



**UNIVERSIDADE FEDERAL DE CAMPINA GRANDE
CENTRO DE TECNOLOGIA E RECURSOS NATURAIS
UNIDADE ACADÊMICA DE ENGENHARIA CIVIL
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA CIVIL E AMBIENTAL**

**MODELAGEM DO DECAIMENTO DE CLORO RESIDUAL LIVRE EM REDES DE
DISTRIBUIÇÃO DE ÁGUA DE COMUNIDADES RURAIS POR MEIO DE REDES
NEURAS ARTIFICIAIS**

GABRIELE DE SOUZA BATISTA

CAMPINA GRANDE – PB

2024

GABRIELE DE SOUZA BATISTA

**MODELAGEM DO DECAIMENTO DE CLORO RESIDUAL LIVRE EM REDES DE
DISTRIBUIÇÃO DE ÁGUA DE COMUNIDADES RURAIS POR MEIO DE REDES
NEURAIS ARTIFICIAIS**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Civil Ambiental – PPGECA da Universidade Federal de Campina Grande – UFCG, como parte dos requisitos necessários para obtenção do grau de Mestra.

Área de concentração: Recursos Hídricos e Saneamento Ambiental

Orientadora: Profa. Dra. Andréa Carla Lima Rodrigues

Coorientadora: Profa. Dra. Marília Marcy Cabral de Araújo

CAMPINA GRANDE – PB

2024

B333m Batista, Gabriele de Souza.
Modelagem do decaimento de cloro residual livre em redes de distribuição de água de comunidades rurais por meio de redes neurais artificiais / Gabriele de Souza Batista – Campina Grande, 2024.
124 f. : il. color.

Dissertação (Mestrado em Engenharia Civil e Ambiental) - Universidade Federal de Campina Grande, Centro de Tecnologia e Recursos Naturais, 2024.
"Orientação: Profa. Dra. Andréa Carla Lima Rodrigues, Profa Dra. Marília Marcy Cabral de Araújo."

Referências.
1. Qualidade da Água. 2. Desinfecção com Cloro. 3. Aprendizagem de Máquina. 4. tSNE. 5. Índice de Sustentabilidade. I. Rodrigues, Andréa Carla Lima. II. Araújo, Marília Marcy Cabral de. III. Título.

CDU 628.16(043)

GABRIELE DE SOUZA BATISTA

**MODELAGEM DO DECAIMENTO DE CLORO RESIDUAL LIVRE EM REDES DE
DISTRIBUIÇÃO DE ÁGUA DE COMUNIDADES RURAIS POR MEIO DE REDES
NEURAS ARTIFICIAIS**

Dissertação aprovada em: 23 de fevereiro de 2024.

COMISSÃO EXAMINADORA

Profa. Dra. Andréa Carla Lima Rodrigues
Universidade Federal de Campina Grande
Presidenta/Orientadora

Profa. Dra. Marília Marcy Cabral de Araújo
Universidade Federal de Campina Grande
Coorientadora

Profa. Dra. Elis Gean Rocha
Universidade Federal de Campina Grande
Examinadora Interna

Profa. Dra. Patrícia Herminio Cunha
Universidade Federal de Campina Grande
Examinadora Externa

Prof. Dr. Gustavo Barbosa Lima da Silva
Universidade Federal da Paraíba
Examinador Externo



MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DE CAMPINA GRANDE
POS-GRADUACAO ENGENHARIA CIVIL AMBIENTAL
Rua Aprigio Veloso, 882, - Bairro Universitario, Campina Grande/PB, CEP 58429-900

REGISTRO DE PRESENÇA E ASSINATURAS

1. **ATA DA DEFESA PARA CONCESSÃO DO GRAU DE MESTRE EM ENGENHARIA CIVIL E AMBIENTAL**

2. **ALUNO(A): GABRIELE DE SOUZA BATISTA / COMISSÃO EXAMINADORA: DR.^a ANDRÉA CARLA LIMA RODRIGUES - PPGECA/UFCG (PRESIDENTE) - ORIENTADORA, DR.^a MARÍLIA MARCY CABRAL DE ARAÚJO - UAEC/UFCG - COORIENTADORA, DR.^a ELIS GEAN ROCHA – EXAMINADORA INTERNA, DR.^a PATRÍCIA HERMÍNIO CUNHA – UAEC/UFCG – EXAMINADORA EXTERNA, DR. GUSTAVO BARBOSA LIMA DA SILVA – UFPB – EXAMINADOR EXTERNO (PORTARIA 05/2024). / TITULO DA DISSERTAÇÃO: “MODELAGEM DO DECAIMENTO DE CLORO RESIDUAL LIVRE EM REDES DE DISTRIBUIÇÃO DE ÁGUA DE COMUNIDADES RURAIS POR MEIO DE REDES NEURAIS ARTIFICIAIS” / ÁREA DE CONCENTRAÇÃO: RECURSOS HÍDRICOS E SANEAMENTO AMBIENTAL / HORA DE INICIO: 09:00 HORAS / NA SALA DA PÓS-GRADUAÇÃO DO LABORATÓRIO DE HIDRÁULICA I, BLOCO CR, CAMPUS SEDE DA UFCG, EM CAMPINA GRANDE, EM FORMATO HÍBRIDO.**

3. **EM SESSÃO REALIZADA EM FORMATO HÍBRIDO, APÓS EXPOSIÇÃO DE CERCA DE 40 MINUTOS, O(A) CANDIDATO(A) FOI ARGUIDO(A) ORALMENTE PELOS MEMBROS DA COMISSÃO EXAMINADORA, TENDO DEMONSTRADO SUFICIÊNCIA DE CONHECIMENTO E CAPACIDADE DE SISTEMATIZAÇÃO NO TEMA DE SUA DISSERTAÇÃO, SENDO-LHE ATRIBUÍDO O CONCEITO “EM EXIGÊNCIA”, SENDO QUE A POSSIBILIDADE DE APROVAÇÃO ESTÁ CONDICIONADA À AVALIAÇÃO DA NOVA VERSÃO DO TRABALHO FINAL, SEGUINDO PROCEDIMENTOS PREVISTOS NA RESOLUÇÃO DO PROGRAMA. O PRESIDENTE DA COMISSÃO EXAMINADORA, OUVIDOS OS DEMAIS MEMBROS, DEVERÁ FICAR RESPONSÁVEL POR ATESTAR QUE AS CORREÇÕES SOLICITADAS NA LISTA DE EXIGÊNCIAS FORAM ATENDIDAS NA VERSÃO FINAL DO TRABALHO. A COMISSÃO EXAMINADORA CUMPRINDO OS PRAZOS REGIMENTAIS, ESTABELECE UM PRAZO MÁXIMO DE 30 DIAS PARA QUE SEJAM FEITAS AS ALTERAÇÕES EXIGIDAS. APÓS O DEPÓSITO FINAL DO DOCUMENTO DE DISSERTAÇÃO, DEVIDAMENTE REVISADO E MEDIANTE ATESTADO DO ORIENTADOR, O CONCEITO “EM EXIGÊNCIA” PASSARÁ IMEDIATAMENTE PARA O DE “APROVADO”. NA FORMA REGULAMENTAR, FOI LAVRADA A PRESENTE ATA, QUE É ASSINADA POR MIM, FLÁVIO PEREIRA DA CUNHA, SECRETÁRIO, ALUNO E OS MEMBROS DA COMISSÃO EXAMINADORA PRESENTES.**

4. **CAMPINA GRANDE, 23 DE FEVEREIRO DE 2024**


5.

6.

7.

8.

9. **GUSTAVO BARBOSA LIMA DA SILVA – EXAMINADOR EXTERNO**

Documento assinado digitalmente
 GUSTAVO BARBOSA LIMA DA SILVA
Data: 26/02/2024 14:22:40-0300
Verifique em <https://validar.iti.gov.br>



Documento assinado eletronicamente por **ANDREA CARLA LIMA RODRIGUES, PROFESSOR(A) DO MAGISTERIO SUPERIOR**, em 23/02/2024, às 12:42, conforme horário oficial de Brasília, com fundamento no art. 8º, caput, da [Portaria SEI nº 002, de 25 de outubro de 2018](#).



Documento assinado eletronicamente por **MARILIA MARCY CABRAL DE ARAUJO, PROFESSOR 3 GRAU**, em 23/02/2024, às 14:15, conforme horário oficial de Brasília, com fundamento no art. 8º, caput, da [Portaria SEI nº 002, de 25 de outubro de 2018](#).



Documento assinado eletronicamente por **Elis Gean Rocha, Usuário Externo**, em 23/02/2024, às 14:20, conforme horário oficial de Brasília, com fundamento no art. 8º, caput, da [Portaria SEI nº 002, de 25 de outubro de 2018](#).



Documento assinado eletronicamente por **Gabriele de Souza Batista, Usuário Externo**, em 23/02/2024, às 15:34, conforme horário oficial de Brasília, com fundamento no art. 8º, caput, da [Portaria SEI nº 002, de 25 de outubro de 2018](#).



Documento assinado eletronicamente por **PATRICIA HERMINIO CUNHA FEITOSA, PROFESSOR 3 GRAU**, em 23/02/2024, às 18:06, conforme horário oficial de Brasília, com fundamento no art. 8º, caput, da [Portaria SEI nº 002, de 25 de outubro de 2018](#).



Documento assinado eletronicamente por **FLAVIO PEREIRA DA CUNHA, SECRETÁRIO (A)**, em 26/02/2024, às 07:34, conforme horário oficial de Brasília, com fundamento no art. 8º, caput, da [Portaria SEI nº 002, de 25 de outubro de 2018](#).



A autenticidade deste documento pode ser conferida no site <https://sei.ufcg.edu.br/autenticidade>, informando o código verificador **4164528** e o código CRC **D5F0568C**.

AGRADECIMENTOS

A Deus por seu amor incondicional e misericórdia e a Virgem Maria por rogar por mim ao Pai nos momentos mais difíceis.

A minha mãe, Claudete, verdadeiramente o maior exemplo da minha vida, que apoiou com muito amor meus estudos e sonhos, sempre lutando para me proporcionar um futuro digno. Obrigada por sempre estar presente, como mãe e pai, e me ensinar a não desistir.

A minha avó, Bernadete, por ter sido tão presente em minha criação, dedicando-se a cuidar de mim com amor e paciência. Ao meu avô (*in memoriam*), Sebastião, pelo seu companheirismo e por sempre se orgulhar de minhas conquistas, você me traz muita saudade.

Ao meu namorado e melhor amigo, Mateus, por ser meu cúmplice e companheiro fiel desde a graduação. Seu amor, sua paciência e sua dedicação em me ajudar foram de grande importância para o desenvolvimento desse trabalho. Este trabalho reflete não apenas meu esforço, mas também a nossa colaboração e sinergia.

A todos os meus amigos, que de perto ou de longe, torciam e me incentivavam a seguir. A amizade de vocês coloriu os dias de estudo e pesquisa com momentos de descontração e alegria, tornando o percurso mais leve e prazeroso.

A minhas orientadoras Andrea e Marília, pelo auxílio, confiança e ensino durante este percurso. A dedicação de Andrea em assegurar a excelência e qualidade do trabalho foi evidente em cada reunião, refletindo no resultado dessa dissertação. À Marília, meu sincero agradecimento por acreditar em meu potencial, me incentivar a prosseguir e me auxiliar com suas correções, palavras e ensinamentos.

A banca avaliadora, composta por Elis, Patricia e Gustavo, por aceitarem o convite e pelas contribuições que enriqueceram e aprimoraram este trabalho.

A toda a equipe envolvida na elaboração dos Planos Municipais de Saneamento Básico (PMSB) e dos Planos de Segurança da Água (PSA), onde participei como engenheira e pesquisadora, por todo conhecimento compartilhado e pela ajuda nessa pesquisa. Em especial à equipe do PSA, que me auxiliou no levantamento dos dados utilizados nessa dissertação. Vocês tornaram nossas viagens de trabalho ao Ceará mais leves, agradáveis e inesquecíveis.

Enfim, a todos que contribuíram para a concretização deste sonho. Muito obrigada!

RESUMO

A desinfecção com cloro visa garantir a integridade sanitária da água e a prevenção de doenças de veiculação hídrica. No entanto, manter as concentrações de cloro residual livre (CRL) é desafiador na gestão de sistemas de abastecimento de água, pois sua concentração pode decair ao longo da rede de distribuição, prejudicando a qualidade da água e a saúde da população. Para compreender o comportamento do CRL ao longo das redes de distribuição, assegurando um desempenho satisfatório da desinfecção, faz-se necessário modelar o seu decaimento. Utilizando os dados de qualidade da água do Sistema de Informação de Vigilância da Qualidade da Água para Consumo Humano (Sisagua) e o método de seleção de variáveis t-Distributed Stochastic Neighbor Embedding (t-SNE), desenvolveu-se uma rede neural artificial (RNA), para avaliação do decaimento do CRL em redes de distribuição de água de comunidades rurais brasileiras. O modelo desenvolvido foi validado com os dados provenientes de amostras de nove Sistemas Simplificados de Abastecimento de Água (SSAA) de comunidades rurais do estado do Ceará. As redes de distribuição dessas localidades foram avaliadas quanto ao atendimento do CRL à Portaria GM/MS nº 888/2021, a partir de indicadores de desempenho e do Índice de Sustentabilidade (IS). Os resultados obtidos evidenciam a relevância dos parâmetros CRL no reservatório, turbidez, pH, comprimento da tubulação e temperatura do dia na previsão do cloro no sistema de distribuição. A RNA desenvolvida a partir dessas variáveis de entrada apresentou elevadas taxas de acerto variando de 85,45% a 99,84%, para os limiares de precisão do método colorimétrico, comumente utilizado em sistemas de abastecimento rurais. Em comparação aos métodos numéricos e computacionais, o modelo pode ser aplicado a qualquer sistema de abastecimento rural, sem a necessidade de determinação da ordem cinética da reação ou estimativa do coeficiente de decaimento de CRL. A aplicação do RNA nos SSAA das comunidades rurais indicou que a robustez do modelo é reduzida para sistemas com captação em manancial superficial, existência de etapa de filtração e reforços de cloro, enquanto obtém resultados satisfatórios para os SSAA com captação por poços, chegando a 100,00% de precisão para o limiar de 0,5 mg/L de CRL. A análise do desempenho das redes de distribuição de água revelou a predominância de uma performance inadequada quanto a concentração do CRL (55,56% dos sistemas), indicando a necessidade de aprimoramento das estratégias de gestão do serviço. De maneira geral, a abordagem desenvolvida oferece uma previsão antecipada de inconformidades relacionadas ao cloro residual livre, além de auxiliar nos ajustes na dosagem do cloro.

Palavras-chave: Qualidade da água; Desinfecção com cloro; Aprendizagem de máquina; t-SNE; Índice de sustentabilidade.

ABSTRACT

Disinfecting water with chlorine aims to guarantee the sanitary integrity of the water and the prevention of waterborne diseases. However, maintaining free residual chlorine (FRC) concentrations is challenging in the management of water supply systems, as its concentration can decline throughout the distribution network, harming water quality and the health of the population. To understand the behavior of FRC throughout distribution networks, ensuring satisfactory disinfection performance, it is necessary to model its decay. Using water quality data from the Sistema de Informação de Vigilância da Qualidade da Água para Consumo Humano (Sisagua) and the t-Distributed Stochastic Neighbor Embedding (t-SNE) variable selection method, an artificial neural network (ANN) was developed, to evaluate the decay of FRC in water distribution networks in Brazilian rural communities. The developed model was validated with data from samples from nine Simplified Water Supply Systems (SWSS) from rural communities in the state of Ceará. The distribution networks of these locations were evaluated regarding FRC compliance with Ordinance No. 888/2021, based on performance indicators and the Sustainability Index (SI). The results obtained highlight the relevance of the FRC parameters in the reservoir, turbidity, pH, pipe length and daytime temperature in predicting chlorine in the distribution system. The ANN developed from these input variables showed high accuracy rates ranging from 85.45% to 99.84%, for the accuracy thresholds of the colorimetric method, commonly used in rural supply systems. Compared to numerical and computational methods, the model can be applied to any rural supply system, without the need to determine the kinetic order of the reaction or estimate the FRC decay coefficient. The application of ANN in the SWSS of rural communities indicated that the robustness of the model is reduced for systems with capture from surface water sources, the existence of a filtration stage and chlorine reinforcements, while obtaining satisfactory results for the SWSS with capture from wells, reaching 100.00% accuracy for the 0.5 mg/L FRC threshold. Analysis of the performance of water distribution networks revealed the predominance of inadequate performance in terms of FRC concentration (55.56% of systems), indicating the need to improve service management strategies. In general, the approach developed offers an early prediction of nonconformities related to free residual chlorine, in addition to assisting in adjustments in chlorine dosage.

Key words: Water quality; Chlorine-disinfection; Machine learning; t-SNE; Sustainability Index.

LISTA DE TABELAS

Tabela 1: Principais agentes químicos de desinfecção e suas características.....	25
Tabela 2: Caracterização dos sistemas de abastecimento de água	55
Tabela 3: Análise descritiva das variáveis quantitativas antes e depois do pré-processamento dos dados	58
Tabela 4: Análise descritiva dos dados das variáveis de entrada do modelo a serem utilizados na etapa de aplicação do modelo	89
Tabela 5: Porcentagens de acerto do modelo para diferentes limiares de erro absoluto	91

LISTA DE QUADROS

Quadro 1: Aplicações de redes neurais artificiais na modelagem do decaimento do cloro	36
Quadro 2: Variáveis consideradas no estudo.....	49
Quadro 3: Codificação das variáveis qualitativas.....	50

LISTA DE FIGURAS

Figura 1: Etapas para potabilização da água em uma estação de tratamento convencional ou de ciclo completo.....	24
Figura 2: Efeito do pH na quantidade relativa de ácido hipocloroso (HOCl) e íon hipoclorito (OCl ⁻) a 20°C.....	29
Figura 3: Curva de <i>breakpoint</i>	30
Figura 4: Reações na massa de água e parede das tubulações que propiciam o consumo de cloro residual livre.....	31
Figura 5: Modelo de RNA com arquitetura do tipo MLP com quatro camadas: uma camada de entrada, duas camadas ocultas e uma camada de saída.....	40
Figura 6: Exemplo de aplicação do t-SNE para diferentes conjuntos de variáveis.....	42
Figura 7: Fluxograma metodológico do estudo.....	46
Figura 8: Distribuição dos sistemas de abastecimento de água rurais com dados registrados no Sisagua por unidade federativa (UF).....	48
Figura 9: Localização das comunidades rurais em estudo.....	54
Figura 10: Classificação de desempenho das redes de distribuição conforme resultado do Índice de Sustentabilidade.....	57
Figura 11: Distribuição de amostras pré-processadas por tipo de captação e existência de filtração.....	60
Figura 12: Distribuição das amostras com cloro residual livre na rede de distribuição de áreas rurais por unidade federativa (UF), entre os anos de 2014 e 2023: a) dados brutos; b) dados pré-processados.....	61
Figura 13: Intervalos de dados para cada categoria considerada no t-SNE.....	63
Figura 14: Fluxograma com as principais combinações testadas.....	64
Figura 15: Testes do t-SNE com combinações do tipo A, entre os parâmetros físico-químicos e a variável temperatura do dia.....	65
Figura 16: Testes do t-SNE com combinações do tipo B, entre os parâmetros físico-químicos, temperatura do dia e economias residenciais.....	67
Figura 17: Testes do t-SNE com combinações do tipo C: a) parâmetros físico-químicos, temperatura do dia e variáveis qualitativas; b) parâmetros físico-químicos, temperatura do dia, economias residenciais e variáveis qualitativas.....	68
Figura 18: Hiperparâmetros definidos para a rede neural artificial.....	69
Figura 19: Erro médio quadrático por repetição para a RNA treinada por EMQ.....	70

Figura 20: Comparação entre os valores preditos e observados de cloro residual livre para a RNA treinada por EMQ.....	71
Figura 21: Função linear de ponderação do erro médio quadrático (EMQ).....	72
Figura 22: Erro médio quadrático por repetição para a RNA treinada por EMQP	72
Figura 23: Comparação entre os valores preditos e observados de cloro residual livre para a RNA treinada por EMQP	73
Figura 24: Porcentagem média de acertos para diferentes limiares de erro absoluto para RNA treinada por EMQ e EMQP	74
Figura 25: Kit comparador de cloro e pH empregado na medição de cloro residual livre.....	76
Figura 26: Mapa da rede de distribuição e localização das amostras do Sistema Lagoa dos Encantos.....	78
Figura 27: Mapa da rede de distribuição e localização das amostras do Sistema Lagoa do Juá	79
Figura 28: Mapa da rede de distribuição e localização das amostras do Sistema Venâncio	80
Figura 29: Mapa da rede de distribuição e localização das amostras do Sistema Carrapateira	81
Figura 30: Mapa da rede de distribuição e localização das amostras do Sistema Realejo	82
Figura 31: Mapa da rede de distribuição e localização das amostras do Sistema Tabocas	83
Figura 32: Mapa da rede de distribuição e localização das amostras do Sistema Acarape.....	84
Figura 33: Mapa da rede de distribuição e localização das amostras do Sistema Calumbi.....	85
Figura 34: Mapa da rede de distribuição e localização das amostras do Sistema Santo Antônio	86
Figura 35: Comparação entre os valores preditos e observados de cloro residual livre para a aplicação da RNA com os dados dos sistemas individualmente.....	92
Figura 36: Comparação entre os valores preditos e observados de cloro residual livre para a aplicação da RNA com os dados dos sistemas com captação em manancial superficial e subterrâneo.....	95
Figura 37: Comparação entre os valores preditos e observados de cloro residual livre para a aplicação da RNA com os dados de todos os sistemas.....	97
Figura 38: Porcentagens de acertos para diferentes limiares de erro absoluto para aplicação do modelo	98
Figura 39: Indicador de confiabilidade para as redes de distribuição de água	100
Figura 40: Indicador de resiliência para as redes de distribuição de água.....	101
Figura 41: Indicador de vulnerabilidade para as redes de distribuição de água	102
Figura 42: Índice de Sustentabilidade para as redes de distribuição de água.....	103

LISTA DE SIGLAS E ABREVIATURAS

ANFIS	<i>Adaptive Neuro-Fuzzy Inference System</i>
BNN	<i>Bayesian Neural Networks</i>
CAGECE	Companhia de Água e Esgoto do Estado do Ceará
CFNN	<i>Cascade Feed Forward Neural Network</i>
Conama	Conselho Nacional do Meio Ambiente
CRC	Cloro Residual Combinado
CRL	Cloro Residual Livre
DIVS	Diretoria de Vigilância Sanitária do Estado de Santa Catarina
EMA	Erro Médio Absoluto
EMQ	Erro Médio Quadrático
EMQP	Erro Médio Quadrático Ponderado
ETA	Estação de Tratamento de Água
FFNN	<i>Feed Forward Neural Network</i>
Funasa	Fundação Nacional de Saúde
GESAR	Gerência de Saneamento Rural
GM	Gabinete Ministerial
GRNN	<i>Generalized Regression Neural Network</i>
HAAS	Ácidos acéticos halogenados
IA	Inteligência Artificial
IBGE	Instituto Brasileiro de Geografia e Estatística
INMET	Instituto Nacional de Meteorologia
IS	Índice de Sustentabilidade
MLP	<i>Multilayer Perceptron</i>
MS	Ministério da Saúde
NTU	Unidade Nefelométrica de Turbidez
ODS	Objetivos de Desenvolvimento Sustentável
ONU	Organização das Nações Unidas
pH	Potencial Hidrogeniônico
PRNN	<i>Pattern Recognition Neural Network</i>
PSA	Plano de Segurança da Água
PVC	Policloreto de Vinila
RBNN	<i>Radial Basis Neural Network</i>
RNA	Rede Neural Artificial
SAA	Sistema de Abastecimento de Água
SDT	Sólidos Dissolvidos Totais

Sisagua	Sistema de Informação de Vigilância da Qualidade da Água para Consumo Humano
Sisar	Sistema Integrado de Saneamento Básico do Ceará
SNIS	Sistema Nacional de Informações sobre Saneamento
SNN	<i>Standard (deterministic) Neural Networks</i>
SOM	<i>Kohonen's Self-Organizing Maps</i>
SSAA	Sistema Simplificado de Abastecimento de Água
TED	Termo de Execução Descentralizada
THM	Trihalometanos
t-SNE	<i>t-Distributed Stochastic Neighbor Embedding</i>
UFCEG	Universidade Federal de Campina Grande
Vigiagua	Programa Nacional de Vigilância da Qualidade da Água para Consumo Humano
VMP	Valor Máximo Permitido

LISTA DE SÍMBOLOS

\hat{y}_i	Valor previsto pelo modelo
B	Viés
C	Confiabilidade
Ca(OCl) ₂	Hipoclorito de cálcio
CaOCl	Cal clorada
Cl ₂	Gás cloro
CRL ₀	Valor observado de cloro residual livre
d(j)	Valor do déficit
F	Função de ativação
HOCl	Ácido hipocloroso
k _b	Coefficiente de velocidade de degradação de desinfetante
L	Padrão de potabilidade conforme a Portaria nº 888/2021
NaOCl	Hipoclorito de sódio
NCl ₃	Tricloramina
NH ₂ Cl	Monocloramina
NH ₃	Amônia
NHCl ₂	Dicloramina
OCl ⁻	Íon hipoclorito
p	Pesos aplicados ao erro médio quadrático
p	Perplexidade
p _{ij}	Probabilidade para os pontos no espaço de maiores dimensões, calculada através da distribuição normal
q _{ij}	Probabilidade para os pontos no espaço de menores dimensões, calculada através da distribuição <i>t-Student</i>
R	Resiliência
R ²	Coefficiente de determinação
t	Período de tempo
t ₁	Hora da coleta da amostra no reservatório
t ₂	Hora da coleta da amostra na rede de distribuição
t _d	Número/Frequência dos déficits
t _r	Número de vezes que um valor satisfatório segue um valor insatisfatório
V	Vulnerabilidade
V _{adm}	Vulnerabilidade adimensional
V _m	Velocidade média
W	Pesos que conectam os neurônios
w _i	Coefficiente de ponderação

X	Variáveis do modelo
x	Dados da amostra para a variável
Y	Resultado da saída da RNA
y_i	Valor real
z	<i>z-score</i>
Δs	Comprimento da tubulação
μ	Média amostral
σ	Desvio padrão amostral

SUMÁRIO

1. INTRODUÇÃO	19
1.1 OBJETIVOS	22
1.1.1 Objetivo Geral	22
1.1.2 Objetivos Específicos	22
2. FUNDAMENTAÇÃO TEÓRICA	23
2.1 TRATAMENTO DA ÁGUA EM SISTEMAS DE ABASTECIMENTO	23
2.2 PROCESSOS DE DESINFECÇÃO	24
2.3.1 O processo de desinfecção por cloro e a legislação brasileira	26
2.3 O CLORO NO PROCESSO DE DESINFECÇÃO DA ÁGUA	27
2.4 DECAIMENTO DE CLORO EM REDES DE DISTRIBUIÇÃO DE ÁGUA	30
2.5 MODELAGEM DO DECAIMENTO DE CLORO EM SISTEMAS DE ABASTECIMENTO	32
2.6 TÉCNICAS DE APRENDIZADO DE MÁQUINA APLICADAS À MODELAGEM DO DECAIMENTO DO CLORO	34
2.6.1 Redes Neurais Artificiais (RNA)	39
2.6.2 t-Distributed Stochastic Neighbor Embedding (t-SNE)	41
2.7 AVALIAÇÃO DE DESEMPENHO DE REDES DE DISTRIBUIÇÃO DE ÁGUA	43
3. METODOLOGIA	46
3.1 COLETA DOS DADOS	47
3.1.1 Delimitação da área de estudo	47
3.1.2 Coleta e pré-processamento dos dados	49
3.2 DESENVOLVIMENTO DA REDE NEURAL ARTIFICIAL (RNA)	51
3.2.1 Seleção dos parâmetros de entrada	51
3.2.2 Treinamento, validação e teste	52
3.3 SIMULAÇÃO DO MODELO	53
3.3.1 Caracterização da área de estudo	53
3.3.2 Aplicação do modelo desenvolvido	56
3.4 AVALIAÇÃO DE DESEMPENHO DAS REDES DE DISTRIBUIÇÃO DE ÁGUA	56
4. RESULTADOS E DISCUSSÕES	58
4.1 COLETA E PRÉ-PROCESSAMENTO DOS DADOS	58
4.2 DESENVOLVIMENTO DA REDE NEURAL ARTIFICIAL (RNA)	63
4.2.1 Seleção das variáveis de entrada do modelo	63

4.2.2 Treinamento, validação e teste	69
4.3 SIMULAÇÃO DO MODELO	76
4.3.1 Caracterização dos sistemas de distribuição de água estudados no estado do Ceará	76
4.3.2 Aplicação do modelo desenvolvido.....	91
4.4 AVALIAÇÃO DE DESEMPENHO DAS REDES DE DISTRIBUIÇÃO DE ÁGUA	99
5. CONCLUSÕES.....	105
5.1 RECOMENDAÇÕES	107
REFERÊNCIAS.....	108
APÊNDICES	122
APÊNDICE A.....	122
APÊNDICE B	124

1. INTRODUÇÃO

Historicamente, as políticas públicas do setor de saneamento básico, que envolvem os serviços de abastecimento de água, priorizam as áreas urbanas em detrimento das rurais. Em 2019, cerca de 13,1% dos domicílios brasileiros estavam localizados na zona rural. Neste mesmo ano, aproximadamente, 71,3% dos domicílios rurais do país eram abastecidos com água por rede de distribuição, poço ou nascente, enquanto 97,8% das residências urbanas eram atendidas pelo serviço (Brasil, 2022). Ainda que o índice de atendimento rural possuísse um percentual elevado, cabe ressaltar que apenas 30,6% desse valor refere-se aos domicílios com rede de distribuição de água (SNIS, 2022).

De acordo com os dados do censo demográfico, em 2010, o estado do Ceará abrigava 6,9% dos domicílios rurais brasileiros, dos quais cerca de 33,7% estavam ligados a redes de abastecimento de água (IBGE, 2011). Na região Nordeste, o estado destaca-se na busca pela universalização do acesso a água para as comunidades rurais, a partir da atuação do Sistema Integrado de Saneamento Básico do Ceará (Sisar-CE). O Sisar é uma organização não governamental, sem fins lucrativos, que colabora com a ampliação da cobertura das redes de distribuição e auxilia no processo de gestão comunitária dos sistemas de abastecimento de água de áreas rurais (Camargo, 2022).

Garantir disponibilidade e gestão sustentável da água para todas as pessoas até o ano de 2030 é um dos objetivos de desenvolvimento sustentável (ODS 6) estabelecidos pela Organização das Nações Unidas (ONU). Nesse sentido, a gestão comunitária da água auxilia na mitigação das dificuldades relacionadas a condições precárias de operação, manutenção e monitoramento de sistemas de abastecimento de água (SAA) de comunidades rurais (Santos; Santana, 2020).

Nas zonas rurais, a escassez hídrica e os conflitos pelos usos da água são problemas pertinentes, que fortalecem os déficits na relação oferta e demanda de recursos hídricos em SAA (Cárcere; Maia-Rodrigues, 2019). Além da importância da disponibilidade hídrica em quantidade suficiente à população, também é necessário que a sua qualidade atenda aos padrões de potabilidade estabelecidos pela Portaria nº 888/2021 do Ministério da Saúde (MS), que altera o Anexo XX da Portaria de Consolidação nº 5/2017. O consumo de água tratada assegura a qualidade de vida das comunidades, por meio da redução de doenças de veiculação hídrica, aumento do desempenho escolar de crianças e da produtividade dos trabalhadores.

Nesse contexto, o cloro é o principal desinfetante empregado no tratamento de água, devido à sua simplicidade de aplicação, ao custo reduzido, ao elevado potencial de remoção de patógenos e à possibilidade de preservar um residual livre até as extremidades da rede. Entretanto, ao longo do percurso pela rede de distribuição, devido a fatores como a idade da tubulação, vazamentos e a distância entre a unidade de tratamento e os pontos de consumo, o cloro reage com substâncias presentes na água e nas paredes das tubulações, podendo ter sua capacidade de desinfecção reduzida ou eliminada (Silva *et al.*, 2019).

Para garantir a segurança microbiológica da água até o consumidor final, o padrão de potabilidade prevê a manutenção de cloro residual livre na água fornecida de no mínimo 0,2 mg/L e no máximo 5,0 mg/L (Brasil, 2021). Em estações de tratamento de água, principalmente em áreas rurais, os operadores que determinam a dosagem de cloro com base em sua experiência, o que pode acarretar ao uso de doses excessivas ou insuficientes. Desta forma, prever e monitorar o processo de decaimento do cloro residual livre ao longo da rede de distribuição é de fundamental importância para auxiliar a gestão do SAA em comunidades rurais e no processo de controle e vigilância da qualidade da água (Kote; Wadkar, 2019).

O decaimento do cloro residual livre é comumente modelado a partir da cinética de primeira ordem simples, sendo os parâmetros cinéticos, geralmente, determinados de forma laboratorial, a partir de campanhas de coletas e análises químicas. Em pequenos municípios, áreas rurais e comunidades tradicionais, existem dificuldades mais acentuadas quanto ao controle e vigilância da qualidade da água, devido à falta de equipamentos e manutenção adequada, o baixo número de técnicos especializados e à limitação de acesso aos laboratórios de referência (Bellotti *et al.*, 2020).

Modelos matemáticos e computacionais têm sido desenvolvidos e incorporados aos sistemas de distribuição para prever o decaimento do cloro ao longo da rede, entre eles modelos cinéticos, rede neural artificial, simuladores numéricos e simuladores hidráulicos, a exemplo dos *softwares* EPANET, AQUASIM e WaterCAD. Conforme Rodrigues e Scalize (2019), existe uma predominância de estudos voltados à modelagem numérica baseada em relações de taxa química fundamental para gerar modelos cinéticos empíricos. Apesar do bom desempenho desses modelos, a necessidade de determinação de um coeficiente de decaimento do cloro para cada aplicação, impede sua replicação em sistemas de abastecimento de água com características diferentes.

Frente a isso, as Redes Neurais Artificiais (RNAs) têm se mostrado eficientes, devido à sua concepção baseada na maneira como o cérebro humano trabalha, adquirindo conhecimento

através de processo de aprendizagem. A utilização de RNAs é vantajosa por não requerer suposições sobre a ordem do modelo ou a estimativa de um coeficiente de decaimento do cloro. As RNAs também podem ser treinadas com dados que representam uma ampla gama de condições operacionais e podem ser retreinadas facilmente com novos dados, permitindo sua aplicação em diversos tipos de sistemas (Santi *et al.*, 2021).

Como modelos estatísticos, as RNAs permitem a previsão de relações empíricas complexas entre o cloro residual livre e os parâmetros que influenciam seu decaimento, sem a necessidade de obtenção de dados extensos e precisos sobre o cloro dentro do sistema de distribuição de água (Onythua, 2022). Na seleção dessas variáveis de entrada, a metodologia adotada deve assegurar o máximo desempenho do modelo, evitando parâmetros redundantes. Conforme o estudo de Zhou *et al.* (2020), destaca-se o método *t-Distributed Stochastic Neighbor Embedding* (t-SNE), proposto por Van der Maaten e Hinton (2008), que permite reduzir a dimensionalidade dos dados e identificar quais os parâmetros são capazes de explicar o decaimento do cloro na RNA com melhor desempenho.

Vários estudos já aplicaram as RNAs para a modelagem do decaimento do cloro, destacando Al-Araimi *et al.* (2019), Kote e Wadkar (2019) e Librantz *et al.* (2018). No entanto, normalmente essas pesquisas estão direcionadas a sistemas de médio a grande porte, instalados em áreas urbanas, deixando uma lacuna quanto a avaliação do método em sistemas de distribuição de comunidades rurais. Nessas áreas, há uma grande dispersão das unidades habitacionais, com a presença de algumas residências próximas ao ponto de cloração, onde ocorre o risco de sobredosagem do cloro, e outras bastante distantes, gerando um maior tempo de percurso da água e a possibilidade de sub-dosagem.

Assim, a presente pesquisa visa desenvolver um modelo de decaimento de cloro residual livre em redes de distribuição de sistemas de abastecimento de comunidades rurais, por meio da aplicação de redes neurais artificiais (RNA). O modelo pretende reduzir os problemas relacionados à necessidade de determinação de um coeficiente de decaimento do cloro para cada aplicação e, conseqüentemente, a dificuldade de replicação em sistemas de abastecimento de água com características diferentes. O sistema proposto ainda permitirá aos responsáveis pelo SAA a previsão antecipada a possíveis inconformidades relacionadas aos padrões de potabilidade vigentes no país. Dessa forma, será possível auxiliar no processo de ajustes na dosagem do cloro e a tomada de decisão quanto à necessidade de instalação de pontos de reforço de cloração.

1.1 OBJETIVOS

1.1.1 Objetivo Geral

Propor, por meio de RNAs, um modelo de decaimento de cloro residual livre em redes de distribuição de água de sistemas de abastecimento de comunidades rurais.

1.1.2 Objetivos Específicos

- Determinar os parâmetros de entrada relevantes na modelagem de decaimento de cloro residual livre a partir da utilização do método *t-Distributed Stochastic Neighbor Embedding* (t-SNE);
- Estudar o comportamento e a robustez do modelo de decaimento de cloro residual livre quando aplicado em redes de distribuição de água de comunidades rurais do estado do Ceará;
- Avaliar o desempenho das redes de distribuição de água de comunidades rurais do estado do Ceará, quanto ao atendimento do cloro residual livre à Portaria de Potabilidade nº 888/2021.

2. FUNDAMENTAÇÃO TEÓRICA

2.1 TRATAMENTO DA ÁGUA EM SISTEMAS DE ABASTECIMENTO

O sistema de abastecimento de água (SAA) para consumo humano consiste em uma instalação composta por um conjunto de infraestruturas, obras civis, materiais e equipamentos, que se destina à produção e ao fornecimento coletivo de água potável. Em geral, um SAA é composto das seguintes unidades: captação, adução, tratamento, reservação, rede de distribuição, estações elevatórias e ramal predial (Brasil, 2019, 2020, 2021).

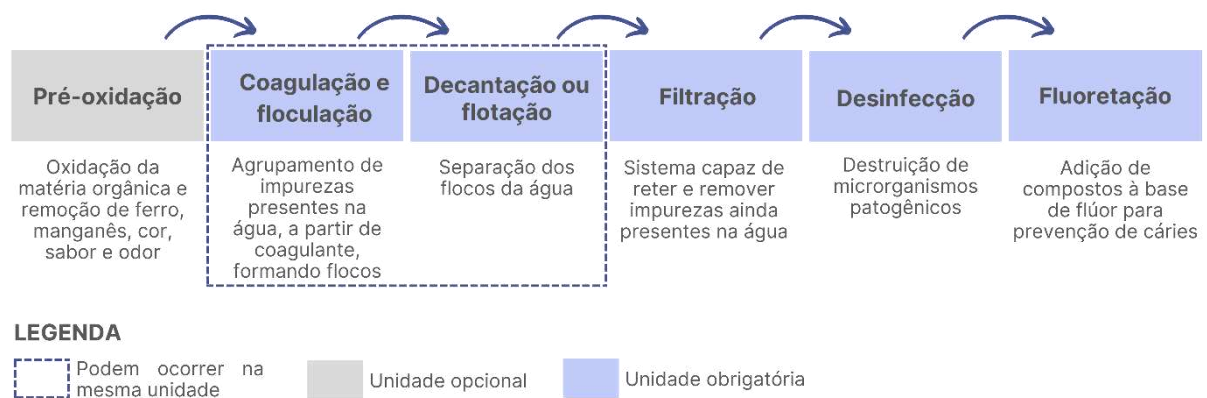
O abastecimento de água é um dos eixos do saneamento básico no Brasil, conforme consta na Lei Federal nº 11.445/2007, atualizada pela Lei nº 14.026/2020. O fornecimento de água adequado, em quantidade e qualidade, é primordial para o desenvolvimento socioeconômico das comunidades, com impactos diretos sobre as condições de saúde e de bem-estar da população. A implantação e gestão eficazes resultam em melhoria na qualidade de vida da população a partir do controle e prevenção de doenças, garantia da prática de hábitos higiênicos, redução da mortalidade infantil, aumento da expectativa de vida e da produtividade econômica (Damke; Pasini, 2020; Fortes *et al.*, 2019).

