico de boxplot (Figura 5.13) observa-se que a mediana (Q2) dos dados de pressão é de 38,5 m.c.a. indicando o ponto central dos dados de pressão, ou seja, metade dos dados está a cima desse valor e a outra metade está a baixo desse valor. Quanto aos outros dois quartis os valores são de 13 m.c.a. para o quartil 1 (Q1), significando que 25% dos dados de pressão estão a abaixo desse valor, e para o quartil 3 (Q3) o valor foi de 40,8 m.c.a., demonstrando que 75% dos dados está a abaixo desse nível. Também foi observado os valores mínimos e máximos de pressão, indicando assim a amplitude dos dados, sendo o maior de 52,94 m.c.a. e o menor de 11,01 m.c.a. Mesmo que um dos nós apresente valores de pressão acima da norma NBR 12218 ele não é considerado um outlier porque ele não desvia significativamente dos demais dados, entrando, assim, como amplitude dos dados.

Figura 5.13 – Boxplot das pressões otimizadas da rede de estudo.



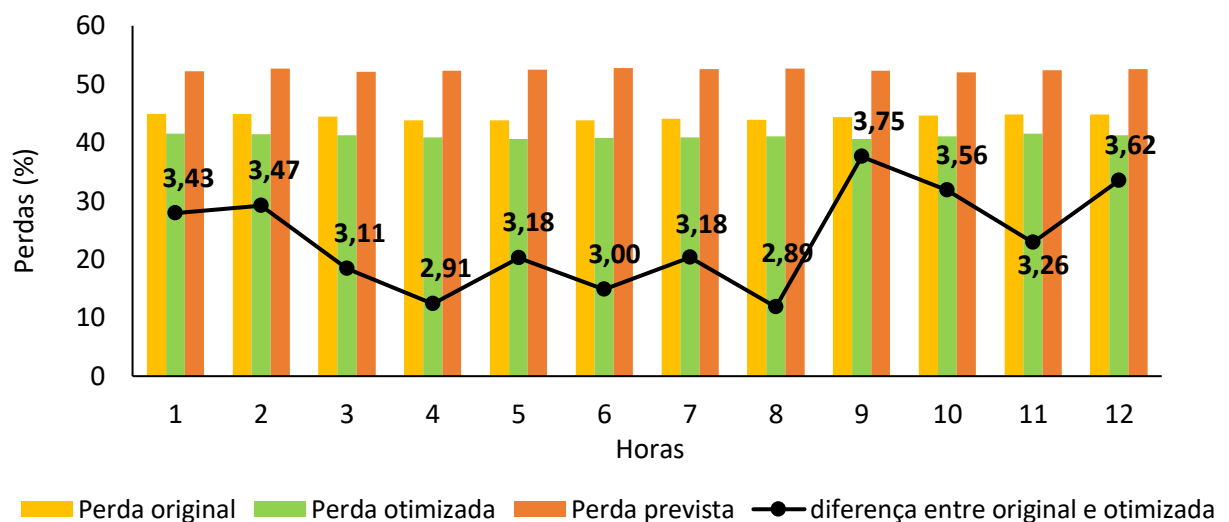
Fonte: Próprio autor (2024).

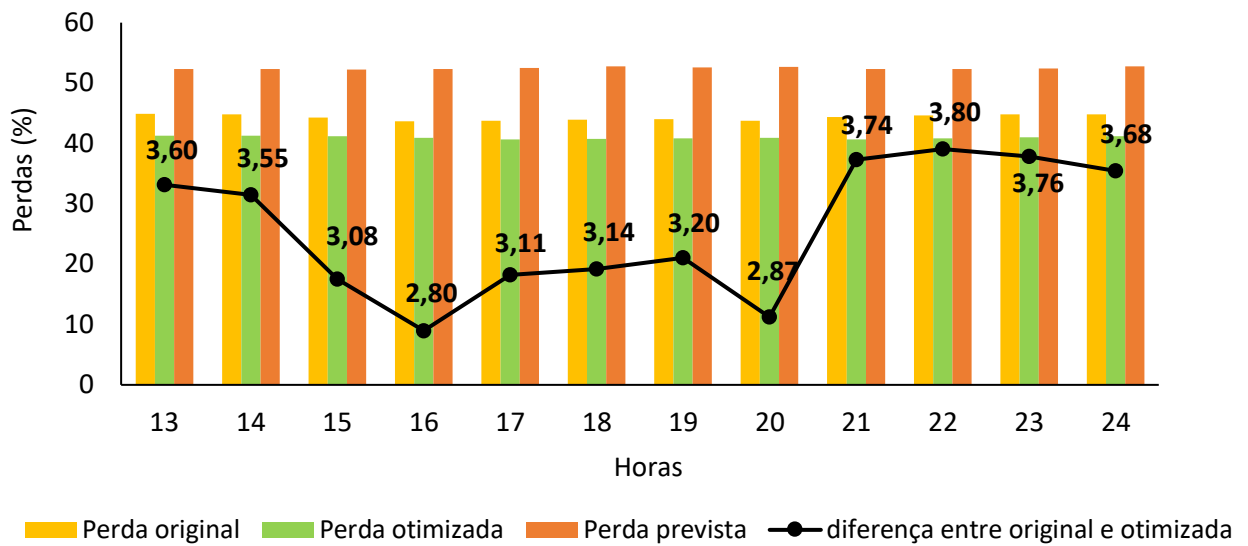
O fato de a rede apresentar esse comportamento com um nó com valores acima da norma e alguns nós com valores próximos do limite mínimo de 10 m.c.a, pode se dar pela configuração da rede, pois ela é constituída por malhas, o que faz com que um nó receba o fluxo de mais de uma direção e com isso, os elementos operacionais como válvulas não tenham uma ação tão efetiva quando visto em um sistema ramificado (Silva, 2023).

Outro fator que pode ter influenciado nesses valores de pressão foi o número de gerações determinado para o modelo de otimização, que apesar de ter encontrado uma geração como ideal dentro do limite determinado, pode não ter sido o suficiente para encontrar valores favoráveis para todos os nós da rede. Além disso, outra questão que pode ter levado a essa ocorrência foram os intervalos determinados aos parâmetros de operação (bomba, VRP e RNF), para que o algoritmo genético pudesse variar para encontrar a otimização ideal para cada hora, pois esses valores são gerados por semente aleatória, o que não permite um controle maior sobre o funcionamento do algoritmo de otimização.

Apesar do sistema apresentar uma redução significativa quando comparado com os piores cenários de operação previstos para a rede, os valores de perdas encontrados na otimização não foram tão significativos quando comparados com a rede original, apresentando uma diferença que varia de 2 a 3 pontos percentuais. A Figura 5.14 traz essa comparação, bem como os valores da diferença entre as perdas originais e as otimizadas para cada hora. Para a melhor visualização dos dados o gráfico foi dividido em duas partes, sendo a primeira com os dados das primeiras doze horas e a segunda parte com os dados das últimas doze horas.

Figura 5.14 – Comparação das perdas originais e otimizadas.





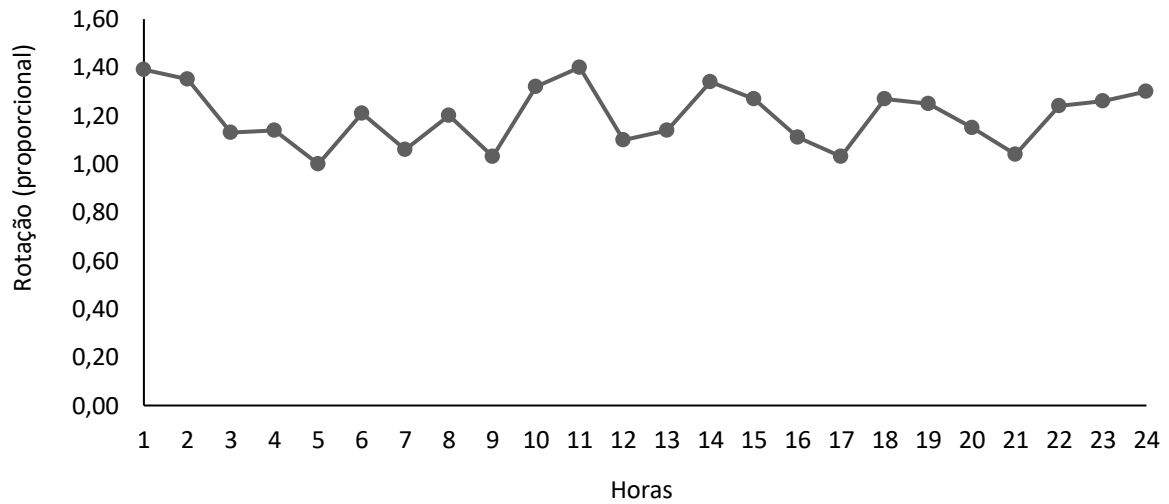
Fonte: Próprio autor (2025).

Mesmo que estes valores da diferença das perdas da rede original com a otimizada sejam considerados discretos, eles são importantes quando se trata de estudos de redes de distribuição de água, onde o número de perdas é sempre elevado, reduzindo a eficiência do sistema e elevando os custos e desperdícios.

Além disso, o algoritmo genético também retornou as configurações dos parâmetros de operação da rede que levaram à otimização das perdas. Esses valores foram definidos pelo intervalo estabelecido para cada um, conforme determinado na metodologia.

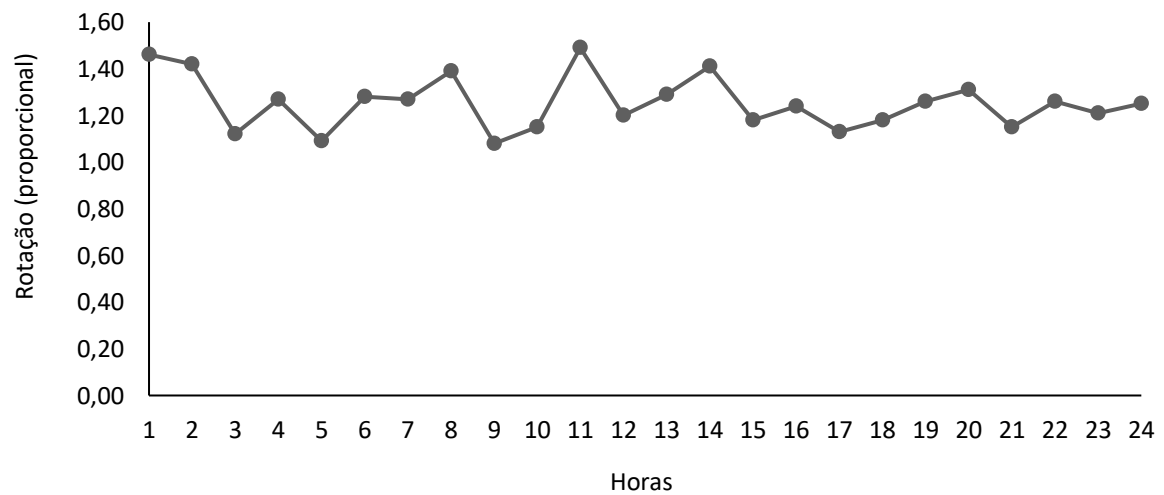
As Figuras 5.15 e 5.16 apresentam os valores encontrados pelo algoritmo de otimização para as duas bombas da rede de estudo para cada uma das horas. Ambas possuem uma oscilação pequena dos valores, variando de 1 a 1,4, mesmo que o intervalo determinado para o algoritmo realizar a busca fosse maior. Além disso, observa-se que o comportamento dos dados das duas bombas apresenta semelhança com os picos de maior e menor valores para as mesmas horas.

Figura 5.15 – Dados da Bomba 7, obtidos pelo algoritmo genético.



Fonte: Próprio autor (2025).

Figura 5.16 – Dados da Bomba 17, obtidos pelo algoritmo genético.

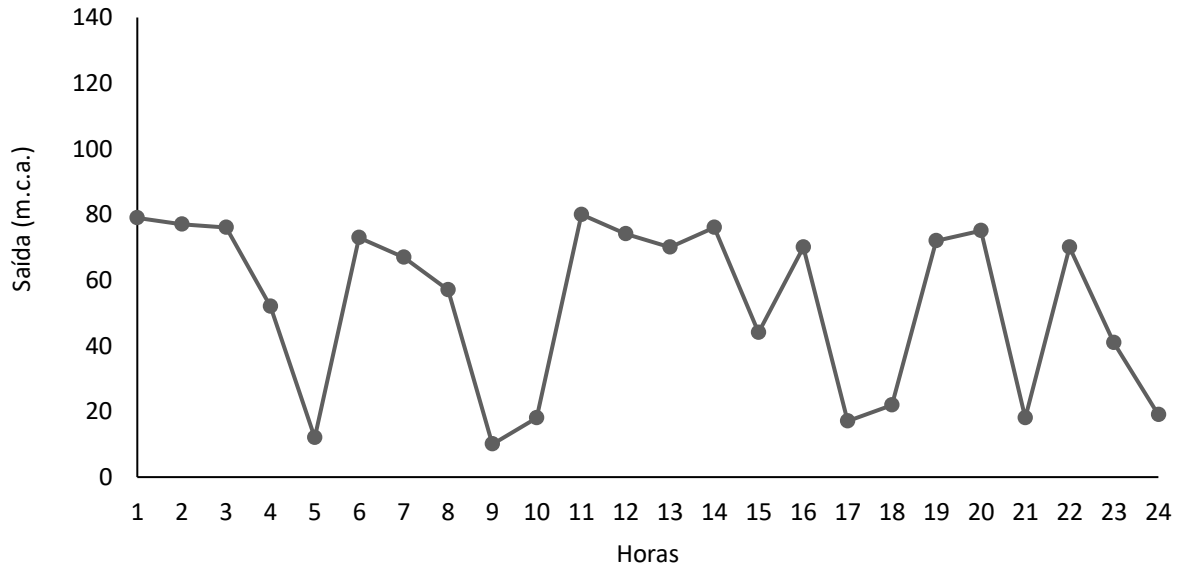


Fonte: Próprio autor (2025).

As Figuras 5.17 a 5.19 mostram a variação ao longo das horas encontradas pelo algoritmo genético para as válvulas 9, 19 e 23 da rede de estudo. O intervalo determinado para que o algoritmo pudesse buscar para encontrar o melhor comportamento da rede para a otimização das perdas foi de 0 a 100% e com isso, observa-se que houve uma grande oscilação nos valores encontrados pelo algoritmo para cada uma das válvulas ao longo da operação. Como as válvulas controlam a pressão a jusante e como a rede possui uma configuração malhada, pode ter influenciado a oscilação das aberturas das válvulas ao longo das horas de operação do sistema. Contudo, apresentam um comportamento semelhante nas horas que apresentaram picos maiores e menores nos valores de abertura das válvulas, como

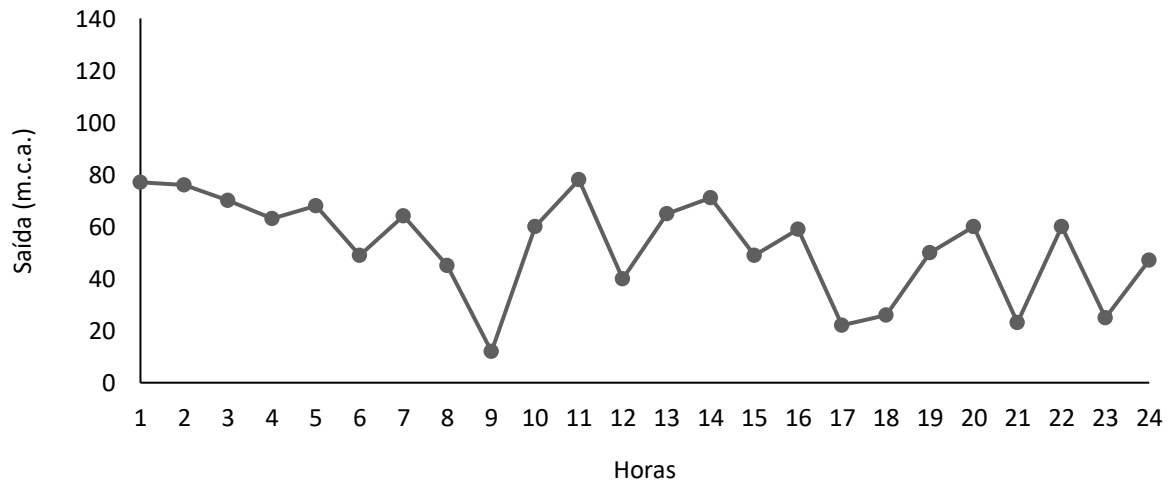
as horas 11 e 14 em que apresentam os maiores valores de abertura das válvulas e as horas 5 e 9 que se observa os menores valores.