Entretanto, a qualidade da água em SAA está sujeita a ameaças desde a captação até a torneira do consumidor. Ao estudarem os riscos de contaminação da água, Corrêa (2020) e Pires (2020) apontam que, em se tratando dos mananciais, a água pode ser contaminada por meio de: excrementos de animais e/ou humanos; resíduos sólidos e seus lixiviados; escoamento de resíduos agrícolas (pesticidas e fertilizantes); fossas rudimentares e lançamento de esgoto não tratado. Sendo assim, é importante assegurar o tratamento dessas águas para o consumo humano, a fim de evitar que esses eventos perigosos contribuam para o aumento de ocorrências de doenças por veiculação hídrica (Manuel *et al.*, 2018; Pereira *et al.*, 2021).

A Estação de Tratamento de Água (ETA) é a unidade do sistema de abastecimento responsável pela adequação da água aos padrões de potabilidade estabelecidos pela Portaria nº 888/2021, melhorando suas características organolépticas, físicas, químicas e microbiológicas. A escolha da tipologia de tratamento a ser empregada têm como fatores definidores as questões técnicas, como as características das águas brutas que chegam ao sistema e fatores de questão econômica, como custos de implantação e operação, disponibilidade de terreno, qualidade de equipamentos e mão de obra que podem ser empregados (Aguiar *et al.*, 2021; Brasil, 2019; Michelin *et al.*, 2019).

O fluxograma da Figura 1 esquematiza as unidades de tratamento convencional ou de ciclo completo da água para consumo. Embora o tratamento convencional geralmente apresente eficiência satisfatória, possui elevado custo de implantação e operação, pois demanda grandes áreas para construção, elevado consumo energético e de produtos químicos. No caso das comunidades rurais brasileiras, nem sempre há recursos disponíveis para tecnologias de tratamento sofisticadas e, nestes casos, é importante considerar sistemas de tratamento econômicos e eficazes (Guedes *et al.*, 2019).

Figura 1: Etapas para potabilização da água em uma estação de tratamento convencional ou de ciclo completo



Fonte: Adaptado de Melo (2018) e Silva (2020).

Brasil (2020), Dias *et al.* (2021) e Oliveira (2021) destacam que os processos de filtração e desinfecção, especialmente por cloro, são amplamente utilizados no tratamento de águas para o abastecimento humano em comunidades rurais e tradicionais. Tais tecnologias devem cumprir o papel em SAA rurais, garantindo a eficiência de tratamento, a partir de remoções satisfatórias de agentes contaminantes e apresentando custos acessíveis. Em aplicações em que se deseja remover algum contaminante específico, esses sistemas podem ser combinados com outros processos de tratamento.

2.2 PROCESSOS DE DESINFECÇÃO

Em um sistema de tratamento de água, a desinfecção possui a função de eliminar ou inativar microrganismos patogênicos presentes na água, por intermédio de agentes físicos e/ou químicos. A ação dos desinfetantes sobre os microrganismos pode ocorrer pelos seguintes mecanismos: destruição da organização estrutural da célula, interferência no nível energético do metabolismo e interferência no crescimento do microrganismo (Brasil, 2019; Silva, 2020; Santos, 2020).

Nos processos físicos, utiliza-se a energia da radiação, de forma que a inativação dos microrganismos se dá a partir da ação do desinfetante sobre os ácidos nucleicos dos microrganismos, promovendo reações fotoquímicas que interrompem a reprodução dos vírus e bactérias. Destacam-se como parte desses processos a radiação ultravioleta, a radiação solar e a fervura (Brasil, 2019).

Dandie *et al.* (2020), Lechevallier *et al.* (2018) e Schorr (2022) explicam que as tecnologias de desinfecção por radiação são eficientes e seguras, tendo em vista a não utilização de produtos químicos, o que inibe a formação de subprodutos no processo. Entretanto, esses processos não têm ação residual na rede de distribuição, propiciando risco de recrudescimento de microrganismos patogênicos. Já a fervura, técnica utilizada em áreas desprovidas de sistema de tratamento de água e de redes de distribuição, também não permite residual e pode ocasionar uma segurança momentânea (Rodrigues, 2018; Silva, 2020).

Quando do emprego de agentes químicos, que apresentam caráter oxidante, o processo de desinfecção se dá pela destruição ou danificação da parede celular e posterior difusão do desinfetante no interior da célula. Os compostos químicos mais empregados no processo de desinfecção são o cloro e seus compostos, o dióxido de cloro e o ozônio (Brasil, 2019; Libânio, 2016). A Tabela 1 exibe o potencial de oxidação e as suas características.

Tabela 1: Principais agentes químicos de desinfecção e suas características

Composto químicos	Fórmula	Potencial de oxidação (V)	Vantagens	Desvantagens
Ozônio	O ₃	2,07	<ul style="list-style-type: none"> ▪ Menor tempo de contato; ▪ Não forma subprodutos de desinfecção. 	<ul style="list-style-type: none"> ▪ Não apresenta ação residual; ▪ Maior consumo de energia; ▪ Custos elevados e operação complexa.
Dióxido de Cloro	ClO ₂	1,91	<ul style="list-style-type: none"> ▪ Possibilita manutenção de concentração residual; ▪ Elevada solubilidade em água. 	<ul style="list-style-type: none"> ▪ Instável; ▪ Forma subprodutos como clorito e clorato; ▪ Custos elevados e operação complexa.
Cloro	Cl ₂	1,36	<ul style="list-style-type: none"> ▪ Amplo espectro de ação germicida; ▪ Deixa residual ativo na água; ▪ Custo razoável e facilidade de transporte, manuseio, armazenamento e aplicação. 	<ul style="list-style-type: none"> ▪ Podem gerar subprodutos, como trihalometanos, uma família de cancerígenos; ▪ Tempo de contato relativamente longo.

Fonte: Adaptado de Brasil (2020), Libânio (2016), Metcalf e Eddie (2016), Rodrigues (2018), Schorr (2022).

A escolha do processo de desinfecção mais adequado é dependente dos condicionalismos locais e da qualidade característica da água bruta, com o objetivo de atingir

o máximo desempenho com o menor custo, atender aos padrões de potabilidade e minimizar a formação de subprodutos da desinfecção (Libânio, 2016).

Apesar do potencial de oxidar moléculas biológicas ser um dos pré-requisitos essenciais para qualquer agente desinfetante, a eficiência do sistema de desinfecção depende de outras fatores, como: dose e tempo de exposição do agente; tipo e concentração dos microrganismos; características físicas e químicas da água; e, unidade de tratamento em que o processo ocorrerá (Libânio, 2016; Metcalf; Eddie, 2016).

Dentre as características físicas da água, Garay *et al.* (2021) e Khedikar *et al.* (2021) apontam que a turbidez influi diretamente sobre a eficiência da desinfecção, devido ao *efeito-escudo* que promove sobre os microrganismos, protegendo-os da ação do desinfetante físico ou químico. Léziart *et al.* (2019) verificaram menor remoção bacteriana de *Escherichia coli* e *Enterococcus faecalis* com compostos de cloro, quando a turbidez se elevava acima de 1 NTU. Além da turbidez, Libânio (2016) e Schorr (2022) apresentam outras características que interferem na eficiência, especialmente de desinfetantes químicos:

- Presença de matéria orgânica, frequentemente associada a cor verdadeira – aderindo aos microrganismos e protegendo-os da ação do desinfetante ou reagindo com este para formar subprodutos;
- Presença de compostos inorgânicos, como ferro e manganês, que reagem com o desinfetante;
- A temperatura da água e o pH, especialmente quando o cloro é empregado como agente desinfetante.

2.3.1 O processo de desinfecção por cloro e a legislação brasileira

No contexto do arcabouço legal brasileiro, relativo à qualidade da água, cabe destacar os seguintes normativos: a Lei Federal nº 14.026, de 15 de julho de 2020; a Resolução Conama nº 357, de 17 de março de 2005; a Resolução Conama nº 396, de 03 de abril de 2008; e, a Portaria GM/MS nº 888, de 04 de maio de 2021.

A Portaria GM/MS nº 888/2021, que altera o Anexo XX da Portaria de Consolidação nº 5/2017, é o principal e mais abrangente instrumento legal relativo à qualidade de água para consumo humano. Esta normativa é a responsável por definir os padrões de potabilidade da água no país e, diferentemente de suas antecessoras, tornou-se mais exigente quanto as restrições de 14 parâmetros e a ampliação de frequência e número de amostras para o monitoramento da qualidade (Brasil, 2021; Soares *et al.*, 2021).

Conforme a Portaria nº 888/2021, em seu Artigo 14, compete ao responsável pelo sistema de abastecimento de água: realizar o monitoramento e exercer o controle da qualidade para consumo humano. Em SAA de comunidades rurais, a gestão comunitária e compartilhada do sistema demonstra-se uma estratégia efetiva para alcançar a justiça e a segurança hídrica, bem como um método legítimo e efetivo de conservação ambiental. Entretanto, caso haja ineficiência de treinamento dos operadores para administração de substâncias de desinfecção, a utilização de doses elevadas de compostos clorados pode constituir risco à saúde da população (Brasil, 2021; Cárceres; Maia-Rodrigues, 2019; Rodrigues, 2018).

No Artigo 24 da Portaria nº 888/2021, destaca-se que toda água para consumo humano fornecida coletivamente deverá passar por processo de desinfecção ou adição de desinfetante para manutenção dos residuais mínimos. Mediante o comportamento do cloro e a sua predisposição a reagir com outros compostos presentes na massa de água, é fundamental o controle e vigilância da água distribuída, a fim de assegurar a proteção da qualidade e evitar a formação de subprodutos da desinfecção.

O Artigo 32 da mesma portaria estabelece que é obrigatória a manutenção de, no mínimo, 0,2 mg/L de cloro residual livre ou de 2,0 mg/L de cloro residual combinado, e no máximo 5,0 mg/L, em toda a extensão do sistema de distribuição (reservatórios e rede) e nos pontos de consumo, assim como, um teor mínimo de 0,5 mg/l de CRL em carros-pipas. Além disso, os Anexos 3 e 4 da portaria citada recomendam que o pH apresente valores entre 6,0 e 9,0 no processo de desinfecção (Brasil, 2021). Essa faixa de pH visa garantir o controle e a ação efetiva das substâncias cloradas durante o tratamento da água.

2.3 O CLORO NO PROCESSO DE DESINFECÇÃO DA ÁGUA

A desinfecção com cloro é comumente usada no tratamento de água para a destruição de patógenos, o controle de microrganismos incômodos, a remoção de ferro e manganês, prevenção de crescimento de algas, remoção de cor e o controle de sabor, odor e desenvolvimento de biofilmes em tubulações. Os principais compostos à base de cloro usados na desinfecção são hipoclorito de sódio (NaOCl), hipoclorito de cálcio (Ca(OCl)_2), cloro (Cl_2), cal clorada (CaOCl) e água sanitária (Brasil, 2019; Libânio, 2016; Stefán *et al.*, 2019; Zhang *et al.*, 2020).

A predominância do uso do cloro na desinfecção da água se dá pelo seu amplo espectro de ação germicida e pela sua capacidade de prevenir contaminação futura, a partir da formação de compostos que permanecem na água, proporcionando residual desinfetante ativo ao longo das

tubulações da rede de distribuição. Além disso, a cloração é um processo relativamente simples e econômico, tornando-o propício para utilização em sistemas de abastecimento mais simples, como os encontrados em comunidades rurais (Brasil, 2019; Libânio, 2016).

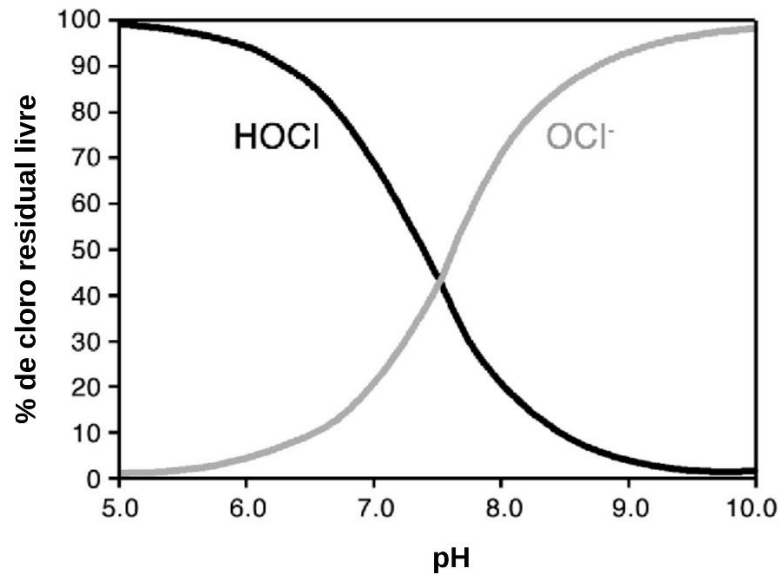
No entanto, em contato com a água, o cloro reage com várias substâncias orgânicas e inorgânicas, podendo formar subprodutos. Os subprodutos normalmente são gerados quando a concentração de cloro livre é alta e o tempo de permanência na rede de distribuição de água é longo. Considerando o risco para a saúde humana (toxicidade e carcinogenicidade) destacam-se os subprodutos: trihalometanos (THMs) e ácidos acéticos halogenados (HAAS) (Onyutha; Kwio-Tamale, 2022; Stefán *et al.*, 2019).

O cloro é um gás liquefeito sob pressão, de coloração verde amarelada e de odor forte. Quando injetado na água, na forma de Cl_2 , o cloro se dissolve formando o ácido hipocloroso (HOCl), conforme descrito na Equação 1. A formação desse ácido depende do pH da água, de forma que o HOCl pode se dissociar e formar o íon hipoclorito (OCl^-), que também é agente de desinfecção, mas que apresenta menor eficiência que o ácido hipocloroso (Equação 2). O somatório das frações de HOCl e OCl^- correspondem ao cloro residual livre (CRL), responsável pela oxidação da matéria orgânica, constituindo um importante indicador do controle e gestão da qualidade da água (Howe *et al.*, 2016; Libânio, 2016; Santos, 2018; Silva *et al.*, 2019).



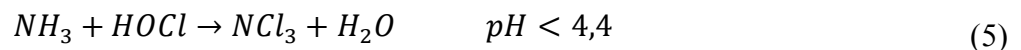
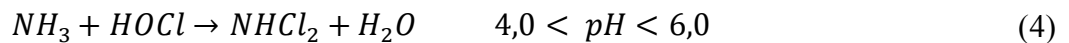
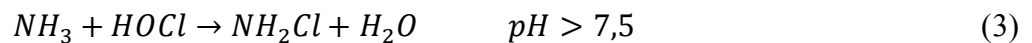
Cabe ressaltar a influência do pH nas reações químicas entre o cloro residual livre e a água, como exibida na Figura 2. O pH inferior a 5 contribui para a predominância do ácido hipocloroso, ao passo que valores de pH superiores a 10 contribuem para a predominância do íon hipoclorito. Assim, tratando-se do gás cloro, um pH de 7 ou menor é desejado visando a desinfecção. Embora o gás cloro (Cl_2) e o hipoclorito de sódio (NaClO) tenham capacidades de desinfecção idênticas, a adição do NaClO na água provoca um aumento adicional do pH e da alcalinidade (Howe *et al.*, 2016; Li *et al.*, 2019; Santos, 2018).

Figura 2: Efeito do pH na quantidade relativa de ácido hipocloroso (HOCl) e íon hipoclorito (OCl⁻) a 20°C

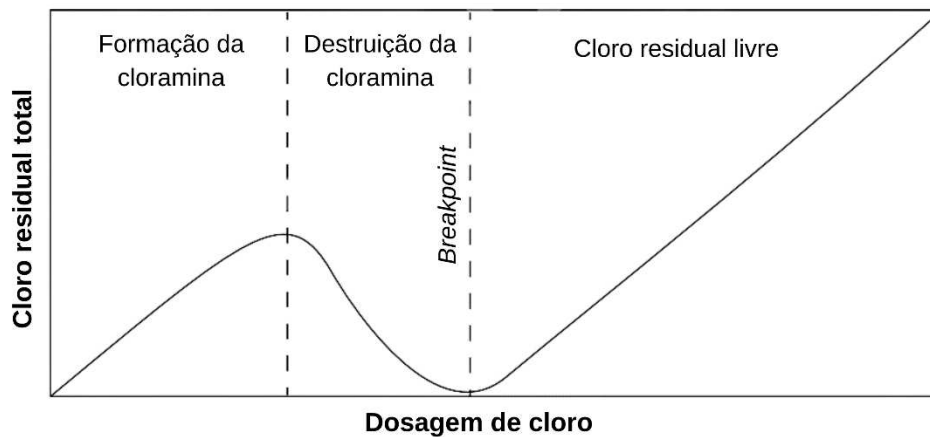


Fonte: Adaptado de Macedo *et al.* (2014).

Na presença de amônia e compostos amoniacais, o cloro livre reage formando compostos clorados ativos, conhecidos como cloraminas ou cloro residual combinado (CRC). O CRC pode se apresentar sob a forma de três espécies de cloramina, de acordo com o valor do pH do meio aquoso: monoclорamina (NH₂Cl), dicloramina (NHCl₂) e tricloramina (NCl₃); com reações descritas nas Equações 3 a 5, respectivamente (Foster, 2022).



Monitorar e controlar a produção de cloro residual combinado é importante pois a formação das cloraminas consome o cloro residual livre e conferem odor e sabor desagradáveis as águas de consumo. Dessa forma, é comum a prática de dosagem de cloro após o *breakpoint* (Figura 3), na qual o CRL é adicionado em uma quantidade suficiente tal que todos os contaminantes e o CRC são degradados dentro da água do sistema de distribuição de água, assegurando uma concentração de cloro residual livre para manutenção da qualidade da água ao longo da rede (Foster, 2022; Wang *et al.*, 2018).

Figura 3: Curva de *breakpoint*

Fonte: Adaptado de Foster (2022).

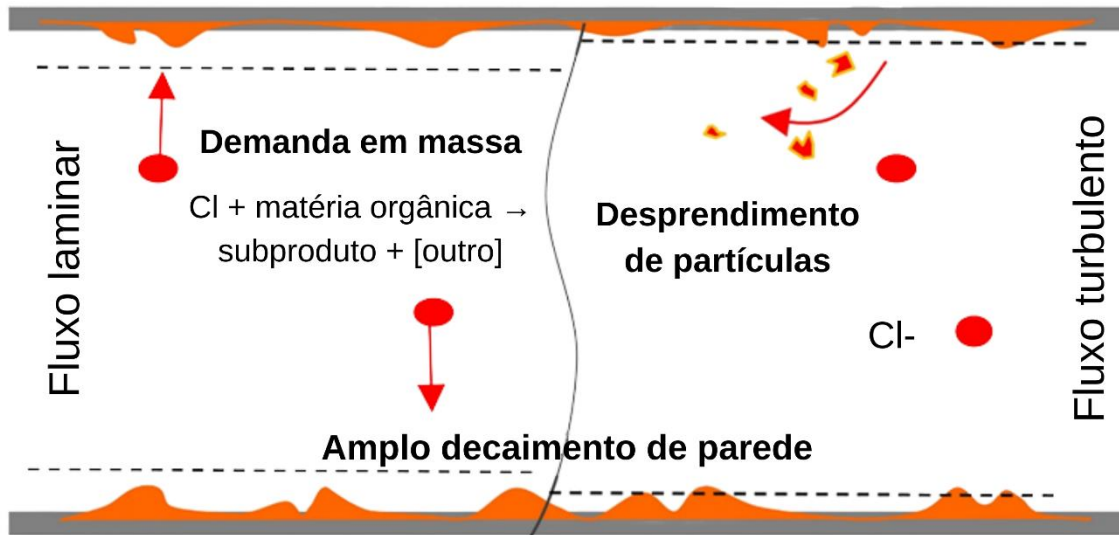
2.4 DECAIMENTO DE CLORO EM REDES DE DISTRIBUIÇÃO DE ÁGUA

No processo de desinfecção da água com o cloro é importante a garantia de uma quantidade residual a fim de que seja assegurada, em toda a rede, a integridade sanitária da água. Entretanto, o cloro não é seletivo, de modo que reage com outros compostos presentes na massa de água. Logo, a busca pela compreensão do seu comportamento e decaimento nas redes de distribuição de água mostra-se fundamental para assegurar a qualidade da água para o consumo humano (Silva *et al.*, 2019).

À medida que a água percorre o sistema de distribuição, a taxa de decaimento do cloro residual tende a aumentar, sendo sua concentração maior na saída da Estação de Tratamento da Água (ETA) e menor nas extremidades da rede, onde parte do residual já poderá ter sido totalmente consumida (Biondo, 2019; García-Ávila *et al.*, 2020; Oliveira *et al.*, 2021; Vargas, 2021).

A estabilidade da concentração de cloro livre ao longo da rede de distribuição está associada a uma série de fatores, como: tempo de percurso da água (engloba o comprimento das tubulações e o número de residências/economias/ligações); idade e material das tubulações; regimes hidrodinâmicos de fluxo nas tubulações; temperatura da água; e, presença de compostos orgânicos e inorgânicos na água (Mazhar *et al.*, 2020; Oliveira, 2018; Oliveira *et al.*, 2021). A ação conjunta desses fatores contribui para a redução da concentração residual de cloro, a partir da degradação em massa e do decaimento da parede (Figura 4).

Figura 4: Reações na massa de água e parede das tubulações que propiciam o consumo de cloro residual livre



Fonte: Adaptado de Zhao *et al.* (2018).

No fluxo laminar na tubulação, onde a trajetória das partículas é bem definida e constante, existe uma propensão para ocorrer reações entre o cloro e as substâncias presentes na água, como a matéria orgânica, o que caracteriza a demanda em massa. A partir da existência de um fluxo turbulento, há um desprendimento dos produtos da corrosão, depósitos e biofilmes que se acumularam ao longo do tempo nas paredes internas das tubulações, os quais reagem com o cloro, resultando no decaimento de parede (Hemdan *et al.*, 2021; Zhao *et al.*, 2018).

O crescimento de biofilmes nas tubulações de redes de distribuição de água pode levar à degradação e corrosão dos materiais que as compõem, facilitando a adesão de mais microrganismos; as alterações nas propriedades organolépticas da água; a redução da seção útil das tubulações; além da diminuição da concentração de desinfetante disponível para inativação de potenciais agentes patogênicos (Hemdan *et al.*, 2021; Vargas, 2021; Zhao *et al.*, 2018).

Em geral, embora a decomposição do cloro dentro de um SAA englobe as parcelas de degradação em massa e na parede, ressalta-se que o predomínio de uma parte em detrimento da outra depende das características do sistema e da qualidade da água bruta. Oliveira *et al.* (2022) e Yun-Hong *et al.* (2019) explicam que em águas com maiores concentrações de matéria orgânica e nutrientes, o decaimento no fluxo de água desempenha papel mais importante no consumo de cloro residual. Bhadula *et al.* (2021), Monteiro *et al.* (2020) e Shi *et al.* (2022) indicam que, no contexto do decaimento do cloro nas paredes dos tubos, as velocidades de fluxo lentas e o envelhecimento e corrosão das tubulações, principalmente daquelas fabricadas em

metal, são fatores que afetam significativamente a redução da concentração residual do desinfetante.

A manutenção das concentrações de cloro residual livre é um desafio para os operadores de água, uma vez que sua concentração pode ser alterada constantemente através da rede de distribuição de água. Caso seja manipulado de forma inadequada e por pessoal incapacitado, a sua dosagem pode apresentar riscos à população (Bhadula *et al.*, 2021; Oliveira, 2018; Tonev; Dimova, 2020).

A sobredosagem de cloro torna a água censurável em termos de sabor e odor e leva à produção de subprodutos de desinfecção que podem causar câncer, distúrbios reprodutivos e defeitos congênitos. Por outro lado, a sub dosagem de cloro pode propiciar a recontaminação da água por agentes microbianos, levando à infecção dos consumidores. Desse modo, para compreender o comportamento do cloro ao longo dos sistemas de abastecimento de água e poder garantir a sua correta dosagem, pode-se modelar o seu decaimento por meio de modelos e ferramentas de simulação (Almeida *et al.*, 2022; Biondo, 2019; Mazhar *et al.*, 2020; Onyutha; Kwio-Tamale, 2022).

2.5 MODELAGEM DO DECAIMENTO DE CLORO EM SISTEMAS DE ABASTECIMENTO

De acordo com Oliveira (2018) e Biondo (2019), a modelagem matemática da qualidade da água representa uma alternativa viável para que se possa conhecer e assegurar um desempenho satisfatório do desinfetante ao longo da rede. Nas últimas décadas, vários estudos foram realizados em diversos países sobre o decaimento do cloro na água potável e variados modelos foram propostos para descrever a cinética de decaimento desse desinfetante (Al-Araimi *et al.*, 2019; Almeida *et al.*, 2022; Biondo, 2019; García-Ávila *et al.*, 2021; Hossain *et al.*, 2021; Mentis *et al.*, 2019).

A modelagem do decaimento do cloro pode ser desenvolvida a partir de modelos baseados em processos ou de modelos estatísticos. A modelagem baseada em processos considera a cinética química para a previsão do decaimento de cloro, cabendo destacar os modelos cinéticos, simuladores numéricos e simuladores hidráulicos – *softwares* EPANET, AQUASIM e WaterCAD. Por outro lado, os modelos estatísticos dependem das relações empíricas, lineares ou não lineares, observadas entre a concentração de cloro e outros parâmetros de qualidade da água, como é o caso das redes neurais artificiais (Onyutha; Kwio-Tamale, 2022; Onyutha, 2022; Rodrigues; Scalize, 2019; Tiruneh *et al.*, 2019).

Os estudos de revisão realizados por Onyutha e Kwio-Tamale (2022) e Rodrigues e Scalize (2019) apontam uma predominância de aplicação de modelos baseados em processos na modelagem do decaimento do cloro. Na aplicação da metodologia, Almeida *et al.* (2022), Biondo (2019), García-Ávila *et al.* (2021) e Mentis *et al.* (2019) reforçam os resultados obtidos por Onyutha e Kwio-Tamale (2022), que assegura que esse tipo de modelagem possui vantagens relacionadas à produção de resultados fáceis de interpretar, a partir das suposições explícitas associadas à sua utilização.

Entretanto, para garantir a confiabilidade das ferramentas de simulação, vários parâmetros devem ser avaliados corretamente. Assim, os modelos baseados em processos exigem dados extensos e precisos, juntamente com modelagem hidráulica e de qualidade da água para determinar o residual de cloro dentro do sistema de distribuição. Dessa forma, é necessária a utilização de matemática complexa e o estabelecimento de várias suposições sobre a ordem cinética da reação, a fim de capturar o comportamento das concentrações de cloro sob diferentes circunstâncias de temperatura, pH, turbidez e outros fatores (Biondo, 2019; Onyutha, 2022; Tiruneh *et al.*, 2019).

Além disso, nesses modelos, as características da água e do sistema de abastecimento (SAA), além dos modos de operação e manutenção, determinam as reações que ocorrem na massa líquida, tornando-se difícil a determinação de um coeficiente de velocidade de degradação de desinfetante (k_b) padrão para qualquer tipo de SAA (SILVA *et al.*, 2019). Tiruneh *et al.* (2019), ao desenvolver uma modelagem da taxa de decaimento de cloro na rede de distribuição de água da cidade de Matsapha, concluiu que a utilização do coeficiente de degradação variável no software EPANET fornece resultados mais precisos em comparação com a modelagem de taxa de decaimento constante, mas eleva a complexidade do método.

Frente a isso, tem-se observado um crescimento da adoção de modelos estatísticos, tanto o aprendizado de máquina quanto as técnicas de aprendizado profundo, na previsão de concentrações de cloro residual livre. Esses modelos, ao preverem relações empíricas complexas entre o cloro e os parâmetros que influenciam no seu decaimento, podem ser utilizados quando não é possível obter dados precisos para os modelos baseados em processos. Além disso, os modelos desenvolvidos podem ser utilizados em regiões e sistemas de abastecimento de água com as mesmas características (Tiruneh *et al.*, 2019; Al-Araimi *et al.*, 2019; Hossain *et al.*, 2021; Onyutha, 2022).

Nos modelos estatísticos, o enfoque está nas relações entre a entrada (parâmetros que influenciam no decaimento) e as saídas (estimativa do cloro residual livre após decaimento),

sem a necessidade de se concentrar no que acontece na geração das relações. Apesar disso, Onyutha e Kwio-Tamale (2022) e Onyutha (2022) ressaltam, em casos em que as previsões de concentração do cloro residual livre são necessárias em um curto tempo disponível, que esse tipo de modelagem pode ser aplicado. No entanto, a aplicabilidade de modelos estatísticos na predição do decaimento do cloro permanece minimamente explorada (Onyutha, 2022).

Ao empregar técnicas de modelagem por rede neural artificial para prever as concentrações de cloro no SAA de Mascate, Al-Araimi *et al.* (2019) obtiveram resultados que demonstram o potencial do modelo na detecção do comportamento completo e não linear entre os dados. Hossain *et al.* (2021), ao avaliarem o decaimento da cloramina, verificaram que o modelo estatístico mostra um desempenho e previsibilidade relativamente maior em comparação com o modelo baseado em processos, destacando que a modelagem orientada por dados pode ser utilizada sem o desenvolvimento de um modelo hidráulico para a rede de abastecimento de água.

2.6 TÉCNICAS DE APRENDIZADO DE MÁQUINA APLICADAS À MODELAGEM DO DECAIMENTO DO CLORO

O aprendizado de máquina é uma abordagem da inteligência artificial (IA) que descreve a capacidade dos sistemas de aprenderem com dados de treinamento específicos do problema, automatizando o processo de construção de modelos analíticos e resolução de tarefas associadas. Dependendo da tarefa de aprendizado, tem-se diversas classes de algoritmos de aprendizado de máquina, incluindo modelos de regressões, algoritmos baseados em instâncias, árvores de decisão, métodos bayesianos e redes neurais artificiais (RNA) (JANIESCH *et al.*, 2021).

Basher e Hajmeer (2000), Braga *et al.* (2000) e Sinz *et al.* (2019) explicam que os modelos artificiais são ferramentas atrativas na solução de problemas tendo em vista a sua capacidade de aprender a partir de exemplos e generalizar a informação aprendida. A aprendizagem e a adaptabilidade permitem que o sistema atualize (modifique) sua estrutura interna em resposta à mudança do ambiente. Já a generalização possibilita a aplicação do modelo a dados não aprendidos, fornecendo respostas coerentes.

As redes neurais artificiais (RNA) são amplamente utilizadas, uma vez que sua estrutura flexível permite que elas sejam modificadas para uma diversidade de contextos. Nesse sentido, as RNA têm sido aplicadas em uma variedade de áreas para modelar problemas complexos em recursos hídricos ao longo dos anos, como demonstram os estudos de Chen *et al.* (2020), Maier

e Dandy (2000), Maier *et al.* (2010), Nourani *et al.* (2014), O'Reilly *et al.* (2018) e Oyeboode e Stretch (2018). Os modelos de RNAs podem ser usados na previsão dos impactos da deterioração da qualidade da água e nos processos de purificação de água potável. Chen *et al.* (2018) e Tabari e Talaei (2015) ressaltam que, na modelagem da qualidade da água, a utilização das redes neurais traz vantagens relacionadas à dispensabilidade de um algoritmo baseado em física para a construção do modelo, tornando a abordagem mais rápida e flexível, e à simplicidade na concepção de relações não-lineares entre os parâmetros.

Al-Araimi *et al.* (2019), Bowden *et al.* (2006), Cordoba *et al.* (2014), Gibbs *et al.* (2006), Kote e Wadkar (2019), Librantz *et al.* (2018), May *et al.* (2008), Santi *et al.* (2021), Soyupak *et al.* (2011), Wadkar e Kote (2017) e Wu *et al.* (2011) já estudaram o desempenho das redes neurais artificiais na modelagem do decaimento do cloro. Os resultados indicaram que o modelo foi consistente e preciso na previsão do cloro residual livre.

Para o desenvolvimento e processamento das RNAs são necessárias três etapas básicas: treinamento, validação e teste, cujas funções são, respectivamente, estabelecer a relação entre dados de entrada e saída, testar a relação criada na fase de treinamento e generalizar o comportamento para situações não aprendidas. Para o bom funcionamento do modelo, é necessária a realização de uma coleta de dados criteriosa, da seleção das variáveis de entrada relevantes, da determinação da arquitetura de rede adequada e da divisão de dados a serem utilizados em cada etapa (Cordoba *et al.*, 2014; Onyutha; Kwio-Tamale, 2022; Santi *et al.*, 2021).

O Quadro 1 apresenta algumas metodologias adotadas na aplicação de redes neurais artificiais na modelagem do decaimento do cloro. De maneira geral, nota-se uma predominância de trabalhos direcionados à sistemas de médio a grande porte, instalados em áreas urbanas, com exceção dos estudos conduzidos por Santi *et al.* (2021) e Wu *et al.* (2011), que avaliam sistemas de distribuição de água em áreas rurais.

Quadro 1: Aplicações de redes neurais artificiais na modelagem do decaimento do cloro

Estudo	Variáveis de entrada	Tipo de arquitetura	Divisão de dados	Área de estudo
Bowden <i>et al.</i> (2006)	Vazão, turbidez, pH, temperatura e cloro	<i>Generalized Regression Neural Network</i> (GRNN)	80% (treinamento e teste) e 20% (validação)	Sistema de distribuição de água de Myponga, Adelaide, Austrália
Gibbs <i>et al.</i> (2006)	Temperatura, vazão, concentração de cloro, carbono orgânico dissolvido, absorvância ultravioleta e tempo de medição	<i>Multilayer Perceptron</i> (MLP), <i>Generalized Regression Neural Network</i> (GRNN) e <i>Kohonen's Self-Organizing Maps</i> (SOM)	80% (treinamento) e 20% (teste)	Sistema de distribuição de água de Hope Valley, Adelaide, Austrália
May <i>et al.</i> (2008)	Sistema de distribuição Cherry Hill-Brushy Plains: cloro, nível do reservatório, vazão da bomba e vazão na tubulação Sistema de distribuição Myponga: cloro livre, pH, turbidez, nível do reservatório, vazão de saída e temperatura	<i>Generalized Regression Neural Network</i> (GRNN)	64% (treinamento), 16% (teste) e 20% (validação)	Sistemas de distribuição de água: Cherry Hill-Brushy Plains, New Haven, Connecticut, Estados Unidos; e, Myponga, Adelaide, Austrália
Soyupak <i>et al.</i> (2011)	pH, condutividade elétrica, turbidez, vazão, temperatura e cloro residual livre	<i>Multilayer Perceptron</i> (MLP)	Não informado	Sistema de distribuição de água, Antália, Turquia
Wu <i>et al.</i> (2011)	Vazão, temperatura e cloro	<i>Generalized Regression Neural Network</i> (GRNN)	60% (treinamento), 20% (teste) e 20% (validação)	Sistema de distribuição de água, Goldfield and Agricultural Water System, Austrália Ocidental
Cordoba <i>et al.</i> (2014)	Temperatura, pH, turbidez, vazão, cloro inicial e cloro residual livre	<i>Multilayer Perceptron</i> (MLP)	50% (treinamento), 25% (teste) e 25% (validação)	Sistema de distribuição de água de Brno-Kohoutovice, distrito de Brno, República Tcheca
Wadkar e Kote (2017)	Grupo I: turbidez da água na entrada da ETA Grupo II: pH e turbidez da água na entrada da ETA Grupo III: pH, turbidez da água na entrada e turbidez da água na saída da ETA Grupo IV: pH, dose de cloro, turbidez da água na entrada e turbidez da água na saída da ETA	<i>Feed Forward Neural Network</i> (FFNN), <i>Pattern Recognition Neural Network</i> (PRNN), <i>Cascade Feed Forward Neural Network</i> (CFNN), <i>Radial Basis Neural Network</i> (RBNN) e <i>Generalized Regression Neural Network</i> (GRNN)	70% (treinamento), 15% (teste) e 15% (validação)	ETA da Cidade de Pimpri Chinchwad, Índia

Quadro 1: Aplicações de redes neurais artificiais na modelagem do decaimento do cloro (continuação)

Estudo	Variáveis de entrada	Tipo de arquitetura	Divisão de dados	Área de estudo
Librantz <i>et al.</i> (2018)	Vazão de água filtrada, vazão do sistema compensador, cloro residual livre na saída da ETA, ponto de ajuste na saída do reservatório e erro de dosagem	Não informado	Não informado	ETA da Região Metropolitana de São Paulo, Brasil
Kote e Wadkar (2019)	Dose de coagulante, turbidez da água na saída da ETA e cloro residual	<i>Radial Basis Neural Network (RBNN), Feed Forward Neural Network (FFNN), Cascade Feed Forward Neural Network (CFNN) e Generalized Regression Neural Network (GRNN)</i>	80% (treinamento) e 20% (teste)	ETA de Maharashtra, Índia
Al-Araimi <i>et al.</i> (2019)	pH, temperatura e concentração de cloro	<i>Multilayer Perceptron (MLP)</i>	Não informado	Sistema de distribuição de água de Mascate, Omã
Santi <i>et al.</i> (2021)	Combinação 1: cloro residual livre no ponto de distribuição e o tempo decorrido entre a medição no ponto de distribuição e no ponto de consumo Combinação 2: variáveis da combinação 1 e parâmetros de qualidade da água no ponto de distribuição (condutividade elétrica, temperatura, pH e turbidez)	<i>Multilayer Perceptron (MLP)</i>	Em geral: 25% (treinamento), 25% (teste) e 50% (validação) Casos específicos: 50% (treinamento), 25% (teste) e 25% (validação); 30% (treinamento), 45% (teste) e 25% (validação)	Sistema de distribuição de água de assentamentos de refugiados no Sudão do Sul, Jordânia e Ruanda

Fonte: A autora (2023).

Além dos tipos de sistemas nos quais foram feitos os estudos com RNAs, o Quadro 1 apresenta as variações das composições das redes. Em relação às variáveis de entrada, Cordoba *et al.* (2014), ao estudarem o sistema de distribuição de um distrito da República Tcheca, indicam que o cloro inicial, a vazão e a temperatura apresentaram influência significativa no decaimento de cloro, com destaque para a temperatura. O estudo de Santi *et al.* (2021), verificou ainda que o pH demonstrou ser um fator importante, que a relação entre turbidez e ponto de distribuição foi menos consistente e que a temperatura apresentou as tendências mais fracas de correlação com o ponto de consumo.

Wadkar e Kote (2017) destacam a importância dos parâmetros turbidez, pH e dose de cloro na modelagem, e, assim como Kote e Wadkar (2019) ressaltam que o desempenho dos modelos de RNA cresceu à medida que o número de variáveis de entrada aumentou. Entretanto, Gibbs *et al.* (2006) salientam a importância de identificação e remoção de parâmetros de entrada redundantes, mesmo que reduza o tamanho do modelo, a fim de diminuir o grau de ruído nos dados e, conseqüentemente, elevar o desempenho da modelagem.

Dessa forma, é importante a utilização de métodos de determinação de variáveis de entrada significativas, como é o caso da metodologia *t-Distributed Stochastic Neighbor Embedding* (t-SNE), empregada por Zhou *et al.* (2020) na seleção de parâmetros eficazes na previsão da qualidade da água. Conforme Jiang *et al.* (2022), o método melhora a precisão da seleção das variáveis e o custo computacional.

Em se tratando da arquitetura do modelo, O'Reilly *et al.* (2018) e Onyutha e Kwio-Tamale (2022) apontam a existência de diversos tipos de redes neurais disponíveis, como: *Adaptive Neuro-Fuzzy Inference System* (ANFIS), *Multilayer Perceptron* (MLP), *Standard (deterministic) Neural Networks* (SNN), *Bayesian Neural Networks* (BNN), *Feed Forward Neural Network* (FFNN), *Pattern Recognition Neural Network* (PRNN), *Cascade Feed Forward Neural Network* (CFNN), *Kohonen's Self-Organizing Maps* (SOM), *Radial Basis Neural Network* (RBNN) e *Generalized Regression Neural Network* (GRNN). Entretanto, conforme exibido no Quadro 1, a MLP e a GRNN são os sistemas neurais mais utilizados.

May *et al.* (2008) ressaltam que o GRNN é mais rápido de desenvolver em relação ao MLP, pois não necessita de um procedimento de treinamento repetitivo, mas possui um maior requisito de computacional e de memória para armazenar os dados de treinamento. Ao avaliar o estudo de Wadkar e Kote (2017), constata-se que as capacidades da MLP e da GRNN, na previsão de cloro residual, mostraram-se boas. Gibbs *et al.* (2006), ao comparar as redes MLP, GRNN e SOM, verificaram que a *Multilayer Perceptron* supera consistentemente as outras

técnicas baseadas em dados, implementadas na modelagem do decaimento do cloro do sistema de distribuição de água de Hope Valley.

Al-Airami *et al.* (2019), Cordoba *et al.* (2014), Santi *et al.* (2021) e Soyupak *et al.* (2011) obtiveram sucesso ao aplicarem esta arquitetura de rede na previsão dos níveis de cloro residual livre. A MLP é a arquitetura de rede mais empregada devido à sua simplicidade de utilização, precisão, acurácia e alta capacidade de identificação de padrões não-lineares existentes entre as variáveis (Castañeda-Miranda; Castaño-Meneses, 2020; Chhipi-Shrestha *et al.*, 2023; Setshedi *et al.*, 2021)

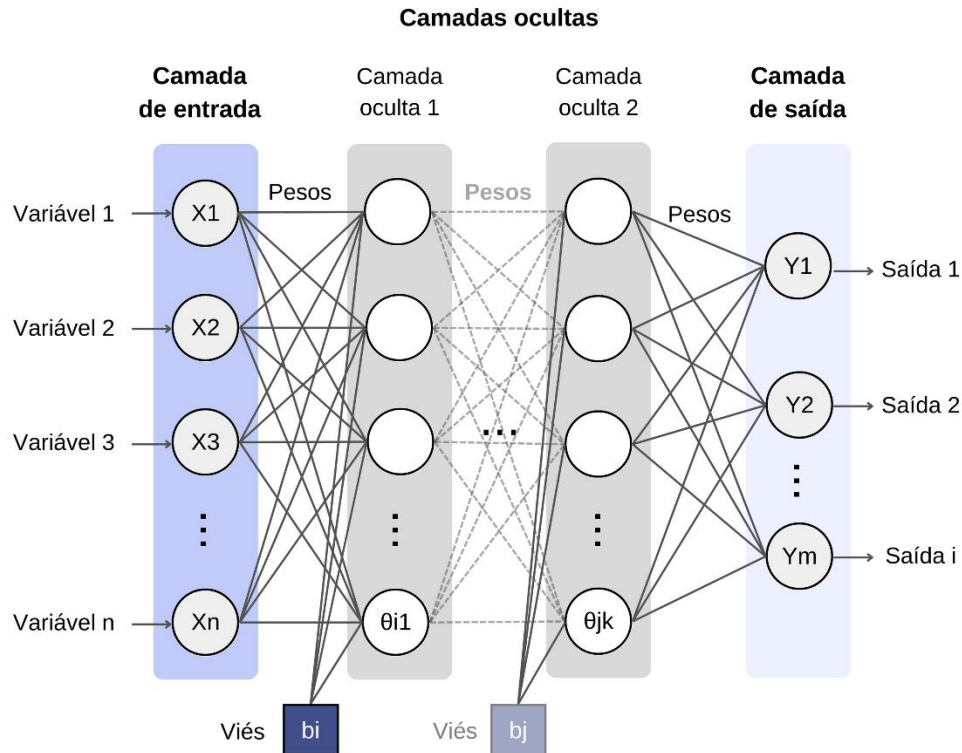
Na divisão dos dados verificou-se que a maioria dos pesquisadores alocou a maior porcentagem dos dados de modelagem para o treinamento, em relação ao teste e validação da rede (Quadro 1). Vale salientar que, o conjunto de treinamento é usado para calibrar o modelo, o conjunto de teste é usado na validação cruzada para evitar o ajuste excessivo e o conjunto de validação é usado para testar o desempenho do modelo treinado (Wu *et al.*, 2011).

2.6.1 Redes Neurais Artificiais (RNA)

As Redes Neurais Artificiais (RNA) consistem em um modelo estatístico formado de elementos de processamento simples, adaptativos e interconectados, denominados neurônios artificiais, que processam informações de maneira paralela e distribuída, por meio de funções matemáticas lineares ou não-lineares. Parte da literatura aproxima as RNAs ao comportamento do cérebro humano, devido à capacidade de resolver problemas complexos por meio de aprendizado baseado em exemplos (Hasson *et al.*, 2020; Nascimento, 2020).

Na sua arquitetura do tipo *Multilayer Perceptron* (MLP) consiste em uma célula nervosa artificial composta por três tipos de camadas: camada de entrada (*input*), camadas ocultas (*hidden*) e camada de saída (*output*) (Figura 5). As variáveis do modelo (X) são introduzidas na camada de entrada, sendo processadas pela rede nas camadas intermediárias, produzindo um resultado (Y) na saída. Os *inputs* são conectados aos nós da *hidden* por meio de pesos (w) e da soma de um viés (b), que representam os coeficientes e termo independente de uma equação, respectivamente. A função de ativação (f) é usada processar o resultado obtido em cada nó, calculando a não-linearidade do modelo. Assim, as variáveis de saída são dadas conforme a Equação 6 (Alam *et al.*, 2022; Marcy *et al.*, 2014; Braga *et al.*, 2000; Chen *et al.*, 2020; Gülcü, 2021; Rustam *et al.*, 2022; Sinz *et al.*, 2019).

Figura 5: Modelo de RNA com arquitetura do tipo MLP com quatro camadas: uma camada de entrada, duas camadas ocultas e uma camada de saída



$$Y_m = f_{k+1} \left(\sum_{m=1}^m w_{mj}^{k+1} \cdot f_k \left(\sum_{j=1}^j w_{ji}^k \cdot f_1 \left(\sum_{i=1}^i w_{in}^1 \cdot X_n + b_i \right) + b_j \right) + b_m \right) \quad (6)$$

O número de camadas e neurônios, a taxa de aprendizado e a função de ativação constituem hiperparâmetros do modelo e devem ser definidos manualmente pelo usuário ou determinados por uma rotina de otimização (Janiesch *et al.*, 2021).

Gülcü (2021) destaca que o treinamento é a parte mais importante e exigente do método, pois envolve encontrar os valores mais adequados para os pesos na arquitetura de rede, consistindo em um problema de otimização desafiador. Nesse sentido, destaca-se a utilização do algoritmo *backpropagation* no processo de aprendizado supervisionado das RNA. Na técnica, baseada na regra delta ou aprendizagem por correção de erro, um supervisor externo direciona o ajuste de pesos, por meio de um algoritmo como o *Levenberg-Marquardt backpropagation*, com vistas à minimização do erro e, assim, à finalização do treinamento com uma possível solução (Marcy, 2017; Onyutha, 2022).

Para garantir melhorias contínuas da previsão é vital que as medidas estatísticas do desempenho do modelo sejam sempre fornecidas. Para a validação da performance das redes neurais aplicadas à modelagem do decaimento de cloro, Onyutha e Tamale (2022) determinaram que a estatística descritiva é comumente aplicada, a partir da determinação do erro médio quadrático (EMQ), do erro médio absoluto (EMA) e do coeficiente de determinação (R^2).

Onyutha (2020) e Onyutha (2021) indicam que o R-quadrado apresenta uma série de problemas em seu uso como medida de desempenho de modelo, pois baseia-se no pressuposto de que a relação entre as séries modeladas é linear, diferentemente do que é observado no decaimento de cloro. Além disso, o R^2 : (i) pode ser baixo para um modelo preciso; (ii) pode ser alto para um modelo imperfeito; (iii) não quantifica viés no modelo; e (iv) tem o mesmo valor nos dois casos em que regredimos x em y e vice-versa.

Embora para a maioria dos modelos a função de perda utilizada no treinamento seja o erro médio quadrático (Equação 7), essa avaliação pode levar a RNA a fornecer previsões que ignoram exemplo raros. Para aliviar o desequilíbrio entre os dados, Peng *et al.* (2021), Santi *et al.* (2021) e Goudarzi (2022) sugerem a aplicação do erro médio quadrático ponderado (EMQP). Neste, atribui-se um coeficiente de ponderação (w_i) à diferença quadrática entre o valor real (y_i) e o valor previsto pelo modelo (\hat{y}_i) (Equação 8), em faixas de valores mais críticas, visando evidenciar os *outliers* para que a rede neural se adapte melhor ao conjunto de dados de treinamento.

$$EMQ = \frac{1}{n} \cdot \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (7)$$

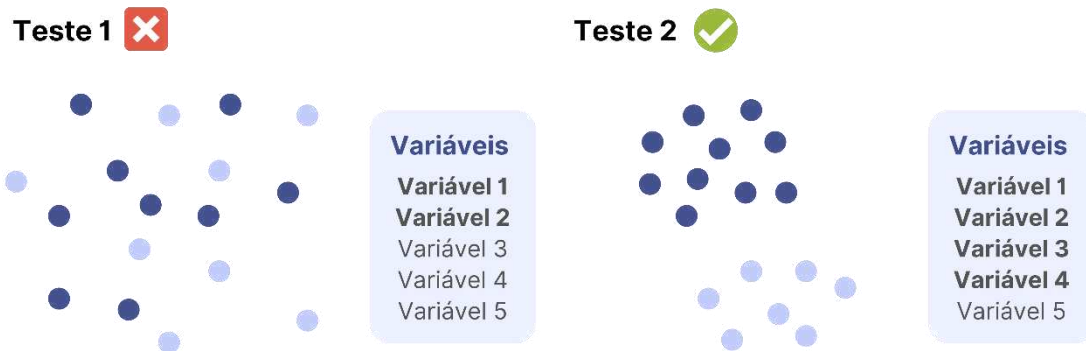
$$EMQP = \frac{1}{n} \cdot \sum_{i=1}^n w_i \cdot (y_i - \hat{y}_i)^2 \quad (8)$$

2.6.2 t-Distributed Stochastic Neighbor Embedding (t-SNE)

O t-SNE, proposto por Van der Maaten e Hinton (2008), é um algoritmo não linear de redução da dimensionalidade dos dados, representando as variáveis em formato de *clusters* no plano ou em mais dimensões. A partir dos agrupamentos formados e a facilidade de identificação dos *clusters* individuais, é possível verificar a similaridades dos dados e

determinar quais são os parâmetros mais adequados para explicar o problema, que neste caso, é o decaimento do cloro residual livre (Figura 6).

Figura 6: Exemplo de aplicação do t-SNE para diferentes conjuntos de variáveis



Fonte: A autora (2023).

A partir da observação da Figura 6 é possível inferir que as variáveis empregadas no teste 2 do t-SNE exibem uma maior interdependência, evidenciada pela presença de agrupamentos mais nitidamente definidos em comparação com o teste 1, onde observa-se uma dispersão das amostras. Os agrupamentos produzidos no teste 2 possibilitam identificar a relação não-linear entre as variáveis e a capacidade de descreverem satisfatoriamente o fenômeno em estudo.

Para tanto, Van der Maaten e Hinton (2008) afirmam que o método se baseia na probabilidade de que a proximidade entre duas amostras no espaço de dimensões elevadas é mantida no espaço de menores dimensões. Essa probabilidade é calculada para todos os pontos no espaço de maiores dimensões através da distribuição normal (p_{ij}) e para os pontos no espaço de menores dimensões por meio da distribuição *t-Student* (q_{ij}). Então, são calculadas as divergências entre as duas probabilidades (C) que, através de uma função gradiente (Equação 9), são atualizadas a cada iteração até serem minimizadas.

$$C = \sum_i \sum_j p_{ij} \cdot \log p_{ij} - p_{ij} \cdot \log q_{ij} \quad (9)$$

Após isso, os dados originais podem ser espacializados em um gráfico de duas ou três dimensões. As posições dos pontos no gráfico t-SNE são definidas de forma a preservar as relações de proximidade entre os pontos do conjunto de dados original. Isso significa que os pontos que eram próximos no espaço de alta dimensão provavelmente estarão próximos no gráfico, e os pontos que eram distantes no espaço de alta dimensão provavelmente estarão mais

separados no gráfico t-SNE (Linderman; Steinerberger, 2019). Isso facilita a visualização e a interpretação das relações entre os pontos de dados em um espaço de dimensionalidade reduzida.

2.7 AVALIAÇÃO DE DESEMPENHO DE REDES DE DISTRIBUIÇÃO DE ÁGUA

O controle e o monitoramento da qualidade da água em sistemas de abastecimento são essenciais para a avaliação da efetividade do tratamento adotado e para a promoção da saúde aos usuários. Nesse sentido, além de prever o decaimento de cloro, é fundamental aferir se o parâmetro apresenta desempenho adequado ao longo da rede de distribuição, em conformidade com os padrões exigidos pela Portaria nº 888/2021 (Araujo *et al.*, 2020; Silva *et al.*, 2019).

Nas últimas décadas, diversos índices e indicadores vêm sendo desenvolvidos com o intuito de mensurar, quantitativamente, as fragilidades hídricas em sistemas de abastecimento de água. A determinação de indicadores de desempenho é primordial para o refinamento da gestão e para auxiliar os processos de tomada de decisão, identificando problemas, promovendo medidas corretivas e apoiando a formulação de ações, políticas e estratégias mais eficientes (Bezerra *et al.*, 2019; Faro, 2020).

Segundo Oliveira (2016), a avaliação de desempenho por meio de indicadores torna-se mais notória pela capacidade de análise e acompanhamento de fenômenos, até mesmo não detectáveis, tendo um significado maior que o fornecido pela observação direta, expressa por gráficos ou formas estatísticas.

Sob esse viés, o estudo de Hashimoto *et al.* (1982) se destaca na aferição do desempenho estatístico de sistemas hídricos, a partir da proposição de indicadores de confiabilidade, resiliência e vulnerabilidade. Ligados aos fatores probabilísticos da operação do sistema, os indicadores demonstram a tendência de como o sistema deverá reagir a eventos de falhas, oferecendo aos gestores um melhor entendimento sobre os riscos (Faro, 2020; Risso *et al.*, 2021).

Inúmeras aplicações dos índices de desempenho em sistemas de recursos hídricos têm sido realizadas, com destaque para Araujo *et al.* (2020), Atashi *et al.* (2020), Castro (2019), Faro (2020) e Hoque *et al.* (2014). Esses trabalhos estão focados no gerenciamento de reservatórios e bacias hidrográficas e na avaliação da eficiência operacional hidráulica de sistemas de abastecimento de água. No entanto, salienta-se a importância de qualificar o

desempenho do SAA quanto à sua capacidade de entregar água em quantidade e qualidade adequadas ao consumo humano (Oliveira, 2016).

A avaliação de desempenho dos SAA quanto à qualidade da água, principalmente após o processo de desinfecção, propicia a determinação das condições da água consumida, identificando anormalidades que possam ocasionar riscos para as condições de saúde da população usuária. Nesse contexto, os conceitos de confiabilidade, resiliência e vulnerabilidade podem ser caracterizados, com base nos estudos de Araujo *et al.* (2020), Faro (2020), Hashimoto *et al.* (1982), Hoque *et al.* (2012) e Loucks (1997), conforme a abordagem da avaliação de desempenho do decaimento de cloro residual livre em sistemas de abastecimento de água.

A confiabilidade (C) pode ser definida como a probabilidade de um sistema atender satisfatoriamente aos valores esperados, levando em conta a frequência dos déficits (t_d) verificados em um determinado período de tempo t (Equação 10). Para efeito da determinação do cloro residual livre ao longo da rede de distribuição, um estado confiável refere-se ao pleno atendimento aos limites estabelecidos na Portaria nº 888/2021.

$$C = 1 - \frac{t_d}{t} \quad (10)$$

Já o indicador resiliência (R) relaciona-se com a probabilidade de que o sistema se recupere após ser submetido a uma determinada perturbação, que tenha gerado uma condição de falha. Pode ser calculado pela Equação 11, a partir da relação entre o número de vezes que um valor satisfatório segue um valor insatisfatório (t_r) e o número de déficits (t_d). No contexto do monitoramento do cloro, essa característica relaciona-se a capacidade do sistema de abastecimento se adaptar as mudanças e retornar aos limites satisfatórios.

$$R = \frac{t_r}{t_d} \quad (11)$$

Por último, a vulnerabilidade (V) descreve o dano médio que ocorre no sistema devido a lapsos em um estado de falha. Para Sandoval-Solis *et al.* (2011), a vulnerabilidade pode se expressa como: i) a falha média; ii) a média das deficiências máximas durante todos os períodos de falha contínua; iii) a probabilidade de ultrapassar um determinado limiar de déficit. Hoque *et al.* (2012) e Hoque *et al.* (2014), ao estudarem a saúde de bacias hidrográficas, verificaram que a vulnerabilidade, relativa à parâmetros de qualidade de água, deve ser calculada como uma

medida proporcional à magnitude das violações dos padrões estabelecidos para um constituinte. Sendo assim, Hoque *et al.* (2012) estabeleceu a Equação 12 para a estimativa desse indicador.

$$V = \frac{1}{t_d} \cdot \sum_{j=1}^M \left(\frac{|d(j) - L|}{L} \cdot |d(j) - L| \right) \quad (12)$$

Onde: t_d é o número de déficits; $d(j)$ é o valor do déficit j ; e, L é o padrão de potabilidade conforme a Portaria nº 888/2021, que em termos do cloro residual livre pode ser igual a 0,2 ou 5,0 mg/L, conforme a Equação 13.

$$L = \begin{cases} 0,2 & \text{se } d(j) < 0,2 \\ 5,0 & \text{se } d(j) > 5,0 \end{cases} \quad (13)$$

Para retirar a dimensionalidade do cálculo da vulnerabilidade, Sandoval-Solis *et al.* (2011) propuseram o quociente entre o indicador e a demanda anual por água. No entanto, essa equação é aplicável apenas para avaliações de desempenho em relação à demanda de água. Para mitigar essa lacuna para a avaliação em relação ao CRL, pode-se adotar a Equação 14, adaptada da Equação 13, que torna a vulnerabilidade adimensional a partir da divisão do seu valor pelo limite (L) estabelecido pela Portaria nº 888/2021.

$$V_{adm} = \frac{1}{t_d} \cdot \sum_{j=1}^M \frac{\left(\frac{|d(j) - L|}{L} \cdot |d(j) - L| \right)}{L} \quad (14)$$

A fim de sintetizar os critérios de desempenho, Sandoval-Solis *et al.* (2011) estabeleceram um Índice de Sustentabilidade (IS), que permite averiguar a performance do sistema, com o intuito de classificar sua operação como satisfatória ou insatisfatória (Equação 15).

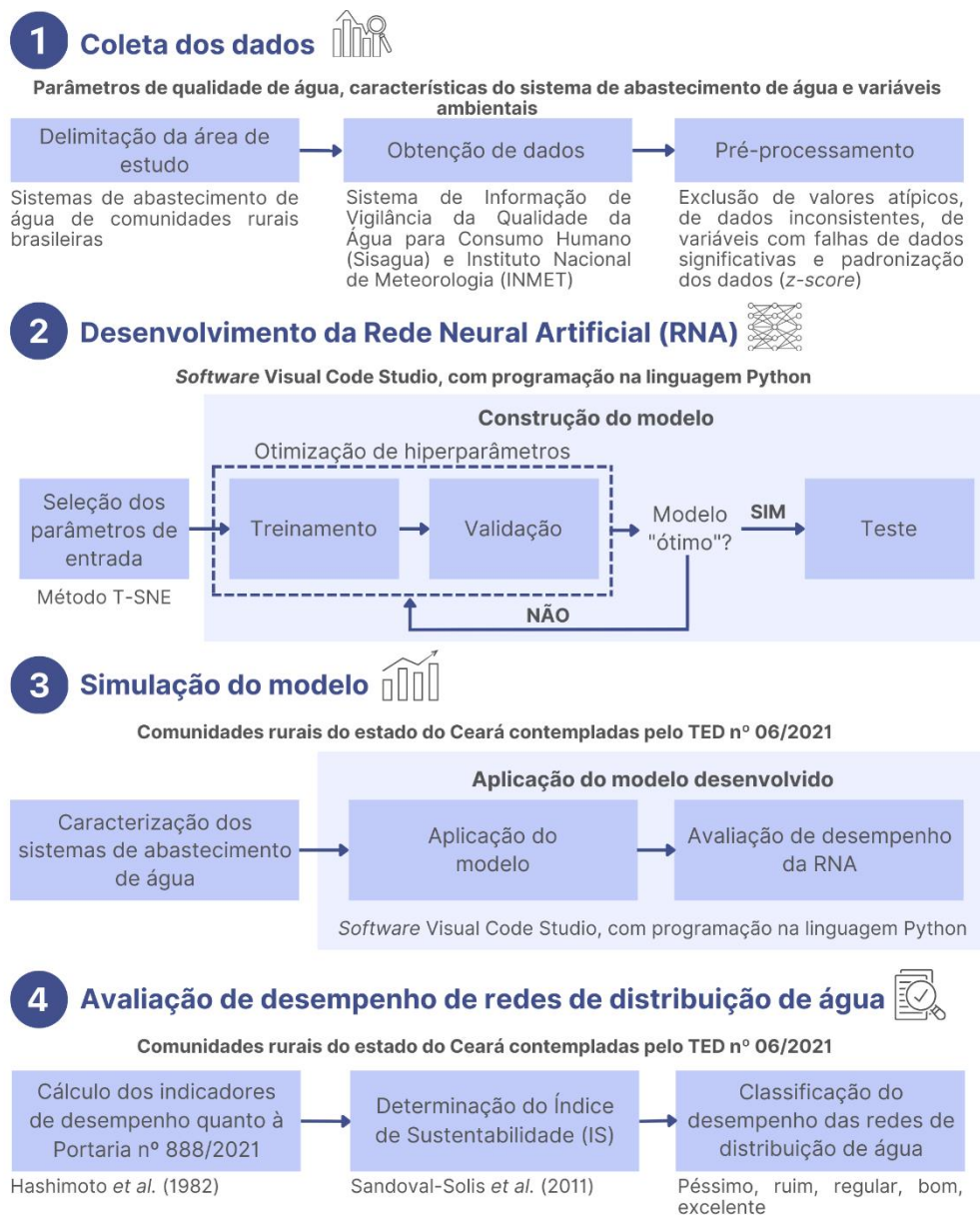
$$IS = \sqrt[3]{C \cdot R \cdot (1 - V_{adm})} \quad (15)$$

Apesar de Loucks (1997) recomendar que a avaliação do desempenho do sistema deve englobar todos os seus componentes, permitindo uma avaliação holística, a aplicação do índice especificamente na avaliação no decaimento de cloro residual livre em SAA é importante para a boa gestão do processo de desinfecção. A metodologia permite diagnosticar as deficiências do sistema quanto ao cloro, auxiliando os processos de ajustes na dosagem do desinfetante e a tomada de decisão quanto à necessidade de instalação de pontos de reforço de cloração (Oliveira, 2018).

3. METODOLOGIA

A metodologia para a elaboração dessa pesquisa baseou-se em quatro etapas principais, detalhadas no fluxograma da Figura 7. São elas: (i) coleta e pré-processamento dos dados; (ii) desenvolvimento do modelo de aprendizado de máquina para a previsão do decaimento de cloro em redes de distribuição de água rurais, a partir da seleção de variáveis pelo método *t-Distributed Stochastic Neighbor Embedding* (t-SNE) e da aplicação da técnica de Redes Neurais Artificiais (RNAs); (iii) simulação do modelo construído em sistemas de abastecimento de água de comunidades rurais do estado do Ceará; e (iv) avaliação de desempenho das redes de distribuição de água quanto ao CRL, por meio dos indicadores de Hashimoto *et al.* (1982).

Figura 7: Fluxograma metodológico do estudo



Fonte: A autora (2023).

3.1 COLETA DOS DADOS

3.1.1 Delimitação da área de estudo

Para a coleta dos dados e posterior desenvolvimento da rede neural artificial, a área de estudo compreende os Sistemas de Abastecimento de Água (SAA) de comunidades rurais brasileiras que possuem dados de qualidade de água registrados no Sistema de Informação de Vigilância da Qualidade da Água para Consumo Humano (Sisagua). O Sisagua é um instrumento do Programa Nacional de Vigilância da Qualidade da Água para Consumo Humano (Vigiagua) que tem como finalidade auxiliar o gerenciamento de riscos à saúde a partir dos dados gerados rotineiramente pelos profissionais do setor de saúde (vigilância) e responsáveis pelos serviços de abastecimento de água (controle) e da geração de informações em tempo hábil para planejamento, tomada de decisão e execução de ações de saúde relacionadas à água para consumo humano (Brasil, 2023).

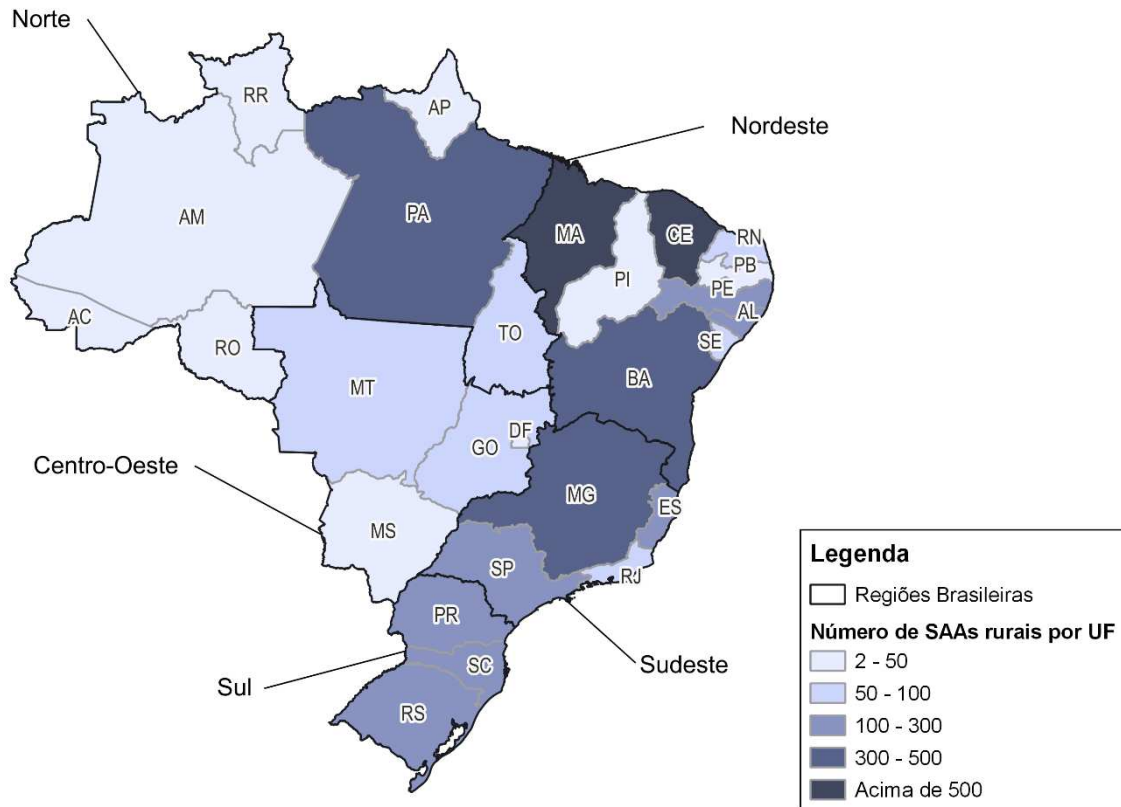
De acordo com o Plano Nacional de Saneamento Rural (Brasil, 2019), apenas 40,5% da população rural brasileira tem atendimento adequado¹ quanto ao serviço de abastecimento de água. Os efeitos da ausência/precariedade dessas infraestruturas em áreas rurais são ainda mais significativos, tendo em vista que estes locais abrigam os maiores percentuais de famílias de baixa renda, altas taxas de analfabetismo e maior incidência de doenças de veiculação hídrica (Bousquat *et al.*, 2022).

O mapa da Figura 8 apresenta a distribuição dos sistemas de abastecimento de água rurais nos estados brasileiros. Ao todo, a área de estudo abriga 4.878 SAAs, que estão concentrados na região Nordeste (52,4%), em especial nos estados do Ceará e Maranhão, que possuem 862 e 648 sistemas, respectivamente.

Conforme dados obtidos no Sisagua (2023), a maior proporção dos sistemas rurais, representando 49,1%, são geridos pelas prefeituras municipais, através das secretarias de Obras, Agricultura ou Saneamento. Esta predominância é marcada na região Norte do país, com destaque para as unidades federativas do Amapá (AP), Pará (PA), Amazonas (AM) e Acre (AC). A prestação do serviço pelo poder público municipal limita o investimento em infraestrutura e no controle e qualidade da água distribuída, por conta da frágil sustentabilidade administrativa, financeira e operacional observada (Brasil, 2019; Silva; Moraes, 2022).

¹ Engloba as populações que: não sofrem com intermitência prolongada ou racionamento; recebem água potável da rede de distribuição, com ou sem canalização interna; ou recebem água potável de poço, nascente ou sistema de captação de água de chuva, com canalização interna (Brasil, 2019).

Figura 8: Distribuição dos sistemas de abastecimento de água rurais com dados registrados no Sisagua por unidade federativa (UF)



Fonte: Adaptado de Sisagua (2023).

Cerca de 36,3% dos SAAs tem o serviço de abastecimento de água prestado por empresas públicas, sociedades de economia mista ou fundações estatais de direito privado, com concentração nas regiões Sul, Sudeste e Centro-Oeste. Santos e Santana (2020) e Sousa (2020) ressaltam que as empresas privadas possuem um menor interesse na prestação do serviço de abastecimento de água em comunidades rurais, devido ao custo para prover bens e infraestruturas a grandes distâncias do meio urbano, a baixa densidade populacional e a reduzida capacidade de pagamento dos usuários.

Os demais sistemas de abastecimento de água, que correspondem a 14,6% do total, são geridos a partir de gestão comunitária, com destaque para o Sistema Integrado de Saneamento Rural (Sisar) que atende a 13,1% dos sistemas. Essa modalidade de gerenciamento dos SAAs é predominante na região Nordeste, principalmente no estado do Ceará, onde 74,0% dos 862 sistemas são atendidos pelo Sisar. Esses sistemas são adequados à realidade rural pois estabelecem responsabilidades na governança e manutenção dos sistemas considerando a realidade local, os valores e os aspectos culturais das comunidades (Santos; Santana, 2020).

3.1.2 Coleta e pré-processamento dos dados

A fim de propiciar a construção de um banco de dados para o desenvolvimento das Redes Neurais Artificiais (RNA), foram coletados de comunidades rurais brasileiras dados históricos com múltiplas variáveis que influenciam no decaimento de cloro. Baseando-se nos trabalhos apresentados no Quadro 1 e levando em conta a influência de variáveis ambientais, das características do sistema de abastecimento de água e da presença de compostos orgânicos e inorgânicos para o decaimento do cloro na água, considerou-se neste estudo as variáveis exibidas no Quadro 2. Com o intuito de aplicar a rede neural artificial em qualquer ponto da rede de distribuição, optou-se pela utilização de dados relacionados aos reservatórios de armazenamento de água tratada, para evitar a necessidade de realização de coletas e análises físico-químicas da água no ponto onde se deseja prever o decaimento do cloro.

Quadro 2: Variáveis consideradas no estudo

Classificação	Variável	Unidade	Tipo
Parâmetros de qualidade de água	Cloro residual livre no reservatório	mg/L	Quantitativa
	Cloro residual livre na rede	mg/L	Quantitativa
	pH no reservatório	-	Quantitativa
	Turbidez no reservatório	uT	Quantitativa
	Condutividade elétrica no reservatório	uS/cm	Quantitativa
	Sólidos dissolvidos totais (SDT) no reservatório	mg/L	Quantitativa
Características do sistema de abastecimento de água	Comprimento da tubulação	m	Quantitativa
	Número de economias residenciais	economias	Quantitativa
	Tipo de captação	-	Qualitativa
	Existência de etapa de filtração	-	Qualitativa
Variável ambiental	Temperatura do dia	°C	Quantitativa

Fonte: A autora (2023).

Os parâmetros de qualidade de água e as informações acerca do sistema de abastecimento foram obtidos para todo o território nacional, por meio do Sistema de Informação de Vigilância da Qualidade da Água para Consumo Humano (Sisagua), no período de 2014 a 2023. Com o objetivo de criar modelos aplicáveis para a previsão do decaimento de cloro, considerou-se dados de parâmetros sentinelas de qualidade da água (cloro residual livre, pH e turbidez), exigidos pela Portaria nº 888/2021 e os parâmetros de condutividade elétrica e sólidos dissolvidos totais (SDT) pela facilidade de obtenção a partir de sensores de baixo custo. A escolha das características do sistema de abastecimento a serem consideradas foi baseada também na influência destas na decomposição do cloro ao longo da rede de distribuição.

Tendo em vista a impossibilidade de obtenção dos projetos dos sistemas de abastecimento de água das comunidades rurais com amostras registradas no Sisagua, utilizou-

se um método indireto para a determinação do comprimento da tubulação. A partir da diferença entre a hora da coleta da amostra no reservatório (t_1) e na rede de distribuição (t_2), obtidos no Sisagua, e considerando uma velocidade média (V_m) de 10 km/h para o deslocamento entre os pontos, estimou-se o valor do comprimento da tubulação (Δs) utilizando a equação da velocidade média (Equação 16). O valor assumido para a velocidade foi calculado com base em relatórios de análise de qualidade de água e mapeamento de SAA de comunidades rurais gerenciados pelo Sistema Integrado de Saneamento Rural (Sisar).

$$V_m = \frac{\Delta s}{|t_2 - t_1|} \quad (16)$$

A maioria dos modelos de Rede Neural Artificial (RNA) não entendem variáveis categóricas do tipo *string* (texto), o que impede que o modelo realize a correlação adequada entre as variáveis (Al-Shehari; Alsowail, 2021). Logo, as informações qualitativas referentes ao tipo de captação e a existência de etapa de filtração foram transformadas em dados numéricos por meio da técnica *One-hot encoder*. Essa codificação altera os valores das categorias existentes na base de dados para valores numéricos binários (0 ou 1), conforme apresentado no Quadro 3.

Quadro 3: Codificação das variáveis qualitativas

Variável	Categorias	Codificação
Tipo de captação	Subterrânea	0
	Superficial	1
Existência de etapa de filtração	Não	0
	Sim	1

Fonte: A autora (2023).

Apesar de diversos estudos utilizarem a temperatura da água como variável de entrada (Quadro 1), esse parâmetro não é contemplado no banco de dados do Sisagua. Para tanto, a fim de considerar o parâmetro no modelo e prever a influência de variáveis ambientais no decaimento do cloro, utilizou-se dados históricos de temperatura média do dia da coleta das amostras, obtidos por meio das estações automáticas do Instituto Nacional de Meteorologia (INMET). Para os municípios sem dados no dia da coleta ou sem estações automáticas em seu território, assumiu-se o valor da temperatura medida na estação mais próxima. Os dados foram processados com o auxílio do *software* Visual Code Studio, utilizando as bibliotecas *GeoPandas*, *Shapely* e *Datetime*.

Com o intuito de reduzir o ruído causado por dados incoerentes e refinar as características dos dados de interesse, realizou-se o pré-processamento dos dados por meio da estatística descritiva (Al-Shehari; Alsowail, 2021; Bhardwaj *et al.*, 2022). Foram excluídas do banco de dados amostras com valores inconsistentes, amostras com o cloro residual livre do reservatório inferior ao cloro na rede de distribuição e amostras com diferença de horário entre as coletas maior que uma hora. Além disso, realizou-se uma pré-seleção das variáveis, eliminando aquelas com falhas significativas (menos de 1.000 dados), com o intuito de garantir que os parâmetros previamente selecionados sejam representativos.

Tendo em vista as escalas e unidades diferentes associadas às variáveis que compõem a base de dados, utilizou-se a técnica de padronização dos dados a fim de garantir comparabilidade entre os parâmetros e o bom desempenho do modelo. Conforme os estudos de Prasad *et al.* (2022) e Wu e Wang (2022), aplicou-se o método de padronização *z-score* para transformar os valores das variáveis em números adimensionais (Equação 17), empregando a biblioteca *StandardScaler* do *software* Visual Code Studio.

$$z = \frac{(x - \mu)}{\sigma} \quad (17)$$

Onde: z é o *z-score*; x é o dado da amostra para a variável; μ é a média amostral; e, σ é o desvio padrão amostral.

3.2 DESENVOLVIMENTO DA REDE NEURAL ARTIFICIAL (RNA)

O desenvolvimento de uma Rede Neural Artificial (RNA) inclui a seleção de parâmetros de entrada significativos seguida do treinamento, validação e teste do modelo. Todo o processamento foi realizado no Visual Code Studio, que é um *software* de código aberto para programação científica na linguagem Python, adequado para análise de dados e aprendizado de máquina.

3.2.1 Seleção dos parâmetros de entrada

Para a determinação das variáveis de entrada, utilizou-se o método *t-Distributed Stochastic Neighbor Embedding* (t-SNE), aplicado à diferentes combinações de parâmetros que influenciam no decaimento do cloro, a fim de verificar qual o conjunto de parâmetros é capaz de explicar o decaimento do cloro na RNA com melhor desempenho.

Empregando a biblioteca de código aberto *Scikit-learn* dividiu-se os dados em categorias, conforme os intervalos do parâmetro de saída – o cloro residual livre na rede de distribuição de água. A quantidade de categorias e os intervalos foram definidos com base no tamanho do banco de dados e na amplitude dos valores da variável, por meio do método de análise exploratória interquartil (gráfico de *boxplot*).

Para garantir o desempenho adequado do t-SNE, definiu-se o número efetivo de vizinhos da amostra, por meio da perplexidade (*p*). O valor desse hiperparâmetro foi determinado com base nos estudos de Van der Maaten e Hinton (2008) e Xiao *et al.* (2023), que estabelecem que os melhores resultados para o método são observados quando a perplexidade está entre 2 e 10% do número de amostras, desde que o valor permaneça dentro do intervalo de 5 a 50.