Figura 5.17 – Dados da válvula 9, obtidos pelo algoritmo genético.



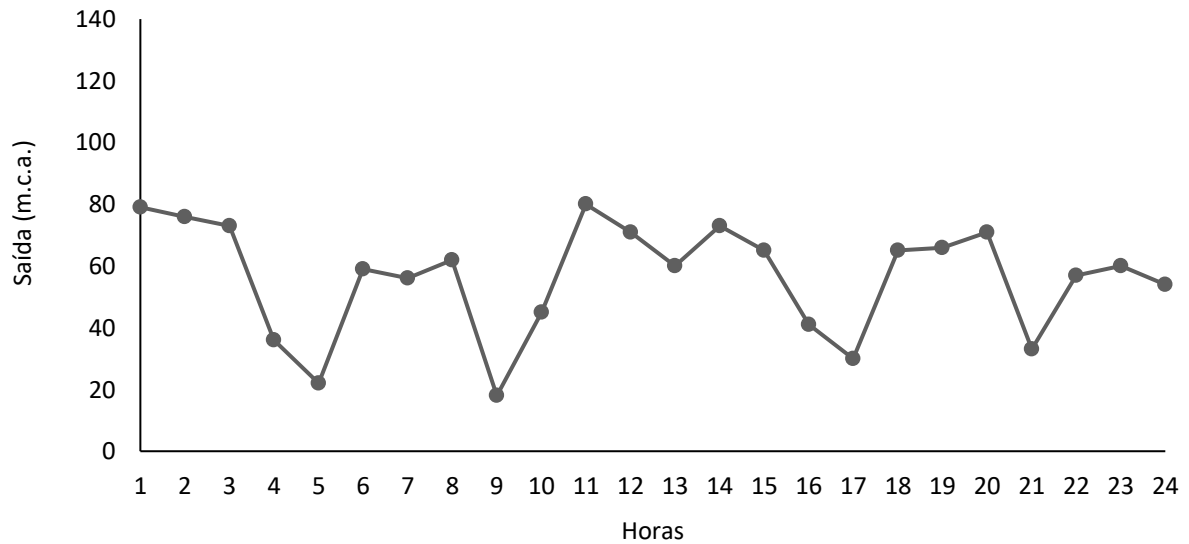
Fonte: Próprio autor (2025).

Figura 5.18 – Dados da válvula 19, obtidos pelo algoritmo genético.



Fonte: Próprio autor (2025).

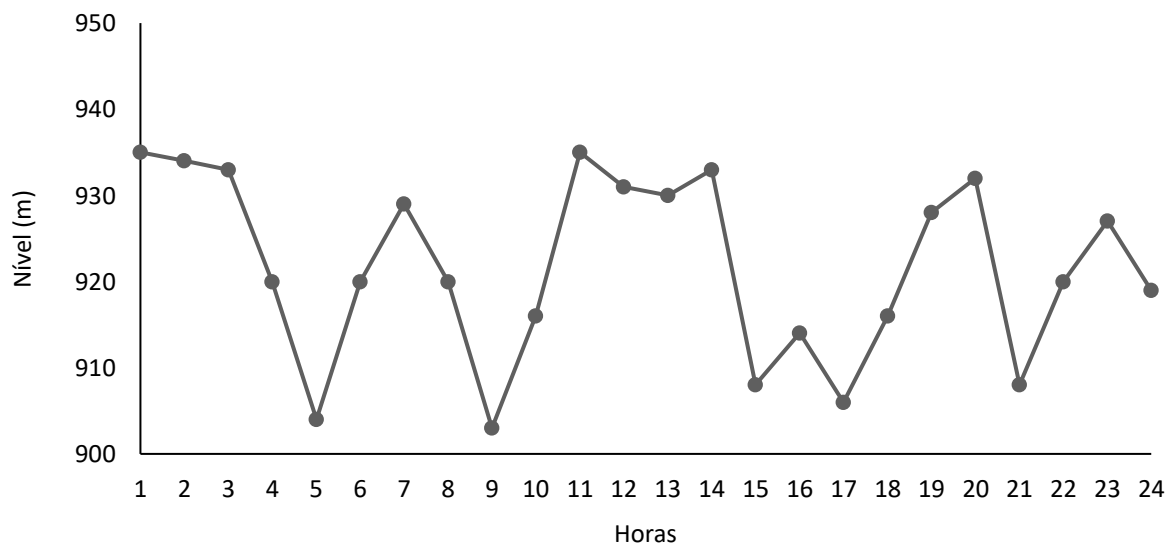
Figura 5.19 – Dados da válvula 23, obtidos pelo algoritmo genético.



Fonte: Próprio autor (2025).

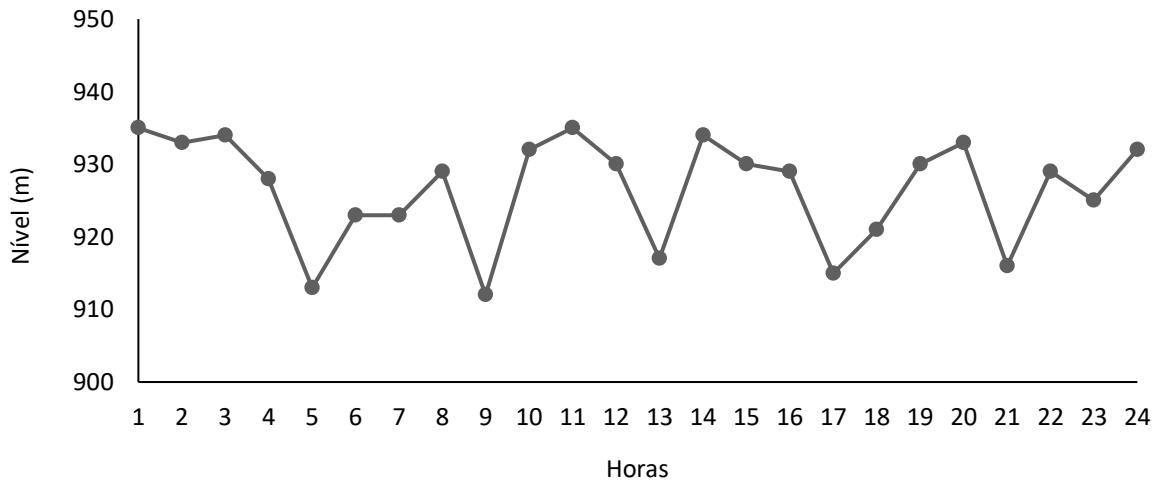
Os reservatórios assim como as válvulas possuem uma oscilação grande como mostra as Figuras 5.20 e 5.21. O intervalo de variação foi de 900 e 940 m.c.a., estabelecido para que o algoritmo pudesse buscar para encontrar a melhor configuração dos reservatórios para cada uma das horas de oscilação. Porém, o reservatório 1 apresenta uma amplitude maior que o reservatório 2, contudo ambos possuem os picos de maiores e menores valores para as mesmas horas, apresentando assim uma certa semelhança entre os dois reservatórios.

Figura 5.20 – Dados do reservatório 1, obtidos pelo algoritmo genético.



Fonte: Próprio autor (2025).

Figura 5.21 – Dados do reservatório 2, obtidos pelo algoritmo genético.

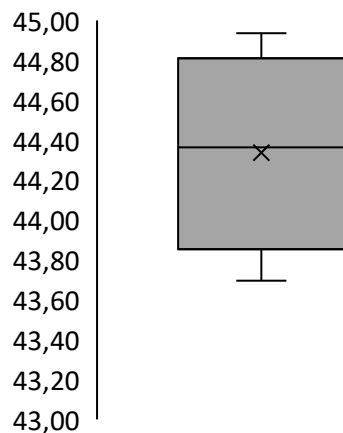


Fonte: Próprio autor (2025).

Conhecendo os dados de perdas encontrados nas etapas do trabalho, foi realizada uma análise das três perdas, original, previstas e otimizadas, por meio da elaboração dos boxplot, para identificar como se deu a distribuição dos valores das perdas e os padrões apresentados por cada uma.

Os dados das perdas originais apresentaram uma mediana de 44,36%, representando o centro dos dados de perdas, onde metade dos dados está acima e a outra metade abaixo desse valor. Os demais quartis possuem valores de 43,85% para o quartil 1 e 44,81% para o quartil 3. Isso significa que 25% dos valores de perdas estão abaixo do valor do quartil 1 e 75% dos dados estão abaixo do valor do quartil 3. Assim, tem-se que 50 % dos valores de perdas estão entre 43,85 e 44,81%. Os valores extremos, mínimos e máximos apresentados foram de 43,70 e 44,94%, correspondendo a uma amplitude total de 1,24 e a amplitude interquartil alcançou um valor de 0,96. A Figura 5.22 traz esse gráfico das perdas originais.

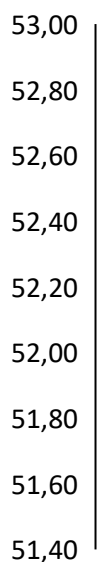
Figura 5.22 – Boxplot das perdas originais.



Fonte: Próprio autor (2025).

Nas perdas previstas a mediana (Q2) possui o valor de 52,38%, correspondendo ao ponto central dos valores das perdas, ou seja, metade dos valores estão a cima desse valor e metade está a baixo desse valor. Nos demais quartis, os valores foram de 52,29% para Q1 e 52,63% para Q3, apresentando uma amplitude de 0,34. E, com isso, é possível inferir que 50% dos dados de perdas previstas estão entre 52,92 e 52,53%. Além disso, a amplitude total foi de 0,81, pois os valores extremos foram de 51,99 e 52,80%. A Figura 5.23 apresenta os valores mencionados para os dados de perdas previstas.

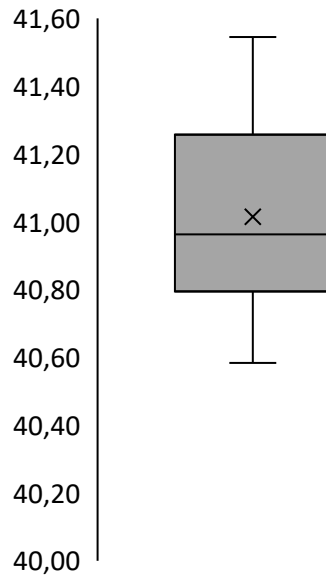
Figura 5.23 – Boxplot das perdas previstas.



Fonte: Próprio autor (2025).

E, por fim, as perdas otimizadas, o valor da mediana (Q2) foi de 40,96%, representando que 50% dos dados estão acima e 50% abaixo desse valor, e o Q1 possui um valor de 40,79% e o Q3 41,26%. Isso corresponde à distribuição da maioria dos dados de perdas otimizadas, ou seja, metade dos valores de perdas está entre os valores apresentados por Q1 e Q3. A Figura 5.24 mostra o que foi mencionado sobre os dados das perdas otimizadas.

Figura 5.24 – Boxplot das perdas otimizadas.



Fonte: Próprio autor (2025).

Dessa forma, é possível sintetizar que os dados de perdas apresentados em cada uma das etapas do estudo, as originais, previstas e otimizadas, possuem uma certa heterogeneidade nos dados, em que as perdas originais possuem a maior variabilidade entre as três perdas analisadas e, além disso, também apresentam uma leve assimetria negativa, indicando que os valores tendem a se concentrar abaixo da mediana. Com isso, tem-se uma oscilação mais intensa no comportamento das perdas, isso devido à disponibilidade maior que as demais, pois foram 48 horas de simulação utilizadas como base para o estudo.

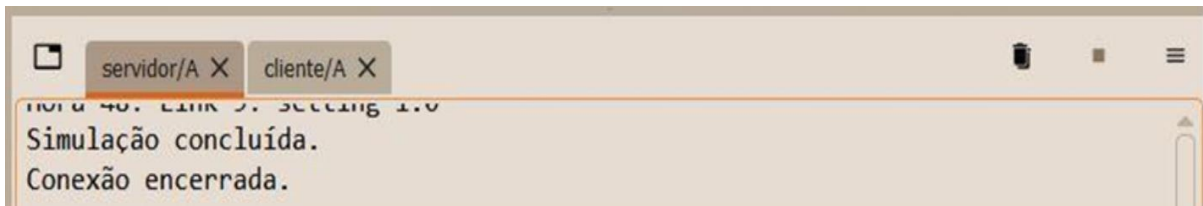
As perdas previstas foram as que apresentaram a menor amplitude total e também a menor amplitude interquartil, o que evidencia uma maior estabilidade nos valores de perdas. Também apresenta uma distribuição aproximadamente simétrica devido à posição de sua mediana. Por fim, a terceira perda analisada, a otimizada, apresentou uma variabilidade moderada, com uma leve assimetria positiva, o que sugere uma maior concentração dos valores acima da mediana.

A diferença de comportamento dos dados de perdas observada em cada etapa da pesquisa pode ser justificada pela quantidade de dados, pois, as perdas originais possuem 48 horas de perdas e as demais somente 24 horas, o que influencia na distribuição e variabilidade dos dados quando analisados pelos gráficos desenvolvidos anteriormente. Além disso, como foram realizados estudos de cenários extremos de operação da rede e, posteriormente sua otimização, esses cenários podem também causar essa diferença na variabilidade dos dados observados nas Figuras 5.22, 5.23, 5.24, das perdas originais, previstas e otimizadas, respectivamente.

Com a finalização da simulação de otimização e conhecendo os dados dos parâmetros operacionais, bombas, PRV e RNF, responsáveis por levar às reduções das perdas dos piores cenários previstos, eles devem ser enviados para o servidor para que ele armazene esses dados, bem como os cenários previstos de perdas e pressão, para que possa controlar a rede caso isso venha a ocorrer no futuro do sistema de distribuição. Assim, ao final da simulação com algoritmo genético, o cliente envia os dados para o servidor para poder ser aplicada a rede quando for necessário.

E assim a comunicação é encerrada pelo cliente e o aviso de fim da conexão é recebido pelo servidor, como mostra a Figura 5.25. Porém, essa comunicação pode ser mantida aberta se mais simulações forem necessárias.

Figura 5.25 – Mensagem recebida pelo servidor no final da comunicação.



Fonte: Próprio autor (2025).

Dessa forma, foram finalizadas todas as simulações, previsão e otimização e principalmente a finalização do protótipo de gêmeo digital adaptativo.

6. CONCLUSÕES

A pesquisa desenvolvida tem como objetivo a aplicação conjunta de três ferramentas: gêmeos digitais, rede neural recorrente LSTM e algoritmo genético para a previsão de cenários adversos de perdas e sua otimização melhorando o conhecimento da rede de distribuição de água, tornando-os mais eficientes, principalmente em tempos de escassez hídrica e mudanças climáticas.

Para isso, foi desenvolvida uma rede de benchmark para que a pesquisa pudesse ser desenvolvida, utilizando também de redes de benchmark da literatura como as redes Anytown (1987), Marques (2023) e Silva (2023), para que o treinamento do modelo de rede neural desenvolvido fosse mais generalista e robusto.

A aplicação do protótipo do gêmeo digital adaptativo foi realizada pelo protocolo de comunicação desenvolvido em Python, com um pacote da própria linguagem, o *WebSocket*. Esta é uma ferramenta de comunicação que permite a troca simultânea entre um servidor e um cliente, em que o servidor possuía os dados da rede, ou seja, como se fosse a própria rede, e o cliente, onde é criada a rede no ambiente virtual, possibilitando a aplicação de ensaios destrutivos no sistema sem que ele sofra qualquer dano, como o realizado na pesquisa de encontrar os cenários mais adversos de perdas que podem ocorrer no sistema.