3.2.2 Treinamento, validação e teste

A construção da Rede Neural Artificial (RNA) foi realizada por meio da elaboração de códigos utilizando as bibliotecas *KerasTuner*, *Scikit-learn* e *TensorFlow*. Em conformidade com os estudos de May *et al.* (2008), Wadkar e Kote (2017) e Wu *et al.* (2011) estabeleceu-se a divisão dos dados dos parâmetros selecionados em 70% para treinamento, 15% para validação e 15% para o teste. Objetivando-se desenvolver um modelo aplicável a qualquer sistema de abastecimento de água rural do Brasil, foram utilizados dados em âmbito nacional.

O modelo foi construído na fase treinamento, onde os dados referentes a este subconjunto foram aplicados à rede neural de arquitetura do tipo *Multilayer Perceptron* (MLP), associada ao algoritmo *back-propagation*, devido ao amplo emprego em sistemas de modelagem de decaimento de cloro (Gibbs *et al.*, 2006; Soyupak *et al.*, 2011; Cordoba *et al.*, 2014; Al-Airami *et al.*, 2019; Santi *et al.*, 2021).

Quanto à estrutura do modelo, definiu-se por meio da biblioteca *KerasTuner*, responsável pela otimização dos hiperparâmetros: a função de ativação, os valores dos parâmetros de calibração do modelo (*momentum* e as taxas de aprendizagem), a quantidade de camadas ocultas e a quantidade de neurônios nas camadas ocultas. Essa biblioteca permite a otimização dos hiperparâmetros de maneira automática, a partir de iterações entre diferentes combinações, buscando o melhor desempenho possível.

A quantidade de lotes de dados (*batch*) e o número de ciclos de treinamento (*epochs*) foram definidos de forma manual. O algoritmo criado foi validado com os dados do subconjunto de validação, obtendo-se um erro médio quadrático (EMQ) que indica a necessidade de reajuste

dos hiperparâmetros manuais. A melhor combinação entre os parâmetros da RNA, ou seja, a melhor rede neural MLP, foi identificada através da obtenção do menor valor de EMQ.

Para obter um melhor ajuste dos dados aplicou-se o erro quadrático médio ponderado (EQMP). Ao adicionar o EQMP atribuiu-se maiores pesos aos erros de certos intervalos dos dados, onde se observou que a previsão é mais crítica, possibilitando uma adequação mais precisa do modelo. As funções de peso foram definidas buscando os melhores ajustes, utilizando como auxílio a plataforma *online* Desmos.

Com o intuito de analisar as variações entre os valores observados e previstos, o modelo gerado foi submetido à etapa de teste, por meio do cálculo de erro médio quadrático (EMQ) e da porcentagem de acertos do modelo para diferentes limiares de erro absoluto, em dez repetições.

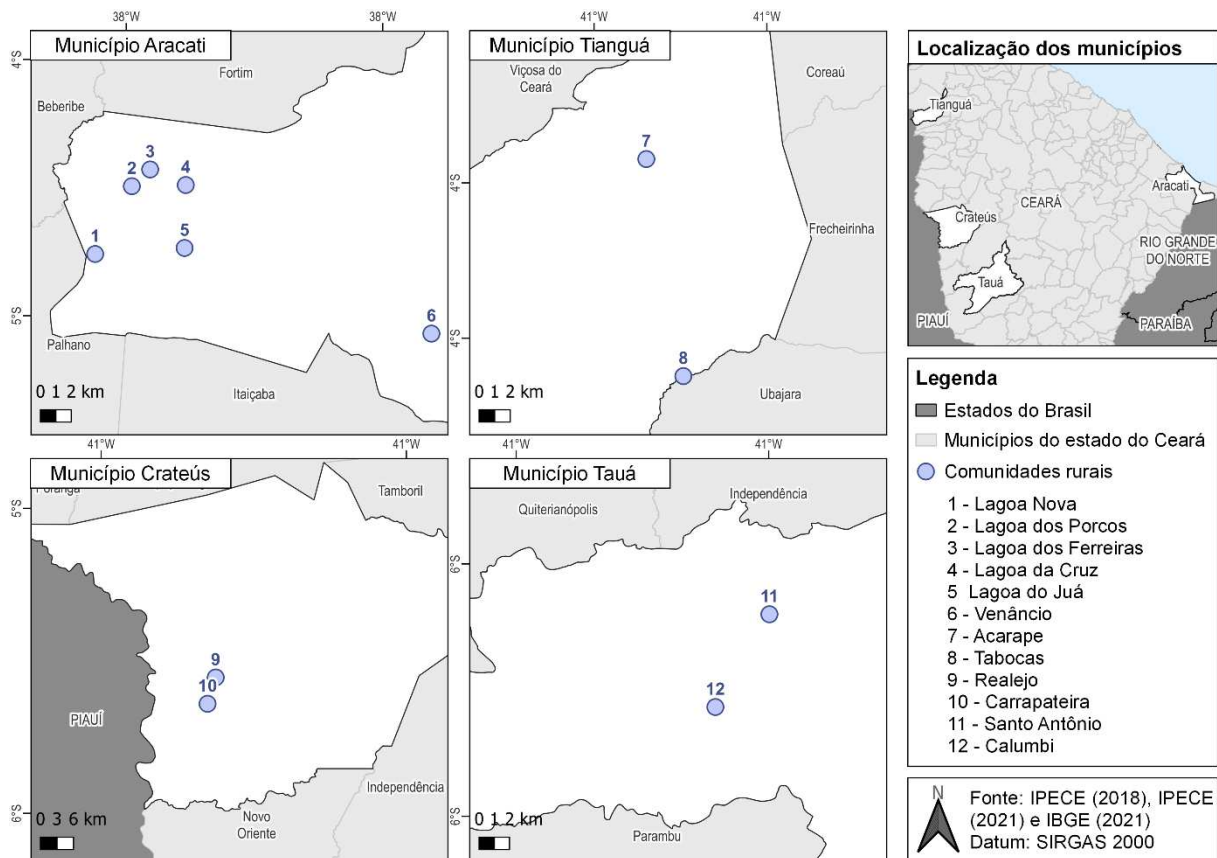
3.3 SIMULAÇÃO DO MODELO

A simulação visa verificar o desempenho do modelo de decaimento de cloro residual livre com a utilização de dados de redes de distribuição de água de comunidades rurais do estado do Ceará que não foram empregados no processo de desenvolvimento da rede neural artificial. Este processo engloba a caracterização da área de estudo e a aplicação do modelo desenvolvido.

3.3.1 Caracterização da área de estudo

A área de estudo compreende as redes de distribuição de água que atendem a 12 comunidades rurais distribuídas em quatro municípios do estado do Ceará, Brasil, apresentadas no mapa da Figura 9. Essas comunidades são atendidas por Sistemas Simplificados de Abastecimento de Água (SSAA). Esses sistemas possuem um menor porte, com processos e equipamentos menos complexos, que incluem as etapas de captação, tratamento, reservação e distribuição de água. O serviço de abastecimento de água das localidades é gerido pelo Sistema Integrado de Saneamento Rural (Sisar), que estabelece uma parceria com as associações comunitárias para a gestão dos sistemas, buscando melhorar a logística e reduzir custos.

Figura 9: Localização das comunidades rurais em estudo



Fonte: A autora (2023).

A Tabela 2 exhibe as informações territoriais, demográficas e sobre o abastecimento de água das localidades estudadas. A área de estudo compreende nove SSAA, onde destaca-se o sistema Lagoa dos Encantos, responsável pelo abastecimento de água em Lagoa Nova, Lagoa da Cruz, Lagoa dos Porcos e Lagoa dos Ferreiras. Em termos de extensão dos sistemas, levando em conta a população, extensão territorial, ligações ativas e extensão da rede de distribuição de água, esse SSAA e os sistemas Realejo e Acarape apresentam o maior porte. Já os sistemas que atendem as localidades de Lagoa do Juá e Santo Antônio possuem menor porte.

Nos SSAA, há uma predominância de uso de mananciais subterrâneos, com captação a partir de poços tubulares. O processo de tratamento, em sua maioria, consiste na simples desinfecção com cloro. Apenas nos sistemas Lagoa dos Encantos e Lagoa do Juá a captação de água é realizada através de um reservatório superficial, o Canal do Trabalhador, e o tratamento de água abrange a etapa de filtração.

Tabela 2: Caracterização dos sistemas de abastecimento de água

Sistema de abastecimento de água	Informações geográficas e demográficas das comunidades atendidas			Informações do sistema de abastecimento de água			
	Comunidade	Área (km ²)	População	Manancial	Processo de tratamento	Ligações ativas	Extensão da rede (km)
Sistema Lagoa dos Encantos	Lagoa Nova	2,7	149	Superficial	Pré-cloração + decantação + filtração + desinfecção	253	40,4
	Lagoa da Cruz	3,1	147				
	Lagoa dos Porcos	0,6	272				
	Lagoa dos Ferreiras	0,7	147				
Sistema Lagoa do Juá	Lagoa do Juá	3,1	195	Superficial	Filtração + desinfecção	44	3,92
Sistema Venâncio	Venâncio	2,3	185	Subterrâneo	Simple desinfecção	81	7,68
Sistema Carrapateira	Carrapateira	2,7	378	Subterrâneo	Simple desinfecção	280	10,14
Sistema Realejo	Realejo	3,7	900	Subterrâneo	Simple desinfecção	404	6,33
Sistema Tabocas	Tabocas	2,4	418	Subterrâneo	Simple desinfecção	168	5,39
Sistema Acarape	Acarape	1,7	1.231	Subterrâneo	Simple desinfecção	452	7,66
Sistema Calumbi	Calumbi	6,5	250	Subterrâneo	Simple desinfecção	95	6,77
Sistema Santo Antônio	Santo Antônio	0,5	210	Subterrâneo	Simple desinfecção	96	2,97

Fonte: Adaptado de França *et al.* (2023a), Funasa e UFCG (2023).

Visando assegurar o controle, potabilidade e segurança da água fornecida para as populações rurais, estas comunidades foram contempladas pelo Termo de Execução Descentralizada (TED) nº 06/2021, parceria Fundação Nacional de Saúde (Funasa) e Universidade Federal de Campina Grande (UFCG). O referido TED visa a elaboração de Planos de Segurança da Água (PSA) nas localidades, a fim de auxiliar os responsáveis pelo serviço na identificação e priorização de perigos e riscos em sistemas de abastecimento de água, desde o manancial até o consumidor (Brasil, 2012).

A caracterização da área de estudo foi realizada a partir dos dados coletados durante a elaboração do TED nº 06/2021, onde identificou-se: a localização dos equipamentos do SSAA (captação, reservatórios, tratamento, válvulas de descarga/limpeza e reforços de cloro); as características hidráulicas da rede de distribuição de água, incluindo o traçado, diâmetros e materiais das tubulações; e os parâmetros de qualidade da água ao longo da rede. Esses dados foram obtidos por meio de visitas de campo ocorridas entre junho e novembro de 2022, e através do monitoramento da qualidade da água realizado entre os meses de janeiro e junho de 2023.

Em consonância com o estabelecido pela Diretriz Nacional do Plano de Amostragem da Vigilância da Qualidade da Água para Consumo Humano (Brasil, 2016), as campanhas de coleta de amostras na rede de distribuição foram conduzidas nas pontas de rede e em locais de grande circulação ou que abrigam populações vulneráveis, como creches, escolas e postos de saúde, quando existiam.

Cada localidade apresentou uma distribuição específica de pontos de coleta. Nos sistemas Tabocas, Acarape, Realejo e Calumbi, foram coletadas amostras na escola e na ponta de rede. Já em Carrapateira e Lagoa do Juá as coletas foram realizadas no ponto de apoio da saúde e na ponta de rede. Em Santo Antônio e Venâncio todas as amostras foram coletadas na ponta de rede. Nas Lagoas, a coleta envolveu as pontas de rede, a escola da comunidade Lagoa Nova e o ponto de apoio da saúde de Lagoas dos Porcos e Lagoa dos Ferreiras.

3.3.2 Aplicação do modelo desenvolvido

Com o intuito de avaliar a capacidade de aplicação do modelo de rede neural artificial (RNA) gerado, foi realizada a validação com os dados provenientes de amostras ao longo das redes de distribuição de água das comunidades rurais estudadas. Para tanto, utilizou-se os dados coletados através de visitas de campo e do monitoramento da água distribuída das 12 comunidades, para a elaboração dos PSA.

O desempenho da RNA foi avaliado a partir do cálculo de erro médio quadrático (EMQ) e da porcentagem de acertos do modelo para diferentes limiares de erro absoluto, conforme realizado na etapa de teste. Dessa maneira, fez-se possível inferir se o modelo da rede neural gerado para prever o decaimento do cloro residual livre é afetado pelos ruídos dos subconjuntos de dados de treinamento, validação e teste e se é adequado na generalização de novos dados.

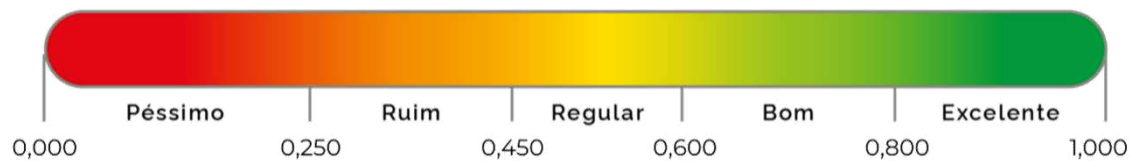
3.4 AVALIAÇÃO DE DESEMPENHO DAS REDES DE DISTRIBUIÇÃO DE ÁGUA

As redes de distribuição de água de comunidades rurais do estado do Ceará tiveram seu desempenho avaliado quanto ao atendimento do parâmetro cloro residual livre (CRL) à Portaria nº 888/2021, a partir dos indicadores de Hashimoto *et al.* (1982): confiabilidade (Equação 10), resiliência (Equação 11) e vulnerabilidade (Equação 14). A performance geral dos sistemas foi determinada pelo Índice de Sustentabilidade (Equação 15), proposto por Sandoval-Solis *et al.* (2011). As equações estão descritas no item 2.7.

A classificação do desempenho das redes de distribuição foi conduzida por meio da atribuição de uma escala nominal vinculada aos intervalos do Índice de Sustentabilidade (Figura

10). Foram estabelecidas cinco classificações nominais e as escalas numéricas correspondentes: excelente, bom, mediano, ruim e péssimo. A definição desses intervalos baseou-se em estudos anteriores na área de abastecimento de água (Cetesb, 2013; Medeiros *et al.*, 2018; Ogata *et al.*, 2016; Silva *et al.*, 2012).

Figura 10: Classificação de desempenho das redes de distribuição conforme resultado do Índice de Sustentabilidade



Fonte: A autora (2023).

A avaliação foi desenvolvida utilizando os dados de CRL coletados durante a elaboração do PSA, em um período de seis meses. O processamento dos dados foi executado em planilhas eletrônicas, empregando-se o *software* Excel©.

4. RESULTADOS E DISCUSSÕES

4.1 COLETA E PRÉ-PROCESSAMENTO DOS DADOS

Dos parâmetros considerados no estudo, exibidos no Quadro 2, excluíram-se a condutividade elétrica e os sólidos dissolvidos totais (SDT) no reservatório devido à insuficiência dos dados. Para a condutividade elétrica, não havia nenhuma amostra para os sistemas de abastecimento de água (SAA) de comunidades rurais. Libânio (2016), indica que existe uma relação linear entre a condutividade elétrica e os SDT, de modo que os sólidos dissolvidos totais correspondem a 64% do valor da condutividade. Entretanto, como existem apenas 441 amostras de sólidos totais dissolvidos de SAA rurais disponíveis na base de dados do Vigilância da Qualidade da Água para Consumo Humano (Sisagua), a representatividade e a validade da rede neural artificial seriam limitadas, caso os parâmetros fossem considerados.

Quanto às variáveis quantitativas pré-selecionadas, a Tabela 3 apresenta uma análise descritiva antes e após o pré-processamento dos dados. Mediante a coleta de informações no Sisagua, percebeu-se a existência de um conjunto de dados robusto, contemplando um número considerável de amostras, que variava entre 6.148, ao se tratar da variável de pH no reservatório, chegando a um valor de 108.874 para o cloro residual livre (CRL) na rede.

Tabela 3: Análise descritiva das variáveis quantitativas antes e depois do pré-processamento dos dados

Variável	Unidade	Pré-processamento	Estatística descritiva				
			Número de amostras	Média	Desvio padrão	Limite inferior	Limite superior
Cloro residual livre no reservatório	mg/L	Antes	9.373	1,45	4,55	0,00	200,00
		Depois	4.776	1,19	1,12	0,00	6,90
Cloro residual livre na rede	mg/L	Antes	108.874	1,03	4,17	0,00	1.600,00
		Depois	4.776	0,71	0,69	0,00	3,00
pH no reservatório	-	Antes	6.148	6,92	1,14	0,00	13,32
		Depois	2.854	7,08	1,05	1,00	13,32
Turbidez no reservatório	uT	Antes	11.742	4,16	101,55	0,00	9.999,00
		Depois	4.317	2,12	5,68	0,00	101,00
Comprimento da tubulação	m	Antes	13.452	6.553,15	9.313,15	0,00	130.333,33
		Depois	4.776	3.812,15	2.428,03	166,67	10.000,00
Número de economias residenciais	economias	Antes	157.097	2.883	18.566	0	900.000
		Depois	4.279	750	1.006	5	4.800
Temperatura do dia	°C	Antes*	-	-	-	-	-
		Depois	4.776	24,44	3,99	2,95	32,21

* A variável de temperatura do dia não apresenta valores antes do pré-processamento pois os dados foram obtidos apenas para as amostras pós-processadas.

Fonte: A autora (2023).

Após a exclusão de valores inconsistentes, com o intuito de diminuir os ruídos no pré-processamento, o número de amostras teve uma redução significativa. Novamente, o parâmetro pH deteve a menor quantidade de amostras, igual a 2.854, e as variáveis de CRL no reservatório, CRL na rede, temperatura do dia e comprimento da tubulação apresentaram 4.776 amostras, equivalendo ao maior número.

A Tabela 3 ainda oferece uma comparação entre a média, o desvio padrão e os limites inferiores e superiores dos dados de cada variável, antes e depois do pré-processamento. Antes, a estatística descritiva exibia valores com ampla variação, existindo valores inconsistentes que poderiam distorcer os padrões observados pela rede neural. Foram encontradas inconformidades como: elevados valores de turbidez e cloro residual livre, que se distanciavam significativamente dos valores máximos estabelecidos pela Portaria nº 888/2021, iguais a 5,0 uT e 5,0 mg/L, respectivamente; pH igual a zero; comprimentos de tubulação acima de 10 km, que destoam das distâncias observadas entre comunidades rurais; e, número de economias residenciais acima de 5.000, que equivaliam a populações de centros urbanos.

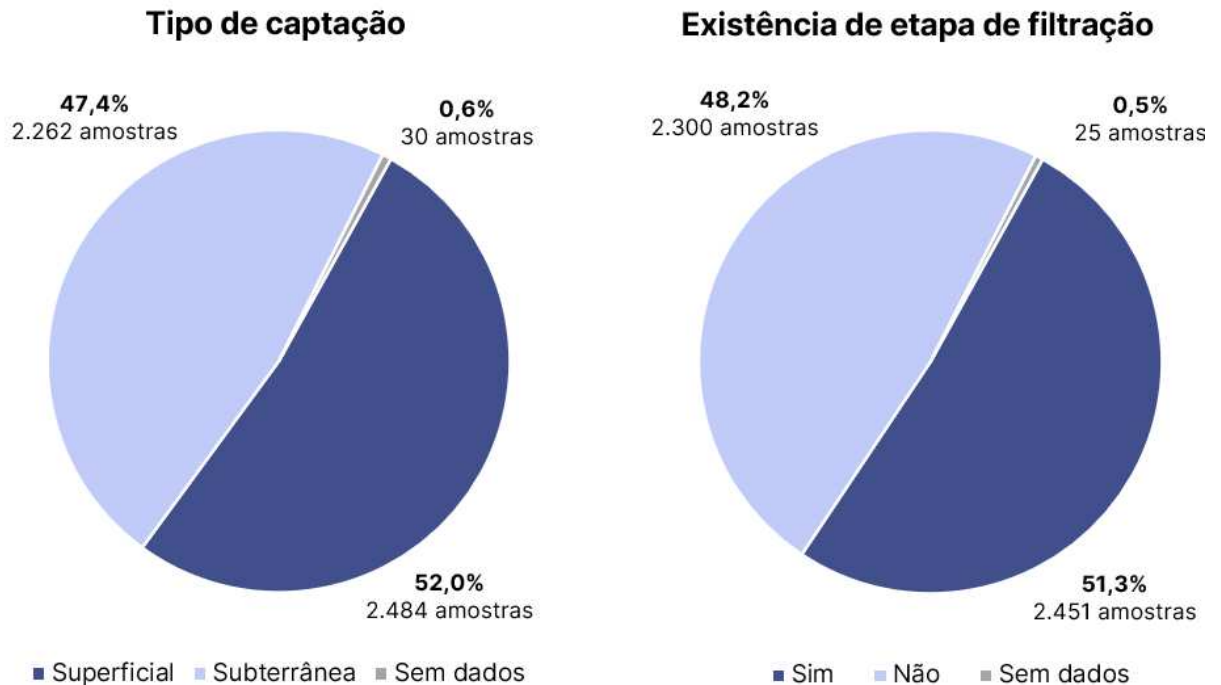
Os dados disponíveis na plataforma Sisagua são inseridos manualmente pelos profissionais da Secretaria de Saúde dos municípios e pelos profissionais que trabalham para os prestadores de serviço de abastecimento de água. Logo, tendo em vista que os SAA rurais são, em sua grande maioria, geridos pela gestão municipal, as incoerências encontradas nos dados coletados têm relação com problemas enfrentados pelos municípios, especialmente aqueles de pequeno porte, quanto a baixa capacidade institucional, deficiências técnicas da equipe, sobreposição de atividades dos agentes responsáveis pelo preenchimento dos dados e estrutura orçamentária reduzida (Moretti *et al.*, 2021; Santana *et al.*, 2021).

Após o pré-processamento (Tabela 3) verifica-se que os dados obtêm uma maior uniformidade, tendo uma redução do desvio padrão e da amplitude dos intervalos. Embora seja observada essa diminuição, para algumas variáveis os desvios padrões estão próximos ao valor da média (cloro residual livre no reservatório, cloro residual livre na rede e comprimento da tubulação) e para outras chegam a ser superior (turbidez e número de economias residenciais). Este fato é benéfico para a aplicação de redes neurais artificiais, tendo em vista a necessidade de garantir a representação de uma gama diversificada de cenários e padrões possíveis que expandem a robustez e a capacidade preditiva da RNA.

Para as variáveis qualitativas, após o procedimento de padronização e pré-processamento dos dados, pode-se verificar a distribuição das amostras por categoria, conforme exibido na Figura 11. Embora haja uma predominância de amostras com tipo de captação

superficial (52,0%) e que possuem etapa de filtração (51,3%), é perceptível que há um equilíbrio entre as categorias de cada variável. Essa equalização das amostras contribui para a mitigação de possíveis vieses na rede neural artificial, decorrentes de desproporções nas categorias originais.

Figura 11: Distribuição de amostras pré-processadas por tipo de captação e existência de filtração

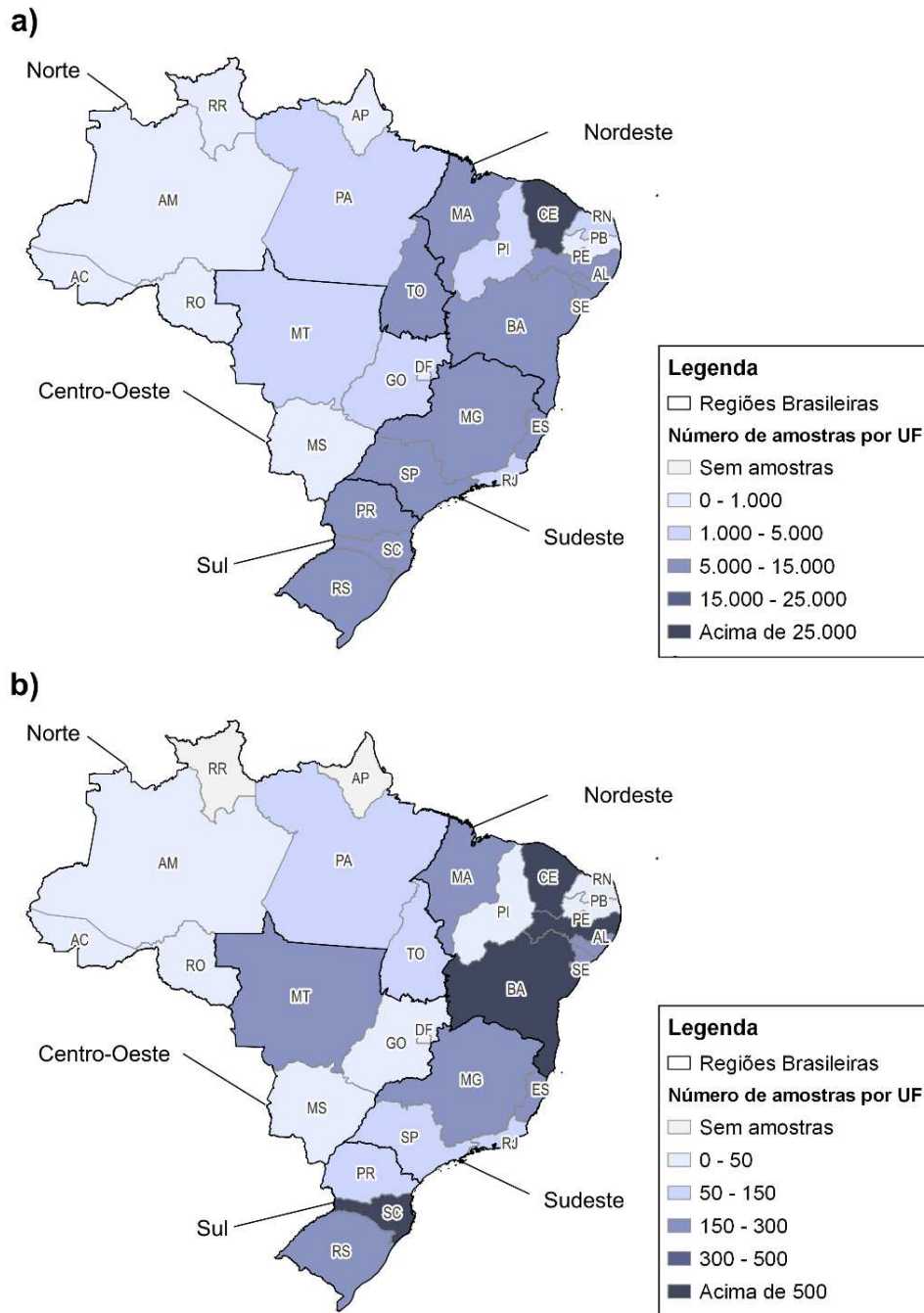


Fonte: A autora (2023).

Os resultados da coleta e pré-processamento dos dados também revelam discussões sobre a qualidade da água em comunidades rurais do território brasileiro, além de propiciar uma avaliação sobre a adequação do modelo em diferentes unidades federativas. Os mapas (a) e (b) da Figura 12 apresentam a distribuição das amostras com cloro residual livre na rede de distribuição de água das áreas rurais nos estados brasileiros antes e depois do pré-tratamento dos dados. Uma vez que o modelo foi treinado com um conjunto de dados que pode ter uma distribuição geográfica desigual, os mapas auxiliam na identificação de estados ou regiões onde o modelo terá um desempenho mais preciso.

Verifica-se uma maior concentração das amostras de cloro residual livre (CRL) nas unidades federativas pertencentes às regiões Nordeste, Sul e Sudeste. Essas regiões detêm as maiores porcentagens de atendimento de populações rurais com rede de água, conforme os dados do SNIS (2022) para o ano de 2020, iguais a 61,9%, 68,0% e 49,2%, respectivamente.

Figura 12: Distribuição das amostras com cloro residual livre na rede de distribuição de áreas rurais por unidade federativa (UF), entre os anos de 2014 e 2023: a) dados brutos; b) dados pré-processados



Fonte: Adaptado de Sisagua (2023).

Ainda sob a ótica da Figura 12, é possível observar que há uma menor quantidade de amostra para os estados da região Norte. Após o pré-processamento de dados, a zona rural das unidades federativas do Roraima (RR) e do Amapá (AP) não dispõem de amostras de CRL. O território da região abriga o menor índice nacional de cobertura de rede de distribuição de água rural (36,7%) (SNIS, 2022). Segundo Albuquerque *et al.* (2019), as fragilidades no controle e

vigilância da qualidade da água nas áreas rurais da região Norte podem estar associadas à insuficiência de recursos humanos e financeiros, além das dificuldades de acesso para as coletas de água, por conta das longas distâncias e pela dependência do transporte fluvial precário.

Com base em Araújo *et al.* (2022) e Bellotti *et al.* (2020), os baixos investimentos, a ausência de pessoal capacitado e de equipamentos e a precariedade na estrutura dos laboratórios de análise de qualidade de água para o consumo humano, reforçam as dificuldades em cumprir as exigências da Portaria nº 888/2021 do Ministério da Saúde, colocando essas populações rurais em situação de risco à saúde.

No Distrito Federal, há uma baixa incidência de amostras, chegando à quantidade nula após o pré-tratamento dos dados (Figura 12). A unidade federativa possui apenas 3% de sua população habitando em áreas rurais, sendo cerca de 15% atendida por rede de abastecimento de água a partir de 61 sistemas de abastecimento independentes (CODEPLAN, 2020). Sua extensa área rural, baixa densidade demográfica e a ampla utilização das terras e da água para o agronegócio são fatores apontados por CODEPLAN (2020) e Derntl (2020), que dificultam a expansão da infraestrutura, o controle e a vigilância da qualidade da água nessas localidades.

Em contrapartida, ao avaliar-se os mapas da Figura 12, há uma clara concentração de amostras de cloro residual livre nos estados do Ceará (CE), Pernambuco (PE), Bahia (BA) e Santa Catarina (SC). Nessas unidades federativas, destaca-se a atuação da gestão comunitária dos sistemas de abastecimento de água, a partir da atuação do Sistema Integrado de Saneamento Básico (Sisar) e da Central Bahia, bem como da Diretoria de Vigilância Sanitária do Estado de Santa Catarina (DIVS).

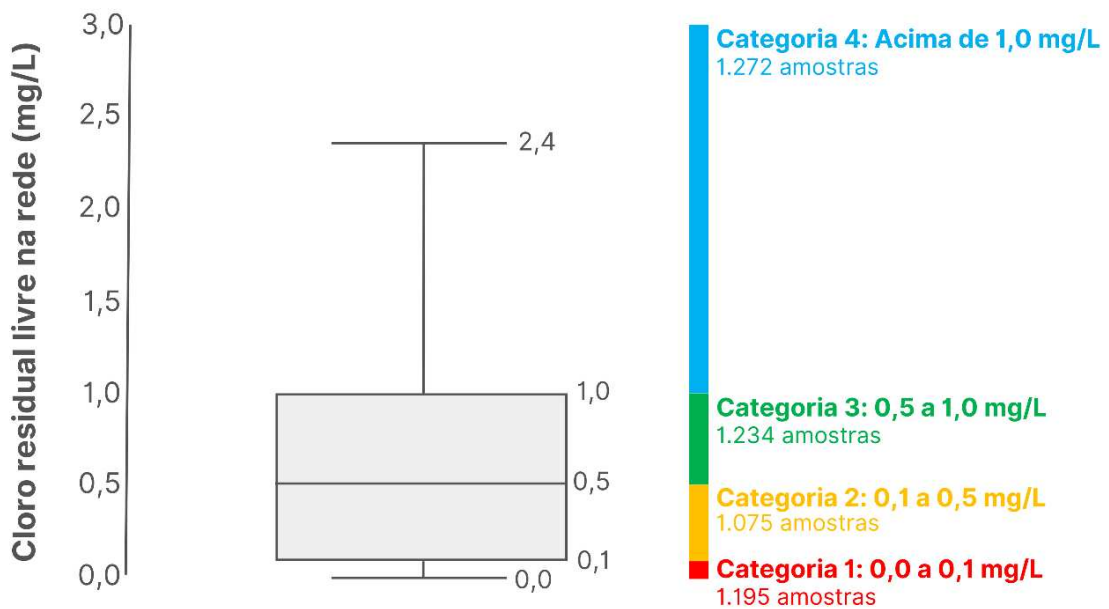
No Ceará, onde obteve-se a maior concentração de amostras de CRL antes e depois do pré-processamento, o Sisar está vinculado a Companhia de Água e Esgoto do Estado do Ceará (CAGECE) por meio da Gerência de Saneamento Rural (GESAR). O GESAR disponibiliza uma equipe multidisciplinar, veículos e estrutura física ao Sisar, garantindo a gestão dos sistemas de abastecimento de águas rurais, incluindo a vigilância da qualidade da água distribuída.

4.2 DESENVOLVIMENTO DA REDE NEURAL ARTIFICIAL (RNA)

4.2.1 Seleção das variáveis de entrada do modelo

A partir da quantidade de amostras disponíveis para a variável cloro residual livre na rede de distribuição, dividiu-se os dados em quatro categorias, definidas por meio de análise exploratória interquartil, conforme demonstrado no gráfico de *boxplot* exibido na Figura 13.

Figura 13: Intervalos de dados para cada categoria considerada no t-SNE



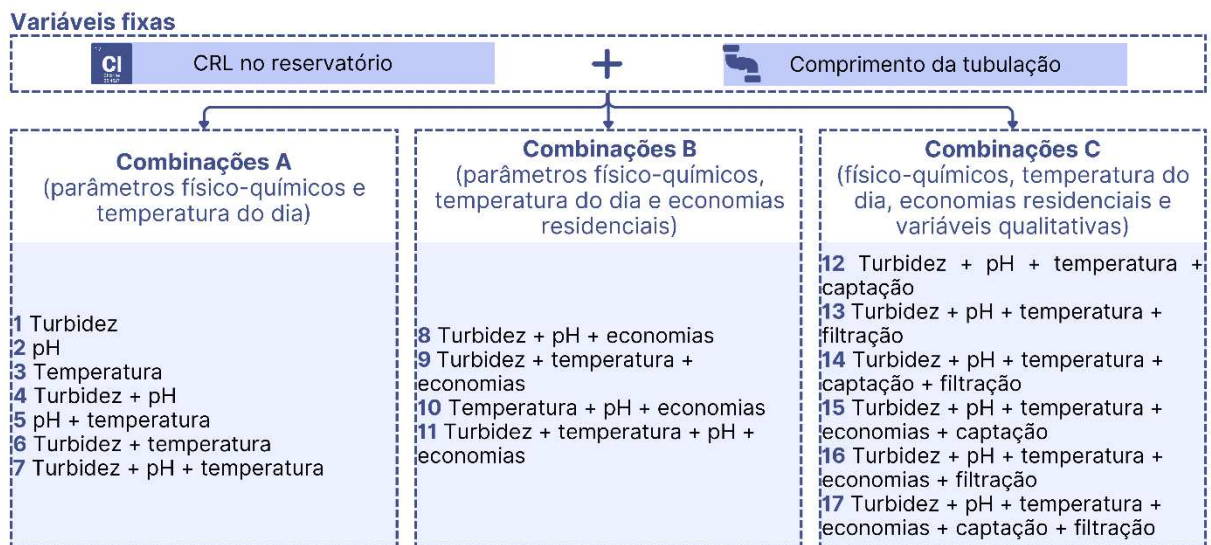
Fonte: A autora (2023).

As quatro categorias e os intervalos associados foram determinados de forma a alocar quantidades similares de amostras em cada categoria, buscando atingir um equilíbrio representativo. Essa alocação estratégica visou evitar vieses que pudessem influenciar indevidamente a seleção de variáveis pelo método *t-Distributed Stochastic Neighbor Embedding* (t-SNE), assegurando a representatividade das análises subsequentes.

Para a aplicação do t-SNE na seleção de variáveis de entrada para construção do modelo de previsão de decaimento de cloro residual livre em redes de distribuição de água, empregou-se as variáveis pré-selecionadas: Cloro residual livre no reservatório, pH no reservatório, turbidez no reservatório, comprimento da tubulação, número de economias residenciais, tipo de captação, existência de etapa de filtração e temperatura do dia. Para garantir que a rede neural artificial possa ser aplicada em qualquer ponto da rede de distribuição, optou-se por manter fixos os parâmetros de CRL no reservatório e de comprimento da tubulação.

A Figura 14 exibe um fluxograma que detalha as principais combinações testadas, designadas como A, B e C. No caso das combinações do tipo "A", foram contemplados parâmetros físico-químicos e de temperatura. Nas combinações do tipo "B", além dos parâmetros abordados nas primeiras combinações (tipo A), incluem-se informações relacionadas à economia residencial. Por último, nas combinações do tipo "C" foram considerados os parâmetros presentes nas combinações A e B, com a adição de variáveis qualitativas. Os resultados referentes à aplicação do t-SNE para as combinações são exibidos nas Figuras 15 a 17, respectivamente. Salienta-se que a perplexidade para cada combinação assumiu valores próximos à 10% do número de amostras disponíveis, assim como orienta Xiao *et al.* (2023).

Figura 14: Fluxograma com as principais combinações testadas



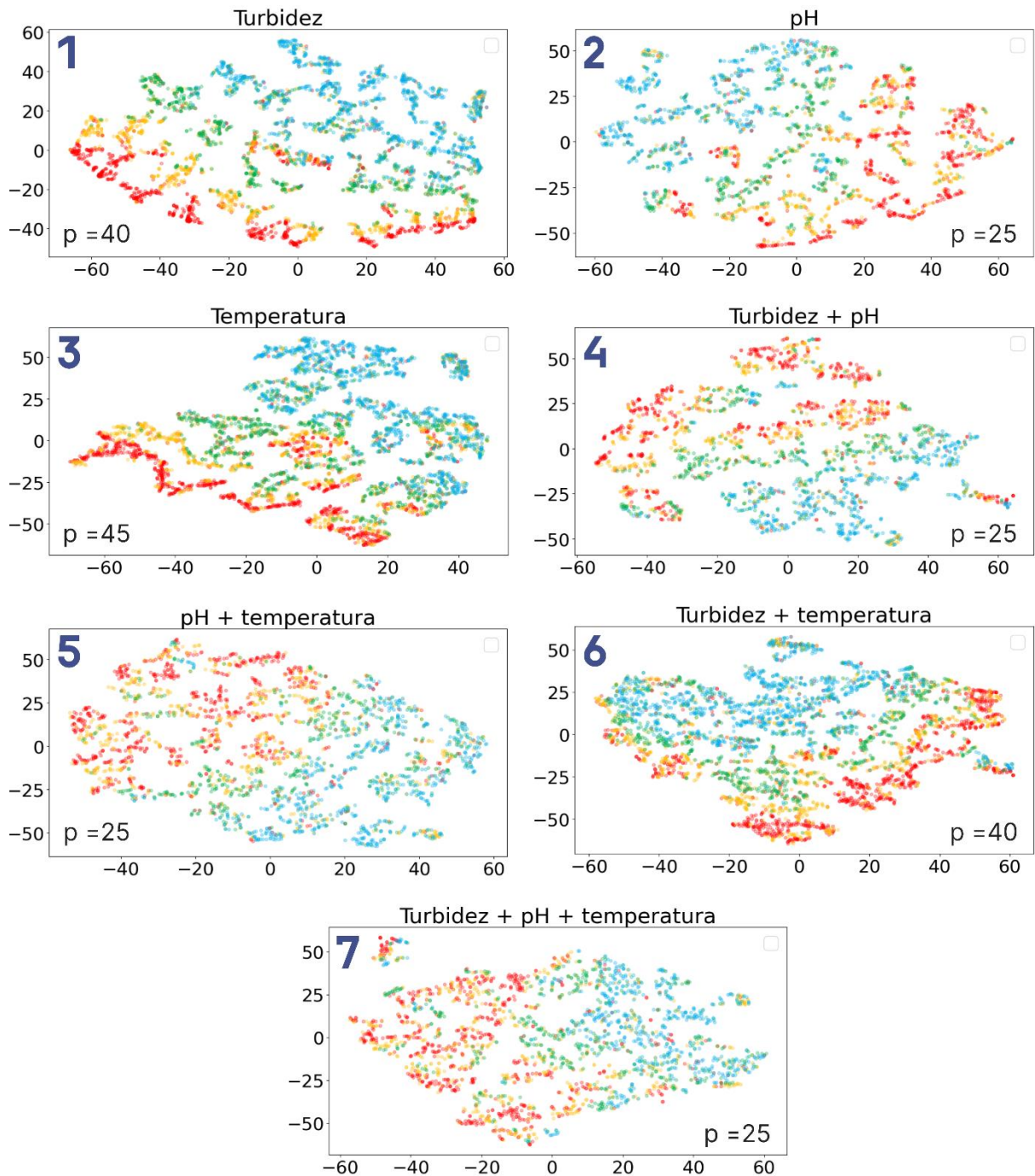
Fonte: A autora (2023).

As combinações do tipo A (Figura 15), demonstram que parâmetros físico-químicos e a temperatura do dia têm significativa relação com o cloro residual livre na rede de distribuição. Individualmente, os parâmetros de pH, temperatura e turbidez, juntamente com as variáveis fixas, geram *clusters* com boa separabilidade entre as categorias (combinações 1, 2 e 3). Os estudos de Cordoba *et al.* (2014), Santi *et al.* (2021) e Wadkar e Kote (2017) já evidenciavam a influência dessas variáveis na capacidade de previsão do decaimento do cloro por redes neurais artificiais.

Há um destaque para a turbidez (combinação 1), que obtém melhor desempenho visual na formação dos agrupamentos, possuindo uma menor interferência entre amostras de classes distintas (Figura 15). Embora o efeito da turbidez na eficácia da desinfecção tenha sido menos

explorado em comparação com o pH e a temperatura da água, Garay *et al.* (2021), Khedikar *et al.* (2021) e Léziart *et al.* (2019) apontam que esse parâmetro influencia a concentração de cloro em diversos níveis, com destaque para a demanda de cloro causada pela matéria orgânica e para o efeito-escudo que promove sobre os microrganismos, protegendo-os da ação do cloro.

Figura 15: Testes do t-SNE com combinações do tipo A, entre os parâmetros físico-químicos e a variável temperatura do dia



Onde: p é a perplexidade. Fonte: A autora (2023).

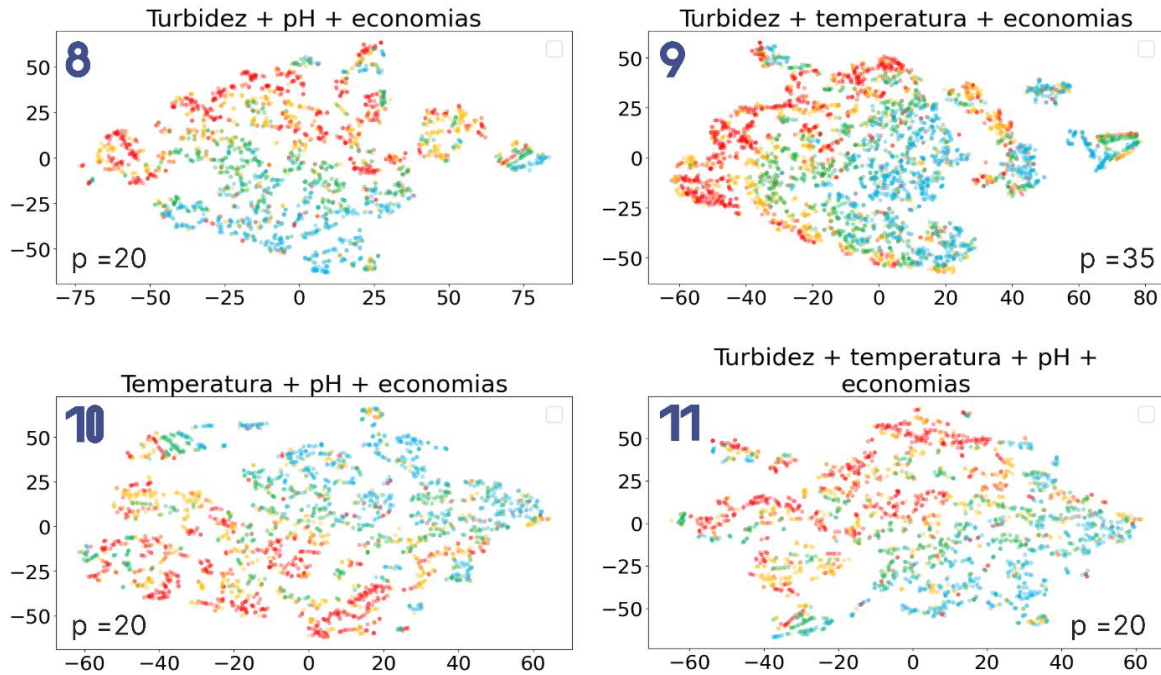
As combinações 4, 5 e 6 ilustradas na Figura 15 também demonstram resultados positivos, destacando-se a capacidade de separação eficaz entre as diferentes categorias na combinação que envolve turbidez e temperatura do dia. Tal fato é apoiado pelo estudo Gholami-Borujeni (2022) que identificou uma relação consistente entre a temperatura, a turbidez e a concentração de cloro residual livre. Apesar da combinação 5 (temperatura + pH) apresentar maiores perturbações entre as categorias, ela ainda exibe uma boa performance, devido à relação intrínseca entre as variáveis e o decaimento do cloro.

No estudo de Jafari *et al.* (2023) e Kote e Wadkar (2019), o desempenho dos modelos de aprendizado de máquina aumentou proporcionalmente com o número de variáveis de entrada. Ainda sob a ótica da Figura 15, destaca-se que a combinação 7, que engloba todos os parâmetros, foi avaliada como a mais eficaz dentro do conjunto, embora não demonstre uma separação completamente precisa entre as categorias. Como o modelo de previsão do decaimento do cloro possui uma natureza não linear, a relação entre as variáveis que afetam o fenômeno não segue um padrão de separabilidade linear ideal ao ser representada pelo t-SNE. A revisão conduzida por Onyutha e Kwio-Tamale (2022) corrobora para o desempenho positivo da combinação, ao constatarem que os parâmetros de turbidez, temperatura e pH são os mais empregados em estudos de modelagem da decomposição do cloro residual livre. Santi *et al.* (2021) verificaram que a utilização dessas variáveis melhorou a capacidade preditiva dos modelos de RNA.

A inclusão do parâmetro de economias residenciais e das variáveis qualitativas é evidenciada nas Figuras 16 e 17. Nota-se que, nos conjuntos de combinações B e C, ocorre uma maior perturbação dos dados em comparação, resultando na formação de agrupamentos que mesclam diferentes classes. O aumento do número de variáveis pode resultar em uma maior probabilidade de incluir parâmetros de entrada redundantes no modelo. Conforme Gibbs *et al.* (2006), considerar variáveis redundantes pode aumentar o grau de ruído nos dados e reduzir o desempenho da previsão.

No conjunto B (Figura 16), a inserção do número de economias residenciais pode conflitar com a variável fixa de comprimento da tubulação, uma vez que ambas estão relacionadas ao porte do sistema de abastecimento de água.

Figura 16: Testes do t-SNE com combinações do tipo B, entre os parâmetros físico-químicos, temperatura do dia e economias residenciais



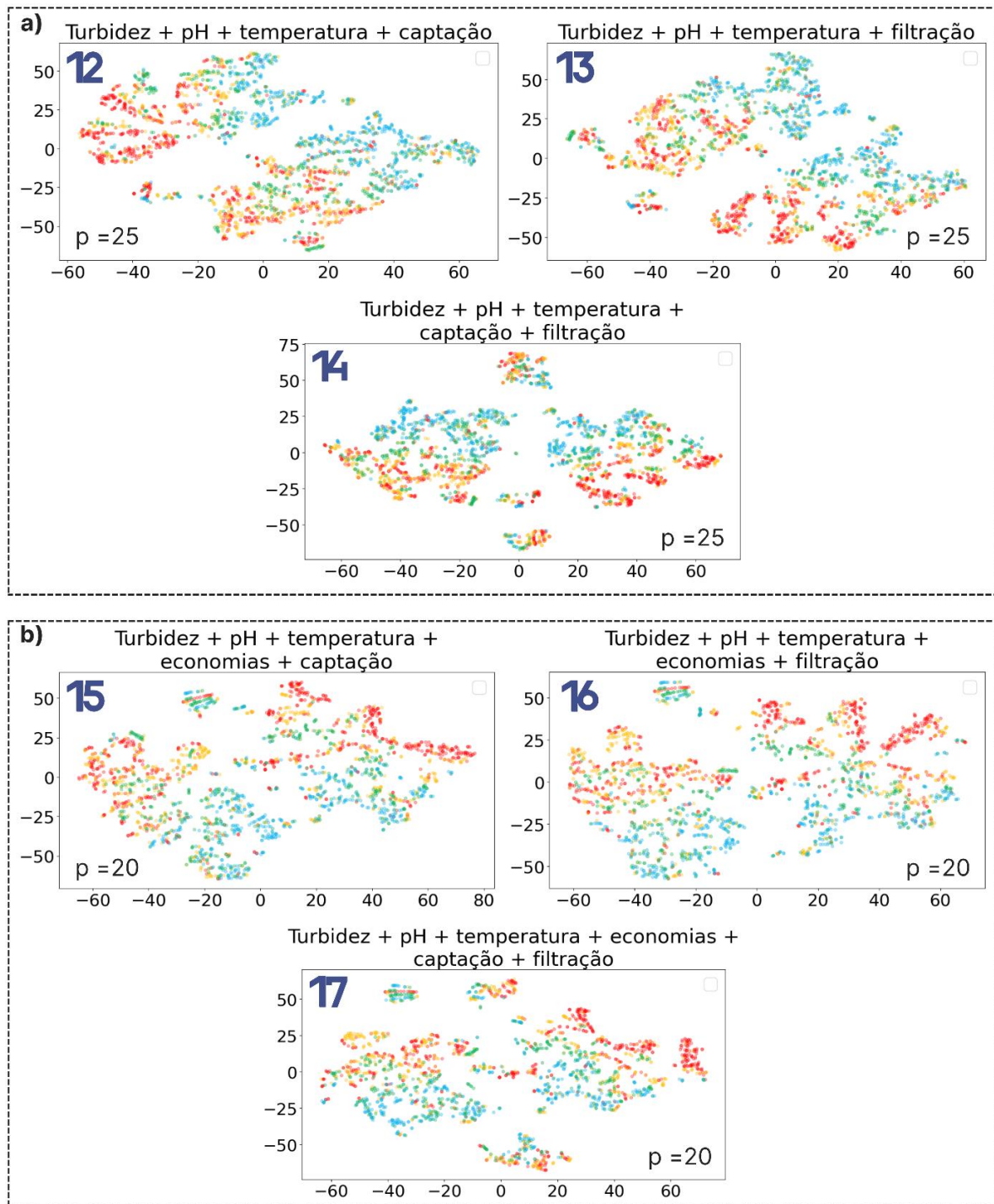
Onde: p é a perplexidade. Fonte: A autora (2023).

No caso do conjunto C (Figura 17), as variáveis qualitativas têm relação direta com a turbidez. Em captações superficiais, é observada uma maior turbidez em comparação com os mananciais subterrâneos. Além disso, a etapa de filtração atua na retenção de sólidos da água, contribuindo diretamente para a redução dos valores do parâmetro.

De maneira geral, é possível observar que as interferências nas combinações ocorrem principalmente na categoria 3 (verde), com variação de 0,5 a 1,0 mg/L de cloro residual livre (CRL), que tende a se misturar com outras classes, o que pode dificultar a distinção precisa dos valores previstos pela rede neural artificial nesse intervalo de valores de CRL. As categorias 1 (vermelho) e 2 (laranja), com intervalos de CRL de 0,0 a 0,5 mg/L, geralmente apresentam sobreposições entre si, indicando possíveis dificuldades de diferenciação entre essas duas faixas de concentração de cloro. Ademais, a categoria 1 apresenta algumas amostras distribuídas ao longo dos agrupamentos da categoria 4 (azul), onde o cloro possui valores acima de 1,0 mg/L. Esses resultados são relevantes para o desenvolvimento e ajuste da rede neural artificial, pois podem impactar a precisão das previsões do cloro residual livre.

Em consonância com os aspectos discutidos, escolheu-se empregar a combinação 7 (tipo A) para o treinamento das redes neurais artificiais, que abrange as seguintes variáveis: cloro residual livre no reservatório, turbidez no reservatório, pH no reservatório, temperatura do dia e comprimento da tubulação.

Figura 17: Testes do t-SNE com combinações do tipo C: a) parâmetros físico-químicos, temperatura do dia e variáveis qualitativas; b) parâmetros físico-químicos, temperatura do dia, economias residenciais e variáveis qualitativas



Onde: p é a perplexidade. Fonte: A autora (2023).

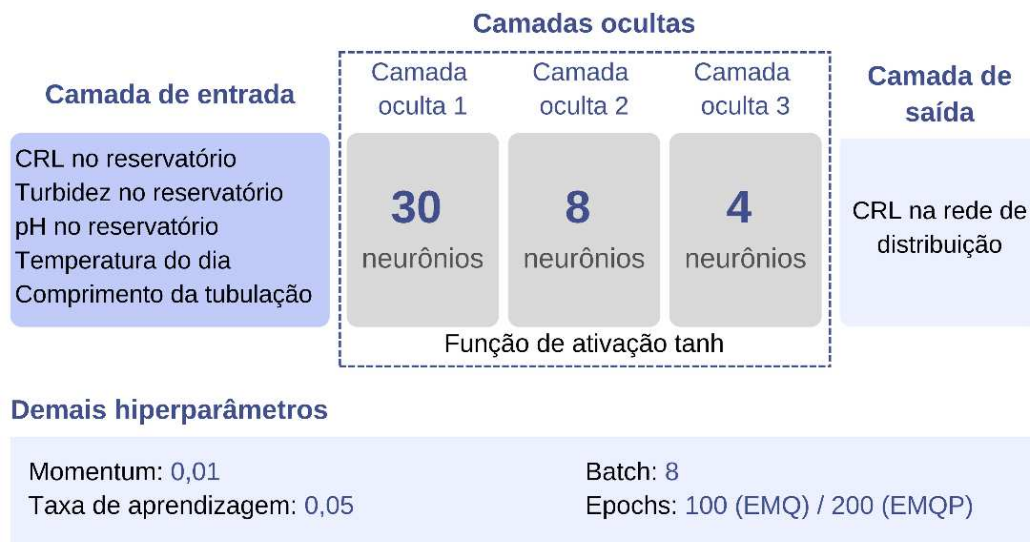
Em contraste com os estudos realizados por May *et al.* (2008), Santi *et al.* (2021) e Wadkar e Kote (2017) que conduziram vários testes aplicando as redes neurais em diferentes combinações de variáveis, a utilização do *t-Distributed Stochastic Neighbor Embedding* (t-SNE) oferece uma abordagem mais eficiente em termos de recursos computacionais. Com o t-SNE,

foi possível analisar diversas combinações de variáveis, destacando as com maior potencial, de maneira otimizada, com economia de tempo e recursos, sem comprometer a qualidade da análise.

4.2.2 Treinamento, validação e teste

Conforme mencionado no tópico 3.2.2, os hiperparâmetros foram definidos a partir da biblioteca *Scikit-learn*, chegando a melhor configuração apresentada na Figura 18, composta por três camadas ocultas com 30, 8 e 4 neurônios, respectivamente, todas utilizando a função de ativação tangente hiperbólica (tanh). Os pesos e vieses para a RNA de melhor desempenho treinada com o erro médio quadrático (EMQ) e com o EMQ ponderado são exibidos nos Apêndices A e B, respectivamente.

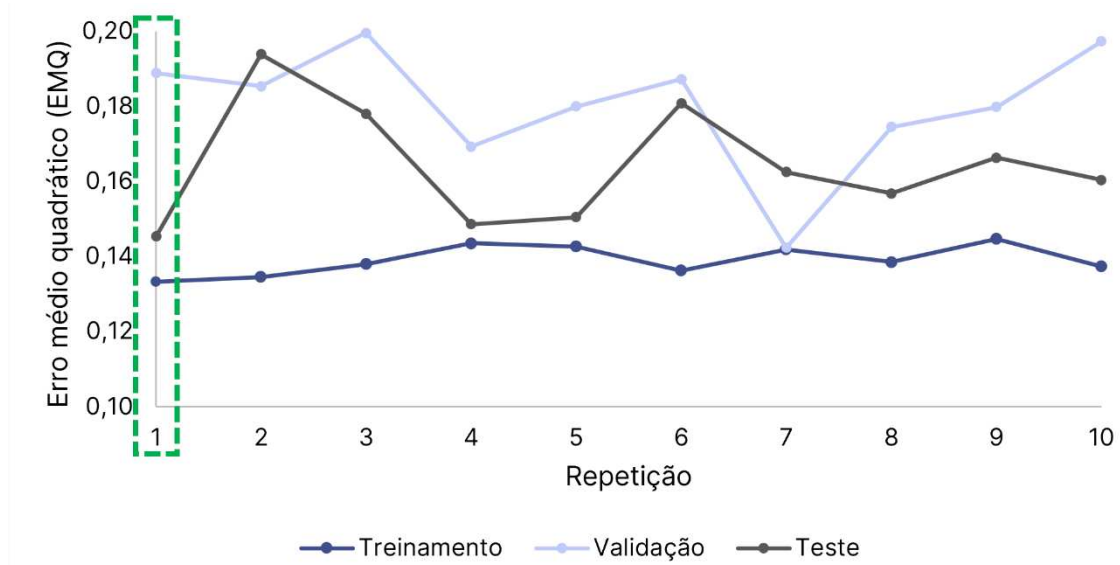
Figura 18: Hiperparâmetros definidos para a rede neural artificial



Fonte: A autora (2023).

Durante o processo de treinamento com o erro médio quadrático, foi possível realizar uma análise do desempenho da rede neural ao longo das dez repetições realizadas como pode ser visualizado na Figura 19. Os resultados indicam que, em média, o EMQ para o conjunto de treinamento foi de 0,1391, para a validação foi de 0,1804 e para o teste foi de 0,1643. Salienta-se que a taxa de aprendizado e o número de épocas utilizadas evitaram o super/subajuste da rede neural aos dados de treino. Embora os erros observados na etapa de validação sejam maiores que no treinamento, não existem diferenças significativas entre os valores, tendo ocorrido uma convergência gradual entre os EMQ ao longo das épocas.

Figura 19: Erro médio quadrático por repetição para a RNA treinada por EMQ

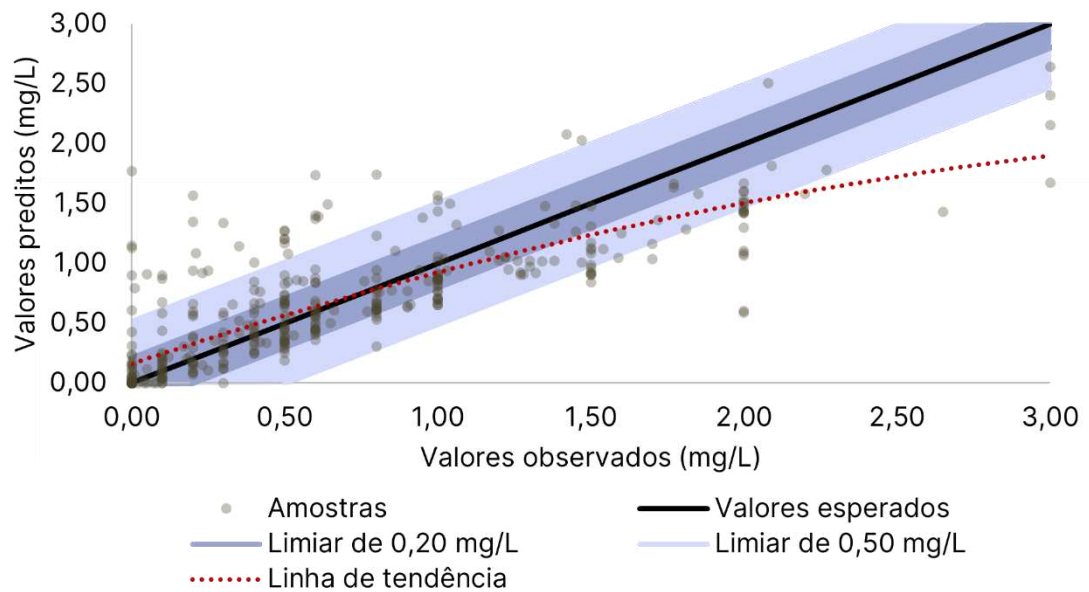


Fonte: A autora (2023).

Conforme a Figura 19, a repetição 1 foi a preferida, por apresentar os menores erros no treinamento e no teste, iguais a 0,1333 e 0,1453, respectivamente. Santi (2021), ao desenvolver redes neurais destinadas à previsão de decaimento de cloro em sistemas de abastecimento de seis assentamentos de refugiados, constatou erros de teste variando entre 0,01 e 0,25. Nos estudos conduzidos por Gibbs (2006) e Kote e Wadkar (2019), de escala local, observou-se uma flutuação do erro médio quadrático entre 0,001 e 1,005. Assim, os valores obtidos destacam-se de maneira positiva em relação às investigações precedentes, uma vez que se refere a um modelo de escopo nacional, abordando uma grande diversidade de padrões.

Para a repetição 1, a comparação entre os valores preditos e observados de cloro residual livre na rede de distribuição pode ser verificada no gráfico da Figura 20. A porcentagem de acerto foi igual a 13,25%. Considerando os limiares de erro absoluto de 0,2 a 0,5 mg/L, que correspondem às faixas de precisão dos equipamentos para medição de CRL, obteve-se uma taxa de acerto variando de 61,04% a 83,64%. É possível visualizar que, para as amostras de até 1,0 mg/L de cloro residual, o modelo reproduz de maneira eficaz o comportamento das observações, com porcentagens de acerto entre 72,79% e 89,05%. Alguns dados nesse intervalo se distanciam dos limiares e são superestimados. Tal fato pode ter relação com a tendência de dispersão da categoria 3 (0,5 a 1,0 mg/L de CRL) e a distribuição de amostras da categoria 1 (0,0 a 0,2 mg/L de CRL) nos *clusters* da categoria 4 (acima de 1,0 mg/L de CRL) verificada a partir da aplicação do método t-SNE.

Figura 20: Comparação entre os valores preditos e observados de cloro residual livre para a RNA treinada por EMQ

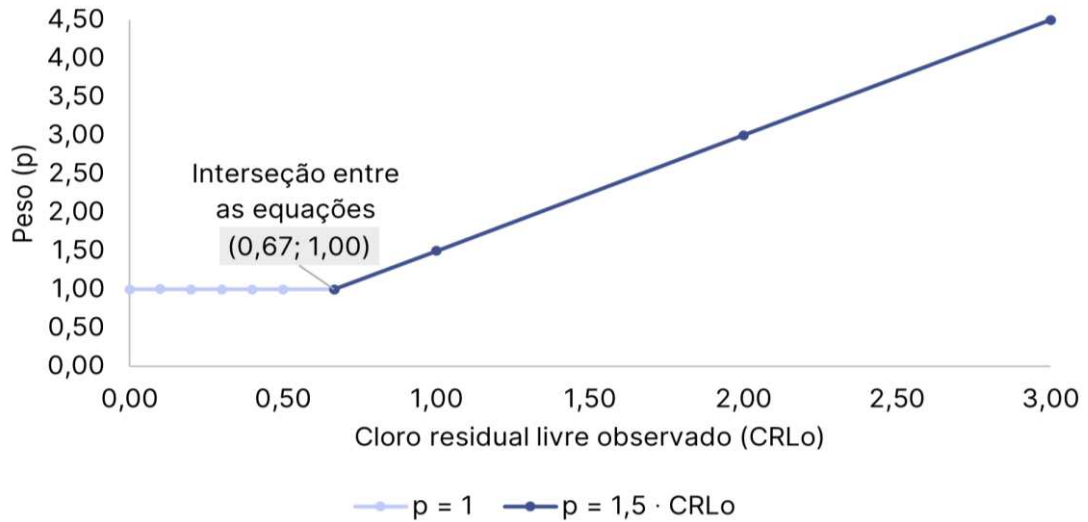


Fonte: A autora (2023).

Detecta-se ainda que há uma propensão de afastamento das amostras com cloro a partir de 1,0 mg/L da reta dos valores esperados, conforme a linha de tendência do gráfico, de modo que são subestimadas, com taxas de acerto variando de 28,43 a 68,63% para limiares de 0,2 a 0,5 mg/L, respectivamente. Estes pontos, correspondem a valores atípicos de cloro, gerando dificuldades de adaptação da rede neural, assim como observado por Gibbs *et al.* (2006) e Santi *et al.* (2021). No estudo de Santi *et al.* (2021), os autores indicaram que os conjuntos de RNA treinados com EMQ não ponderado produziram previsões subdispersas, com o percentual de captura variando de 0 a 50%. Embora comparando-se com este estudo, o modelo gerado tenha obtido êxito, o emprego do erro médio quadrático ponderado (EMQP) pôde melhorar a previsão para esse intervalo crítico.

No treinamento com o método do EMQP, aplicou-se um peso (p) ao erro médio quadrático por meio de uma função linear que corresponde ao máximo entre as funções $p = 1$ e $p = 1,5 \cdot CRL_0$, onde CRL_0 é o valor observado de cloro residual livre (Figura 21). Esta equação demonstrou ser a mais eficaz no ajuste dos dados, uma vez que não ponderou o EMQ para concentrações de cloro mais baixas, onde já se observava um bom ajuste, e atribuiu maior ponderação ao erro à medida que o valor do cloro aumentava. A elevação progressiva do peso permitiu um ajuste mais preciso para as maiores concentrações de cloro, em que se verificava um maior distanciamento entre os valores previstos e os valores observados.

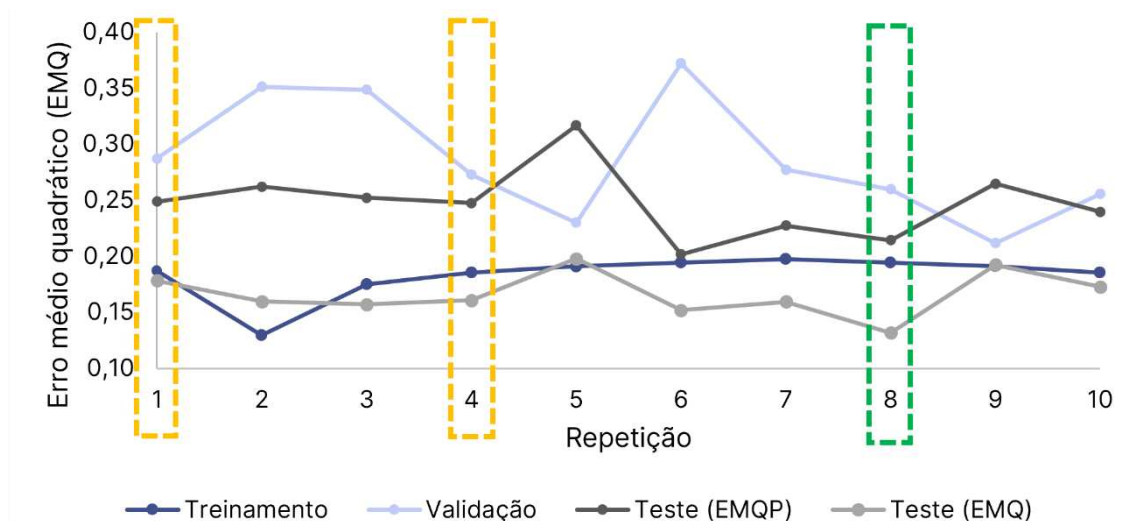
Figura 21: Função linear de ponderação do erro médio quadrático (EMQ)



Fonte: A autora (2023).

Os erros médios quadráticos (EMQ) obtidos para os conjuntos de teste e o EMQP para o treinamento, validação e teste, após as dez repetições, estão representados na Figura 22. A média dos EMQP foi de 0,1832 para o conjunto de treinamento, 0,2867 para a validação e 0,1662 para o teste. Observou-se que os valores de erro no treino e na validação foram superiores em relação ao treinamento pelo EMQ, devido à ponderação utilizada no método. Com o maior número de épocas utilizado, conforme apresentado na Figura 18, a diferença média entre as perdas no treinamento e validação aumentou de 0,04139 para 0,10358, em comparação à avaliação sem os pesos. Entretanto, ao analisar que os EMQP do teste estão próximos aos do treino, é possível verificar que esse fato não afeta a capacidade de generalização da rede.

Figura 22: Erro médio quadrático por repetição para a RNA treinada por EMQP

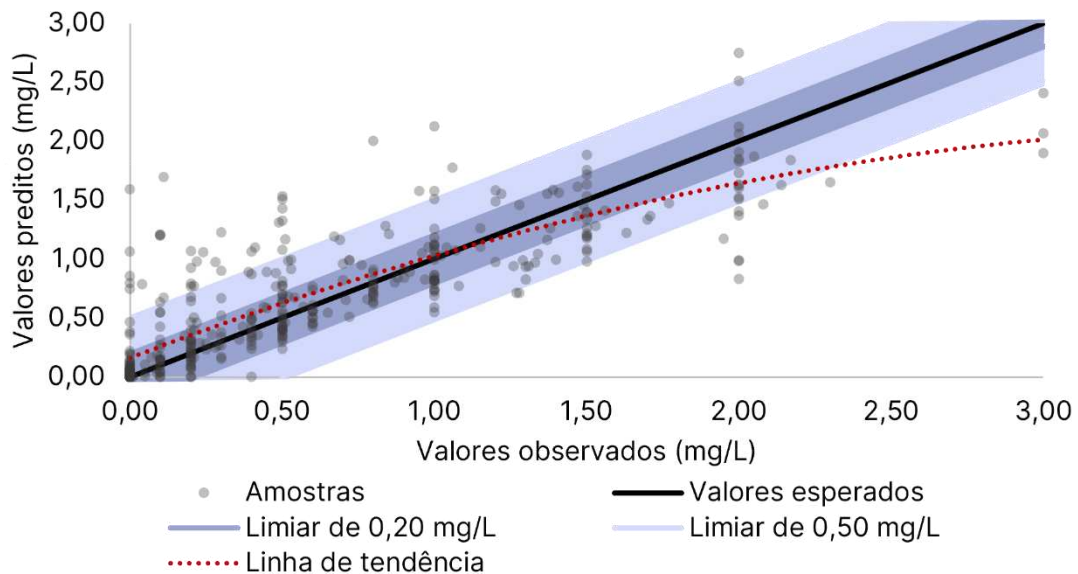


Fonte: A autora (2023).

Avaliando a Figura 22, nota-se que a repetição 8 obteve melhor desempenho por apresentar o menor erro médio quadrático de teste, além da tendência de convergência entre os EMQP dos conjuntos de treinamento e validação. Apesar disso, as repetições 1 e 4 apresentaram as melhores performances para os limiares de erro absoluto de 0,2 e 0,5 mg/L, respectivamente, com porcentagens de acerto de 66,23% para a repetição 1 e de 87,27% para a repetição 4.

Para a repetição 8, a Figura 23 exibe a comparação entre os valores previstos e observados de cloro residual livre na rede de distribuição. A taxa de acerto global para esta repetição foi de 8,57%. Novamente para os limiares de erro absoluto situados entre 0,2 e 0,5 mg/L, verificou-se uma taxa de acerto oscilando de 64,68% a 85,71%. Com a ponderação do EMQ durante o treinamento, houve um melhor ajuste dos dados acima de 1,0 mg/L de CRL aos limiares considerados, obtendo-se porcentagens de acerto variando entre 38,10% e 80,00%. Em contrapartida, algumas amostras de cloro no intervalo de 0,0 a 0,5 mg/L acabaram se distanciando da reta dos valores esperados, onde a taxa de acerto no intervalo ficou entre 74,78% e 86,73%. Isso pode ser atribuído à influência dos pesos conferidos ao erro. Ao dar mais peso aos erros quadráticos maiores, o processo de treinamento prioriza a correção das amostras com discrepâncias significativas.

Figura 23: Comparação entre os valores preditos e observados de cloro residual livre para a RNA treinada por EMQP



Fonte: A autora (2023).

De forma geral, os pontos que excederam os limiares de 0,2 e 0,5 mg/L de cloro residual livre correspondem às unidades federativas da Bahia e de Santa Catarina. Isto se deve à alta concentração de amostras de cloro residual livre nesses estados, após o pré-processamento dos

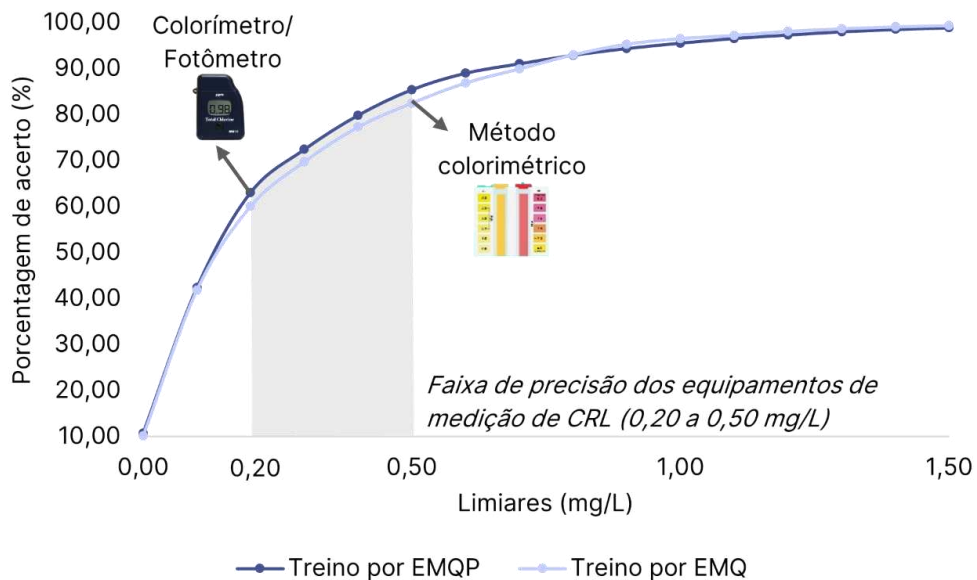
dados, como evidenciado no mapa da Figura 12b. Essa ocorrência é esperada devido à elevada incidência de amostras nesses locais. Além disso, é relevante mencionar que os dados de entrada dessas amostras não demonstraram a presença de valores atípicos em comparação com a estatística descritiva apresentada na Tabela 3.

A ocorrência de concentrações de cloro residual livre na rede de distribuição iguais aos níveis registrados no reservatório também contribuíram para a dispersão de alguns pontos no gráfico da Figura 23. Essas amostras, que não refletem adequadamente o decaimento real do cloro, estão presentes principalmente nos valores observados de 1,0, 2,0 e 3,0 mg/L de CRL.

Adicionalmente, foi observada que a predominância de pontos fora dos limiares corresponde a amostras de água provenientes de mananciais superficiais e/ou foram submetidas à etapa de filtração. Essas características do sistema contribuem para variações significativas da turbidez, condicionadas aos tipos de partículas e impurezas presentes na água, à variabilidade climática e a eficiência do tratamento, podendo gerar perturbações na rede neural artificial.

O gráfico apresentado na Figura 24 ilustra a porcentagem média de acertos para diferentes limiares de erro absoluto, comparando o desempenho de uma rede neural artificial treinada antes e depois da aplicação da ponderação do erro médio quadrático. Assim como verificado por Santi (2021), o treinamento de RNAs utilizando o EMQP resultou em melhorias significativas no desempenho da previsão em diversos indicadores de dispersão e confiabilidade.

Figura 24: Porcentagem média de acertos para diferentes limiares de erro absoluto para RNA treinada por EMQ e EMQP



Fonte: A autora (2023).

Observa-se uma certa elevação no desempenho, especialmente para limiares compreendidos entre 0,1 e 0,7 mg/L de cloro, indicando uma otimização mais eficaz na previsão nesse intervalo (Figura 24). Destaca-se, particularmente, o limiar de 0,2 mg/L de cloro residual livre, onde a aplicação do método elevou a taxa de acerto média de 60,10% para 63,12%. Para o limiar de 0,5 m/L de CRL, a porcentagem passou de 82,52% para 85,45%, correspondendo ao segundo maior aumento. Embora nenhuma das abordagens tenha conduzido a uma captura perfeita, a utilização do EMQP evidencia um aprimoramento da capacidade preditiva da RNA.

Assim como previamente abordado e demonstrado no gráfico da Figura 24, os limiares de 0,2 a 0,5 mg/L de cloro residual livre correspondem à faixa típica de precisão dos equipamentos utilizados na medição de CRL. O limite inferior (0,2 mg/L de CRL) está associado aos equipamentos como colorímetros/fotômetros de bolso, enquanto o limite superior (0,5 mg/L de CRL) está relacionado a métodos colorimétricos.

Apesar desses métodos tradicionais serem amplamente adotados, sua utilização implica em um maior consumo de recursos. Conforme cotação realizada no ano de 2023, o colorímetro/fotômetro tem os maiores custos associados, variando entre 370,50 a 724,00 reais. Em contrapartida, a colorimetria, é uma alternativa indicada pelo Manual Prático de Análise de Água da Fundação Nacional de Saúde (Brasil, 2013) por ser mais econômica, com custos oscilando entre 39,06 e 79,50 reais, tendo soluções de teste suficientes para cerca de 50 análises.

Contudo, segundo Soares *et al.* (2016), o uso da colorimetria pode resultar em uma redução na precisão da determinação da concentração de cloro, devido à sua susceptibilidade a interferências externas. Isso se deve, à possibilidade de ocorrerem erros na leitura visual, nos quais fatores como a luminosidade e a subjetividade na avaliação visual do operador podem influenciar a interpretação da cor, afetando a acuidade da medição.

O método colorimétrico é o mais empregado na vigilância e controle da qualidade da água de sistemas de abastecimento de comunidades rurais. A Figura 25 apresenta o kit comparador de cloro e pH utilizado pelos operadores do Sistema Integrado de Saneamento Rural (Sisar) na medição diária do cloro residual livre na rede de distribuição de água, conforme indica Sisar (2018).

Figura 25: Kit comparador de cloro e pH empregado na medição de cloro residual livre



Fonte: A autora (2023).