Dessa forma, para realizar a previsão dos piores cenários para a rede estudada, foi utilizada a rede neural recorrente LSTM, por meio das bibliotecas do Python *Scikit-Learn*, *Keras (Tensor Flow)*. A escolha desse tipo de rede neural foi pelo fato de serem melhores em trabalhar com dados temporais e por terem duas memórias, o que melhora sua capacidade de previsão, podendo ser mais ágeis nesse tipo de atividade.

As previsões foram realizadas com base em dados de 48 de operação das redes de distribuição para a previsão dos cenários adversos que poderiam ocorrer nas 24 horas seguintes de operação do sistema. Os resultados encontrados pela rede neural empregada se mostraram satisfatório, pois encontrou perdas que variaram entre 51 e 53 %, que é um valor extremamente elevado para sistema de distribuição de água, o que representa em média um aumento de 15,44% em comparação com as perdas médias apresentadas pela rede antes das simulações realizadas.

Assim, a realização de testes destrutivos nesse tipo de sistema é viável, pois foi possível conhecer cenários adversos que podem vir a ocorrer e, com isso, desenvolver medidas de prevenção ou um plano tático para mitigar esse tipo de ocorrência.

Para a prevenção desses cenários adversos e para conhecer como reduzi-los caso venham a ocorrer, foi desenvolvido um algoritmo genético com a função de minimizar essas perdas utilizando o controle dos parâmetros operacionais da rede, como as bombas, válvulas e os reservatórios. Sendo utilizado o pacote do Python para algoritmos genéticos, o *DEAP*.

Os resultados encontrados por essa aplicação foram favoráveis, encontrando uma redução média para as horas estudadas de 11,42 p.p. Sendo que a hora 18 apresentou maior redução, com 12,02 p.p., e a menor redução foi na hora 1, com um valor de 10,71 p.p. Apesar de a redução das perdas ser significativa o algoritmo não conseguiu alcançar valores ideais de pressão para um dos nós da rede, ficando com valores de aproximadamente 2 m.c.a. acima do valor estabelecido por norma para o nó 10.

Isso pode ser justificado pelo número de gerações aplicadas ao algoritmo, que foi de 300, estabelecido com base na literatura. Porém, ocorreu um fato que também limitou esse valor: o tempo computacional, pois com esse número de gerações o algoritmo levou 6 horas para realizar a simulação completa para todas as 24 horas de simulação. Há também a questão de a inicialização do algoritmo ter sido realizada por meio de uma semente aleatória, levando-o a realizar buscas dentro de intervalos previamente estabelecidos para cada parâmetro, o que pode ter comprometido sua capacidade de encontrar valores ideais de pressão para esse nó.

Além disso, a configuração da rede de estudo pode ter interferido, pois ela é uma rede malhada, possibilitando que o fluxo de água se dê em mais de uma direção, o que pode alterar o valor de pressão de um nó, mesmo que uma válvula ou bomba esteja em uma das tubulações que chegam a esse nó.

Como o gêmeo digital possibilita a troca de comunicação de maneira simultânea entre ativos, como o apresentado pelo trabalho, servidor/cliente, os dados de perdas e as pressões dos cenários previstos, bem como os dados da configuração dos parâmetros da rede depois da otimização, foram enviados para o servidor para que a rede fosse capaz de evitar e/ou mitigar os cenários adversos por meio do controle das configurações dos parâmetros operacionais da rede.

Dessa forma, o trabalho desenvolvido mostrou que a utilização das ferramentas apresentadas de maneira conjunta permitiu conhecer melhor o sistema estudado, permitindo que a criação de cenários destrutivos pudesse ser realizada sem que o sistema sofresse com algum dano. Além disso, a otimização dos cenários adversos possibilitou o controle ao sistema, uma vez que a configuração dos parâmetros operacionais da rede tornou-se conhecida com essa atividade.

7. RECOMENDAÇÕES

Como limitação do presente estudo, destaca-se o fato de que a pesquisa foi conduzida utilizando uma rede de distribuição de água de benchmark de pequeno porte. Embora esse tipo de rede represente, em certa medida, o comportamento de sistemas reais, sua simplicidade estrutural e operacional pode não reproduzir com fidelidade a dinâmica observada em redes reais de abastecimento, especialmente no que se refere à variabilidade de consumo, à ocorrência de falhas operacionais e à complexidade topológica.

Dessa forma, recomenda-se que trabalhos futuros realizem a aplicação da metodologia proposta em sistemas reais de distribuição de água, preferencialmente de maior porte e complexidade. A inclusão de redes com maior número de nós, diferentes zonas de pressão, múltiplos reservatórios e dispositivos operacionais (como válvulas redutoras de pressão e bombas) pode proporcionar uma avaliação mais robusta das ferramentas desenvolvidas, sobretudo em cenários operacionais mais desafiadores.

Adicionalmente, sugere-se a ampliação dos parâmetros computacionais utilizados nos modelos de otimização e aprendizado. No caso das redes neurais, a utilização de um maior número de épocas de treinamento pode contribuir para a melhoria da capacidade de generalização e precisão dos modelos. De forma análoga, o aumento do número de gerações no algoritmo genético pode favorecer a obtenção de soluções melhores. Ressalta-se que tais ampliações não foram exploradas no presente estudo em virtude de limitações computacionais, o que abre espaço para investigações futuras com maior capacidade de processamento.

Outra recomendação relevante consiste na avaliação do desempenho da metodologia sob diferentes condições de incerteza, erros de medição e falhas em sensores. A incorporação de análises de sensibilidade e de cenários estocásticos pode contribuir para a compreensão da robustez do modelo proposto frente a condições adversas.

Por fim, recomenda-se a investigação de outras técnicas de inteligência artificial e otimização, como métodos baseados em aprendizado profundo, algoritmos híbridos ou abordagens multiobjetivo, visando à comparação de desempenho e à possível melhoria dos resultados obtidos na presente pesquisa.

REFERÊNCIAS

- ABADI, Martín et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems. Google Brain. <<https://www.tensorflow.org/>>
- ALVES, J. M. G. **Big Data Analytics na Gestão de uma Rede de Distribuição de Água**. 61 p. Dissertação (Mestrado em Business Analytics), Instituto Universitário de Lisboa (ISCTE), Lisboa - Portugal, 2022. Disponível em: <https://www.iscte-iul.pt/tese/13258>. Acesso em: 10 maio. 2023.
- AL-WASHALI, T. *et al.* Assessment of water losses in distribution networks: Methods, applications, uncertainties, and implications in intermittent supply. **Resources Conservation and Recycling**, v. 152, 2020. DOI: <https://doi.org/10.1016/j.resconrec.2019.104515>.
- ANDRADE SOBRINHO, R.; BORJA, P. C. Gestão das perdas de água e energia em sistema de abastecimento de água da Embasa: Um estudo dos fatores intervenientes na RMS. **Engenharia Sanitária e Ambiental**, v. 21, n. 4, p. 783–795, 2016. DOI: <https://doi.org/10.1590/S1413-41522016116037>.
- ARAÚJO, R. L.; SILVA, M. F. E. Gêmeos digitais e a Gestão de Riscos de Inundações Urbanas em Salvador: Estudo de caso em bacia de retenção. **In: XXV Simpósio Brasileiro de Recursos Hídricos-ABRHidro**, Salvador, 2023.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 12218: Projeto de rede de distribuição de água para abastecimento público**. Rio de Janeiro, 2017.
- AUGUSTIN, Aymeric et al. **Websockets 15.0.1 documentation**, 2013. <<https://websockets.readthedocs.io/en/stable/project/support.html>>
- AURÉLIEN, G. **Hands-On Machine Learning with Scikit-Learn, Keras & TensorFlow**. 2ª edição, Sebastopol, CA: O'Reilly Media, 2019.
- AYYASH, F.; ZHANG, C.; JAVADI, A. A.; FARMANI, R. Optimal Operation of Intermittent Water Supply Systems under Water Scarcity. **Journal of Water Resources Planning and Management**, v. 150, n. 3, 2024. Doi: <https://doi.org/10.1061/JWRMD5.WRENG-6227>.
- AZEVEDO NETTO, J. A.; FERNANDEZ, M. F. **Manual de Hidráulica**. 8. ed. São Paulo: Editora Edgard Blücher, 1998.
- BAI, J.; ZHU, J.; SONG, Y.; ZHAO, L.; HOU, Z.; DU, R.; LI, H. A3T-GCN: Attention Temporal Graph Convolutional Network for Traffic Forecasting. **ISPRS Int. J. Geo-Inf.** v.10, n. 485, 2021. DOI: <https://doi.org/10.3390/ijgi10070485>.
- BALEKELAYI, N.; ZERAEBRUK, K. N.; TEKLEMARIAM, M.; TESFAMARIAM, S. Optimization of Water Distribution System Operation with Multiple Tanks and Pumps: Application for Asmara, Eritrea's Water Supply System. **Journal of Pipeline Systems Engineering and Practice**, v. 12, n. 4, 2021. Doi: [https://doi.org/10.1061/\(ASCE\)PS.1949-1204.0000580](https://doi.org/10.1061/(ASCE)PS.1949-1204.0000580).

BARBEDO, M. D. G. **Estudos de calibração de redes de abastecimento de água utilizando o acoplamento dos softwares R e EPANET**. Dissertação (Mestrado em Ciências em Meio Ambiente e Recursos Hídricos). Universidade Federal de Itajubá, Itajubá-MG, 2022.

BAYILMIŞ, C; EBLEME, M. A.; ÇAVUŞOĞLU, U.; KÜÇÜK K.; SEVIN, A. A survey on communication protocols and performance evaluations for Internet of Things. **Digital Communication and Networks**, v. 8, n. 6, p. 1094-1104, 2022. DOI: <https://doi.org/10.1016/j.dcan.2022.03.013>.

BERG, C. V. Drivers of non-revenue water: A cross-national analysis. **Utilities Policy**, v. 36, pp. 71-78, 2015. Doi: <https://doi.org/10.1016/j.jup.2015.07.005>.

BONILLA, C. A.; BRENTAN, B.; MONTALVO, I.; AYALA-CABRERA, D.; IZQUIERDO, J. Assessing the Impacts of Failures on Monitoring Systems in Real-Time Data-Driven State Estimation Models Using GCN-LSTM for Water Distribution Networks. **Water**, v. 17, n. 46, 2025. DOI: <https://doi.org/10.3390/w17010046202>.

BRAHMBHATT, P.; MAHESHWARI, A.; GUDI, R. D. Digital twin assistência decision support system for quality regulation and leak localization tasks in large-scale water distribution networks. **Digital Chemical Engineering**, v. 9, 2023. DOI: <https://doi.org/10.1016/j.dche.2023.100127>.

BRASIL. Ministério da Saúde. **Fundação Nacional da Saúde. Redução de perdas em sistemas de abastecimento de água**. 2ª ed. Brasília: Funasa, 2014.

BRASIL. Ministério do Desenvolvimento Regional. Secretaria Nacional de Saneamento – SNS. Diagnóstico Temático Serviço de Água e Esgoto – Visão Geral. Brasília, 2021.

BRASIL. CONFEDERAÇÃO NACIONAL DOS MUNICÍPIOS. **Brasil é um dos 20 países que mais perde água tratada**. Ag. Senado, Brasília - DF, 2021.

BRASIL. Ministério do Desenvolvimento Regional. Secretaria Nacional de Saneamento – SNS. Diagnóstico Temático Serviço de Água e Esgoto – Visão Geral, ano de referência 2022. Brasília, 2023.

BRASIL. Secretaria Nacional de Saneamento e Ambiental, Ministério das Cidades. **Sistema Nacional de Informações em Saneamento Básico - Relatório dos Serviços de Abastecimento de Água**. Brasília, 2024.

CAMPESATO, O. Artificial Intelligence, Machine Learning, and Deep Learning. **Mercury Learning and Information**, 2020.

CARIÇO, N.; RAMOS, H. M. **REDES INTELIGENTES E MACHINE LEARNING NA GESTÃO EFICIENTE DOS SISTEMAS DE DISTRIBUIÇÃO**. XVII Congresso Iberoamericano sobre Sistemas de Abastecimiento, Saneamiento y Riego - SEREA23, 2023.

CHAGAS, E. M. P. F.; RODRIGUES, D. L.; TAVARES, J. M. R. S. Método de segmentação de objectos em imagens baseado em contornos activos e algoritmo genético. **In: Congresso de Métodos Numéricos en Ingeniería**, n. May, p. 491–507, 2014.

CHAVES, B. B. **ESTUDO DO ALGORITMO ADABOOST DE APRENDIZAGEM DE MÁQUINA APLICADO A SENSORES E SISTEMAS EMBARCADOS**. Dissertação (Mestrado em Engenharia Mecânica) Escola Politécnica da Universidade de São Paulo, 2012.

CHEN, W. *et al.* Genetic optimization toward operation of water intake-supply pump stations system. **Journal of Cleaner Production**, v. 279, january 2021. DOI: <https://doi.org/10.1016/j.jclepro.2020.123573>.

CHENG, M.; LI, J.; WANG, C.; YE, C.; CHANG, Z. Graph Laplace Regularization-based pressure sensor placement strategy for leak localization in the water distribution networks under joint hydraulic and topological feature spaces. **Water Research**, v. 257, 2024. DOI: <https://doi.org/10.1016/j.watres.2024.121666>.

CHINI, C. M.; STILLWELL, A. S. The State of U.S. Urban Water: Data and the Energy-Water Nexus. **Water Resources Research**, v. 54, n. 3, p. 1796–1811, 2018. DOI: <https://doi.org/10.1002/2017WR022265>.

CHOLLET, François et al. Keras: Deep Learning for Humans. <<https://keras.io/>>

CIMINO, C.; NEGRI, E.; FUMAGALLI, L. Review of digital twin applications in manufacturing. **Computers in Industry**, v. 113, 2019. DOI: <https://doi.org/10.1016/j.compind.2019.103130>.

CONTARINI, G. **REDES NEURAIAS LSTM E MODELO GARCH: UMA ABORDAGEM CONJUNTA PARA PREVISÃO DE RETORNOS**. Monografia (Bacharelado em Economia) Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2020. Disponível em: <https://pantheon.ufrj.br/bitstream/11422/14800/1/GContarini.pdf>. Acesso em: 11 mar. 2024.

CORREIA, V. P.; DANTAS, L.; GUEDES, L. A.; FERNADES, M. **ANALISE DE DESEMPENHO DE REDES NEURAIAS LSTM COM TECNICAS DE PRUNING PARA DETECÇÃO DE FALHAS EM PROCESSOS INDUSTRIAS**, XV Simpósio Brasileiro de Automação Inteligente - SBAI 2021. DOI: 10.20906/sbai.v1i1.2602

COSTA, L. H. M.; FROTA, A. F. **Tutorial da ferramenta toolkit do Epanet para programadores**. In: SIMPÓSIO ÍTALO BRASILEIRO DE ENGENHARIA SANITÁRIA E AMBIENTAL- SIBESA, XIV., 18 a 20 jun. 2018, Foz do Iguaçu/PR. Anais [...], Foz do Iguaçu/PR, 2018.

COURNAPEAU, David et al. scikit-learn: Machine Learning in Python. <<https://scikit-learn.org/>>

DEPARTAMENTO DE ÁGUA E ESGOTO DE JUNDIAÍ. **Controle de Perdas**. CIJUN, Jundiaí - SP, 2024. <<https://daejundiai.com.br/indicadores-dae-jundiai/controle-de-perdas/>>

DALARMELINA, N. V. **Uma abordagem Ensemble Learning para modelos de detecção de intrusão para redes industriais**. Dissertação (Mestrado em Ciências de Computação e Matemática Computacional), Universidade de São Paulo, São Carlos - SP, 2023. DOI: <https://doi.org/10.11606/D.55.2023.tde-12012024-154719>.

DANESHGAR, S.; POLESEL, F.; BORZOOEI, S. SORENSSEN, H.; PEETRS, R.; WEIJERS, S. NOPENS, I.; TORFS, E. A full-scale operational digital twin for a water resource recovery facility—A case study of Eindhoven Water Resource Recovery Facility. **Water Environment Research**, v. 96, n. 3, 2024. Doi: <https://doi.org/10.1002/wer.11016>.