Observa-se na imagem que a precisão de 0,5 mg/L se aplica somente ao intervalo de CRL até 2,0 mg/L, limiar no qual a rede neural artificial treinada com erro médio quadrático ponderado (EMQP) demonstrou uma taxa de acerto média de 85,45%. A partir desse ponto até 3,0 mg/L de CRL, é observada uma precisão de 1,0 mg/L e acima desse valor é igual a 2,0 mg/L, limiares em que a RNA apresentou uma porcentagem de acerto de 95,58% e 99,84%, respectivamente. As altas taxas de acerto, associadas as precisões correspondentes ao método colorimétrico, demonstram que o modelo desenvolvido representa uma abordagem promissora na previsão dos níveis de cloro residual livre em sistemas de distribuição de água rurais.

4.3 SIMULAÇÃO DO MODELO

4.3.1 Caracterização dos sistemas de distribuição de água estudados no estado do Ceará

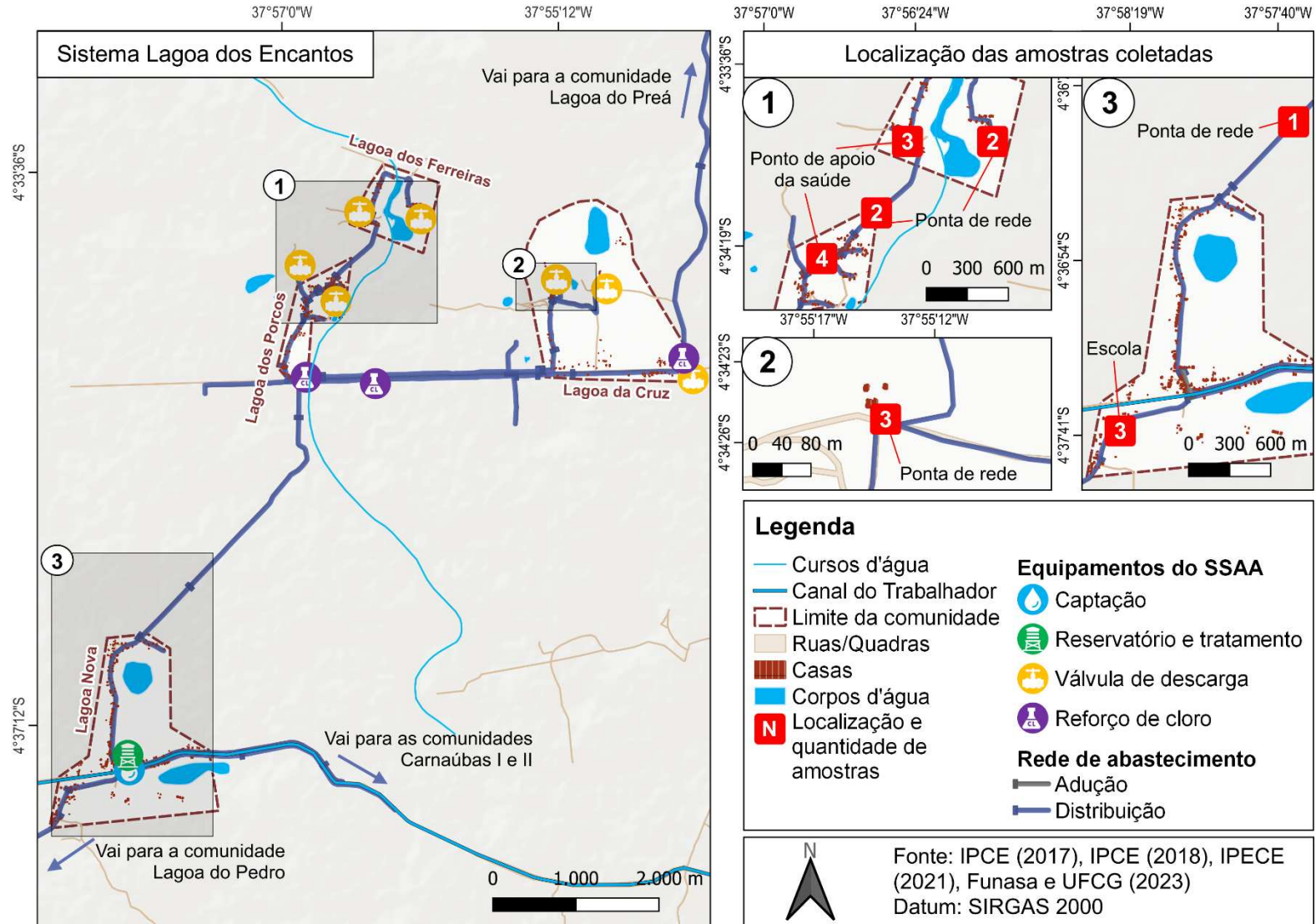
Os mapas das Figuras 26 a 34 apresentam o mapeamento da rede de distribuição de água dos sistemas, incluindo informações sobre os pontos e as quantidades de amostras de água coletadas. As tubulações das redes de distribuição representadas nos mapas são constituídas de Policloreto de Vinila (PVC) e polietileno e apresentam diâmetros que variam de 32 a 150 mm, sendo o diâmetro de 60 mm predominantemente utilizado.

Conforme os dados coletados por Funasa e UFCG (2023), os pontos de quebras frequentes ocorrem nas tubulações de polietileno, utilizadas para conduzir água da rede principal para as residências, devido ao tráfego de veículos e ao ressecamento causado pela exposição ao sol. Nas comunidades Carrapateiras, Calumbi, Realejo e Santo Antônio algumas tubulações são danificadas quando há o nivelamento e terraplanagem das estradas após o período chuvoso.

Em relação às válvulas de descarga, apenas os sistemas Lagoa dos Encantos (Figura 26), Lagoa do Juá (Figura 27) e Realejo (Figura 30) possuem quantidades adequadas destes dispositivos. Nos sistemas de Venâncio (Figura 28) e Carrapateira (Figura 29), identificou-se apenas uma válvula de descarga instalada, enquanto nos sistemas de Tabocas (Figura 31), Acarape (Figura 32), Calumbi (Figura 33) e Santo Antônio (Figura 34), não foram identificadas a presença desses dispositivos. As válvulas de descarga são projetadas para facilitar a limpeza das tubulações, permitindo o esvaziamento quando necessário, para a realização de reparos e manutenção geral do sistema de abastecimento de água.

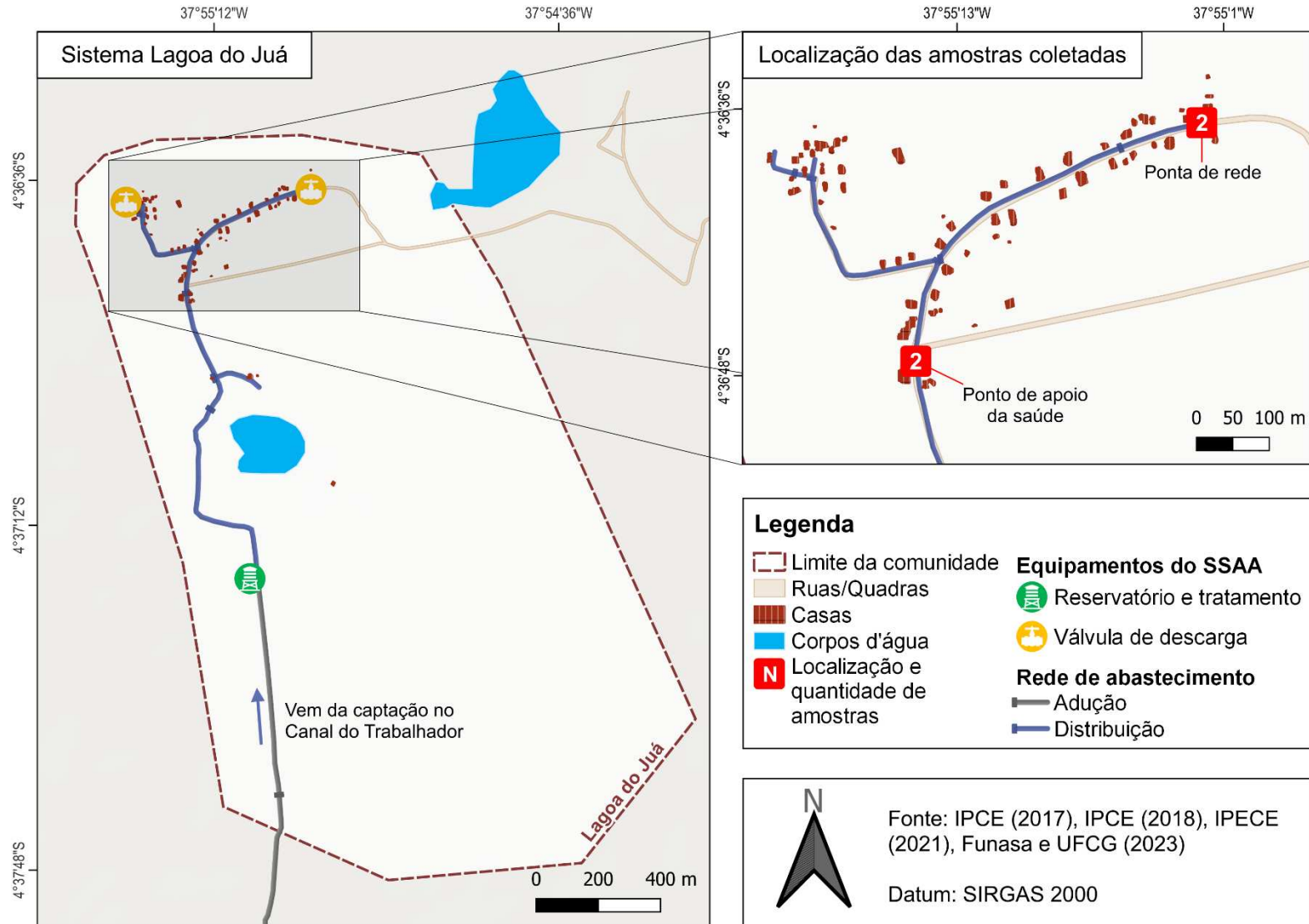
As quebras frequentes e a ausência de limpeza adequada nas tubulações da rede de distribuição de água dessas comunidades podem favorecer a contaminação da água. A presença de depósitos e biofilmes e a entrada de contaminantes externos (partículas do solo, microrganismos e outros poluentes) nas tubulações corroboram para redução da concentração do cloro residual livre ao longo do percurso da água, comprometendo sua capacidade de manter a qualidade microbiológica adequada até a torneira do consumidor.

Figura 26: Mapa da rede de distribuição e localização das amostras do Sistema Lagoa dos Encantos



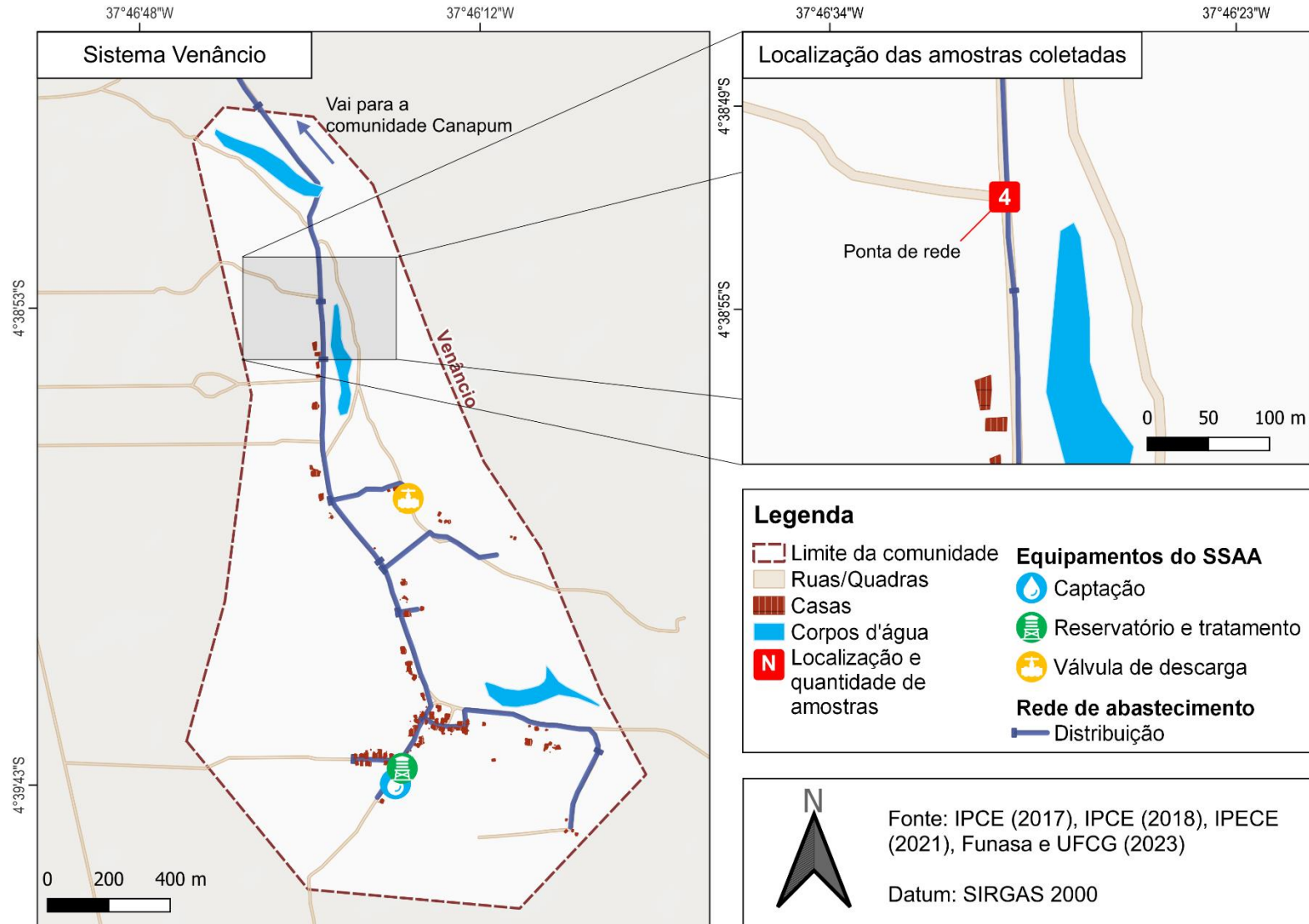
Fonte: Adaptado de Funasa e UFCG (2023).

Figura 27: Mapa da rede de distribuição e localização das amostras do Sistema Lagoa do Juá



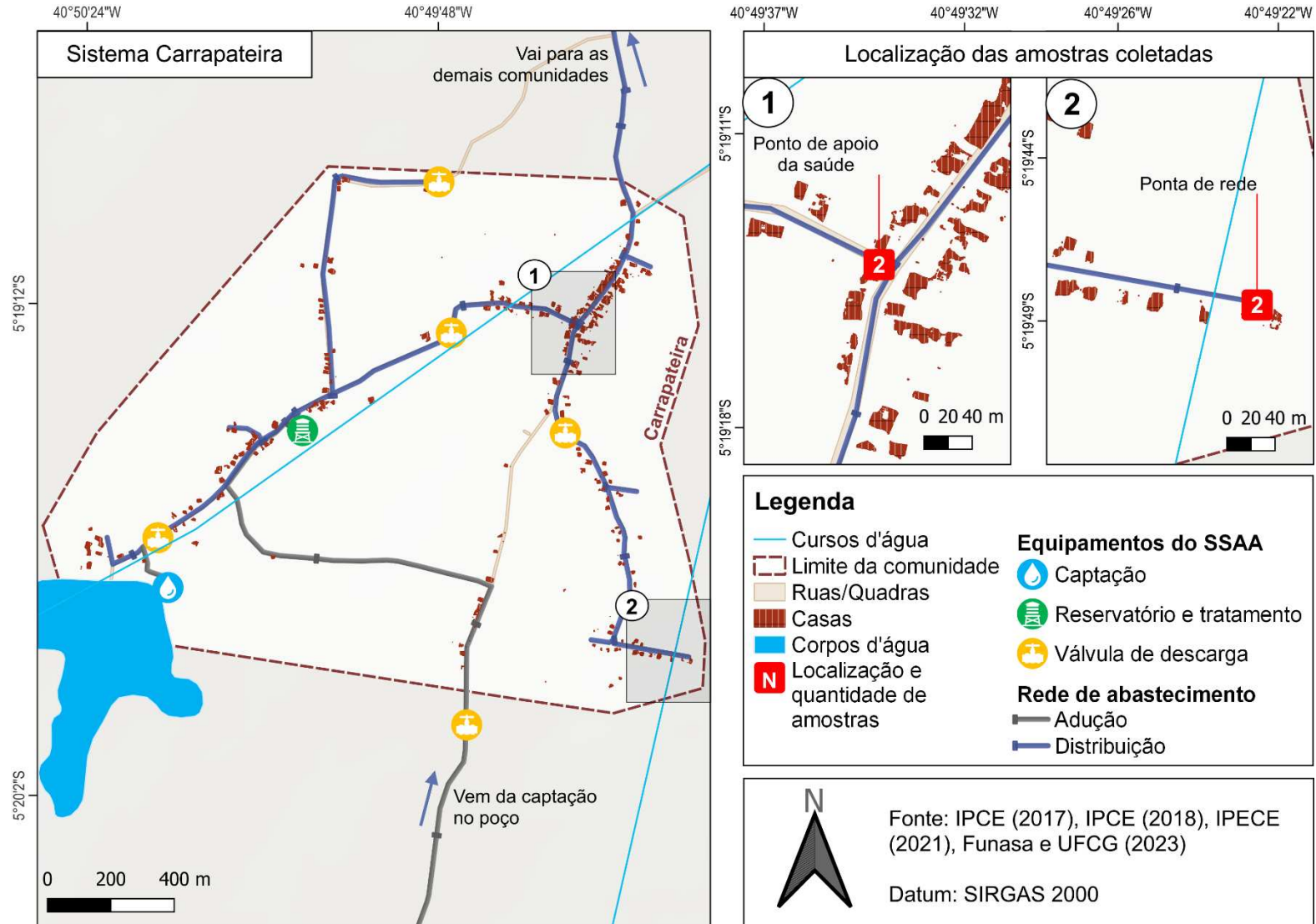
Fonte: Adaptado de Funasa e UFCG (2023).

Figura 28: Mapa da rede de distribuição e localização das amostras do Sistema Venâncio



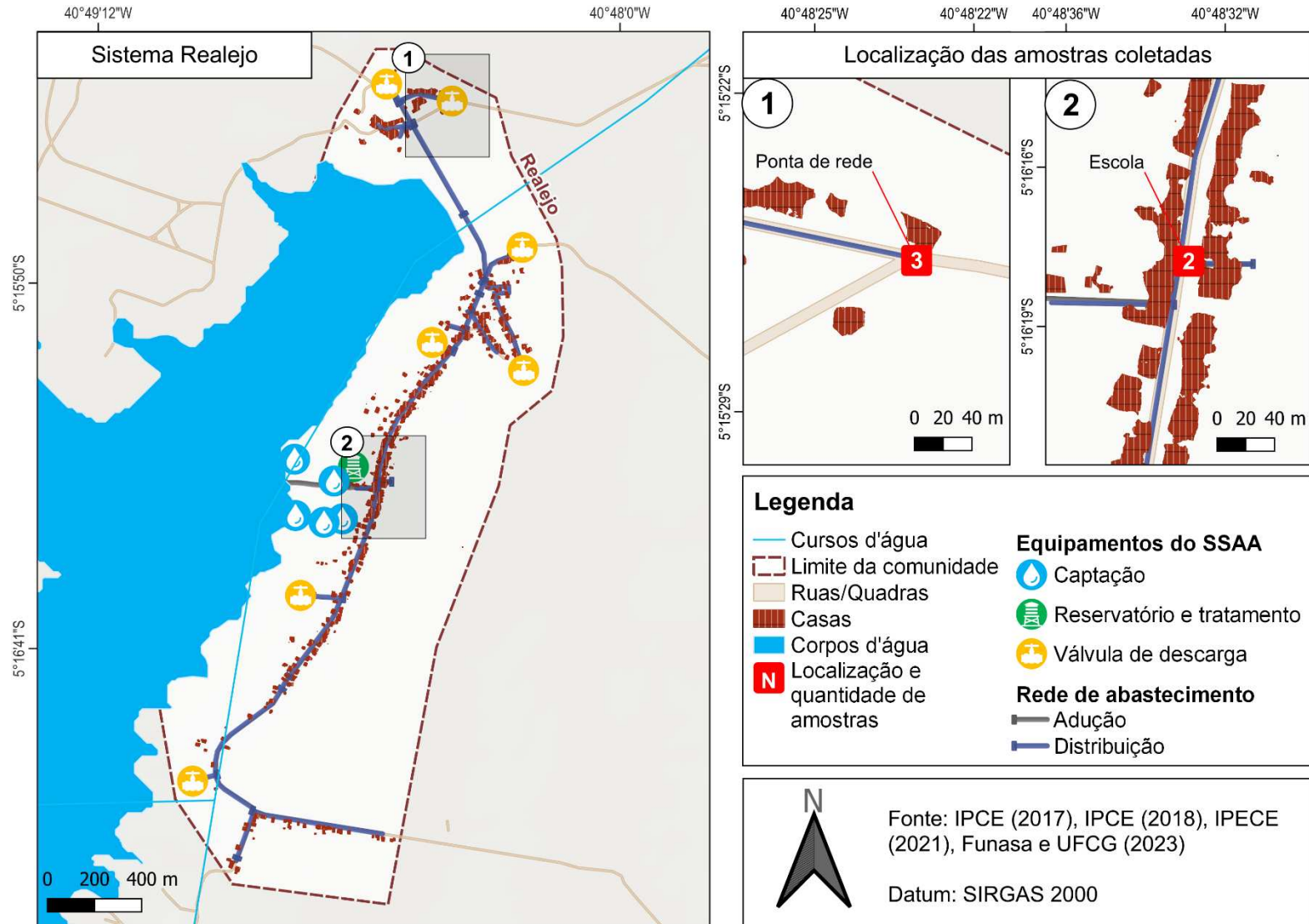
Fonte: Adaptado de Funasa e UFCG (2023).

Figura 29: Mapa da rede de distribuição e localização das amostras do Sistema Carrapateira



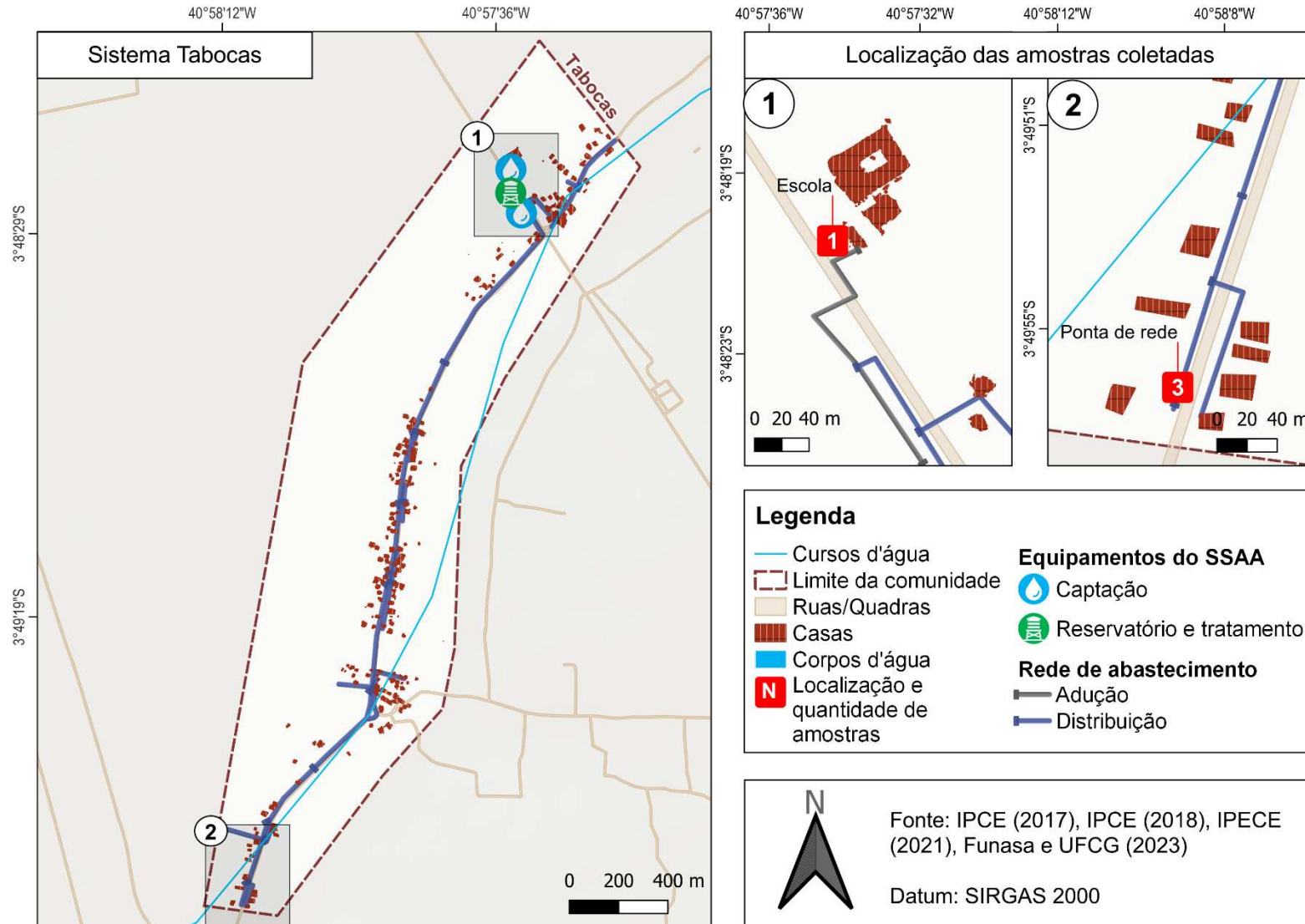
Fonte: Adaptado de Funasa e UFCG (2023).

Figura 30: Mapa da rede de distribuição e localização das amostras do Sistema Realejo



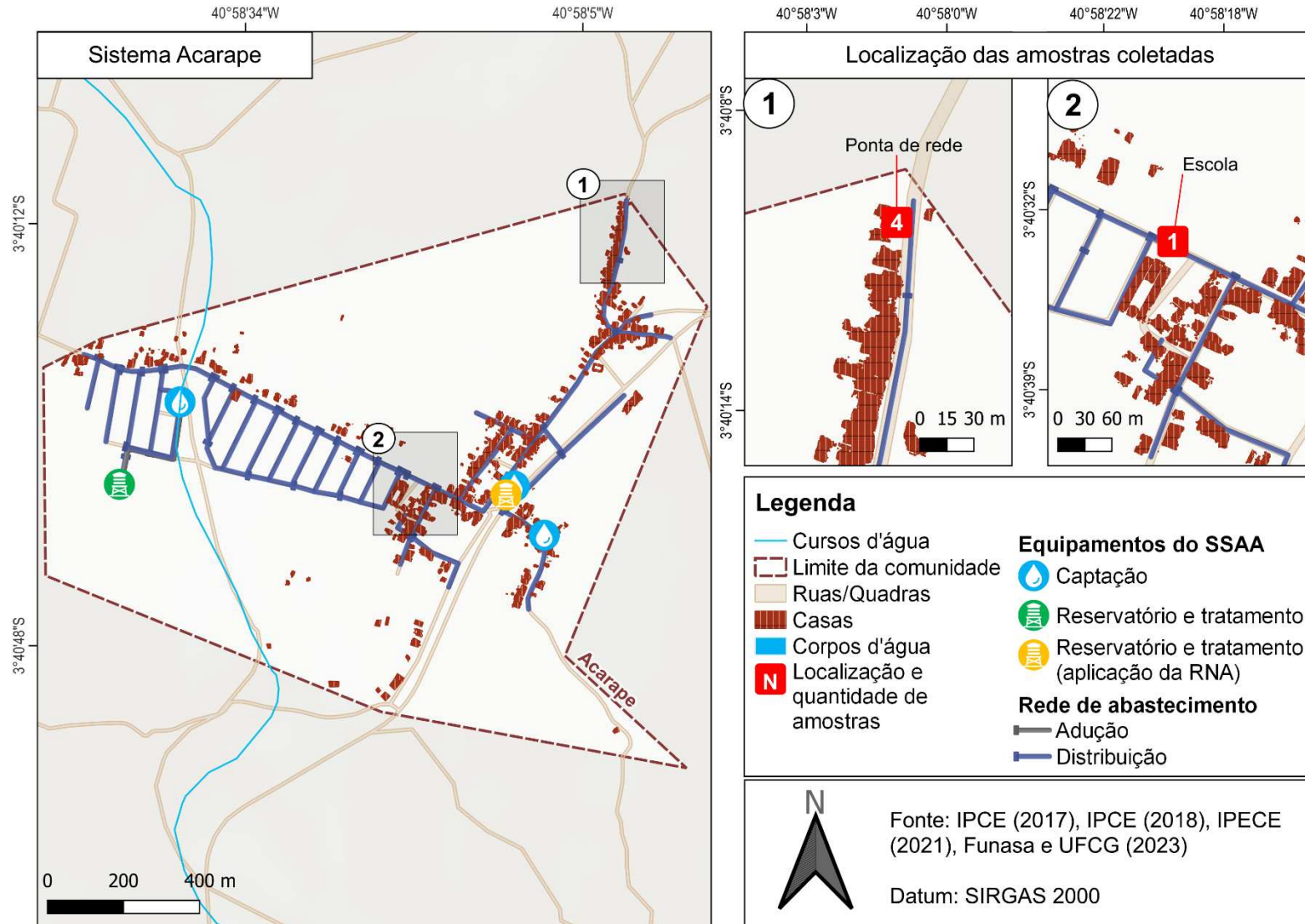
Fonte: Adaptado de Funasa e UFCG (2023).

Figura 31: Mapa da rede de distribuição e localização das amostras do Sistema Tabocas



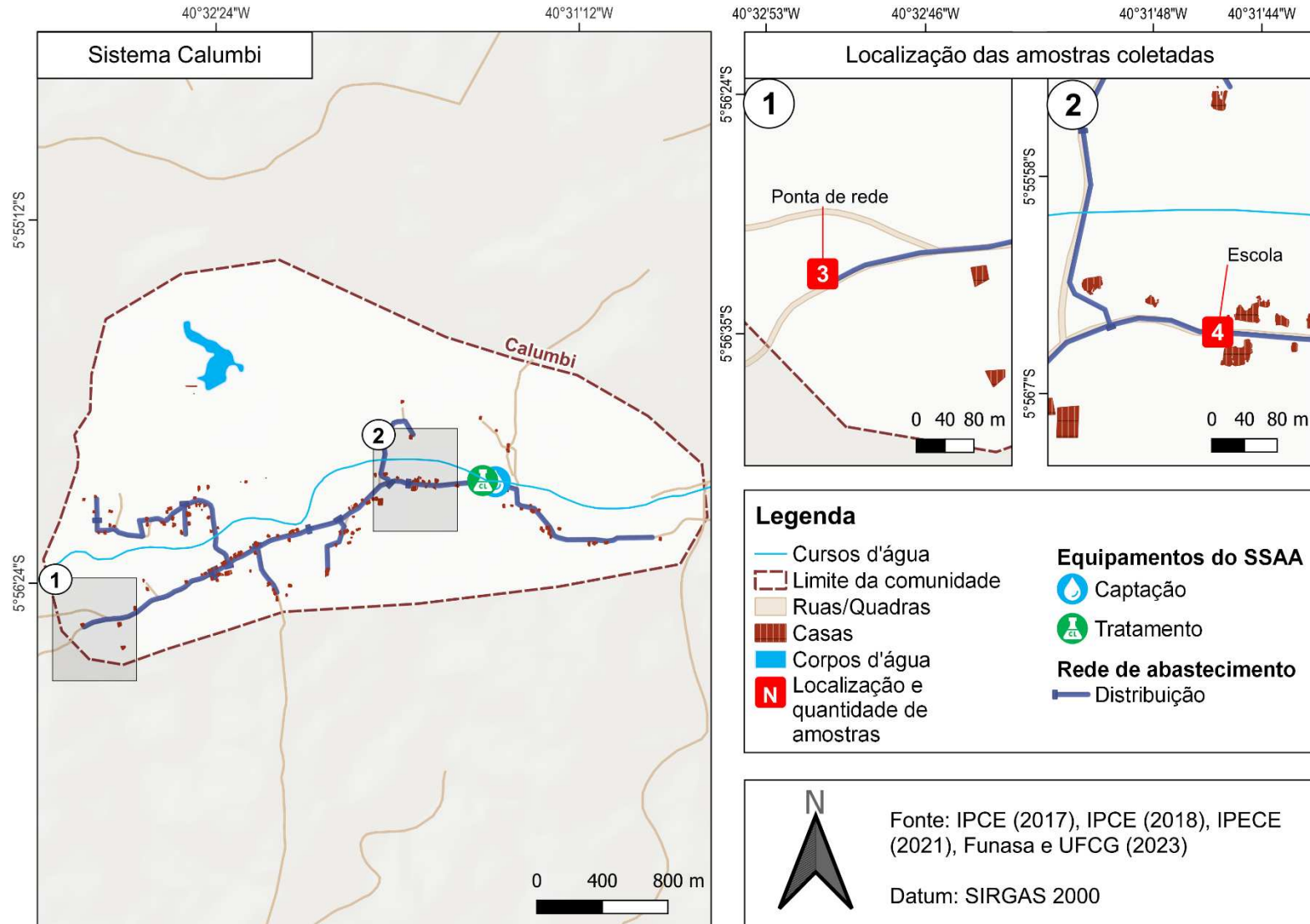
Fonte: Adaptado de Funasa e UFCG (2023).

Figura 32: Mapa da rede de distribuição e localização das amostras do Sistema Acarape



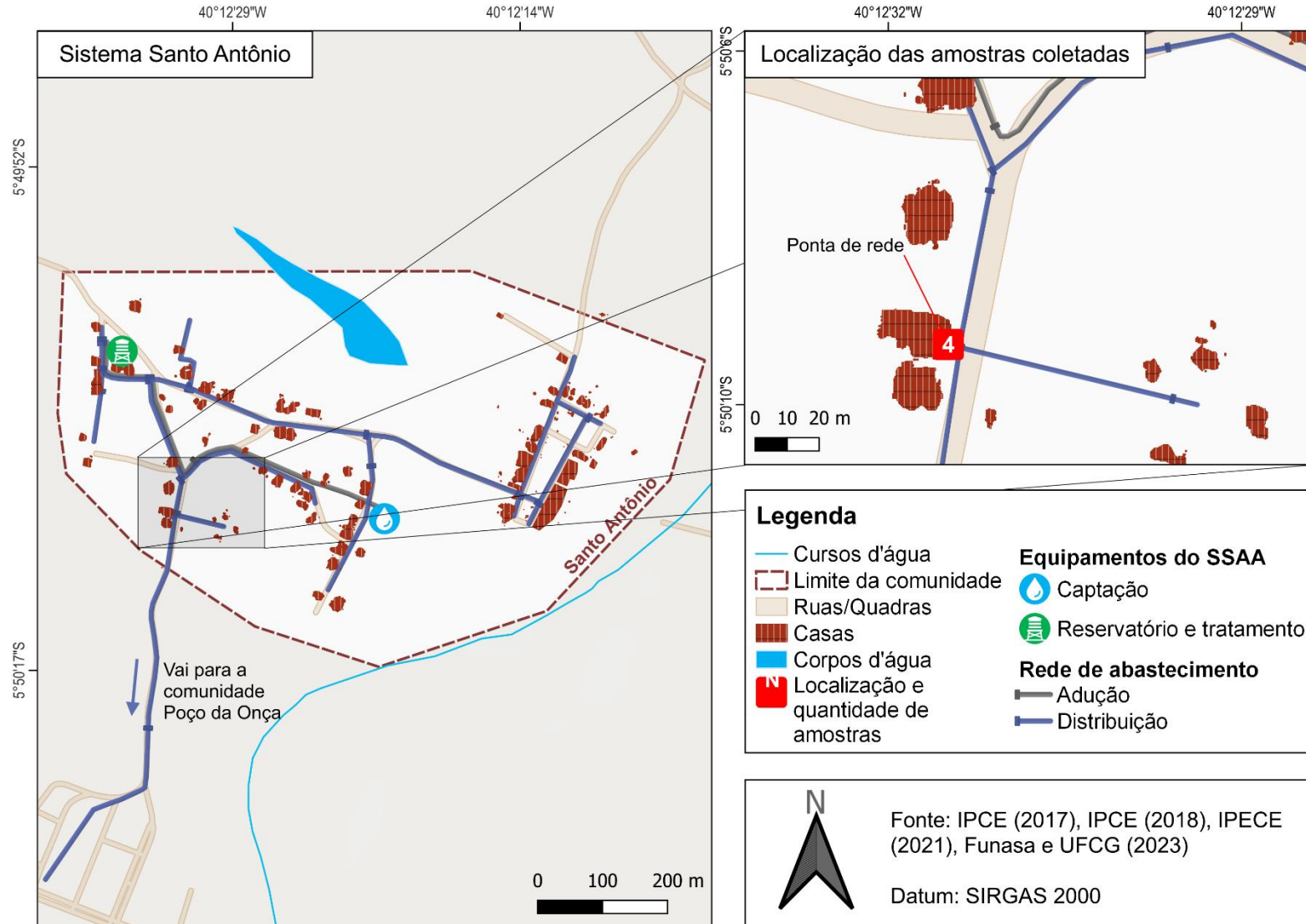
Fonte: Adaptado de Funasa e UFCG (2023).

Figura 33: Mapa da rede de distribuição e localização das amostras do Sistema Calumbi



Fonte: Adaptado de Funasa e UFCG (2023).

Figura 34: Mapa da rede de distribuição e localização das amostras do Sistema Santo Antônio



Fonte: Adaptado de Funasa e UFCG (2023).

A Tabela 4 apresenta a estatística descritiva dos dados das amostras coletadas, utilizadas como variáveis de entrada na aplicação do modelo desenvolvido (item 4.3.2). Os dados apresentados foram pré-processados, a partir da exclusão das amostras inconsistentes. As inconformidades encontradas estavam relacionadas à ocorrência de valores de cloro residual livre (CRL) iguais ou inferiores no reservatório, em comparação com a rede de distribuição.

Ao analisar os dados da Tabela 4, identificaram-se amostras com valores discrepantes em relação aos padrões de potabilidade estabelecidos pela Portaria nº 888/2021. A variável que apresentou a maior quantidade de amostras inconformes foi o CRL na rede (40%), com predominância de valores abaixo do limite estabelecido pela portaria (0,20 mg/L). Essas inconsistências foram observadas em todos os sistemas, com exceção do sistema Lagoa do Juá.

Nos sistemas estudados, a dosagem de cloro é estabelecida pela experiência do operador, aumentando os riscos de sub e sobredosagem do desinfetante. Além disso, a população das comunidades exerce influência nessas dosagens, pois há uma pressão sobre o operador do SSAA, com reclamações frequentes sobre o excesso de cloro na água. Essa dinâmica pode resultar em desafios adicionais na manutenção dos níveis ideais de cloro residual livre na água tratada.

No sistema Lagoa dos Encantos (Figura 26), a extensa rede de distribuição contribui para o decaimento do CRL. Para assegurar a potabilidade da água em todas as residências atendidas pelo sistema, foram instalados dois cloradores para o reforço do cloro. Um deles atende às localidades Lagoas dos Porcos e Lagoa dos Ferreiras e o outro a comunidade Lagoa da Cruz. Nesse sistema, utilizou-se os dados do reforço de cloro, ao invés dos valores relativos ao reservatório, para as redes de distribuição das comunidades Lagoa dos Porcos e Lagoa dos Ferreiras. Já para a rede de Lagoa da Cruz considerou-se os dados do reservatório. Essa escolha objetivou avaliar a robustez do modelo em diferentes cenários e a capacidade de adaptação da rede neural artificial a sistemas com reforço de cloro.