DANGETI, P. Techniques for exploring supervised, unsupervised, and reinforcement learning models with Python and R. **Packt Publishing Ltd.** Birmingham-UK, 2017. Disponível em: <https://pdfcoffee.com/statistics-for-machine-learning-pdf-free.html>. Acesso em: 20 out. 2024.

DATA SCIENCE ACADEMY. **Deep Learning Book Brasil**. Disponível em: <https://www.datascienceacademy.com.br>

DRAGICEVIC, A.; LOBIANCO, A.; LEBLOIS, A. Forest planning and productivity-risk trade-off through the Markowitz mean-variance model. **Forest Policy and Economics**, v. 64, pp. 25-34, 2016. Doi: <https://doi.org/10.1016/j.forpol.2015.12.010>.

DUI, H.; CAO, T.; WANG, F. Digital twin-based resilience evaluation and intelligent strategies of smart urban water distribution networks for emergency management. **Resilient Cities and Structures**, v. 4, n.1, pp. 41-52, 2025. Doi: <https://doi.org/10.1016/j.rcns.2025.02.001>.

EGITO, T. B. **OTIMIZAÇÃO DE SISTEMAS DE DISTRIBUIÇÃO DE ÁGUA COM ÊNFASE NA EFICIÊNCIA ENERGÉTICA DE BOMBAS E OPERAÇÃO DE RESERVATÓRIOS**. Tese (Doutorado em Engenharia Civil) Universidade Federal de Pernambuco, Recife - PE, 2022. Disponível em: https://bdtd.ibict.br/vufind/Record/UFPE_cb0c75c23b47da8abe734a072ab1af20. Acesso em: 16 jun. 2024.

EVANGELISTA, L. A. **Avaliação da eficiência de redes de distribuição de água ao longo da sua vida útil**. Dissertação (Mestrado em Saneamento, Meio Ambiente e Recursos Hídricos), Universidade Federal de Minas Gerais, Belo Horizonte-MG, 2022. Disponível em: <https://repositorio.ufmg.br/server/api/core/bitstreams/4b35b153-4a35-40b1-908d-5898d802b084/content>. Acesso em: 20 out. 2025.

FARGHADAN, A.; ZAMANI, M. S.; GHAZIZADEH, M. J. A fault-injection-based approach to leak localization in water distribution networks using an ensemble model of Bayesian classifiers. **Journal of Process Control**, v. 132, 2023. DOI: <https://doi.org/10.1016/j.jprocont.2023.103110>.

FERREIRA, A. B. Avaliação de Operadores de Algoritmo Genéticos em Otimização Multidimensional. Dissertação (Mestrado em Engenharia Mecânica), Universidade Estadual Paulista – UNESP, Faculdade de Engenharia de Ilha Solteira, p. 199, 2007.
FAWCETT, T. Introduction to ROC analysis. **Pattern Recognition Letters**, v. 27, p. 861-874, 2006. DOI: <http://dx.doi.org/10.1016/j.patrec.2005.10.010>.

FETTE, I.; MELNIKOV, A. **The WebSocket Protocol**. RFC 6455. Internet Engineering Task Force (IETF), dez. 2011.

FOGEL, D.B. An introduction to simulated evolutionary optimization. **IEEE Transactions on Neural Networks**, v. 5, n.1, p. 3-14, 1994. DOI: 10.1109/72.265956.

FONTANA, N.; GIUGNI, M.; PORTOLANO, D. Losses reduction and energy production in water-distribution networks. **Journal of Water Resources Planning and Management Reston**, v. 138, n. 3, p. 237–244, 2012. DOI: [https://doi.org/10.1061/\(ASCE\)WR.1943-](https://doi.org/10.1061/(ASCE)WR.1943-)

FORMIGA, K. T. M. **OTIMIZAÇÃO MULTI OBJETIVO DE PROJETOS DE REDES DE DISTRIBUIÇÃO DE ÁGUA**. Tese (Doutorado em Hidráulica e Saneamento), Universidade de São Paulo, Escola de Engenharia de São Carlos, São Carlos - SP, 2005. DOI: <https://doi.org/10.11606/T.18.2016.tde-29012016-125410>.

FORTIN, F. A.; RAINVILLE, F. M.; GARDNER, M. A.; PARIZEAU, M.; GAGNÉ, C. DEAP: Evolutionary Algorithms Made Easy. **Journal of Machine Learning Research**, v. 13, p. 2171-2175, 2012. DOI: https://www.researchgate.net/publication/235707001_DEAP_Evolutionary_algorithms_made_easy.

FU, M.; ZHANG, Q.; RONG, K.; YASEEN, Z. M.; ZHENG, L.; ZHENG, J. Integrated dynamic multi-threshold pattern recognition with graph attention long short-term neural memory network for water distribution network losses prediction: An automated expert system. **Engineering Applications of Artificial Intelligence**, v. 127, 2024. DOI: <https://doi.org/10.1016/j.engappai.2023.107277>

FUERTES, P. C.; ALZAMORA F. M.; CAROT, M. H.; CAMPOS, J. C. A. Building and exploiting a Digital Twin for the management of drinking water distribution. **Urban Water Journal**, v. 17, n.8, 2020. DOI: <https://doi.org/10.1080/1573062X.2020.1771382>.

GALVÃO, C. O. et al. **Sistemas inteligentes - aplicações a recursos hídricos e ciências ambientais**. Associação Brasileira de Recursos Hídricos. Editora da Universidade Federal do Rio Grande do Sul, Volume 7, Porto Alegre, 1999.

GEMANOPOULOS, G.; JOWITT, P. W. Leakage reduction by excess pressure minimization in a water supply network. **Proceedings of the Institution of Civil Engineers**, v. 87, n. 2, p. 195-214, 1989. DOI: <https://doi.org/10.1680/iicep.1989.2003>.

GOBBO, F. M.; NOGUEIRA, A. A.; SILVA, A. M. M.; PAIVA, P. S.; ISMAIL, I. A. L.; AGNESINI, M. V. Automatização de Válvula Redutora de Pressão (VRP) em sistema de abastecimento de água. **Contribuciones a Las Ciencias Sociales**, São José dos Pinhais, v.17, n.1, p.3440-3456, 2024. DOI: 10.55905/revconv.17n.1-204.

GODASE, V.; MULANI, A.; GHODAKE, R., TAKALE, S. PLC-Assisted Smart Water Distribution with Rapid Leakage Detection and Isolation. **Journal of Control Systems and Converters**, v. 1, n. 3, pp. 1-13, 2025. Doi: <http://dx.doi.org/10.2139/ssrn.5417135>.

GOMES, H. P. **Sistemas de Abastecimento de Água: Dimensionamento Econômico e Operação de Redes e Elevatórias**. 3. ed. João Pessoa: Editora Universitária/UFPB, 2009. 277 p.

GÓMEZ-CORONEL, L.; SANTOS-RUIZ, I.; BLESÁ, J.; PUIG, V.; LÓPEZ-ESTRADA, F. R. Leak localization in an urban water distribution network using a LSTM deep neural network, **IFAC-PapersOnLine**, v. 58, n. 4, pp. 79-84, 2024. DOI: <https://doi.org/10.1016/j.ifacol.2024.07.197>.

GOMEZ-CORONEL, L.; BLESÁ, J.; SANTOS-RUIZ, I.; LÓPEZ-ESTRADA F. R.; PUIG, V. Neural network-based leak localization in water distribution networks using the gravity center of pressure measurements. **Journal of Water Process Engineering**, v. 77, 2025. DOI: <https://doi.org/10.1016/j.jwpe.2025.108348>.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. Deep learning. **MIT press**, 2016. DOI: https://www.deeplearningbook.org/front_matter.pdf.

GOUVEIA, C. G. N. **TÉCNICAS DE APRENDIZADO DE MÁQUINA APLICADAS À PREVISÃO DE VAZAMENTO EM RAMAIS DE REDES DE DISTRIBUIÇÃO DE ÁGUA**. Dissertação (Mestrado em tecnologia Ambiental e Recursos Hídricos), Universidade de Brasília, 2022. Disponível em: <http://repositorio.unb.br/handle/10482/43766>. Acesso em: 20 out. 2024.

GRIEVES, M. Digital twin: manufacturing excellence through virtual factory replication. **White paper**, v. 1, p. 1-7, 2014. Disponível em: https://www.researchgate.net/publication/275211047_Digital_Twin_Manufacturing_Excellence_through_Virtual_Factory_Replication. Acesso em: 13 out. 2023.

GRIEVES, M.; VICKERS, J. Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems. In *Transdisciplinary Perspectives on Complex Systems*. **Springer, Cham Switzerland**, p. 85–113, 2017. DOI: https://doi.org/10.1007/978-3-319-38756-7_4.

GUANAIS, A. L. R.; COHIM, E. B.; MEDEIROS, D. L. **Avaliação energética de um sistema integrado de abastecimento de água**. **Artigo Técnico**. 2016. DOI: 10.1590/S1413-41522017146180.

GÜMÜS, D.; ALVER, A.; AKBAL, F. Development of a digital twin framework for hybrid adsorption-ultrafiltration systems in drinking water treatment. **Process Safety and Environmental Protection**, v. 205, 2026. Doi: <https://doi.org/10.1016/j.psep.2025.108182>.

HELLER, L; PÁDUA, V. L. **Abastecimento de água para o consumo humano**. 3. ed. Belo Horizonte: Editora UFMG, 2010.

HERNADEZ-SAMANIEGO, E.; NAVARRO-GOMEZ, C.; SÁNCHEZ, D. H.; SÁNCHEZ-NAVARRO, J. R. Coefficients and curves of hourly and daily variations of water demand for improved operation of potable water distribution systems: a case study of Chihuahua City, Mexico. **Water Practice and Technology**, v. 18, n. 8, p. 1991-2001, 2023. DOI: <https://doi.org/10.2166/wpt.2023.117>.

HIENINGER, T.; GOPPELT, F.; SCHMIDT-VOLLUS, R.; SCHLUCKER, E. Energy-saving potential for centrifugal pump storage operation using optimized control schemes. **Energy Efficiency**, v. 14, n. 23, 2021. Doi: <https://doi.org/10.1007/s12053-021-09932-5>.

HIRNER, W.; LAMBERT, A. Losses from water supply systems: Standart terminology and recommended performance measures. **International Water Association**, p. 13, London, 2000. Disponível em: <https://www.ircwash.org/resources/losses-water-supply-systems-standard-terminology-and-recommended-performance-measures>. Acesso em: 10 abr. 2023.

HOCHREITER, S.; SCHMIDHUBER, J. Long Short-Term Memory. **Neural Computation**, v. 9, n. 8, p. 1735–1780, 1997. DOI: 10.1162/neco.1997.9.8.1735

HOSSAIN, S.; HEWA, G.A.; CHOW, C.W.K.; COOK, D. Modelling and Incorporating the Variable Demand Patterns to the Calibration of Water Distribution System Hydraulic Model. **Water**, v. 13, n. 20, 2021. DOI: <https://doi.org/10.3390/w13202890>. <https://www.sciencedirect.com/science/article/pii/S2405896324002817>.

HOLLAND, J. H. **Adaptation in Natural and Artificial Systems**. Ann Arbor: University of Michigan Press, 1975. Disponível em: http://repo.darmajaya.ac.id/3794/1/Adaptation%20in%20Natural%20and%20Artificial%20Systems_%20An%20Introductory%20Analysis%20with%20Applications%20to%20Biology%20C%20Control%20and%20Artificial%20Intelligence%20%28A%20Bradford%20Book%29%20%28%20PDFDrive%20%29.pdf. Acesso em: 20 out. 2024.

ISLAM, M. R.; AZAM, S.; SHANMUGAM, B.; MATHUR, D. A Review on Current Technologies and Future Direction of Water Leakage Detection in Water Distribution Network. **IEEE Access**, v. 10, pp. 107177-107201, 2022. Doi: 10.1109/ACCESS.2022.3212769.

JANIESCH, C.; ZSCHECH, P.; HEINRICH, K. Machine learning and deep learning. **Electronic Markets**, v. 31, p. 685-695, 2021. DOI: <https://doi.org/10.1007/s12525-021-00475-2>.

JOAQUIM, E. A. A; PIQUEIRO, F. O.; KAI, J. P. DESENVOLVIMENTO DE SISTEMAS DE GESTÃO DE INFRAESTRUTURAS PARA ÁGUAS PLUVIAIS EM ZONAS URBANAS ATRAVÉS DA UTILIZAÇÃO DE GÊMEOS DIGITAIS. **Revista Aracê**, v. 7, n.6, pp. 31047-31068, 2025. Doi: <https://doi.org/10.56238/arev7n6-115>.

KANAKOUDIS, V.; MUHAMMETOGLU, H. Urban water pipe networks management towards non-revenue water reduction: two case studies from Greece and Turkey. **Clean - Soil Air Water**, v. 42, n. 7, p. 880–892, 2014. DOI: <http://dx.doi.org/10.1002/clen.201300138>.

KERAS. **Dropout layer**, 2021. <https://keras.io/api/layers/regularization_layers/dropout/>.

KINGMA, D. P.; BA, J. L. **ADAM: A METHOD FOR STOCHASTIC OPTIMIZATION**. Published as a conference paper at the 3rd International Conference for Learning Representations, San Diego, 2015. DOI: <https://arxiv.org/abs/1412.6980>.

KLISE, K. A.; HART, D. B.; BYNUM, M.; HOGGE, J.; HAXTON, T.; MURRAY, R.; BURKHARDT, J. Water Network Tool for Resilience (WNTR) User Manual: Version 0.2.3. U.S. EPA Office of Research and Development, p. 82, Washington-DC, 2020.

KRITZINGER, W.; KARNER, M.; TRAAR, G.; HENJES, J.; SIHN, W. Digital Twin in manufacturing: A categorical literature review and classification. **IFAC-PapersOnLine**, v.51, n.11, p. 1016-1022, 2018. DOI: <https://doi.org/10.1016/j.ifacol.2018.08.474>.

KURIAN, V. *et al.* Optimal operation of water distribution networks with intermediate storage facilities. **Computers and Chemical Engineering**, v. 119, p. 215–227, 2018. DOI: <https://doi.org/10.1016/j.compchemeng.2018.04.017>.

KYRIAKOU, M. S.; DEMETRIADES, M.; VRACHIMIS, S. G.; ELIADES, D. G.; POLYCARPOU, M. M. EPyT: An EPANET-Python Toolkit for Smart Water Networks Simulations. **The Journal of Open Source Software**, v. 8, n. 92, 2023. DOI: <https://doi.org/10.21105/joss.05947>

LACERDA, E. G. M.; CARVALHO, A. C. P. L. F. **Introdução aos Algoritmos Genéticos - Sistemas Inteligentes - Aplicações a Recursos Hídricos e Ciências Ambientais**. 1. ed. Porto Alegre: Editora da Universidade Federal do Rio Grande do Sul, 1999.

LARA, L. L. D.; SILVA, F. G. B.; SILVA, A. T. Y. L.; MARQUES, S. M.; BARBEDO, M. D. G.; REIS, J. A. T. “Proposal of a methodology for adjusting hydraulic parameters and pressures using variations of genetic algorithm operators for optimization and reduction of losses in water distribution networks”. **Engenharia Sanitária Ambiental**, v. 29, p. 1-9, 2024. DOI: <https://doi.org/10.1590/S1413-415220240032>.

LE, X. H.; HO, H.V.; LEE, G.; JUNG, S. Application of Long Short-Term Memory (LSTM) Neural Network for Flood Forecasting. **Water**, v. 11, n. 7, 2019. DOI: <https://doi.org/10.3390/w11071387>

LEE, C.W.; YOO, D.G. Development of Leakage Detection Model and Its Application for Water Distribution Networks Using RNN-LSTM. **Sustainability**, v. 13, 2021. DOI: <https://doi.org/10.3390/su13169262>.