Para o parâmetro de turbidez, a Portaria nº 888/2021 estabelece um Valor Máximo Permitido (VMP) de 5,0 uT na saída do tratamento. A presença de valores acima desse padrão pode ser atribuída a resíduos, matéria orgânica, substâncias ou microrganismos na água, que podem reduzir a eficácia da desinfecção (Garay *et al.*, 2021; Khedikar *et al.*, 2021). Nos sistemas estudados (Tabela 4), a turbidez excedeu o VMP em apenas duas amostras coletadas na rede de distribuição do Sistema Lagoa do Juá. Esta condição pode indicar a necessidade de manutenções ou alterações nas etapas de tratamento.

Durante visitas técnicas realizadas em novembro de 2022, identificou-se um vazamento no fundo do filtro do sistema, o que pode ter permitido a passagem de partículas e sedimentação na água, comprometendo a eficiência do tratamento. Embora o sistema Lagoa dos Encantos compartilhe o mesmo manancial, não foram observadas amostras inconformes, indicando que a filtração garantiu a manutenção da turbidez dentro dos padrões estabelecidos.

Ressalta-se que, nos sistemas com captação por manancial subterrâneo, a turbidez geralmente permanece baixa, mesmo na ausência de processo de filtração. Isso se deve à menor suscetibilidade da água à presença de impurezas, além da filtração natural proporcionada pelo solo e pela rocha subterrânea, como apontam Correa *et al.* (2008) e Brasil (2019).

Quanto ao pH da água destinada ao consumo humano, a portaria de potabilidade indica uma faixa entre 6,0 e 9,0, para evitar a corrosividade da água e garantir a eficácia do cloro residual livre. A partir da Tabela 4, constata-se que todas as amostras dos sistemas Venâncio, Tabocas e Acarape apresentam pH abaixo do intervalo recomendado, o que pode inibir a ação do agente desinfetante sobre os microrganismos patogênicos, como verificado por Seymour *et al.* (2020) e Sundas *et al.* (2021). Ademais, na presença de amônia e compostos amoniacais, o pH baixo contribui para a formação do cloro residual combinado, que consome o cloro residual livre e confere odor e sabor desagradáveis à água (Foster, 2022).

As temperaturas mais baixas são observadas nos sistemas Tabocas e Acarape, localizados na região da serra, no município de Tianguá/CE (Tabela 4). As temperaturas mais elevadas são registradas nos demais sistemas, localizados nos municípios de Aracati/CE, Crateús/CE e Tauá/CE. É relevante mencionar que a variação térmica entre essas localidades não é significativa, devido à proximidade geográfica e a natureza climática do Semiárido Brasileiro onde esses sistemas estão inseridos.

Ao analisar o desvio padrão e a amplitude dos intervalos (Tabela 4), nota-se que os valores não apresentam discrepâncias significativas em comparação com os dados utilizados no desenvolvimento da rede neural artificial. Isso assegura a representação de uma ampla gama de cenários nos dados de teste, garantindo uma avaliação mais precisa da confiabilidade e robustez do modelo em aplicações reais.

Tabela 4: Análise descritiva dos dados das variáveis de entrada do modelo a serem utilizados na etapa de aplicação do modelo

Sistema de abastecimento de água (SAA)	Variável/ Unidade	Estatística descritiva					
		Número de amostras	Número de amostras inconformes*	Média	Desvio padrão	Limite inferior	Limite superior
Sistema Lagoa dos Encantos	Cloro residual livre no reservatório (mg/L)	18	0	2,82	1,17	0,50	4,25
	Cloro residual livre na rede (mg/L)		4	0,74	0,75	0,00	2,50
	pH no reservatório		0	7,28	0,13	7,08	7,49
	Turbidez no reservatório (uT)		0	0,11	0,16	0,00	0,39
	Comprimento da tubulação (m)		-	3.317,50	3.018,03	906,57	9.553,63
	Temperatura do dia (°C)		-	26,53	0,75	24,87	27,48
Sistema Lagoa do Juá	Cloro residual livre no reservatório (mg/L)	4	0	3,71	1,97	0,85	5,00
	Cloro residual livre na rede (mg/L)		0	0,74	1,93	0,56	4,00
	pH no reservatório		0	6,76	0,17	6,62	7,01
	Turbidez no reservatório (uT)		2	6,86	7,30	0,62	14,50
	Comprimento da tubulação (m)		-	1.389,39	327,62	1.105,66	1.673,12
	Temperatura do dia (°C)		-	26,10	0,99	24,87	26,90
Sistema Venâncio	Cloro residual livre no reservatório (mg/L)	4	1	1,02	0,77	0,12	2,00
	Cloro residual livre na rede (mg/L)		1	0,62	0,61	0,09	1,50
	pH no reservatório		4	5,55	0,07	5,51	5,65
	Turbidez no reservatório (uT)		0	0,25	0,48	0,00	0,96
	Comprimento da tubulação (m)		-	1.759,17	0,00	1.759,17	1.759,17
	Temperatura do dia (°C)		-	26,68	0,73	25,72	27,48
Sistema Carrapateira	Cloro residual livre no reservatório (mg/L)	4	2	1,13	1,32	0,00	2,52
	Cloro residual livre na rede (mg/L)		2	0,95	1,15	0,00	2,31
	pH no reservatório		0	7,26	0,21	7,06	7,44
	Turbidez no reservatório (uT)		0	0,05	0,07	0,00	0,15
	Comprimento da tubulação (m)		-	1.724,48	780,83	1.048,26	2.400,70
	Temperatura do dia (°C)		-	26,49	0,52	25,89	27,16
Sistema Realejo	Cloro residual livre no reservatório (mg/L)	5	2	1,29	1,00	0,00	2,00
	Cloro residual livre na rede (mg/L)		3	1,02	0,80	0,00	1,50
	pH no reservatório		0	6,74	0,16	6,44	6,83
	Turbidez no reservatório (uT)		0	0,07	0,08	0,00	0,16
	Comprimento da tubulação (m)		-	2.073,38	1.795,65	197,92	3.479,97
	Temperatura do dia (°C)		-	26,62	0,30	25,89	26,60

Tabela 4: Análise descritiva dos dados das variáveis de entrada do modelo a serem utilizados na etapa de aplicação do modelo (continuação)

Sistema de abastecimento de água (SAA)	Variável/ Unidade	Estatística descritiva					
		Número de amostras	Número de amostras inconformes*	Média	Desvio padrão	Limite inferior	Limite superior
Sistema Tabocas	Cloro residual livre no reservatório (mg/L)	4	0	1,88	1,13	0,48	3,00
	Cloro residual livre na rede (mg/L)		1	1,02	0,82	0,00	2,00
	pH no reservatório		4	4,67	0,27	4,30	4,84
	Turbidez no reservatório (uT)		0	0,53	1,06	0,00	2,11
	Comprimento da tubulação (m)		-	2.041,67	750,00	1.666,67	2.666,67
	Temperatura do dia (°C)		-	23,18	0,28	22,82	23,48
Sistema Acarape	Cloro residual livre no reservatório (mg/L)	5	0	0,91	1,17	0,20	3,00
	Cloro residual livre na rede (mg/L)		2	0,54	0,84	0,00	2,00
	pH no reservatório		5	3,75	0,13	3,66	3,97
	Turbidez no reservatório (uT)		0	0,20	0,24	0,00	0,46
	Comprimento da tubulação (m)		-	1.988,63	525,88	1.047,90	2.223,81
	Temperatura do dia (°C)		-	23,20	0,52	22,82	23,99
Sistema Calumbi	Cloro residual livre no reservatório (mg/L)	7	2	0,49	0,70	0,00	2,00
	Cloro residual livre na rede (mg/L)		6	0,21	0,57	0,00	1,50
	pH no reservatório		0	6,96	0,05	6,90	7,00
	Turbidez no reservatório (uT)		0	0,51	0,69	0,00	1,50
	Comprimento da tubulação (m)		-	1.396,44	1.192,56	440,27	2.671,34
	Temperatura do dia (°C)		-	26,26	0,96	24,90	27,16
Sistema Santo Antônio	Cloro residual livre no reservatório (mg/L)	4	2	0,41	0,63	0,00	1,33
	Cloro residual livre na rede (mg/L)		3	0,23	0,45	0,00	0,90
	pH no reservatório		0	7,23	0,09	7,09	7,30
	Turbidez no reservatório (uT)		0	0,09	0,18	0,00	0,36
	Comprimento da tubulação (m)		-	363,34	0,00	363,34	363,34
	Temperatura do dia (°C)		-	26,93	0,48	26,46	27,50
Todos os sistemas	Cloro residual livre no reservatório (mg/L)	55	9	1,49	1,51	0,00	5,00
	Cloro residual livre na rede (mg/L)		22	0,81	0,96	0,00	4,00
	pH no reservatório		13	6,24	1,16	3,66	7,49
	Turbidez no reservatório (uT)		2	0,96	2,48	0,00	14,50
	Comprimento da tubulação (m)		-	1.794,20	2.047,28	197,92	9.553,63
	Temperatura do dia (°C)		-	25,75	1,40	22,82	27,50

* De acordo com a Portaria nº 888/2021 do MS, que altera o Anexo XX da Portaria de Consolidação nº 5/2017. Legenda: **Parâmetros inconformes**. Fonte: A autora (2023).

4.3.2 Aplicação do modelo desenvolvido

Ao empregar os dados provenientes dos nove Sistemas Simplificados de Abastecimento de Água (SSAA) estudados, pôde-se avaliar a capacidade de generalização da rede neural artificial (RNA) desenvolvida. A Tabela 5 apresenta as porcentagens de acerto do modelo para os limiares de erro absoluto de 0,2 a 0,5 mg/L, que correspondem às faixas de precisão dos equipamentos para medição de cloro residual livre (CRL). Já os gráficos da Figura 35 exibem uma análise comparativa entre os valores preditos e observados de cloro residual livre para cada rede de distribuição de água, juntamente com os erros médios quadráticos (EMQ) correspondentes.

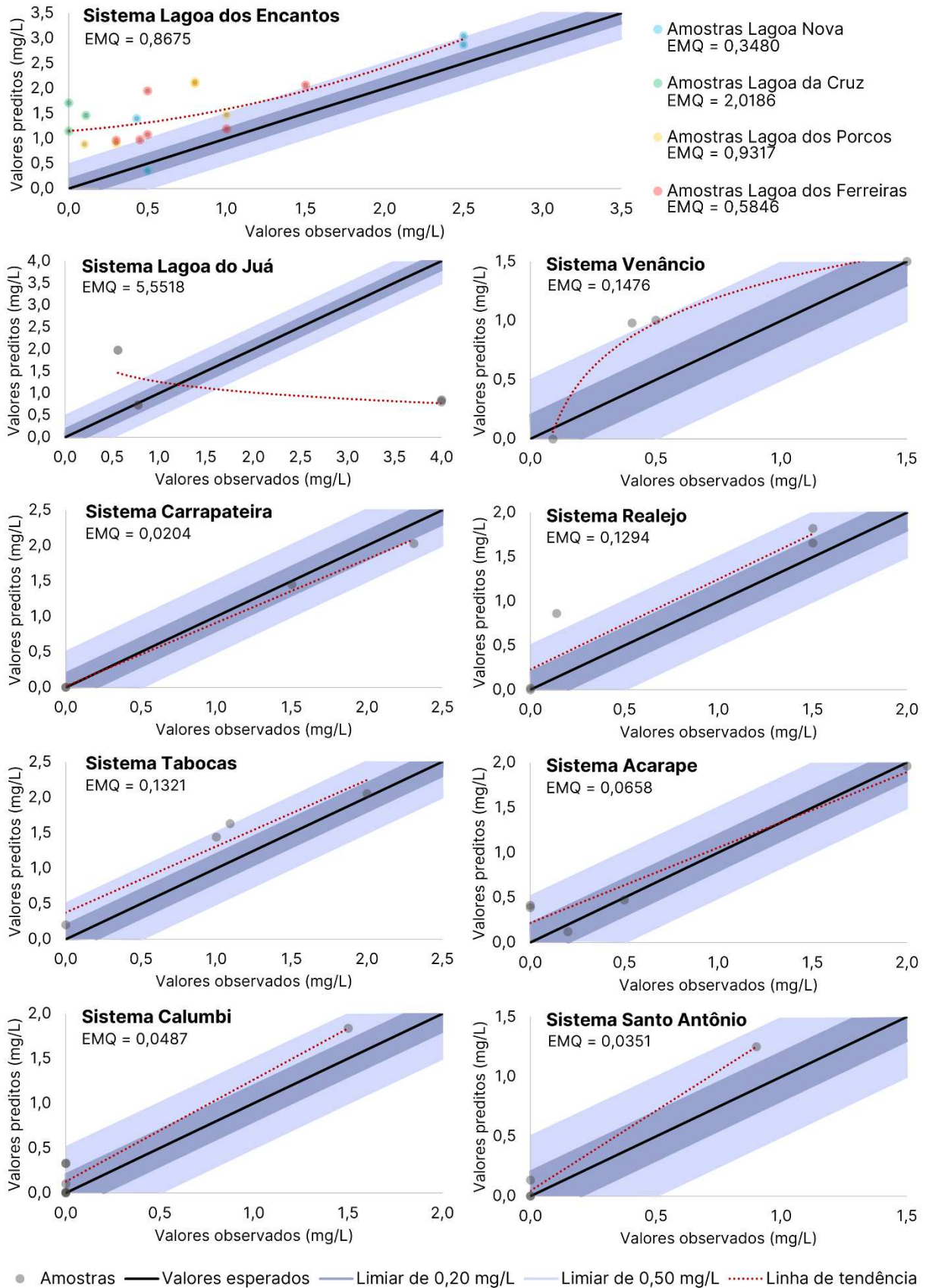
Tabela 5: Porcentagens de acerto do modelo para diferentes limiares de erro absoluto

Município	Sistema de abastecimento de água	Comunidade	Porcentagens de acerto (%)	
			Limiar de 0,20 mg/L*	Limiar de 0,50 mg/L**
Aracati	Sistema Lagoa dos Encantos	Lagoa Nova	25,00	50,00
		Lagoa da Cruz	0,00	0,00
		Lagoa dos Porcos	0,00	20,00
		Lagoa dos Ferreiras	16,67	16,67
		Todas	11,11	22,22
	Sistema Lagoa do Juá	Lagoa do Juá	25,00	25,00
	Sistema Venâncio	Venâncio	50,00	50,00
Crateús	Sistema Carrapateira	Carrapateira	75,00	100,00
	Sistema Realejo	Realejo	60,00	80,00
Tiangúá	Sistema Tabocas	Tabocas	25,00	75,00
	Sistema Acarape	Acarape	60,00	100,00
Tauá	Sistema Calumbi	Calumbi	57,14	100,00
	Sistema Santo Antônio	Santo Antônio	75,00	100,00

* Faixa de precisão do colorímetro/fotômetro. ** Faixa de precisão do método colorimétrico.
Fonte: A autora (2023).

Ao avaliar o EMQ, identificou-se que os sistemas de Lagoa do Juá (5,5518) e Lagoa dos Encantos (0,8675) apresentaram o pior desempenho, enquanto os sistemas de Santo Antônio (0,0351) e Carrapateira (0,0204) detiveram os melhores resultados (Figura 35). Durante o desenvolvimento da RNA, conforme abordado na seção 4.2.2, obteve-se uma média de EMQ de 0,1662 na fase de teste, servindo como parâmetro de referência para a eficácia do modelo. Os valores obtidos na aplicação da RNA para os sistemas em análise estiveram abaixo dessa média, com exceção dos sistemas Lagoa do Juá e Lagoa dos Encantos, onde a média foi superada.

Figura 35: Comparação entre os valores preditos e observados de cloro residual livre para a aplicação da RNA com os dados dos sistemas individualmente



Fonte: A autora (2023).

Os sistemas Venâncio, Realejo, Tabocas, Acarape e Calumbi detiveram um desempenho regular em comparação com os demais, com EMQ variando de 0,05 a 0,15, aproximadamente (Figura 35). Esses valores estão dentro dos intervalos obtidos por Gibbs (2006), Kote e Wadkar (2019), May *et al.* (2008) e Santi (2021), com variação do EMQ entre 0,001 e 1,005. Em comparação com o estudo de Onyutha (2022), que alcançou um erro médio quadrático de 0,07, apenas Acarape (0,0658) e Calumbi (0,0487) obtiveram performance satisfatória.

Conforme a Tabela 5, 100% das amostras dos SSAA Carrapateira, Acarape, Calumbi e Santo Antônio tiveram seus valores preditos dentro dos limiares de erro absoluto de 0,2 e 0,5 mg/L de CRL. No sistema Lagoa dos Encantos e Lagoa do Juá, a porcentagem de acertos para o limiar de 0,5 mg/L não ultrapassou 25,00%. Resultados intermediários foram observados em Venâncio, Realejo e Tabocas, onde o acerto variou de 50,00% a 80,00% para o limiar de 0,5 mg/L.

Nos sistemas com melhor desempenho, Santo Antônio e Carrapateira, obtiveram-se porcentagens de acerto de 75,00% e 100,00% para os limiares de 0,2 e 0,5 mg/L de cloro residual livre, respectivamente (Tabela 5). Essas taxas são aproximadamente 18% maiores que a média de acertos obtida na etapa de teste durante a construção da RNA. É importante ressaltar que esses sistemas não apresentaram inconformidades em relação aos parâmetros de pH e turbidez, com base na Portaria nº 888/2021. Isso influencia positivamente na capacidade de generalização da rede neural, pois os dados utilizados na etapa de treinamento são predominantemente provenientes de amostras que atendem aos padrões de potabilidade, semelhantemente aos estudos de Jafari *et al.* (2023) e Santi *et al.* (2021).

Ao avaliar SSAA com inconformidades relacionadas ao pH, observa-se um desempenho mais variado. Dos três sistemas com pH abaixo do intervalo estabelecido pela Portaria nº 888/2021, os sistemas Venâncio e Tabocas apresentaram um desempenho regular atingindo 50,00% e 25,00% de acerto para o limiar de 0,2 mg/L, e 50,00% e 75,00% para o limiar de 0,5 mg/L, respectivamente (Tabela 5).

No sistema Lagoa do Juá (Figura 35 e Tabela 5), onde foi registrada a maior taxa de EMQ e baixas porcentagens de acerto para ambos os limiares (25,00%), tem-se inconformidades associadas à turbidez. O parâmetro ultrapassou os limites estabelecidos pela portaria de potabilidade em 50% das amostras avaliadas. Esse cenário ressalta a influência de amostras atípicas ou inconformes no desempenho da rede neural artificial, assim como observado por Santi *et al.* (2021).

Em relação ao sistema Lagoa dos Encantos (Figura 35), a presença dos reforços de cloro ao longo do SSAA pode ter influenciado negativamente na performance da rede neural. Os melhores resultados em termos de predição do cloro residual livre foram alcançados na rede de distribuição da comunidade Lagoa Nova, onde não há recloração, conforme evidenciado na Tabela 5. Nessa localidade, as taxas de acerto variaram de 25,00% a 50,00% para os limiares de erro absoluto situados entre 0,2 e 0,5 mg/L, apresentando o menor erro médio quadrático do sistema.

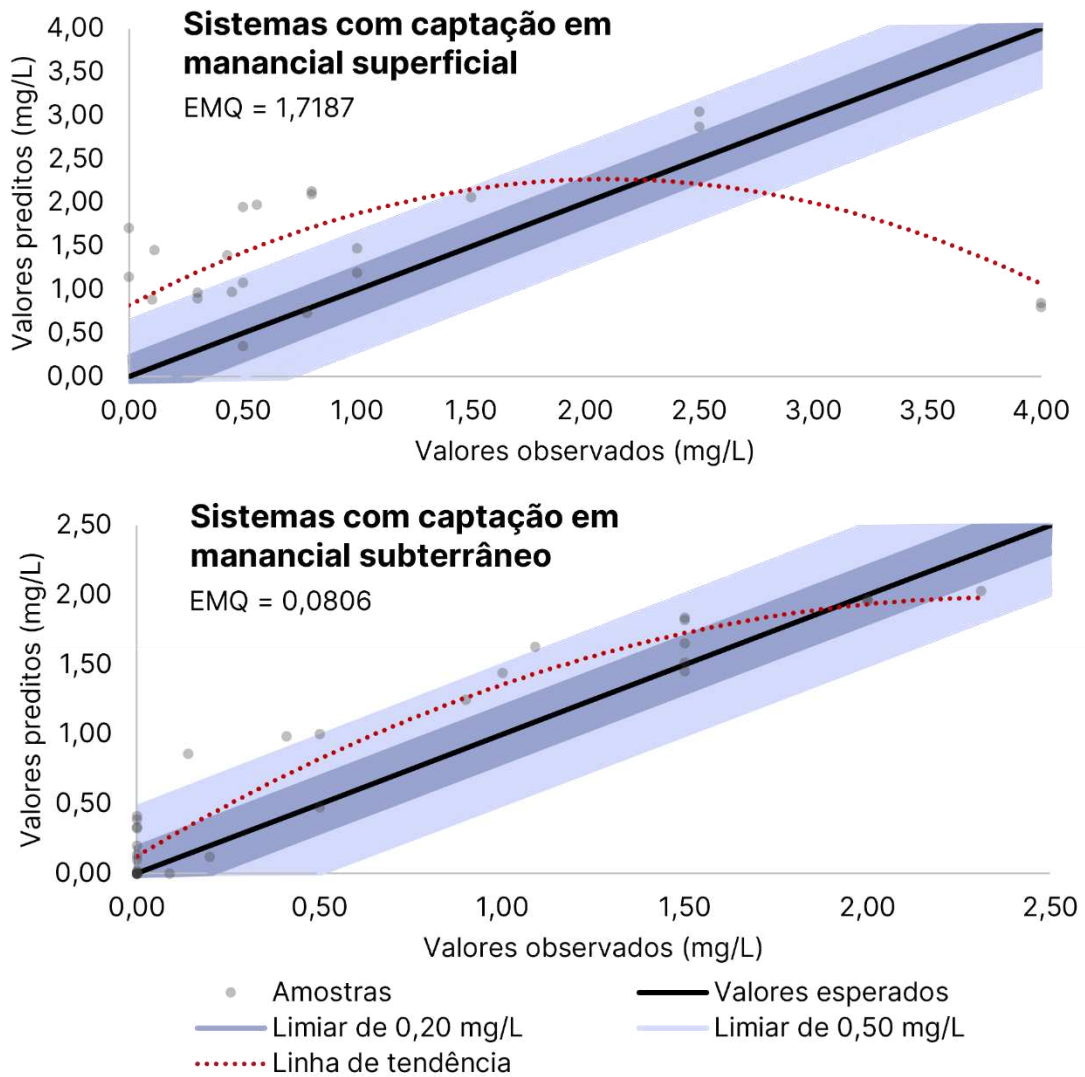
Em contrapartida, a comunidade Lagoa da Cruz registrou os piores resultados, com todas as amostras situadas fora dos limiares de 0,2 e 0,5 mg/L e EMQ igual a 2,0186 (Figura 35). Nas localidades Lagoa dos Porcos e Lagoa dos Ferreiras, embora o desempenho tenha sido inferior, ainda superou o observado em Lagoa da Cruz (Tabela 5). Isso indica que a utilização dos dados de reforço de cloro, ao invés dos dados do reservatório, pode contribuir para a melhoria dos resultados.

Conforme destacado por Bello *et al.* (2019) e Kurajica *et al.* (2021), a recloração pode ocasionar a formação de subprodutos de desinfecção. Isso tem o potencial de modificar o comportamento cinético e a ação do cloro residual livre (Zhao *et al.*, 2018), introduzindo não-linearidades adicionais ao modelo. Essa dinâmica intrínseca ao sistema pode afetar significativamente a performance da rede neural artificial.

É interessante observar que as variáveis de comprimento da tubulação e temperatura do dia não exercem uma influência significativa no desempenho do modelo, na aplicação desenvolvida. Este fato é notado ao considerar os sistemas Lagoa do Juá e Santo Antônio, por exemplo, que compartilham características similares em termos de extensão da rede de distribuição e condições climáticas, mas exibem performances distintas.

Nota-se que os sistemas Lagoa dos Encantos e Lagoa do Juá, que apresentaram o pior desempenho, possuem captação superficial e etapa de filtração. Por outro lado, os melhores resultados foram observados em sistemas com captação do tipo subterrânea, proveniente de poços, com tratamento por simples desinfecção com cloro. Os gráficos da Figura 36 fornecem uma comparação visual entre os valores estimados e reais de CRL para a aplicação da RNA em sistemas com captação em manancial superficial e subterrâneo, com os EMQ associados.

Figura 36: Comparação entre os valores preditos e observados de cloro residual livre para a aplicação da RNA com os dados dos sistemas com captação em manancial superficial e subterrâneo



Fonte: A autora (2023).

Para os sistemas com captação superficial, a porcentagem de acerto atingiu 13,64% e 22,73% para os limiares de erro absoluto de 0,2 e 0,5 mg/L, respectivamente, ficando abaixo cerca de 75% das taxas de acerto alcançadas durante a fase de teste. Em contraste, para os SSAA com captação em manancial subterrâneo, a porcentagem de acerto elevou-se para 57,58% e 87,88%, para os respectivos limiares, superando em 2,8% a taxa de acerto para o limiar de 0,5 mg/L, obtida durante a etapa de teste. Analisando os EMQ, também foram obtidos resultados favoráveis para os SSAA com captação por poços, apresentando um valor 52% menor em comparação com a etapa de teste, enquanto nos sistemas com captação no Canal do Trabalhador, esse erro é 934% superior.

A discrepância de desempenho entre os sistemas com diferentes mananciais e com existência de filtração pode ser atribuída à natureza variada do parâmetro da turbidez. Esse parâmetro pode interagir de maneiras distintas com o cloro residual livre, de acordo com o tipo de partícula presente na água (Garay *et al.*, 2021; Khedikar *et al.*, 2021). Nos mananciais superficiais, predomina o material orgânico, que tem maior demanda por cloro. A turbidez nesses corpos d'água varia ao longo do ano, de acordo com as estações secas e chuvosas, o que implica em mudanças no processo de tratamento. Nas águas subterrâneas, as argilas são os principais compostos causadores de turbidez, não exercendo uma forte demanda por cloro.

Isso afeta a capacidade de generalização da RNA, pois tem-se a introdução de não-linearidades na relação entre o CRL e a turbidez que talvez não tenham sido incorporadas na etapa de treinamento. Bowden *et al.* (2006), Santi *et al.* (2021), Tinelli e Juran (2019) ressaltam a vantagem de utilização de RNA pela facilidade de retreinar os modelos, permitindo sua adaptação a condições de qualidade da água potencialmente dinâmicas. Assim, a rede neural pode ser alimentada com uma diversidade de amostras, incluindo principalmente os casos em que o desempenho da RNA foi subótimo, a fim de enriquecer a robustez do modelo.

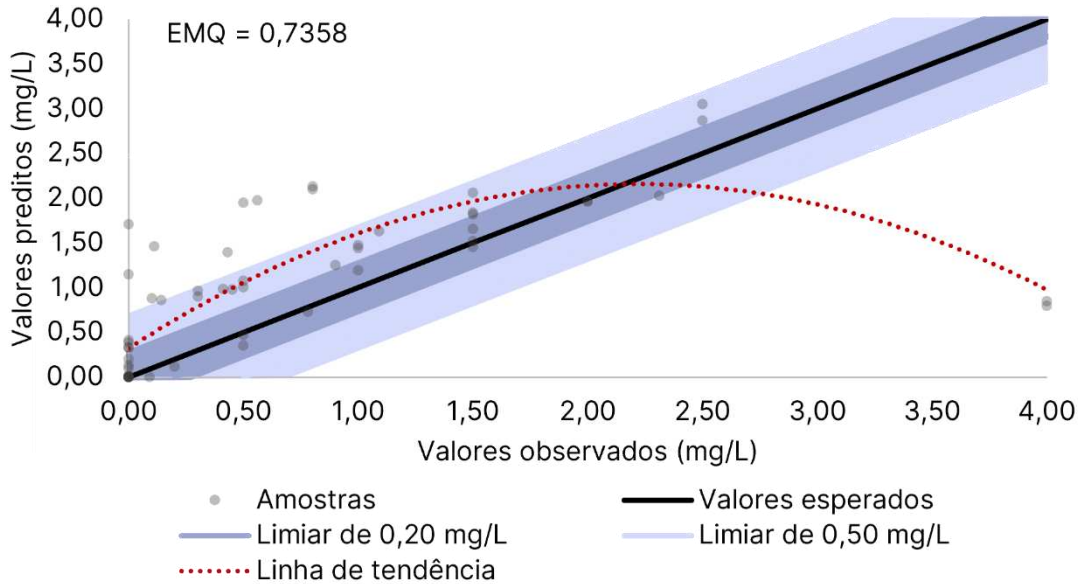
Para isso, é essencial o monitoramento contínuo da qualidade da água nos SSAA. Durante as visitas técnicas realizadas e a partir do estudo conduzido por França *et al.* (2023a), foram identificados problemas relacionados à escassez ou inexistência de dados relacionados as análises físico-químicas de amostras de água. Nos sistemas localizados no município de Aracati/CE, não são realizadas coletas na rede de distribuição. Já nos SSAA das localidades Carrapateira, Realejo, Acarape e Calumbi não foram encontrados dados no Sistema de Informação de Vigilância da Qualidade da Água para Consumo Humano (Sisagua).

Os métodos convencionais de monitoramento da qualidade da água para consumo humano são manuais e exigem tempo, pessoal capacitado e altos custos financeiros. Nas zonas rurais, a utilização dessas metodologias é mais desafiadora, tendo em vista às maiores distâncias para coleta de amostras, onerando o processo. Frente a isso, tem-se a utilização de sistemas inteligentes de monitoramento da qualidade da água, compostos por sensores, microcontroladores e *software*, como os desenvolvidos por Bogdan *et al.* (2023), Lacerda (*no prelo*) e Pasika e Gandla (2020). Estes proporcionam o acompanhamento em tempo real dos parâmetros exigidos pela Portaria nº 888/2021, contribuindo positivamente para a alimentação e retreino da RNA.

A Figura 37 exibe a comparação entre os valores previstos e observados de cloro residual livre para a aplicação do modelo considerando os dados de todos os sistemas. O erro

médio quadrático foi de 0,7358, 343% mais elevado em relação ao observado durante o teste, devido ao baixo desempenho dos sistemas com captação em manancial superficial.

Figura 37: Comparação entre os valores preditos e observados de cloro residual livre para a aplicação da RNA com os dados de todos os sistemas



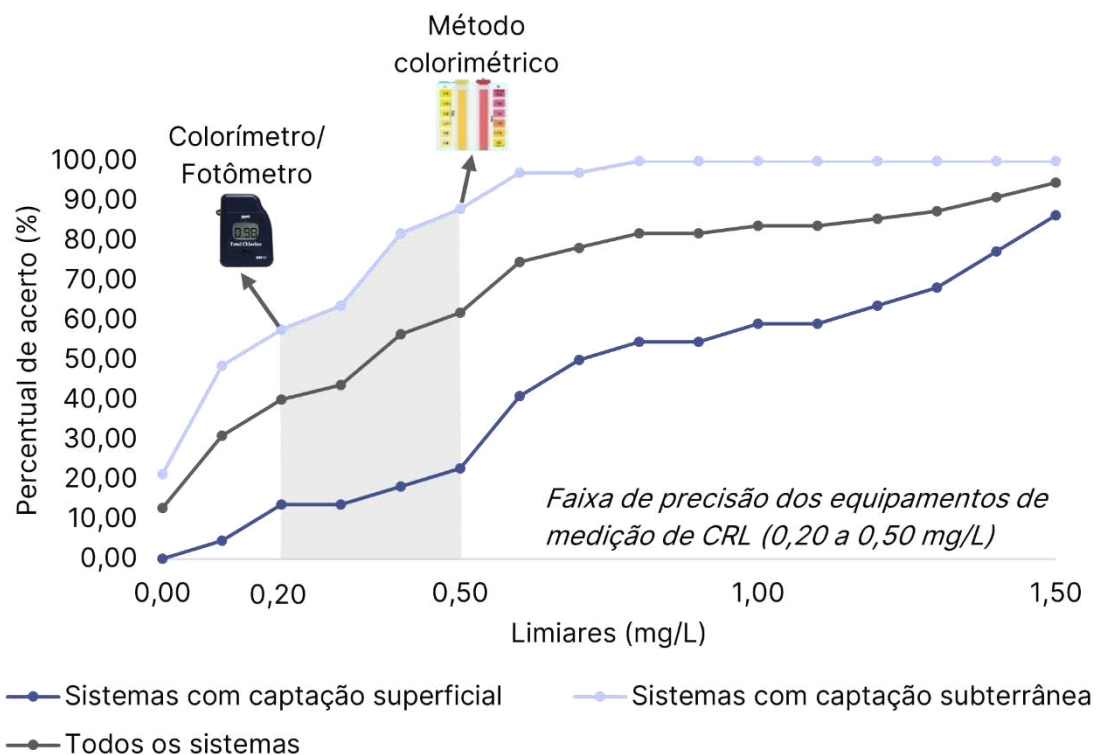
Fonte: A autora (2023).

Comparando com os estudos de Gibbs (2006), Kote e Wadkar (2019) e May *et al.* (2008), que apresentaram uma variação do EMQ entre 0,001 e 1,005, o erro obtido permanece dentro do intervalo. Contudo, o erro médio quadrático verificado é superior ao relacionar com o EMQ alcançado por Al-Araimi *et al.* (2019), inferior a 0,01, por Onyutha (2022), igual a 0,07 e por Santi (2021), que variou entre 0,01 e 0,25. Vale salientar que, ao considerar apenas os sistemas com manancial subterrâneo (Figura 36), que representam a maioria dos casos, o EMQ permanece dentro do intervalo obtido por Santi (2021).

Sob a ótica da Figura 37, analisando os limiares de erro absoluto situados entre 0,2 e 0,5 mg/L, constatou-se uma taxa de acerto oscilando de 40,00% a 61,82%. Em geral, as amostras que ultrapassam os intervalos de precisão dos equipamentos de determinação de CRL estão localizados abaixo de 1,0 mg/L e acima de 3,0 mg/L, onde obtém-se porcentagens de acerto variando entre 40,00% e 55,00%. Para os valores entre 1,0 e 3,0 mg/L de CRL, há um acerto de 100,00% para o limiar de 0,5 mg/L. Analogamente ao observado na construção da RNA, os pontos com valores mais baixos de CRL são superestimados, enquanto os valores mais altos são subestimados. Isso evidencia a dificuldade do modelo em prever valores mais extremos, assim como observado por Gibbs *et al.* (2006) e Santi *et al.* (2021).

O gráfico da Figura 38 representa as porcentagens de acerto para vários limiares de erro absoluto, permitindo uma comparação do desempenho da RNA quando aplicada a sistemas com captação superficial, subterrânea e utilizando todos os dados disponíveis. Nota-se que para os sistemas que possuem captação em manancial subterrâneo, a RNA alcança uma taxa de 100,00% de acerto para o limiar de 0,8 mg/L. Esse valor indica um excelente desempenho do modelo quando comparado ao método colorimétrico, empregado na rotina operacional dos sistemas estudados, onde a precisão do equipamento pode chegar até 2,0 mg/L (Figura 25).

Figura 38: Porcentagens de acertos para diferentes limiares de erro absoluto para aplicação do modelo



Fonte: A autora (2023).

Contudo, para os sistemas em geral e com captação superficial, a taxa de acerto de 100,00% é atingida apenas para o limiar de 3,3 mg/L. Para os limiares de 1,8 mg/L e 1,4 mg/L, respectivamente, esses sistemas alcançam mais de 90% de acerto, indicando uma performance robusta em faixas de precisão próximas ao limite máximo do método colorimétrico (2,0 mg/L). Esses resultados evidenciam a eficácia do modelo desenvolvido em comparação com as práticas convencionais de determinação do cloro residual livre.

Koo *et al.* (2008) obtiveram um percentual de acertos de 95,42% na validação da RNA desenvolvida para um sistema de distribuição de uma cidade da Coréia do Sul. Bowden *et al.* (2006) alcançaram uma taxa de 98,00% para o limiar de 0,2 mg/L de erro absoluto, ao construir

e aplicar uma rede neural artificial na ETA de Myponga, no sul de Adelaide, Austrália do Sul. Apesar desses trabalhos apresentarem modelos com performance superior ao formulado neste estudo, essas RNAs são específicas para os sistemas de abastecimento de água estudados, enquanto a rede neural construída possui abrangência nacional.

Comparando com os estudos de Gibbs *et al.* (2006), Santi *et al.* (2021) e Santi *et al.* (2022), também de caráter local, a capacidade de predição do modelo treinado com dados nacionais apresentou resultados satisfatórios. Para Santi *et al.* (2021) e Santi *et al.* (2022), as porcentagens de captura variaram de 27% a 65%, enquanto Gibbs *et al.* (2006) obtiveram apenas 14% de acurácia. Dessa forma, considerando as limitações dos métodos convencionais e a eficácia demonstrada pela RNA, sua aplicação representa uma contribuição para a gestão eficiente e segura da qualidade da água em comunidades rurais brasileiras.

4.4 AVALIAÇÃO DE DESEMPENHO DAS REDES DE DISTRIBUIÇÃO DE ÁGUA

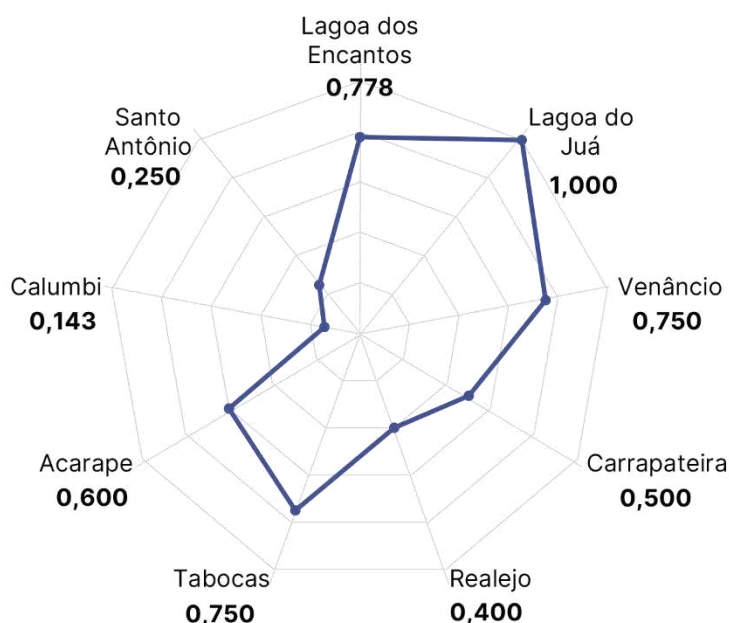
Os gráficos das Figuras 39 a 41 apresentam a performance das redes de distribuição de água estudadas quanto ao atendimento do parâmetro cloro residual livre (CRL) à Portaria nº 888/2021, com base nos indicadores propostos por Hashimoto *et al.* (1982): confiabilidade, resiliência e vulnerabilidade.