LI, J.; ZHENG, W.; WANG, C.; CHENG, M. Optimal sensor placement for leak location in water distribution networks based on EGAE clustering algorithm. **Journal of Cleaner Production**, v. 426, 2023. DOI: <https://doi.org/10.1016/j.jclepro.2023.139175>.

LI, X.; LUO, J.; LI, Y.; WANG, W.; HONG, W.; LIU, M.; LI, X.; LV, Z. Application of effective water-energy management based on digital twins technology in sustainable cities construction. **Sustainable Cities and Society**, v. 87, 2022. Doi: <https://doi.org/10.1016/j.scs.2022.104241>.

MA, Z.; ZHU, Y.; CHEN, C.; LI, T.; LI, Y.; LI, X.; WANG, Y.; WAITE, T. D.; GUAN, J. Towards the digitalization of water treatment facilities: A case study on machine learning-enabled digital twins. **Journal of Water Process Engineering**, v. 77, 2025. Doi: <https://doi.org/10.1016/j.jwpe.2025.108316>.

MADNI, A.M.; MADNI, C.C.; LUCERO, S.D. Leveraging digital twin technology in model-based systems engineering. **Systems**, v.7, e.1, 2019. DOI: <https://doi.org/10.3390/systems7010007>.

MAHDI, N. M. et al. Leak detection and localization in water distribution systems using advanced feature analysis and an Artificial Neural Network. **Desalination and Water Treatment**, 2024. DOI: [10.1016/j.dwt.2024.100685](https://doi.org/10.1016/j.dwt.2024.100685)

MAKISHA, N; KAZIMIROVA, T. **Principles of energy saving in water supply and sewage systems**. In: INTERNATIONAL CONFERENCE ON RESEARCH IN MECHANICAL ENGINEERING SCIENCES, 2017, Manipal. **Proceedings**, p. 1-6, 2018. DOI: <https://doi.org/10.1051/matecconf/201814404013>.

MARC, T. J. J.; ZAIN, N. The improbability of detecting trade-offs and some practical solutions, **Journal of Evolutionary Biology**, v. 37, n. 10, pp. 1205–1214, 2024. Doi: <https://doi.org/10.1093/jeb/voae096>

MARQUES, S. M. **Aplicação de algoritmos genéticos em operação de redes de distribuição de água com o uso de softwares R e EPANET visando ao controle de perdas de água – estudo em uma rede teórica**. Dissertação (Mestrado em Meio Ambiente e Recursos Hídricos), Universidade Federal de Itajubá, Itajubá-MG, 2023. Disponível em: . Acesso em: 25 abr. 2023.

MARQUES, S. M.; SILVA, F. G. B.; SILVA, A. T. Y. L.; BARBEDO, M. D. G.; MARCONDES, M. C.; ALVES, S. C. R.; REIS, J. A. T. Evaluation of hydraulic behavior of water distribution network varying reservoirs levels, roughness, and diameters with the use of R and EPANET. **Revista Ambiente & Água**, V. 18, 2023. DOI: <https://doi.org/10.4136/ambi-agua.2893>.

MARSLAND, S. **MACHINE LEARNING: An Algorithm Perspective**. Second Edition. [s.i]: Chapman & Hall/CRC, [s.d]. Disponível em: https://doc.lagout.org/science/Artificial%20Intelligence/Machine%20learning/Machine%20Learning_%20An%20Algorithmic%20Perspective%20%282nd%20ed.%29%20%5BMarsland%202014-10-08%5D.pdf. Acesso em: 20 out. 2024.

MARTÍNEZ-BAHENA, B.; CRUZ-CHÁVEZ, M.A.; ÁVILA-MELGAR, E.Y.; CRUZ-ROSALES, M.H.; RIVERA-LOPEZ, R. Using a Genetic Algorithm with a Mathematical Programming Solver to Optimize a Real Water Distribution System. **Water**, v. 10, n. 1318, 2018. DOI: <https://doi.org/10.3390/w10101318>.

MARTINHO, W. C. S. **An improved simulation-based iterated local search metaheuristic for gravity fed water distribution network design optimization**. Dissertação (Mestrado em Ciência da Computação) Universidade Federal da Bahia. Salvador-BA, 2020. DOI: <https://repositoriohml.ufba.br/handle/ri/33639>.

MARTINHO, W. C. S.; MELO, R. A.; SÖRENSEN, K. An enhanced simulation-based iterated local search metaheuristic for gravity fed water distribution network design optimization. **Computers and Operations Research**, v. 135, 2021. DOI: <https://doi.org/10.1016/j.cor.2021.105429>.

MASSAGO, S. **Introdução ao Algoritmo Genético**. Universidade Federal de São Carlos, 2013.

MICROSOFT. **Netstat and Taskkill commands**. Redmond: Microsoft, 2016. <<https://docs.microsoft.com/en-us/windows-server/administration/windows-commands/netstat>> e <<https://docs.microsoft.com/en-us/windows-server/administration/windows-commands/taskkill>>

MOHRI, M.; ROSTAMIZADEH, A.; TALWALKAR, A. **Foundations of Machine Learning**. MIT Press, Massachusetts Institute of Technology, 2012. Disponível em: https://www.hlevkin.com/hlevkin/45MachineDeepLearning/ML/Foundations_of_Machine_Learning.pdf. Acesso em: 20 out. 2024.

MÓLLER, D. S. **LOCALIZAÇÃO OTIMIZADA DE ESTAÇÕES ELEVATÓRIAS DE REFORÇO (BOOSTERS) EM REDES DE DISTRIBUIÇÃO DE ÁGUA**. Dissertação (Mestrado em Saneamento, Meio Ambiente e Recursos Hídricos), Universidade Federal de Minas Gerais - UFMG, Belo Horizonte - MG, 2021. Disponível em: <https://hdl.handle.net/1843/55747>. Acesso em: 19 fev. 2025.

MÜCKE, N. T.; PANDEY, P.; JAIN, S.; BOHTÉ, S. M.; OOSTERLEE, C. W. A Probabilistic Digital Twin for Leak Localization in Water Distribution Networks Using Generative Deep Learning. **Sensors**, v.23, n. 6179, 2023, 23. Doi: <https://doi.org/10.3390/s23136179>.

MUTIKANGA, H. E.; SHARMA, S. K.; VAIRAVAMOORTHY, K. Methods and tools for managing losses in water distribution systems. **Journal of Water Resources Planning and Management, Reston**, v. 139, n. 2, p. 166–174, 2013. DOI: 10.1061/(ASCE)WR.1943-5452.0000245.

MYKONIATIS, K. & HARRIS, G.A. A digital twin emulator of a modular production system using a data-driven hybrid modeling and simulation approach. **Journal of Intelligent Manufacturing**, v. 32, p. 1899-1911, 2021. DOI: <https://doi.org/10.1007/s10845-020-01724-5>.

NASCIMENTO, C. M.; SILVA NETTO, A. V.; VILLANUEVA, J. M. M.; MACÊDO, E. C. T. **Metodologia Aplicada ao Desenvolvimento de Gêmeos Digitais de Sistemas de Bombeamento de Água**. Sociedade Brasileira de Automática (SBA) - XXIV Congresso Brasileiro de Automática (CBA), 2022. Fortaleza - CE. DOI: 10.20906/CBA2022/3463.

NASCIMENTO, C. M. **GÊMEO DIGITAL DE UM SISTEMA DE BOMBEAMENTO DE ÁGUA**. Dissertação (Mestrado em Engenharia Elétrica) Universidade Federal da Paraíba, João Pessoa-PB, 2022. Disponível em: <https://repositorio.ufpb.br/jspui/handle/123456789/25126> Acesso em: 21 abr. 2024.

NETO, R. E. P. **Avaliação da codificação real e binária na identificação das rugosidades de tubulações aplicando algoritmo genético com o método transiente inverso**. Dissertação (Mestrado em Engenharia Civil) Universidade Federal do Ceará, Fortaleza-CE, 2021. 90p. Disponível em: <http://repositorio.ufc.br/handle/riufc/58450>. Acesso em: 15 jun. 2025.

OLIVEIRA, A.; GOMES, A.; BAPTISTA, R. MODELO PREDITIVO DE FUGAS NO SISTEMA DE DISTRIBUIÇÃO MUNICIPAL DE ÁGUA DE VILA NOVA DE GAIA. **XIII CONGRESSO DA GEOGRAFIA PORTUGUESA**, p.174-178, Coimbra - Portugal, 2021. Disponível em: https://www.apgeo.pt/sites/default/files/xiiicgp_livroatas_15062022.pdf. Acesso em: 23 jun. 2023.

OLIVEIRA, A. L. M. B. **Estimativa da Rugosidade em Sistema Abastecimento de Água utilizando Algoritmo Genético e Gêmeo Digital**. Dissertação (Mestrado em Engenharia

Mecânica) Universidade Federal da Paraíba, João Pessoa-PB, 2022. Disponível em: <https://repositorio.ufpb.br/jspui/handle/123456789/26111>. Acesso em: 13 mar. 2024.

OLIVEIRA, G.; MARCATO, F. S.; SCAZUFCA, P.; MARGULIES, B. N. **PERDAS DE ÁGUA 2020 (SNIS 2018): DESAFIOS PARA DISPONIBILIZADADE HÍDRICA E AVANÇO DA EFICIÊNCIA DO SANEAMENTO BÁSICO**. GO Associados, São Paulo, 2020. Disponível em: https://tratabrasil.org.br/wp-content/uploads/2022/09/Relatorio_Final_-_Estudo_de_Perdas_2020_-_JUNHO_2020.pdf. Acesso em: 23 jun. 2023.

OLIVEIRA, G.; SCAZUFCA, P.; SAYON, P. L. **ESTUDO DE PERDAS DE ÁGUA DO INSTITUTO TRATA BRASIL DE 2023 - SNIS (2021): DESAFIOS PARA A DISPONIBILIDADE HÍDRICA E AVANÇO DA EFICIÊNCIA DO SANEAMENTO BÁSICO NO BRASIL**. GO Associados, São Paulo, 2023. Disponível em: https://tratabrasil.org.br/wp-content/uploads/2023/06/Estudo-de-Perdas-de-Agua_2023.pdf. Acesso em: 23 jun. 2023.

ORTIZ, J. D.; ALMARAZ, H. B.; GOMEZ, R. O.; CAPETILLO, C. B. Water distribution system calibration using the Finite Element Method coupled to a Genetic Algorithm. **Journal of Water Process Engineering**, v. 54, 2023. DOI: <https://doi.org/10.1016/j.jwpe.2023.104017>.

OUGUI, J. Y. **Estudo da operação otimizada de sistema de bombeamento de água**. Dissertação (Mestrado em Hidráulica e Saneamento) – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2003. Disponível em: <http://repositorio.eesc.usp.br/handle/RIEESC/5348>. Acesso em: 22 out. 2023.

PARK, J. Y.; KIM, K. J.; JEUNG, M.; JANG, I. S.; YU, J. W.; KANG, M. S.; BAE, H. S.; JEONG, C.; BAEK, S. S. Digital-twin tool for a drinking water distribution system using augmented reality and EPANET. **Environmental Modelling & Software**, v. 197, 2026. Doi: <https://doi.org/10.1016/j.envsoft.2025.106829>.

PARVAZE, S.; KUMAR, R.; KHAN, J.N.; AL-ANSARI, N.; PARVAZE, S.; VISHWAKARMA, D. K.; ELBELTAGI, A.; KURIQI, A. Optimization of Water Distribution Systems Using Genetic Algorithms: A Review. **Archives of Computational Methods in Engineering**, v. 30, pp. 4209–4244, 2023. Doi: <https://doi.org/10.1007/s11831-023-09944-7>.

PATHIRANE, A.; KAZAMA, S.; TAKIZAWA, S. Dynamic analysis of non-revenue water in district metered areas under varying water consumption conditions owing to COVID-19. **Heliyon**, v. 10, n. 1, 2024. Doi: <https://doi.org/10.1016/j.heliyon.2023.e23516>.

PAVANSIMHA, M.N.; YUSUF, J. Optimization of water distribution network by genetic algorithm with EPANET-Python Toolkit. **Journal of Earth System Science**, v. 134, n. 205, 2025. Doi: <https://doi.org/10.1007/s12040-025-02651-w>.

PERSANTEZ, J. E.; ALGHAMDI, F.; SABU, S.; MAHINTHAKUMAR, G.; BERGLUND, E. Z. Using a digital twin to explore water infrastructure impacts during the COVID-19 pandemic. **Sustainable Cities and Society**, v. 77, 2022. DOI: <https://doi.org/10.1016/j.scs.2021.103520>.

PHILIP, S.; MARAKKATH, N.; GANASEN, N. Soft computing techniques for predicting the compressive strength properties of fly ash geopolymer concrete using regression-based machine learning approaches. **Journal of Building Pathology and Rehabilitation**, 2024. DOI: 10.1007/s41024-024-00461-z.

PIRES, C. A. G. **COMPARAÇÃO DE MODELOS DE PREVISÃO DE SÉRIES TEMPORAIS: Aplicação a Caudais de Água Distribuída em Rede de Abastecimento**. Dissertação (Mestrado em Informática) Politécnico de Portalegre - Escola Superior de Tecnologia e Gestão, Porto Alegre - RS, 2024. Disponível em: <https://comum.rcaap.pt/bitstream/10400.26/52511/1/BCTFC427.pdf>. Acesso em: 10 mar. 2025.

POMPEU, M. L. F.; NEPOMUCENO, N. **OTIMIZAÇÃO ENERGÉTICA DA OPERAÇÃO DE BOMBAS EM REDES DE DISTRIBUIÇÃO DE ÁGUA USANDO UM GÊMEO DIGITAL NO APOIO À DECISÃO**. In: ANAIS DO SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL, 2023, São José dos Campos. Anais eletrônicos. Campinas, Galoá, 2023.

PORTO, R. de M. **Hidráulica Básica**. 4. ed. São Carlos: EESC-USP, 2006. Disponível em: <https://pt.scribd.com/document/401794712/Livro-Hidraulica-Basica-Rodrigo-Porto-4%C2%AA-Edicao-pdf>. Acesso em: 05 ago. 2023.

RAMOS, H. M.; MORANI, M. C.; CARAVETTA, A.; FECARROTTA, O.; ADEYEYE, K.; LÓPEZ-JIMÉNEZ, A.; PÉREZ-SANCHEZ, M. New Challenges towards Smart Systems' Efficiency by Digital Twin in Water Distribution Networks. **Water**, v.14, n.8, 2022. DOI: <https://doi.org/10.3390/w14081304>.

RECA, J.; MARTÍNEZ, J.; LÓPEZ, R. A Hybrid Water Distribution Networks Design Optimization Method Based on a Search Space Reduction Approach and a Genetic Algorithm. **Water**, v. 9, n. 845, 2017. DOI: <https://doi.org/10.3390/w9110845>.

ROLLER, R. P.; MONTEIRO, L. N.; TOMAZINI, L. R.; GODOY, E. P. Detecção de vazamento em redes hidráulicas utilizando Redes Neurais de Grafos. **XXIV Congresso Brasileiro de Automática - CBA**, 2022. DOI: 10.20906/CBA2022/3688.

ROMERO-BEN, L.; CEMBRANO, G.; PUIG, V.; BLESAS, J. Model-free Sensor Placement for Water Distribution Networks using Genetic Algorithms and Clustering. **IFAC-PapersOnline**, v.15, n. 33, 202. DOI: <https://doi.org/10.1016/j.ifacol.2022.11.009>.

ROSSMAN, L. A. EPANET programmer's Toolkit Manual. Water Supply and Water Resources Division, National Risk Management Research Laboratory, US Environmental Protection Agency: Cincinnati, 2000. Disponível em: https://lopez-ibanez.eu/doc/toolkit_help.pdf. Acesso em: 22 out. 2023.

ROSSMAN, L. A. EPANET 2.0 Manual do Usuário. traduzido e adaptado por Laboratório de Eficiência Energética e Hidráulica em Saneamento. Universidade Federal da Paraíba, 2009.

SANCHEZ, A. S; OLIVEIRA-Esquerre, K. P. Internet of things for a smart campus: on-line monitoring of water consumption in University Buildings. **International Journal of**

Engineering Research & Technology, v. 7, p. 341-345, 2018. DOI: <http://dx.doi.org/10.17577/IJERTV7IS030187>.

SANGROULA, U.; HAN, K. H.; KOO, K. M.; GNAWALI, K.; YUM, K. T. Optimization of Water Distribution Networks Using Genetic Algorithm Based SOP–WDN Program. **Water**, v. 14, n. 851, 2022. Doi: <https://doi.org/10.3390/w14060851>.

SANTOS, R. I.; LÓPEZ, E. F. R.; PUIG, V.; VALENCIA, P. G.; HERNÁNDEZ, H. R. Pressure Sensor Placement for Leak Localization in Water Distribution Networks Using Information Theory. **Sensors**, v. 22, n. 443, 2022. Doi: <https://doi.org/10.3390/s22020443>.

SANTOS, R. F. V. C. **Python: Guia prático do básico ao avançado**. E-book, [s.l.], 2ª Edição, 2020. Disponível em: <https://pt.scribd.com/document/615553433/Python-Guia-pratico-do-basico-Rafael-F-V-C-Santos>. Acesso em: 20 mar. 2023.

SCACCIA, K. **Validação Cruzada Aninhada com Scikit-learn**. Disponível em: <https://dataml.com.br/validacao-cruzada-aninhada-com-scikit-learn/>

SCRUCCA, L. “GA: A Package for Genetic Algorithms in R.” **Journal of Statistical Software**, 53(4), 1–37, 2013. DOI:10.18637/jss.v053.i04.

SHI, L.; ZHANG, J.; YU, X.; FU, D. Artificial neural network-based water distribution scheme in real-time in long-distance water supply systems. **AQUA - Water Infrastructure, Ecosystems and Society**, v. 73, n. 8, pp. 1611-1620, 2024. Doi: <https://doi.org/10.2166/aqua.2024.087>.

SHIDDIQI, A. M.; ZA'IN, C.; LATHIFAH, A.; AHMAD, T.; PURWITASARI, D. GA-Sense: Sensor placement strategy for detecting leaks in water distribution networks based on time series flow and genetic algorithm. **MethodsX**, v. 12, 2024. DOI: <https://doi.org/10.1016/j.mex.2024.102612>.

SHIRDEL, M.; ASADI, R.; DO, D.; Hintlian, M. DEEP LEARNING WITH KERNEL FLOW REGULARIZATION FOR TIME SERIES FORECASTING. **arXiv**, v. 1, 2021. DOI: <https://doi.org/10.48550/arXiv.2109.11649>.

SHIU, C-C; CHUNG, C-C; CHIANG, T. Enhancing the EPANET hydraulic model through genetic algorithm optimization of pipe roughness coefficients. **Water Resources Management**, v. 38, n. 1, p. 323-341, 2024. Doi: <http://dx.doi.org/10.1007/s11269-023-03672-0>.

SILQUEIRA, M. G.; SILVA, F. G. B.; SILVA, A. T. Y. L.; LARA, L. L. D.; MARQUES, S. M.; BARBEDO, M. D. G.; REIS, J. A. T. Correlation of Water and Energy Parameters in a South of Minas Gerais Water Supply System. **Revista De Gestão Social E Ambiental - RGSA**, v. 18, n. 10, 2024. Doi: <https://doi.org/10.24857/rgsa.v18n10-200>.

SILVA, F. G. B. **Estudos de calibração de redes de distribuição de água através de algoritmos genéticos**. Tese (Doutorado em Hidráulica e Saneamento) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2003. DOI: 10.11606/T.18.2016.tde-23112016-163209.

SILVA, A. T. Y. L. **Aplicação de Redes Neurais Multicamadas Associadas a Algoritmos Genéticos Aplicadas a Operação de Redes de Distribuição de Água com Vistas a Eficiência Hidroenergética em Cidades Inteligentes**. Tese (Doutorado em Ciência em Meio Ambiente e Recursos Hídricos). Universidade Federal de Itajubá, Itajubá - MG, 2023.

Disponível em:

https://repositorio.unifei.edu.br/jspui/bitstream/123456789/3530/1/Tese_2023009.pdf. Acesso em: 20 out. 2024.

SILVA, C. V. C.; RODRIGUES, G. P. W. **APLICAÇÃO DE METODOLOGIA DE PENALIZAÇÃO ADAPTATIVA COM ALGORITMO GENÉTICO PARA A OTIMIZAÇÃO DO DIMENSIONAMENTO DE REDES DE DISTRIBUIÇÃO DE ÁGUA**. 32º Congresso da ABES, Congresso Brasileiro de Engenharia Sanitária e Ambiental. Disponível em: https://abes-dn.org.br/anaiselronicos/32cbesa/817_tema_i.pdf. Acesso em: 25 mar. 2025

SILVA, A. T. Y. L.; SILVA, F. G. B.; SILVA, A. C.; REIS, J. A. T.; FREITAS, C. L.; VALERIO, V. E. M. Proposal of optimal operation strategy applied to water distribution network with statistical approach. **Revista Ambiente & Água**, v. 15, n. 2, 2020. DOI: <https://doi.org/10.4136/ambi-agua.2455>.

SILVA, B. A. A.; OLIVEIRA, M. S.; REIS, A.; AIÃO, M.; COSTA, R. **OTIMIZAÇÃO DOS CUSTOS DE DISTRIBUIÇÃO DE ÁGUA COM SMART DIGITAL TWINS: UM ESTUDO DE CASO**. 32º Congresso da ABES, Belo Horizonte - MG, 2023.

SILVA, A. C.; SILVA, F. G. B.; VALÉRIO, V. E. M.; SILVA, A. T. Y. L.; MARQUES, S. M.; REIS, J. A. T. Application of data prediction models in a real water supply network: comparison between arima and artificial neural networks. **Brazilian Journal of Water Resources**, v. 29, n. 12, 2024. DOI: <https://doi.org/10.1590/2318-0331.292420230057>.

SILVA JÚNIOR, J. F. **Detecção de Perdas em Sistemas de Distribuição de Água através de Redes de Sensores Sem Fio**. Dissertação (Mestrado em Engenharia Elétrica) Centro de Tecnologia e Geociências, Universidade Federal de Pernambuco, Recife, 2017. Disponível em: <https://repositorio.ufpe.br/handle/123456789/29444>. Acesso em: 11 maio 2025.

SIMPSON, A.R.; DANDY, G.C.; MURPHY, L. Genetic Algorithms Compared to Other Techniques for Pipe Optimization. **Journal of Water Resources Planning and Management**, v.120, n.4, p.423-443, 1994. DOI: [https://doi.org/10.1061/\(ASCE\)0733-9496\(1994\)120:4\(423\)](https://doi.org/10.1061/(ASCE)0733-9496(1994)120:4(423)).

SINGH, M.; FUENMAYOR, E.; HINCHY, E.; QIAO, Y.; MURRAY, N.; DEVINE, D. Digital Twin: Origin to Future. **Applied System Innovation**, v.4, n. 2, 2021. DOI: <https://doi.org/10.3390/asi4020036>.

SINGH, A.; MAHESHWARI, A.; SINGH, S. Digital Twin Framework for Leakages Detection in Large-scale Water Distribution Systems: A Case Study of IIT-Jodhpur Campus. **IFAC-PapersOnline**, v. 57, p. 280-285, 2024. DOI: <https://doi.org/10.1016/j.ifacol.2024.05.048>.

SOBRINHO, R. A; BORJA, P. C. Gestão das perdas de água e energia em sistema de abastecimento de água da Embasa: Um estudo dos fatores intervenientes na RMS.

Engenharia Sanitária e Ambiental, n. 4, v. 21, p. 783-795, 2016. DOI: <https://doi.org/10.1590/S1413-41522016116037>.

SOUSA, A. C.; SOARES, A. K. Modelo para otimização hidroenergética da operação de bombeamento em sistemas de distribuição de água. **Engenharia Sanitária e Ambiental**, n. 3, p. 429–440, 2020. DOI: <https://doi.org/10.1590/S1413-415220190255>.

SOUSA, D. L.; BEZERRA, A. A.; CASTRO, M. A. H.; NETO, G. C. R.; ARAUJO, R. B. A. Desenvolvimento de um modelo híbrido Migha-RNA para calibração de rugosidades absolutas em redes de distribuição de água. **Engenharia Sanitária Ambiental**, v.27, n.6, p.1199-1207, 2022. DOI: <https://doi.org/10.1590/S1413-415220220151>.

SOUZA, R. G. M. **AVALIAÇÃO HIDROENERGÉTICA DE SISTEMAS DE DISTRIBUIÇÃO DE ÁGUA: suprimento contínuo versus intermitente**. Dissertação (Mestrado em Saneamento, Meio Ambiente e Recursos Hídricos) Universidade Federal de Minas Gerais - UFMG, Belo Horizonte - MG, 2021. Disponível em: <https://repositorio.ufmg.br/server/api/core/bitstreams/74ee2408-b679-40de-bd29-9a453e0472a3/content>. Acesso em: 13 abr. 2024.

SRINIVAS, M.; PATNAIK, L. M. Adaptive Probabilities of Crossover and Mutation in Genetic Algorithms. **IEEE Transactions on Systems, Man and Cybernetics**, v. 24, n. 4, p. 656-667, 1994. DOI: <https://doi.org/10.1109/21.286385>.

SRINIVASAN, L.; SCHARNAGL, J.; SCHILLING, K. Analysis of Websockets as the New Age Protocol for Remote Robot Teleoperation. **IFAC Proceedings Volumes**, v. 46, n. 29, p. 83-88, 2013. DOI: <https://doi.org/10.3182/20131111-3-KR-2043.00032>.

SRIVASTAVA, N.; HINTON, G.; KRIZHEVSKYK, A.; SUTSKEVER, I.; SALAKHUTDINOV, R. Dropout: a simple way to prevent neural networks from overfitting. **Journal of Machine Learning Research**, v. 15, p. 1929–1958, 2014. DOI: <https://www.cs.toronto.edu/~rsalakhu/papers/srivastava14a.pdf>

STEARNS, S. C.; RUSLAN, M. 'Trade-offs', 2nd edn, Oxford Academic, 2024. Doi: <https://doi.org/10.1093/oso/9780192871978.003.0004>.

SYED, T.A.; KHAN, M.Y.; JAN, S.; ALBOUQ, S.; ALQAHTANY, S.S.; NAQASH, M.T. Integrating Digital Twins and Artificial Intelligence Multi-Modal Transformers into Water Resource Management: Overview and Advanced Predictive Framework. **AI**, v. 5, pp. 1977-2017, 2024. Doi: <https://doi.org/10.3390/ai5040098>

TARDELLI FILHO, J. Aspectos relevantes do controle de perdas em sistemas públicos de abastecimento de água. *Revista DAE*, v. 64, n. 201, p. 6–20, 2015. DOI: 10.4322/dae.2015.012.

TEIXEIRA, R.; SILVA, D.; JUNIOR, H. M.; MENDOZA, L. F. Previsão de Séries Temporais de Velocidade do Vento Utilizando Redes Neurais Artificiais e Métodos Estatísticos na Região de Arraial do Cabo - RJ. Congresso Brasileiro de Inteligência Computacional, 2020. DOI: <https://doi.org/10.21528/CBIC2019-101>.

TOSCANO, W., & CABRAL, E. L. L. Contextualização de redes neurais recorrentes. *Exacta*, v. 4, p.65–66, 2008. DOI: <https://doi.org/10.5585/exacta.v4iesp.730>.

TSUTIYA, M. T. Abastecimento de água. 3 ed. Departamento de Engenharia Hidráulica e Sanitária da Escola Politécnica da Universidade de São Paulo: São Paulo, 2006. Disponível em: <https://pt.scribd.com/doc/127473795/Abastecimento-de-Agua-Tsutiya>. Acesso em: 05 ago. 2023.

UNADKAT, S. B.; CIOCOIU, M. M.; MEDSKER, L. R. Recurrent Neural Networks: Design and Applications. Boca Raton: **CRC Press**, 2001. Disponível em: https://books.google.com.br/books?id=ME1SAkN0PyMC&printsec=frontcover&redir_esc=y#v=onepage&q&f=false. Acesso em: 20 out. 2024.

UNPINGO, J. **Python for Probability, Statistics, and Machine Learning**. Extras Online. [s.i]. Springer, 2016. Disponível em: <http://54.243.252.9/engr-1330-webroot/3-Readings/Python-for-Probability-Statistics-And-Machine-Learning.pdf>. Acesso em: 10 dez. 2024.

VIEIRA, L. T. DA S. **Análise e avaliação do comportamento de parâmetros hidráulicos de uma rede de distribuição de água do Sul de Minas Gerais**. Dissertação (Mestrado em Meio Ambiente e Recursos Hídricos) - Universidade Federal de Itajubá, Itajubá-MG, 2019. Disponível em: https://repositorio.unifei.edu.br/jspui/bitstream/123456789/1967/1/dissertacao_2019076.pdf. Acesso em: 02 maio 2023.

WALSKI, T. M.; BRILL, E. D.; GESSLER, J.; GOULTER, I. C.; JEPSON, R. M. ; LANSEY, K.; LEE, H. L.; LIEBMAN, J. C.; MORGAN, D. R.; ORMSBEE, L.. "Battle of the Network Models: Epilogue." **Journal of Water Resources Planning and Management**, v. 113, n. 2, p. 191-203, 1987. DOI: [https://doi-org.ezproxy.uky.edu/10.1061/\(ASCE\)0733-9496\(1987\)113:2\(191\)](https://doi-org.ezproxy.uky.edu/10.1061/(ASCE)0733-9496(1987)113:2(191)).

WAQAS, M.; HUMPHRIES, U. W. A critical review of RNN and LSTM variants in hydrological time series predictions. **MethodsX**, v.13, 2024. DOI: <https://doi.org/10.1016/j.mex.2024.102946>.

WU, X.; LU, G.; WU, Z. Remote Sensing Technology in the Construction of Digital Twin Basins: Applications and Prospects. **Water**, v. 15, n. 11, 2023. Doi: <https://doi.org/10.3390/w15112040>.

XU, Z.; WANG, Z.; GAO, C.; ZHANG, K.; LV, J.; LIU, L.; A digital twin system for centrifugal pump fault diagnosis driven by transfer learning based on graph convolutional neural networks. **Computers in Industry**, v.163, 2024. DOI: <https://doi.org/10.1016/j.compind.2024.104155>.

YU, H.; LV, H.; YANG, Y.; ZHAO, R. Water Demand Prediction Model of University Park Based on BP-LSTM Neural Network. **Water**, v. 17, n. 18, 2025. DOI: <https://doi.org/10.3390/w17182729>.

ZANFEI, A.; BRENTAN B. M.; MENAPACE, A.; RIGHETTI, M. A short-term water demand forecasting model using multivariate long short-term memory with meteorological

data. **Journal of Hydroinformatics**, v. 11, n. 5, p. 1053-1065, 2022. DOI: <https://doi.org/10.2166/hydro.2022.055>.

ZEKRI, S.; JABEUR, N.; GHARRAD, H. Smart Water Management Using Intelligent Digital Twins. **Computing and Informatics**, v. 41, n.1, pp. 135–153, 2022. Doi: https://doi.org/10.31577/cai_2022_1_135.

ZHANG, Q.; LIU, B.; ZHOU, F.; WANG, Q.; KONG, J. **State-of-charge estimation method of lithium-ion batteries based on long short term memory network State-of-charge estimation method of lithium-ion batteries based on long-short term memory network**. I IOP Conference Series: Earth and Environmental Science, v. 208, Conferência Internacional sobre Poluição do Ar e Engenharia Ambiental (APEE 2018), Hong Kong-China, 2018. DOI: 10.1088/1755-1315/208/1/012001.

ZHANG, Y.; LI, S; ZHENG, Y.; ZOU, Y. Multi-model based pressure optimization for large-scale water distribution networks. **Control Engineering Practice**, v. 95, 2020. DOI: <https://doi.org/10.1016/j.conengprac.2019.104232>.

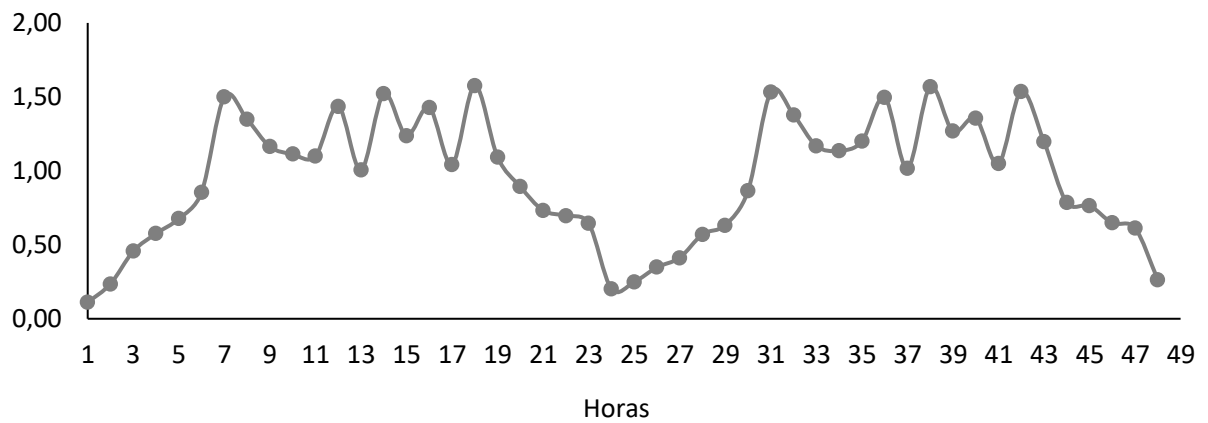
ZHANG, C.; MA, Y. **Ensemble Machine Learning, Methods and Applications**. Springer, London, 2012. Disponível em: <https://link.springer.com/book/10.1007/978-1-4419-9326-7>. Acesso em: 20 out. 2024.

ZHONG, N., LI, Y. P., LI, X. Z., GUO, C. X. AND WU, T. Accurate prediction of salmon storage time using improved raman spectroscopy. **Journal of Food Engineering**, v. 293, 2021. DOI: <https://doi.org/10.1016/j.jfoodeng.2020.110378>.

ANEXO

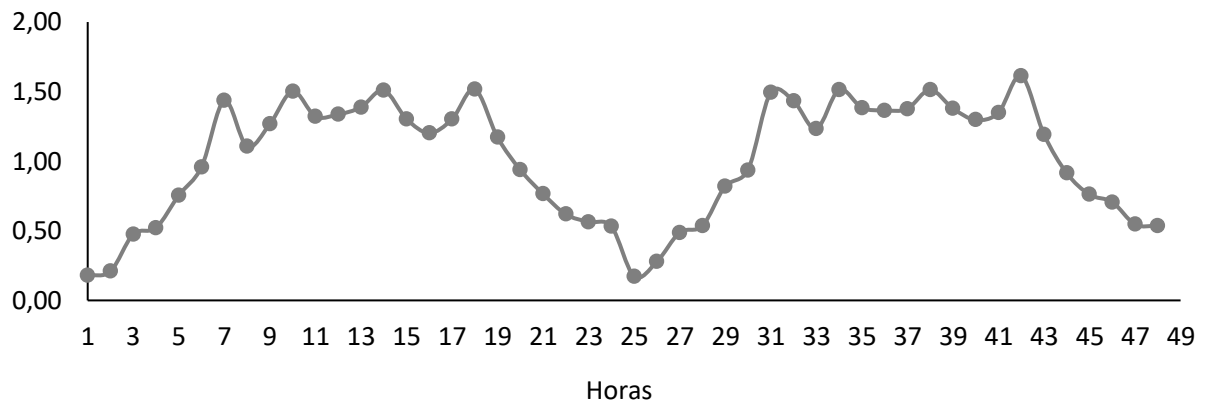
Para a utilização das redes auxiliares no estudo para que o treinamento da rede neural desenvolvida fosse mais robusto, possibilitando que ela seja mais generalista, foi aplicado a curva de demanda para cada uma das redes auxiliares, assim como foi feito na rede de estudo. Os valores obtidos para cada uma das redes estão nas Figuras 1 a 3.

Figura 1 – Curva de demanda da rede Marques (2023).



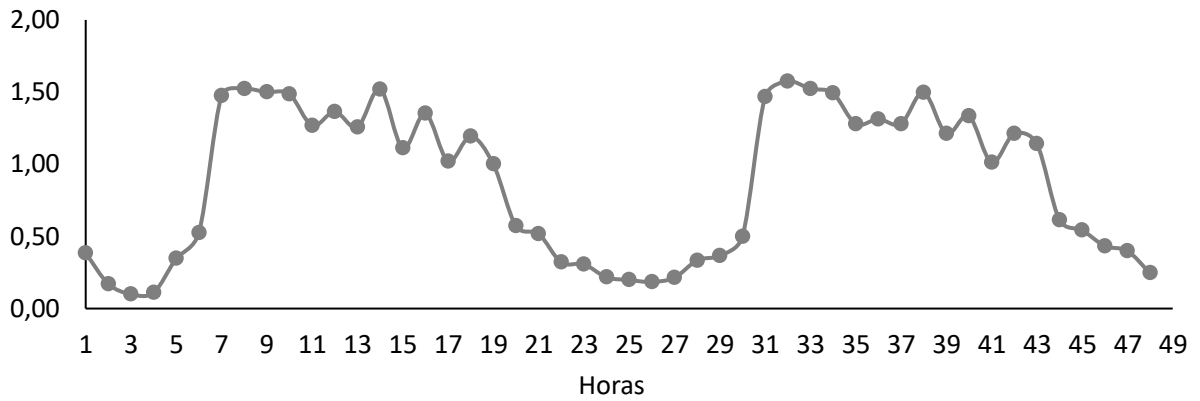
Fonte: Próprio autor (2025).

Figura 2 – Curva de demanda da rede Silva (2023).



Fonte: Próprio autor (2025).

Figura 3 – Curva de demanda da rede Anytown.



Fonte: Próprio autor (2025).

Com isso, foi encontrado os valores de pressão e perdas de cada uma das redes que estão apresentados nas Tabelas 1 a 6, sendo que as Tabelas 1, 3 e 5 são as pressões e as Tabelas 2, 4 e 6 os valores de perdas para a rede Marques (2023), Silva (2023) e Anytown (Walski *et al.*, 1987), respectivamente.

Tabela 1 - Pressões da rede Marques (2023).

(continuação)

Horas	Nó												
	3	4	5	6	7	8	9	10	11	12	13	14	15
1	44,76	49,63	49,42	59,68	64,62	59,87	79,29	83,74	74,69	86,79	80,97	24,23	28,43
2	44,85	49,76	49,62	59,79	64,75	59,92	79,54	84,18	74,80	87,92	82,38	24,36	28,65
3	44,34	48,96	48,36	59,11	63,93	59,65	78,00	81,43	74,14	80,80	73,51	23,53	27,30
4	44,58	49,35	48,98	59,44	64,33	59,78	78,75	82,77	74,46	84,29	74,86	23,93	27,96
5	44,63	49,41	49,08	59,50	64,39	59,80	78,88	83,00	74,51	84,87	78,58	24,00	28,07
6	44,68	49,50	49,22	59,57	64,48	59,83	79,04	83,30	74,59	85,65	79,55	24,09	28,21
7	44,53	49,27	48,85	59,37	64,25	59,75	78,60	82,50	74,40	83,59	76,98	23,85	27,82
8	44,66	49,47	49,16	59,54	64,45	59,82	78,98	83,18	74,56	85,35	79,17	24,06	28,16
9	44,63	49,42	49,09	59,50	64,40	59,80	78,89	83,01	74,52	84,92	78,64	24,01	28,08
10	44,82	49,72	49,56	59,76	64,71	59,90	79,46	84,05	74,77	87,58	81,96	24,32	28,58
11	44,74	49,60	49,37	59,65	64,58	59,86	79,23	83,63	74,67	86,52	80,63	24,19	28,38
12	44,98	49,96	49,94	59,97	64,96	59,99	79,93	84,88	74,97	89,69	84,61	24,57	28,98
13	44,53	49,27	48,85	59,37	64,28	59,75	78,60	82,50	74,40	83,59	76,98	23,85	27,82
14	44,66	49,47	49,16	59,54	64,45	59,82	78,98	83,18	74,56	85,35	79,17	24,06	28,16
15	44,46	49,46	49,07	59,64	64,43	59,79	78,97	83,10	74,98	84,97	78,96	24,18	28,40
16	44,82	49,72	49,56	59,76	64,71	59,90	79,46	84,05	74,77	87,58	81,96	24,32	28,58
17	44,86	49,78	49,50	59,81	64,77	59,92	79,57	84,24	74,82	88,08	82,59	24,38	28,68
18	44,78	49,68	49,78	59,68	64,78	59,89	79,05	83,41	74,59	85,64	79,45	24,13	28,14
19	44,35	49,98	48,68	59,79	64,64	59,57	78,98	82,64	74,64	83,71	76,97	23,45	27,98
20	44,64	49,87	49,46	59,48	64,65	59,72	78,98	83,17	74,98	85,64	79,39	24,04	28,63
21	44,67	49,62	49,97	59,42	64,63	59,87	78,98	83,04	74,59	84,98	78,98	24,64	28,12
22	44,97	49,74	49,93	59,97	64,30	59,41	79,64	84,97	41,35	87,34	81,98	24,69	28,91
23	44,86	49,78	49,65	59,81	64,77	59,92	79,57	84,24	74,82	88,05	82,59	24,38	28,86
24	44,86	49,05	49,22	59,75	64,84	59,38	79,40	83,03	74,95	85,56	79,55	24,09	28,21
25	44,94	49,90	49,85	59,91	64,90	59,97	79,81	84,66	74,92	89,15	83,93	24,51	28,88
26	44,86	49,79	49,67	59,82	64,78	59,93	79,59	84,27	74,82	88,16	82,68	24,39	28,96

Fonte: Próprio autor (2025).

Tabela 1 - Pressões da rede Marques (2023).

(conclusão)

Nó													
Horas	3	4	5	6	7	8	9	10	11	12	13	14	15
27	44,31	48,92	48,30	59,07	63,89	59,63	77,92	81,29	74,11	80,44	73,05	23,48	27,23
28	44,58	49,34	48,97	59,44	64,32	59,78	78,74	82,75	74,46	84,25	77,80	23,93	27,95
29	44,97	49,78	48,97	59,64	64,31	59,78	78,61	82,91	74,31	83,98	77,15	23,35	27,91
30	44,68	49,49	49,20	59,56	64,47	59,83	79,02	83,26	74,58	58,56	79,43	24,08	28,20
31	44,23	49,64	48,21	59,34	64,21	59,37	78,64	82,39	74,31	83,97	76,69	23,84	27,98
32	44,92	49,14	49,32	59,64	64,12	59,31	78,31	83,91	74,12	85,67	79,31	24,93	28,98
33	44,56	49,31	48,92	59,41	64,29	59,77	78,69	82,65	74,43	83,98	77,47	23,90	27,90
34	44,14	49,71	49,61	59,74	64,37	59,31	79,31	83,97	74,33	87,99	81,41	24,64	28,97
35	44,78	49,14	49,78	59,44	64,78	59,14	79,64	84,12	74,98	88,36	82,34	24,81	48,74
36	44,97	49,96	49,94	59,96	64,96	59,99	79,92	84,86	74,97	89,66	84,57	24,57	28,98
37	44,64	49,97	49,32	59,78	64,28	59,31	79,31	83,14	74,64	85,69	79,49	24,98	27,98
38	44,51	49,24	48,80	59,34	64,21	59,74	78,54	82,39	74,37	83,30	76,62	23,82	27,77
39	44,66	49,46	49,16	59,54	64,44	59,82	78,97	82,17	74,56	85,31	79,13	24,05	27,15
40	44,56	49,31	48,92	59,41	64,29	59,77	78,69	82,65	74,43	83,98	77,47	23,90	27,90
41	44,81	49,70	49,53	59,74	64,69	59,90	79,41	83,97	74,75	87,38	81,72	24,30	28,54
42	44,87	49,80	49,69	59,83	64,79	59,93	79,61	84,31	74,92	88,26	82,81	24,40	28,71
43	44,68	49,49	49,20	59,56	64,47	59,83	79,02	83,26	74,58	85,56	79,43	24,08	28,20
44	44,51	49,24	48,80	59,34	64,21	59,74	78,54	82,39	74,37	83,30	76,62	23,82	27,77
45	44,66	49,46	49,16	59,54	64,44	59,82	78,97	83,17	74,56	85,31	79,13	24,05	28,15
46	44,56	49,31	48,92	59,41	64,29	59,77	78,69	82,65	74,43	83,98	77,47	23,90	27,90
47	44,87	49,80	49,69	59,83	64,79	59,93	79,61	81,31	74,83	88,26	82,81	24,40	28,71
48	44,68	49,49	49,20	59,56	64,47	59,83	79,02	83,26	74,58	85,56	79,43	24,08	28,20

Fonte: Próprio autor (2025).

Tabela 2 - Perdas da rede Marques (2023).

(continuação)

Horas	Perda
1	56,53
2	56,69
3	55,70
4	56,08
5	56,27
6	56,37
7	56,09
8	56,33
9	56,27
10	56,64
11	56,50
12	56,94
13	56,09
14	56,33
15	56,33
16	56,64
17	56,71
18	56,42
19	56,15
20	56,41
21	56,37

Fonte: Próprio autor (2025).

Tabela 2 - Perdas da rede Marques (2023).

(conclusão)

Horas	Perda
22	55,48
23	56,72
24	56,38
25	56,86
26	56,73
27	55,65
28	56,18
29	56,16
30	55,38
31	56,05
32	56,37
33	56,14
34	56,58
35	57,39
36	56,93
37	56,40
38	56,05
39	56,26
40	56,14
41	56,62
42	56,74
43	56,36
44	56,05
45	56,33
46	56,14
47	56,63
48	56,36

Fonte: Próprio autor (2025).

Tabela 3 - Pressões da rede Silva (2023).

(continuação)

Horas	Nó												
	3	4	5	6	7	8	9	10	11	12	13	14	15
1	44,73	49,50	54,52	59,56	64,44	59,83	79,56	84,44	74,75	89,46	84,75	15,48	10,48
2	44,91	49,83	54,84	59,85	64,81	59,94	79,85	84,81	74,92	89,82	84,92	15,83	10,83
3	44,43	48,95	53,99	59,07	63,84	59,65	79,07	83,84	74,49	88,88	84,49	14,91	9,89
4	44,55	49,16	54,19	59,26	64,07	59,72	79,26	84,07	74,59	89,11	84,59	15,13	10,12
5	44,51	49,09	54,13	59,20	64,00	59,70	79,20	84,00	74,56	89,03	84,56	15,06	10,05
6	44,46	49,00	54,04	59,12	63,90	59,67	79,12	83,90	74,51	88,94	84,51	14,97	9,95
7	44,52	49,12	54,15	59,22	64,02	59,71	79,22	84,02	74,57	89,06	84,57	15,08	10,07
8	44,52	49,11	54,15	59,22	64,02	59,71	79,22	84,02	74,57	89,05	84,57	15,08	10,07
9	44,61	49,27	54,30	59,36	64,20	59,76	79,36	84,20	74,65	89,23	84,65	15,25	10,24
10	44,90	49,81	54,82	59,83	64,79	59,94	79,83	84,79	74,91	89,80	84,91	15,81	10,81
11	44,82	49,67	54,68	59,71	64,63	59,89	79,71	84,63	74,84	89,64	84,84	15,66	10,65
12	44,99	49,97	54,98	59,98	64,97	59,99	79,98	84,97	74,99	89,97	84,99	15,97	10,64
13	44,73	49,50	54,52	59,56	64,44	59,83	79,56	84,44	74,75	89,46	84,75	15,48	10,48
14	44,91	49,83	54,84	59,85	64,81	59,94	79,85	84,81	74,92	89,82	84,92	15,83	10,83
15	44,43	48,95	53,99	59,07	63,84	59,65	79,07	83,84	74,49	88,88	84,49	14,91	9,89
16	44,55	49,16	54,19	59,26	64,07	59,72	79,26	84,07	74,59	89,11	84,59	15,13	10,12

Fonte: Próprio autor (2025).

Tabela 3 - Pressões da rede Silva (2023).

(conclusão)

Nó													
Horas	3	4	5	6	7	8	9	10	11	12	13	14	15
17	44,51	49,09	54,13	59,20	64,00	59,70	79,20	84,00	74,56	89,03	84,56	15,06	10,05
18	44,46	49,00	54,04	59,12	63,90	59,67	79,12	83,90	74,51	88,94	84,51	14,97	9,95
19	44,52	49,12	54,15	59,22	64,02	59,71	79,22	84,02	74,57	89,06	84,57	15,08	10,07
20	44,52	49,11	54,15	59,22	64,02	59,71	79,22	84,02	74,57	89,05	84,57	15,08	10,07
21	44,61	49,27	54,30	59,36	64,20	59,76	79,36	84,20	74,65	89,23	84,65	15,25	10,24
22	44,90	49,81	54,82	59,83	64,79	59,94	79,83	84,79	74,91	89,80	84,91	5,81	10,81
23	44,82	49,67	54,68	59,71	64,63	59,89	79,71	84,63	74,84	89,64	84,84	15,66	10,65
24	44,99	49,97	54,98	59,98	64,97	59,99	79,98	84,97	74,99	89,97	84,99	15,97	10,97
25	44,99	49,98	54,98	59,89	64,97	59,99	79,98	84,97	74,99	89,98	84,99	15,98	10,98
26	44,92	49,85	54,86	59,87	64,84	59,95	79,87	84,84	74,93	89,84	84,93	15,85	10,85
27	44,79	49,62	54,64	59,66	64,58	59,87	79,66	84,58	74,81	89,59	84,81	15,61	10,61
28	44,38	48,86	53,91	58,99	63,74	59,62	78,99	83,74	74,45	88,79	84,45	14,82	9,80
29	44,57	49,20	54,23	59,30	64,12	59,74	79,30	84,12	74,61	89,15	84,61	15,17	10,16
30	44,48	49,03	54,07	59,15	63,93	59,68	79,15	83,93	74,53	88,97	84,53	15,00	9,98
31	44,47	49,02	54,06	59,14	63,92	59,68	79,14	83,92	74,52	88,96	84,52	14,98	9,97
32	44,48	49,04	54,08	59,15	63,94	59,68	79,15	83,94	74,53	88,98	84,53	15,01	9,99
33	44,49	49,05	54,09	59,17	63,96	59,69	79,17	83,96	74,54	88,99	84,54	15,02	10,01
34	44,59	49,25	54,28	59,34	64,17	59,75	79,34	84,17	74,63	89,20	84,63	15,22	10,21
35	44,82	49,67	54,68	59,71	64,63	59,89	79,71	84,63	74,84	89,65	84,84	15,66	10,66
36	44,90	49,82	54,82	59,84	64,80	59,94	79,84	84,80	74,91	89,80	84,91	15,81	10,81
37	44,99	49,98	54,98	59,98	64,97	59,99	79,98	84,97	74,90	89,98	84,99	15,96	10,98
38	44,92	49,85	54,86	59,87	64,84	59,95	79,87	84,84	74,93	89,84	84,93	15,85	10,85
39	44,79	49,62	54,64	59,66	64,85	59,87	79,66	84,58	74,81	89,59	75,81	15,61	10,61
40	44,38	48,86	53,91	58,99	63,74	59,62	78,99	83,74	74,45	88,79	84,45	14,85	9,80
41	44,57	49,20	54,23	59,30	64,12	59,74	79,30	84,12	74,61	89,15	84,61	15,17	10,16
42	44,48	49,03	54,07	59,15	63,93	59,68	79,15	83,93	74,53	88,97	84,53	15,00	9,98
43	44,47	49,02	54,06	59,14	63,92	59,68	79,14	83,92	74,52	88,96	84,52	14,98	9,97
44	44,48	49,04	54,08	59,15	63,94	59,68	79,15	83,94	74,53	88,98	84,53	15,01	9,99
45	44,49	49,05	54,09	59,17	63,99	59,69	79,17	83,96	74,54	88,99	84,54	15,02	10,01
46	44,59	49,25	54,28	59,34	64,17	59,75	79,34	84,17	74,63	89,20	84,63	15,22	10,21
47	44,82	49,67	54,68	59,71	64,63	59,89	79,71	84,63	74,84	89,65	84,84	15,66	10,66
48	44,90	49,82	54,82	59,84	64,80	59,94	79,84	84,80	74,91	89,80	84,91	15,81	10,81

Fonte: Próprio autor (2025).

Tabela 4 - Perdas da rede Silva (2023).

(continuação)

Horas	Perda
1	54,12
2	54,17
3	53,52
4	53,87
5	53,90
6	53,58
7	53,93
8	54,14
9	54,10
10	54,45
11	54,83

Fonte: Próprio autor (2025).

Tabela 4 - Perdas da rede Silva (2023).

(conclusão)

Horas	Perda
12	54,73
13	54,42
14	54,37
15	54,37
16	54,41
17	54,66
18	54,73
19	54,61
20	54,52
21	54,72
22	54,58
23	55,02
24	54,98
25	53,37
26	53,45
27	53,23
28	53,23
29	53,32
30	53,27
31	53,38
32	53,45
33	53,51
34	53,71
35	53,86
36	53,95
37	53,9
38	53,97
39	53,71
40	53,79
41	53,91
42	53,92
43	53,96
44	53,99
45	54,08
46	54,09
47	54,36
48	54,47

Fonte: Próprio autor (2025).

Tabela 5 - Pressões da rede Anytown.

(continuação)

Horas	Nó																		
	20	30	40	50	55	60	70	75	80	90	100	110	115	120	130	140	150	160	170
1	100,50	72,20	72,08	72,00	58,91	71,76	72,26	58,82	71,80	71,83	71,90	72,04	58,80	41,28	41,25	58,79	41,48	41,26	41,24
2	101,04	72,52	72,41	72,32	59,21	71,95	72,57	59,11	72,08	72,13	72,19	72,32	59,08	41,44	41,40	59,06	41,77	41,40	41,40
3	101,20	72,62	72,51	72,42	59,31	72,02	72,67	59,20	72,18	72,23	72,29	72,41	59,18	41,49	41,46	59,15	41,87	41,44	41,46
4	98,16	71,33	71,08	71,05	58,02	71,31	71,38	58,00	70,99	70,97	71,02	71,17	57,99	40,74	40,60	57,98	40,66	40,80	40,32
5	98,10	71,30	71,05	71,02	57,99	71,29	71,35	57,97	70,96	70,93	70,99	71,14	57,96	40,71	40,57	57,95	40,63	40,78	40,29
6	98,62	71,52	71,31	71,29	58,26	71,43	71,59	58,23	71,22	71,21	71,26	71,39	58,22	40,92	40,81	58,21	40,89	40,96	40,61
7	98,65	71,53	71,33	71,30	58,27	71,43	71,60	58,24	71,23	71,22	71,27	71,40	58,23	40,93	40,82	58,23	40,90	40,97	40,62
8	98,96	71,64	71,45	71,42	58,39	71,48	71,71	58,36	71,35	71,34	71,39	71,53	58,35	41,03	40,94	58,34	41,02	41,05	40,78

Fonte: Próprio autor (2025).

Tabela 5 - Pressões da rede Anytown.

(conclusão)

Nó																			
Horas	20	30	40	50	55	60	70	75	80	90	100	110	115	120	130	140	150	160	170
9	99,14	71,69	71,52	71,49	58,45	71,49	71,76	58,41	71,41	71,40	71,44	71,59	58,41	41,08	41,00	58,40	41,07	41,09	40,86
10	99,18	71,70	71,53	71,50	58,46	71,49	71,77	58,42	71,42	71,41	71,45	71,61	58,42	41,09	41,01	58,41	41,08	41,10	40,88
11	100,18	72,05	71,91	71,84	58,77	71,66	72,11	58,69	71,67	71,69	71,75	71,91	58,67	41,21	41,18	58,66	41,35	41,20	41,15
12	100,05	71,99	71,84	71,78	58,71	71,62	72,05	58,64	71,63	71,64	71,70	71,86	58,63	41,19	41,16	58,61	41,30	41,19	41,12
13	98,98	71,64	71,46	71,43	58,40	71,48	71,72	58,36	71,36	71,35	71,39	71,54	58,36	41,04	40,95	58,35	41,03	41,06	40,79
14	100,59	72,25	72,13	72,05	58,96	71,79	72,31	58,86	71,85	71,87	71,94	72,09	58,85	41,30	41,28	58,83	41,28	41,27	41,40
15	100,10	72,26	72,15	72,24	59,13	72,40	72,48	59,41	72,81	72,37	72,97	72,34	59,84	40,91	41,33	58,90	41,63	41,44	41,46
16	99,16	71,20	71,94	71,50	57,46	71,71	71,43	58,10	71,41	71,34	70,87	71,71	57,14	40,93	40,51	57,64	40,45	40,64	40,14
17	98,10	71,24	71,50	71,43	57,13	71,39	71,45	57,39	70,96	70,46	70,87	71,65	57,67	40,31	40,64	57,47	40,32	40,87	41,64
18	98,45	71,65	71,23	71,34	58,89	71,46	71,65	58,67	71,67	71,64	71,23	71,46	58,21	40,49	40,79	58,74	40,98	40,35	40,16
19	98,32	71,31	71,74	71,64	58,34	71,64	71,45	58,46	71,23	71,64	71,72	71,04	58,64	40,32	10,27	58,00	40,48	40,97	40,46
20	98,96	71,47	71,45	71,64	58,34	71,73	71,72	58,47	71,45	71,64	71,98	71,26	58,47	41,07	40,58	58,67	41,67	40,49	40,78
21	99,14	70,60	51,74	71,97	58,64	71,64	71,97	58,84	71,78	71,95	71,78	71,97	58,78	41,70	41,30	58,34	41,78	41,88	10,64
22	100,00	79,40	71,64	71,97	58,48	71,68	71,48	58,67	71,15	71,67	71,48	71,97	58,78	41,17	41,97	58,77	41,84	41,98	40,78
23	100,15	72,41	71,89	71,32	58,78	71,69	72,44	58,98	71,48	71,68	71,78	71,18	58,80	41,92	41,78	58,99	41,12	41,78	40,00
24	100,11	71,89	71,68	71,98	58,78	71,23	72,87	58,64	71,68	71,98	71,78	71,69	58,48	41,99	41,78	58,78	41,65	41,26	41,98
25	102,00	72,65	72,64	72,98	58,97	71,98	72,57	59,11	72,18	72,19	72,95	72,98	59,19	41,99	41,97	59,31	41,99	41,78	41,98
26	100,50	72,69	72,46	72,98	59,34	72,09	72,45	59,37	72,78	72,39	72,49	72,33	59,48	41,98	41,89	59,46	41,98	41,98	41,94
27	97,15	71,98	71,56	71,06	57,94	71,90	71,60	58,00	70,89	71,00	71,04	71,65	57,97	41,94	41,46	57,49	40,78	40,84	40,15
28	98,10	71,94	71,94	71,98	57,46	71,64	71,45	57,94	70,75	70,45	70,97	71,14	57,64	40,46	40,68	57,48	40,64	40,67	40,48
29	98,45	71,45	71,64	71,98	58,48	71,48	71,98	58,97	71,78	71,27	71,98	71,74	58,97	40,97	40,23	58,87	40,89	40,78	40,48
30	98,45	71,68	71,78	71,98	58,67	71,78	71,89	58,69	71,67	71,97	71,68	71,15	58,69	40,97	40,98	58,36	40,97	40,74	40,90
31	98,67	71,68	71,97	71,84	58,74	71,95	71,57	58,75	71,69	71,48	71,98	71,54	58,56	41,08	40,69	58,97	40,97	41,97	40,47
32	100,10	71,48	71,65	71,58	58,64	71,89	71,67	58,67	71,98	71,48	71,42	71,97	58,79	41,09	41,97	58,90	41,00	41,97	40,46
33	99,18	71,98	71,48	71,65	58,78	71,98	71,65	58,48	71,98	71,74	71,84	71,68	58,12	41,30	41,54	58,96	41,10	41,36	40,99
34	100,70	72,06	71,59	71,89	58,98	71,64	72,64	58,98	71,78	71,96	71,69	71,48	58,71	41,32	41,98	58,78	41,69	41,30	41,15
35	100,07	71,20	71,01	71,98	58,97	71,64	72,95	58,98	71,46	71,89	71,90	71,78	58,97	41,31	41,54	58,74	41,31	41,87	41,97
36	100,50	72,48	72,09	72,31	58,74	71,98	72,61	58,48	71,97	71,32	71,97	72,06	58,14	41,36	41,25	58,79	41,49	41,32	41,04
37	101,04	72,52	72,41	72,32	59,21	71,95	72,57	59,11	72,08	72,13	72,19	72,32	59,08	41,44	41,40	59,06	41,77	41,40	41,40
38	101,20	72,62	72,51	72,42	59,31	72,02	72,67	59,20	72,18	72,23	72,29	72,41	59,18	41,49	41,46	59,15	41,87	41,44	41,46
39	98,16	71,33	71,08	71,05	58,02	71,31	71,38	58,00	70,99	70,97	71,02	71,17	57,99	40,74	40,60	57,98	40,66	40,80	40,32
40	98,10	71,30	71,05	71,02	57,99	71,29	71,35	57,97	70,96	70,93	70,99	71,14	57,96	40,71	40,57	57,95	40,63	40,78	40,29
41	98,62	71,52	71,31	71,29	58,26	71,43	71,59	58,23	71,22	71,21	71,26	71,39	58,22	40,92	40,81	58,21	40,89	40,96	40,61
42	98,65	71,53	71,33	71,30	58,27	71,43	71,60	58,24	71,23	71,22	71,27	71,40	58,23	40,93	40,82	58,23	40,90	40,97	40,62
43	98,96	71,64	71,45	71,42	58,39	71,48	71,71	58,36	71,35	71,34	71,39	71,53	58,35	41,03	40,94	58,34	41,02	41,05	40,78
44	99,14	71,69	71,52	71,49	58,45	71,49	71,76	58,41	71,41	71,40	71,44	71,59	58,41	41,08	41,00	58,40	41,07	41,09	40,86
45	99,18	71,70	71,53	71,50	58,46	71,49	71,77	58,42	71,42	71,41	71,45	71,61	58,42	41,09	41,01	58,41	41,08	41,10	40,88
46	100,18	72,05	71,91	71,84	58,77	71,66	72,11	58,69	71,67	71,69	71,75	71,91	58,67	41,21	41,18	58,66	41,35	41,20	41,15
47	100,05	71,99	71,84	71,78	58,71	71,62	72,05	58,64	71,63	71,64	71,70	71,86	58,63	41,19	41,16	58,61	41,30	41,19	41,12
48	100,50	72,20	72,08	72,00	58,91	71,76	72,26	58,82	71,80	71,83	71,90	72,04	58,80	41,28	41,25	58,79	41,48	41,26	41,24

Fonte: Próprio autor (2025).

Tabela 6 - Perdas da rede Anytown.

(continuação)

Horas	Perda
1	57,54
2	57,67
3	57,71
4	57,14
5	57,12
6	57,24
7	57,25
8	57,31
9	57,33
10	57,34
11	57,48
12	57,46

Fonte: Próprio autor (2025).

Tabela 6 - Perdas da rede Anytown.

(conclusão)

Horas	Perda
13	57,31
14	57,56
15	57,68
16	57,18
17	57,11
18	57,27
19	56,5
20	57,34
21	56,23
22	57,65
23	57,49
24	57,54
25	57,81
26	57,79
27	57,21
28	57,14
29	57,35
30	57,37
31	57,40
32	57,46
33	57,43
34	57,54
35	57,52
36	57,53
37	57,67
38	57,71
39	57,14
40	57,12
41	57,24
42	57,25
43	57,31
44	57,33
45	57,34
46	57,48
47	57,46
48	57,54

Fonte: Próprio autor (2025).