Em relação a confiabilidade (Figura 39), a média alcançada pelos Sistemas Simplificados de Abastecimento de Água (SSAA) avaliados foi de 0,575. O sistema Lagoa do Juá deteve o melhor desempenho, tendo em vista a ausência de inconformidades relacionadas ao CRL na rede de distribuição, como indicado na Tabela 4. Nos sistemas Santo Antônio e Calumbi, onde se identificou a maior porcentagem de amostras fora dos limites estabelecidos pela Portaria nº 888/2021, atingindo 75,00% e 85,71%, respectivamente, o indicador registrou os valores mais baixos.

Cerca de 63,34% das amostras inconformes de CRL concentram-se nas extremidades da rede de distribuição, devido à distância do ponto de desinfecção, que contribui para o decaimento do cloro. Esses pontos influenciam diretamente na diminuição do indicador de confiabilidade do sistema em relação ao cloro, tornando-os prioritários para o monitoramento da qualidade da água, conforme preconiza a Diretriz Nacional do Plano de Amostragem da Vigilância da Qualidade da Água para Consumo Humano (Brasil, 2016).

Figura 39: Indicador de confiabilidade para as redes de distribuição de água

Confiabilidade (C)



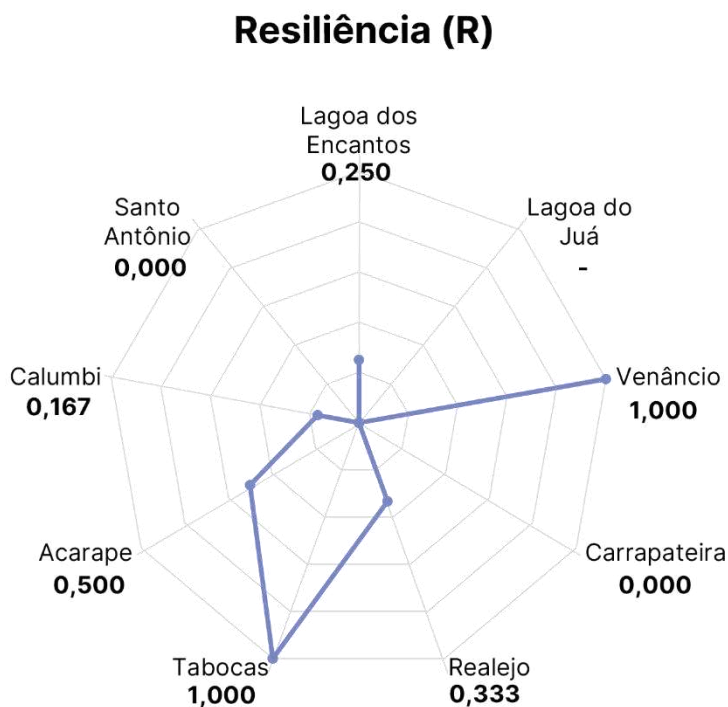
Fonte: A autora (2023).

Apesar de apresentar o segundo melhor desempenho para o indicador de confiabilidade, o sistema Lagoa dos Encantos apresenta um ponto de atenção. Todas as amostras da comunidade Lagoa da Cruz e uma amostra de Lagoa dos Porcos estavam inconformes, mesmo considerando os reforços de cloro implementados. Atualmente, os operadores realizam abastecimentos semanais no clorador, utilizando 3,0 kg de cloro pastilha. Contudo, a dosagem baseada na experiência dos operadores pode sugerir que a quantidade do desinfetante ou a frequência de reposição não esteja sendo suficiente.

A capacidade de recuperação das redes de distribuição quanto ao cloro residual livre foi avaliada a partir do indicador de resiliência (Figura 40). É importante destacar que, conforme indicação Silva Filho *et al.* (2023), a resiliência e a vulnerabilidade não foram calculadas para o sistema Lagoa do Juá, pela ausência de eventos de falha registrados no período avaliado.

Em geral, obteve-se uma média de 0,406, sugerindo um maior tempo de reestabelecimento do fornecimento de água potável após um evento de falha. Os SSAA Venâncio e Tabocas destacaram-se pela excelente capacidade de recuperação no período, atingindo o valor máximo do indicador (1,00). Isso indica que após qualquer deficiência relacionada ao cloro residual livre, a amostra subsequente estava dentro dos padrões de potabilidade estabelecidos pela Portaria nº 888/2021.

Figura 40: Indicador de resiliência para as redes de distribuição de água

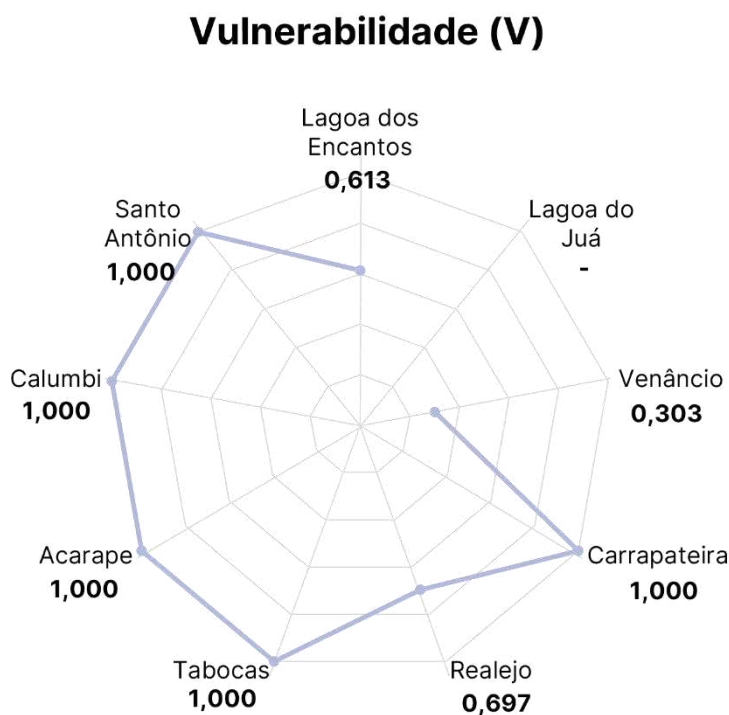


Fonte: A autora (2023).

Em Santo Antônio, Carrapateira e nas amostras da comunidade Lagoa da Cruz (SSAA Lagoa dos Encantos) verificou-se as piores performances, sinalizando que não se recuperam após apresentarem amostras com concentração de CRL fora dos limites estabelecidos. Para assegurar a resiliência nesses sistemas, é essencial realizar monitoramento e controle de qualidade da água. Essa abordagem possibilita aos gestores e operadores dos SSAA a implementação de medidas de mitigação, preparação e resposta em eventos de falha, permitindo reduzir a probabilidade de falha, as consequências e o tempo de recuperação dos níveis desejados de cloro (Balaei *et al.*, 2021; Quitana *et al.*, 2020).

A análise da vulnerabilidade (Figura 41) dos sistemas, que está relacionada à magnitude das falhas, revelou uma média elevada, atingindo o valor de 0,827. Isso deve-se à máxima vulnerabilidade identificada nos sistemas Santo Antônio, Calumbi, Acarape, Tabocas e Carrapateira. Nesses SSAA, todas as inconformidades do CRL estavam associadas à ausência do desinfetante na água, colocando em risco as comunidades atendidas. Em comunidades rurais e tradicionais, os impactos decorrentes da ingestão de água contaminada são mais acentuados devido ao somatório de condições como as disposições inadequadas de resíduos e a falta de tratamento sanitário. Esses fatores favorecem um ciclo em que as condições sanitárias precárias, aliadas às altas prevalências de parasitoses e desnutrição, diminuem a capacidade de resposta fisiológica humana (Assis *et al.*, 2020).

Figura 41: Indicador de vulnerabilidade para as redes de distribuição de água



Fonte: A autora (2023).

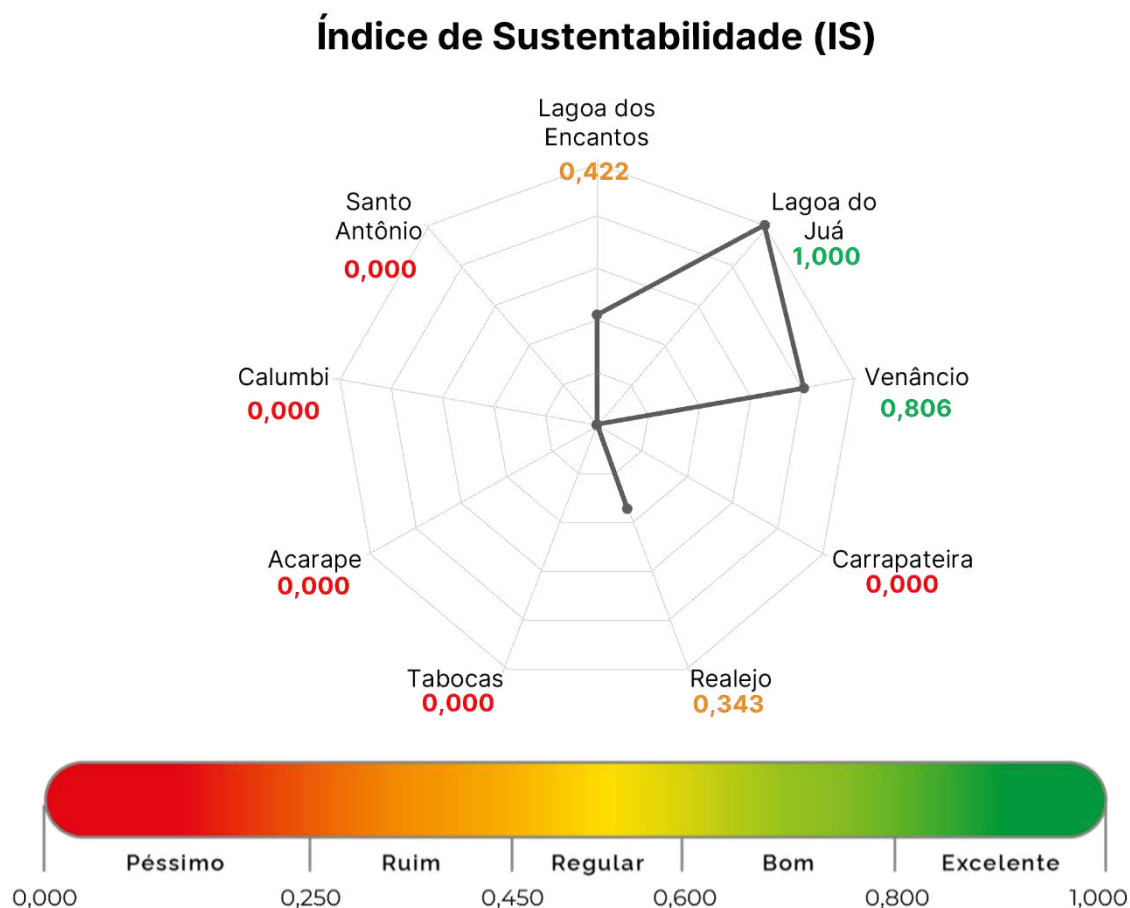
Para assegurar a redução das vulnerabilidades do sistema em relação à concentração de cloro residual livre, é fundamental a adoção de uma gestão eficiente nos SSAA. Isso pode ser alcançado por meio da implementação de Planos de Segurança da Água (PSA), que estão sendo desenvolvidos para os sistemas estudados, a partir do TED nº 06/2021 parceria Funasa/UFCG. Os PSAs avaliam as vulnerabilidades em todo o sistema de abastecimento de água, desde a captação até o consumidor final, organizando e sistematizando as práticas de gerenciamento voltadas ao consumo humano (Brasil, 2012). Segundo França *et al.* (2023b), os documentos que norteiam o saneamento básico no país enfatizam a necessidade da elaboração do PSA em sistemas e soluções rurais, para identificar e gerenciar os riscos e perigos e indicar as ações de emergência e contingência necessárias.

Para sintetizar a avaliação de desempenho, a Figura 42 apresenta o Índice de Sustentabilidade (IS) proposto por Sandoval-Solis *et al.* (2011) para as redes de distribuição de água em estudo, que engloba os indicadores analisados.

Ao averiguar a distribuição percentual dos sistemas em cada categoria, observa-se que 55,56% foram classificados como “péssimos”, 22,22% como “ruins” e 22,22% como “excelentes”. Nenhuma rede de distribuição apresentou classificação “regular” ou “boa”. A predominância de desempenho inadequado foi influenciada pelos sistemas Santo Antônio,

Calumbi, Acarape, Tabocas e Carrapateira, que atingiram o valor máximo no indicador de vulnerabilidade (Figura 41), e pelos sistemas Santo Antônio e Carrapateira que registraram resiliência (Figura 40) igual a zero.

Figura 42: Índice de Sustentabilidade para as redes de distribuição de água



Por outro lado, os sistemas Lagoa do Juá e Venâncio destacaram-se pelo desempenho excelente em relação ao IS. No caso de Lagoa do Juá, isso deve-se à ausência de amostras inconformes no período avaliado, evidenciando uma ótima performance na manutenção dos padrões estabelecidos para o CRL. Já o sistema Venâncio obteve essa classificação devido à baixa taxa de amostras inconformes (25,00%), sua eficiente capacidade de recuperação após falhas e a reduzida magnitude de violações dos padrões estabelecidos para o CRL.

Em simulações realizadas por Ren *et al.* (2020) e Zhang *et al.* (2020), os autores verificaram que há uma relação complexa entre os indicadores de vulnerabilidade, resiliência e confiabilidade, sendo os dois últimos inversamente proporcionais ao primeiro. Ou seja, assim como observado no presente estudo, quando o desempenho do sistema é muito confiável, o

SSAA tem uma rápida recuperação e baixa severidade em relação as inconformidades associadas ao CRL.

Os demais sistemas analisados enquadraram-se na categoria de desempenho “ruim”. No SSAA Lagoa dos Encantos, a resiliência frente a eventos de falha, de apenas 0,250, foi um fator determinante para essa classificação. No sistema Realejo, os baixos valores dos indicadores de confiabilidade (0,400) e resiliência (0,333) influenciaram a performance “ruim” obtida.

Os resultados encontrados evidenciam as vulnerabilidades das redes de distribuição no que diz respeito ao CLR, auxiliando os gestores do serviço de abastecimento de água. Esses dados podem ser empregados na estimativa da eficácia do processo de desinfecção, englobando a entrega de níveis aceitáveis de cloro residual livre e os procedimentos de recuperação após a falha. Para realizar essa avaliação de forma mais abrangente, é recomendável que ela seja conduzida de maneira rotineira e utilizando uma quantidade mais ampla de dados, especialmente amostras provenientes dos pontos preconizados pela Diretriz Nacional do Plano de Amostragem da Vigilância da Qualidade da Água para Consumo Humano (Brasil, 2016).

Ventura *et al.* (2022) destacam alguns desafios relacionados a essa prática, como a confiabilidade dos dados e a complexidade do modelo a ser implementado, especialmente em comunidades rurais. Nesse contexto, a aplicação do modelo desenvolvido para a previsão do decaimento de cloro residual livre, a partir de RNA, torna-se fundamental. Essa abordagem evita a necessidade de realização de inúmeras coletas ao longo da rede, simplificando o processo, reduzindo os custos associados e viabilizando o monitoramento eficiente do CRL.

5. CONCLUSÕES

Este estudo desenvolveu um modelo de decaimento de cloro residual livre (CRL) em redes de distribuição de água de comunidades rurais brasileiras. Para superar os problemas relacionados à necessidade de determinação da ordem cinética da reação e a estimativa do coeficiente de decaimento do FRC para cada sistema de abastecimento de água, realizadas de maneira laboratorial, utilizou-se métodos de aprendizado de máquina: Redes Neurais Artificiais (RNA) e *t-Distributed Stochastic Neighbor Embedding* (t-SNE).

A aplicação do t-SNE para a seleção de variáveis, utilizando dados com abrangência nacional, indicou a relevância dos parâmetros físico-químicos de qualidade de água considerados e da temperatura do dia. A combinação 7 foi avaliada como a mais eficaz, sendo empregada para o treinamento das RNA, abrangendo as variáveis: cloro residual livre no reservatório, turbidez no reservatório, pH no reservatório, temperatura do dia e comprimento da tubulação. O emprego deste método representou uma abordagem otimizada para explorar as diversas combinações de variáveis, assegurando uma gestão eficaz de tempo e recursos, sem prejudicar a precisão da avaliação.

Na construção da RNA, o treinamento com o erro médio quadrático ponderado (EMQP) mostrou-se mais preciso nos intervalos críticos de medição, com aumentos de 5,0% e 3,6% na taxa de acerto média para os limiares de 0,2 e 0,5 mg/L, respectivamente, em comparação com o treino por erro médio quadrático (EMQ). Em contraste com os métodos tradicionais de previsão de cloro residual livre, a metodologia proposta apresentou desempenho superior, com taxas de acerto de 85,45%, 95,58% e 99,84% para as precisões do equipamento de 0,5 mg/L, 1,0 mg/L e 2,0 mg/L, respectivamente. Ao ser comparado com modelos para sistemas de abastecimento locais desenvolvidos em estudos semelhantes, a rede neural construída em âmbito nacional demonstrou desempenho positivo, com um EMQ médio durante o teste 83,46% menor que o máximo observado em outros trabalhos (1,005).

A aplicação do modelo nos nove Sistemas Simplificados de Abastecimento de Água (SSAA) de comunidades rurais do estado do Ceará estudados, indicou que as porcentagens de acerto da rede neural oscilaram entre 40,00% e 61,82%, para as precisões de 0,2 e 0,5 mg/L, respectivamente. Identificou-se que a robustez do modelo varia de acordo com o tipo de manancial de captação, a existência de etapa de filtração e de reforços de cloro.

Nos sistemas de Lagoa do Juá e Lagoa dos Encantos, obteve-se os piores desempenhos, com acertos variando de 11,11% a 25,00% ao considerar os limiares de 0,2 e 0,5 mg/L de CRL.

A captação de água no Canal do Trabalhador e a utilização de filtro no tratamento da água desses SSAA podem ter influenciado negativamente a performance da rede neural, devido à interação variada do parâmetro turbidez com o cloro residual livre. Particularmente para o SSAA Lagoa dos Encantos, a presença de reforços de cloro também contribuiu para o baixo desempenho. Os SSAA Santo Antônio e Carrapateira, com captação por poços em manancial subterrâneo, alcançaram taxas de acerto 18% maiores que a média obtida na etapa de teste durante a construção da RNA, chegando a 100,00% de precisão para o limiar de 0,5 mg/L de cloro residual livre.

A análise do desempenho das redes de distribuição de água revelou a predominância de uma performance inadequada quanto ao atendimento do cloro residual livre aos padrões estabelecidos pela Portaria nº 888/2021. Dos sistemas avaliados, 55,56% obtiveram classificação “péssima” quanto ao Índice de Sustentabilidade (IS), resultado influenciado pela vulnerabilidade máxima (1,000) alcançada pelos sistemas Calumbi, Acarape, Tabocas, Carrapateira e Santo Antônio, e pela resiliência mínima (0,000) que os dois últimos sistemas apresentaram.

Um desempenho “excelente” foi obtido pelos sistemas Lagoa do Juá e Venâncio, tendo em vista à alta confiabilidade e resiliência associadas a baixa vulnerabilidade. Os SSAA Lagoa dos Encantos e Realejo enquadraram-se na categoria de desempenho “ruim”, devido à baixa capacidade de recuperação após falhas e reduzida confiabilidade observada em Realejo (0,400).

Os resultados obtidos evidenciam as fragilidades dos SSAA quanto ao CRL. Este fato, destaca a necessidade de aprimorar as estratégias de gestão do serviço de abastecimento de água, implementando o Plano de Segurança da Água (PSA) e conduzindo as avaliações de desempenho de forma regular. Para tanto, garantir o controle e a qualidade da água é essencial, sendo viável a aplicação do modelo desenvolvido para a previsão do decaimento de cloro residual livre, a partir de RNA.

A implementação do modelo desenvolvido em sistemas de distribuição de água de comunidades rurais poderá promover a otimização dos recursos financeiros e os requisitos de tempo e pessoal, proporcionando uma performance superior em comparação aos métodos atualmente empregados. O modelo pode ser integrado ao reservatório da estação de tratamento de água, utilizando dados provenientes de sensores de baixo custo, auxiliando na determinação da taxa ótima de dosagem de cloro e na previsão de inconformidade relacionadas aos padrões de potabilidade estabelecidos pela Portaria nº 888/2021 do Ministério da Saúde, por meio de um *software*.

5.1 RECOMENDAÇÕES

Em função dos resultados apresentados neste estudo é sugerido para trabalhos futuros:

- Realizar o retreino da Rede Neural Artificial (RNA), a fim de considerar as condições dinâmicas dos sistemas e da qualidade da água, buscando aprimorar a eficácia do modelo desenvolvido. Este processo deverá levar em consideração as características particulares das amostras nas quais a RNA apresentou desempenho subótimo, especialmente aquelas provenientes de sistemas com captação em manancial superficial, etapa de filtração e recloração;
- Desenvolver um modelo de RNA de classificação, que verifique a conformidade com a Portaria nº 888/2021, indicando se o CRL está conforme ou inconforme, ao invés de prever valores específicos de cloro na rede de distribuição. Essa abordagem poderá possibilitar um aumento na performance do modelo, por reduzir a complexidade da saída da rede neural;
- Elaborar um programa computacional para calcular os indicadores de desempenho (confiabilidade, resiliência e vulnerabilidade), assim como o índice de sustentabilidade, utilizando os dados previstos pela RNA. O *software* poderá avaliar em tempo real a performance das redes de distribuição de água quanto ao CRL, auxiliando na gestão dos sistemas de abastecimento de água de comunidades rurais.

REFERÊNCIAS

- AGUIAR, T. D.; ARRUDA, P.; SCALIZE, P. S. Tecnologias de tratamento de água para abastecimento público empregadas no estado de Goiás. In: Congresso da Água, 15., 2021, Rio de Janeiro. **Anais eletrônicos** [...] Lisboa: APRH, 2021. p. 1-4. Disponível em: <https://www.aprh.pt/congressoagua2021/docs/15ca_96.pdf>. Acesso em: 08 out. 2022.
- ALAM, G.; IHSANULLAH, I.; NAUSHAD, M.; SILLANPÄÄ, M. Applications of artificial intelligence in water treatment for optimization and automation of adsorption processes: Recent advances and prospects. **Chemical Engineering Journal**, v. 427, p. 1-19, jan. 2022. DOI: 10.1016/j.cej.2021.130011.
- AL-ARAIMI, M. M. JOY, V. M.; FERRO, S. Optimization and Assessment of Residual Chlorine using Response Surface Methodology (RSM) and Artificial Neural Network (ANN) Modeling. **International Journal of Recent Technology and Engineering**, v. 8, n. 3, p. 258-263, 2019. DOI: 10.35940/ijrte.C4122.098319.
- ALBUQUERQUE, A. C.; CÉSSE, E. Â. P.; FELISBERTO, E.; SAMICO, I. C.; FRIAS, P. G. Avaliação de desempenho da regionalização da vigilância em saúde em seis Regiões de Saúde brasileiras. **Cadernos de Saúde Pública**, v. 35, n. 2, p. 1-16, 2019. DOI: 10.1590/0102-311X00065218.
- ALMEIDA, M. E.; ALVIM, C. A. N.; MAIOLA, M. R. A.; CAINZOS, R. L. P. Simulação computacional do decaimento de cloro residual livre na rede de distribuição de água de uma comunidade de Londrina. **Revista Terra & Cultura: Cadernos de Ensino e Pesquisa**, v. 38, n. especial, p. 19-31, Londrina, 2022. Disponível em: <<http://periodicos.unifil.br/index.php/Revistateste/article/view/2629>>. Acesso em: 15 out. 2022.
- AL-SHEHARI, T.; ALSOWAIL, R. A. An Insider Data Leakage Detection Using One-Hot Encoding, Synthetic Minority Oversampling and Machine Learning Techniques. **Entropy**, v. 23, n. 10, p. 1-24, 2021. DOI: 10.3390/e23101258.
- ARAUJO, A. A. Q.; LIMA, A. R. B.; SANTANA, C. F. D.; OLIVEIRA, S. H.; NASCIMENTO, R. S.; OLIVEIRA, R. Desempenho de rede de abastecimento de água a partir da simulação computacional utilizando o EPANET 2.0. **Brazil Journal of Development**, v. 6, n. 8, p. 55257-55269, ago. 2020. DOI: 10.34117/bjdv6n8-084.
- ARAUJO, L. F.; CAMARGO, F. P.; NETTO, A. T.; VERNIN, N. S.; ANDRADE, R. C. Análise da cobertura de abastecimento e da qualidade da água distribuída em diferentes regiões do Brasil no ano de 2019. **Ciência & Saúde Coletiva**, v. 27, n. 7, p. 2935-2947, 2022. DOI: 10.1590/1413-81232022277.16472021.
- ASSIS, E. M.; SANTOS, E. M.; FARIA, M. C. S.; RODRIGUES, J. L.; GARCEZ, A.; BOMFETI, C. A.; BARCELLOS, N. T. A vulnerabilidade de populações indígenas: qualidade da água consumida pela comunidade Maxakali, Minas Gerais, Brasil. **Sociedade & Natureza**, v. 32, p. 279-290, 2020. DOI: 10.14393/SN-v32-2020-43436.
- ATASHI, M.; KHODASHENAS, S.R.; ZIAEI, A. N.; FARMANI, R. Resilience analysis under simultaneous failure of pipes in water distribution network (Case study in one of the cities of Khorasan Razavi). **Journal of Hydraulics**, v. 15, n. 3, p. 93-105, set. 2020. DOI: 10.30482/JHYD.2020.235645.1464.
- BALEI, B.; NOY, I.; WILKINSON, S.; POTANGAROA, R. Economic factors affecting water supply resilience to disasters. **Socio-Economic Planning Sciences**, v. 76, p. 1-9, ago. 2021. DOI: 10.1016/j.seps.2020.100961.

- BASHEER, I A.; HAJMEER, M. Artificial neural networks: fundamentals, computing, design, and application. **Journal of Microbiological Methods**, v. 43, n. 1, p. 3-31, DEZ. 2000. DOI: 10.1016/S0167-7012(00)00201-3.
- BELLO, O.; ABU-MAHFOUZ, A. M.; HAMAM, Y.; PAGE, P. R.; ADEDEJI, K. B.; PILLER, O. Solving Management Problems in Water Distribution Networks: A Survey of Approaches and Mathematical Models. **Water**, v. 11, n. 3, p. 1-29, mar. 2019. DOI: 10.3390/w11030562.
- BELLOTTI, L.; BRANDÃO, S. R.; PACHECO, K. T. S.; FRAZÃO, P.; ESPOSTI, C. D. D. Vigilância da qualidade da água para consumo humano: potencialidades e limitações com relação à fluoretação segundo os trabalhadores. **Saúde em Debate**, v. 43, n. spe3, p. 51-62, dez. 2020. DOI: 10.1590/0103-11042019S304.
- BEZERRA, S. T. M.; PERTEL, M.; MACÊDO, J. E. S. Avaliação de desempenho dos sistemas de abastecimento de água do Agreste brasileiro. **Ambiente Construído**, v. 19, n. 3, p. 249-258, jul./set. 2019. DOI: 10.1590/s1678-86212019000300336.
- BHADULA, R. C.; KALA, V. N.; PANT, R.; KHOLIYA, D.; SINGH, S. J. Ageing of pipe induced chlorine concentration decay: A one dimensional mass transfer modeling. **Materialstoday: Proceedings**, v. 46, n. 20, p. 10761-10765, 2021. DOI: 10.1016/j.matpr.2021.01.670.
- BHARDWAJ, A.; DAGAR, V.; KHAN, M. O.; AGGARWAL, A.; ALVARADO, R.; KUMAR, M.; IRFAN, M.; PROSHA, R. Smart IoT and Machine Learning-based Framework for Water Quality Assessment and Device Component Monitoring. **Environmental Science and Pollution Research**, v. 29, p. 46018-46036, 2022. DOI: 10.1007/s11356-022-19014-3.
- BIONDO, F. D. **Modelo de gerenciamento da qualidade da água: estudo do decaimento de cloro na rede de distribuição**. Dissertação (Mestrado em Recursos Hídricos e Saneamento Ambiental) – Centro de Ciências Exatas e Tecnológicas, Universidade Estadual do Oeste do Paraná, Cascavel, 2019.
- BOUSQUAT, A.; FAUSTO, M. C. R.; ALMEIDA, P. F.; LIMA, J. G.; SEIDL, H.; SOUSA, A. B. L.; GIOVANELLA, L. Remoto ou remotos: a saúde e o uso do território nos municípios rurais brasileiros. **Revista de Saúde Pública**, v. 56, p. 1-11, ago. 2022. DOI: 10.11606/s1518-8787.2022056003914.
- BOWDEN, G. J.; NIXON, J. B.; DANDY, G. C.; MAIER, H. R.; HOLMES, M. Forecasting chlorine residuals in a water distribution system using a general regression neural network. **Mathematical and Computer Modelling**, v. 44, n. 5-6, p. 469-484, set. 2006. DOI: 10.1016/j.mcm.2006.01.006.
- BRAGA, A. P.; LUDERMIR, T. B.; CARVALHO, A. C. P. L. F. **Redes Neurais Artificiais: Teoria e Aplicações**. 1. ed. Rio de Janeiro: LTC, 2000.
- BRASIL. Conselho Nacional do Meio Ambiente. **Resolução CONAMA nº 357, de 17 de março de 2005**. Dispõe sobre a classificação dos corpos de água e diretrizes ambientais para o seu enquadramento, bem como estabelece as condições e padrões de lançamento de efluentes, e dá outras providências. Brasília: Diário Oficial da União, 2005.
- BRASIL. Conselho Nacional do Meio Ambiente. **Resolução CONAMA nº 396, de 3 de abril de 2008**. Dispõe sobre a classificação e diretrizes ambientais para o enquadramento das águas subterrâneas e dá outras providências. Brasília: Diário Oficial da União, 2008.
- BRASIL. Ministério da Economia. Instituto de Pesquisa Econômica Aplicada. **Cadernos ODS: ODS 6 – Assegurar a disponibilidade e gestão sustentável da água e saneamento para todas e todos**. Brasília: IPEA, 2019.

BRASIL. Ministério da Saúde. Fundação Nacional de Saúde. **Manual de Saneamento**. 5. ed. Brasília: Funasa, 2019.

BRASIL. Ministério da Saúde. Fundação Nacional de Saúde. **Manual Prático de Análise de Água**. 4. ed. Brasília: Funasa, 2013.

BRASIL. Ministério da Saúde. Gabinete do Ministro. **Portaria nº 888, de 04 de maio de 2021**. Altera o Anexo XX da Portaria de Consolidação GM/MS nº 5, de 28 de setembro de 2017, para dispor sobre os procedimentos de controle e de vigilância da qualidade da água para consumo humano e seu padrão de potabilidade. Brasília: Diário Oficial da União, 2021.

BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. Departamento de Vigilância em Saúde ambiental e Saúde do Trabalhador. **Diretriz Nacional do Plano de Amostragem da Vigilância da Qualidade da Água para Consumo Humano**. 1. ed. Brasília: Ministério da Saúde, 2016.

BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. Departamento de Vigilância em Saúde ambiental e Saúde do Trabalhador. **Plano de segurança da água: garantindo a qualidade e promovendo a saúde – um olhar do SUS**. 1. ed. Brasília: Ministério da Saúde, 2012.

BRASIL. Ministério do Desenvolvimento Regional. Secretaria Nacional de Saneamento. **Plano Nacional de Saneamento Básico: Relatório de Avaliação Anual 2020**. Brasília: Secretaria Nacional de Saneamento, 2022.

BRASIL. Programa Vigiagua. 2023. Disponível em: <<https://sisagua.saude.gov.br/sisagua/paginaExterna.jsf>>. Acesso em: 05 set. 2023.

BRASIL. Secretaria-Geral. Subchefia para Assuntos Jurídicos. **Lei Federal nº 14.026, de 15 de julho de 2020**. Atualiza o marco legal do saneamento básico e dá outras providências. Brasília: Diário Oficial da União, 2020.

CÁRCERE, N. D.; MAIA-RODRIGUES, B. H. A gestão comunitária da água: Caminhos para promoção da justiça hídrica e mitigação de conflitos socioambientais. **Cadernos do Leste**, Belo Horizonte, v. 19, n. 19, p. 21-34, jan./dez., 2019. DOI: 10.29327/248949.19.19-2.

CASTAÑEDA-MIRANDA, A.; CASTAÑO-MENESES, V. M. Smart frost measurement for anti-disaster intelligent control in greenhouses via embedding IoT and hybrid AI methods. **Measurement**, v. 164, p. 1-13, nov. 2020. DOI: 10.1016/j.measurement.2020.108043.

CASTRO, C. M. S. N. **Emprego de indicadores e índices de sustentabilidade na avaliação de medidas emergenciais de gestão da água: caso da bacia do Rio Descoberto, DF**. Dissertação (Mestrado Profissional em Rede Nacional em Gestão e Regulação de Recursos Hídricos) – Instituto de Pesquisas Hidráulicas, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2019.

CETESB. Companhia Ambiental do Estado de São Paulo. **Apêndice C: Índices de Qualidade das Águas**. São Paulo: CETESB, 2013.

CHEN, S.; FANG, G.; HUANG, X.; ZHANG, Y. Water Quality Prediction Model of a Water Diversion Project Based on the Improved Artificial Bee Colony–Backpropagation Neural Network. **Water**, v. 10, n. 6, p. 1-19, jun. 2018. DOI: 10.3390/w10060806.

CHEN, Y.; SONG, L.; LIU, Y.; YANG, L.; LI, D. A Review of the Artificial Neural Network Models for Water Quality Prediction. **Applied Sciences**, v. 10, n. 17, p. 1-49, ago. 2020. DOI: 10.3390/app10175776.

CODEPLAN. Companhia de Planejamento do Distrito Federal. **Um Panorama das Águas no Distrito Federal**. Brasília: CODEPLAN, 2020.

CORDOBA, G. A. C.; TUHOVČÁK, L.; TAUŠ, M. Using Artificial Neural Network Models to Assess Water Quality in Water Distribution Networks. **Procedia Engineering**, v. 70, p. 399-408, 2014. DOI: 10.1016/j.proeng.2014.02.045.

CORREA, R. F. M. **Elaboração de modelo conceitual para monitoramento de riscos à contaminação da água em comunidades rurais**. Dissertação (Mestrado em Engenharia Urbana) – Centro de Ciências Exatas e de Tecnologia, Universidade Federal de São Carlos, São Carlos, 2020.

CORREIA, A.; BARROS, E.; SILVA, J.; RAMALHO, J. Tecnologias de tratamento de água para abastecimento público empregadas no estado de Goiás. In: Encontro Regional de Matemática Aplicada e Computacional, 8., 2008, Natal. **Anais eletrônicos [...]** Natal: UFRN, 2008. p. 1-5. Disponível em: <https://dimap.ufrn.br/~sbmac/ermac2008/Anais/Resumos%20Estendidos/Analise%20da%20urbidez_Aislan%20Correia.pdf>. Acesso em: 11 fev. 2024.

DAMKE, T.; PASINI, F. A importância da potabilidade da água no saneamento básico para a promoção da saúde pública no Brasil. **Revista Teccen**, v. 13, n. 1, p. 08-15, jan./jul., 2020. DOI: 10.21727/teccen.v13i1.2200.

DANDIE, C. E.; OGUNNIYI, A. D.; FERRO, S.; HALL, B.; DRIGO, B.; CHOW, C. W. K.; VENTER, H.; MYERS, B.; DEO, P.; DONNER, E.; LOMBI, E. Disinfection options for irrigation water: reducing the risk of fresh produce contamination with human pathogens. **Critical Reviews in Environmental Science and Technology**, v. 50, n. 20, p. 2144-2174, 2020. DOI: 10.1080/10643389.2019.1704172.

DERNTL, M. F. Brasília e suas unidades rurais: planos e projetos para o território do Distrito Federal entre fins da década de 1950 e início da década de 1960. **Anais do Museu Paulista**, v. 28, p. 1-32, 2020. DOI: 10.1590/1982-02672020v28e30.

DIAS, J. P.; LÉLIS, J. L. F.; DIAS, I. D. **Saneamento rural**. 1. ed. Brasília: Emater-DF, 2021.

FARO, G. T. C. **Avaliação de sistemas de recursos hídricos complexos por meio de indicadores de desempenho**. Dissertação (Mestrado em Ciências) – Departamento de Engenharia de Hidráulica e Ambiental, Escola Politécnica da Universidade de São Paulo, São Paulo, 2020.

FORTES, A. C. C.; BARROCAS, P. R. G.; KLIGERMAN, D. C. A vigilância da qualidade da água e o papel da informação na garantia do acesso. **Saúde Debate**, v. 43, n. especial 3, p. 20-34, dez., 2019. DOI: 10.1590/0103-11042019S302.

FOSTER, A. R. **Evaluation of magnetic fields and the impacts of chloramines on water treatment**. Tese (Doutorado em Filosofia) – Department of Environmental Science, University of Arizona, 2022.

FRANÇA, L. R. C.; SOUSA, A. F. N.; BATISTA, G. S.; CUNHA, P. H.; CARNEIRO, D. Q. Vigilância da qualidade da água para consumo humano nos sistemas de abastecimento de água e soluções coletivas em comunidades rurais do estado do Ceará. In: Congresso Brasileiro de Engenharia Sanitária e Ambiental, 32., 2023, Belo Horizonte. **Anais eletrônicos [...]** Belo Horizonte: ABES, 2023a. p. 1-13. Disponível em: <https://cbesa.sigotech.online/storage/trabalhos/arquivos/completo/718_tema_xii.pdf>. Acesso em: 20 dez. 2023.

FRANÇA, L. R. C.; SOUSA, A. F. N.; CUNHA, P. H.; ARAÚJO, M. M. C.; GRADVOL, S. T. S. Planos de Segurança da Água em comunidades rurais: análise de metodologias de implementação no Brasil e no mundo. In: Congresso Brasileiro de Engenharia Sanitária e Ambiental, 32., 2023, Belo Horizonte. **Anais eletrônicos** [...] Belo Horizonte: ABES, 2023b. p. 1-8. Disponível em: <https://cbesa.sigotech.online/storage/trabalhos/arquivos/completo/714_tema_xii.pdf>. Acesso em: 07 jan. 2024.

FUNASA. Fundação Nacional de Saúde. UFCG. Universidade Federal de Campina Grande. **Produto 03 – Relatório de Diagnóstico do SSAA/SAC**. Campina Grande: UFCG, 2023.

GARAY, D. F. A.; VOLTOLINI, L. C.; BASTOS, R. G.; SOUZA, C. F. Effect of turbidity on ultraviolet disinfection of domestic wastewater for agricultural reuse. **Rev. Ambient. Água**, v. 16, n. 6, p. 1-12, Taubaté, 2021. DOI: 10.4136/ambi-agua.2766.

GÁRCIA-ÁVILA, F.; SÁNCHEZ-ALVARRACIN, C.; CADME-GALABAY, M.; CONCHADO-MARTÍNEZ, J.; GARCÍA-MEIRA, G.; ZHINDÓN-ARÉVALO, C. Relationship between chlorine decay and temperature in the drinking water. **MethodsX**, v. 7, p. 1-9, 2020. DOI: 10.1016/j.mex.2020.101002.

GHOLAMI-BORUJENI, F. Relationship Between Environment Temperature, TOC, Turbidity, and HPC on Free Residual Chlorine in Municipal Drinking Water Distribution Networks (WDNs): A Case Study (Abadan, Iran). In: International Conference on Advances in Civil and Ecological Engineering Research, 4., 2022, Singapore. **Anais eletrônicos** [...] Singapore: Springer, 2022. DOI: 10.1007/978-981-19-5783-3_24.

GIBBS, M. S.; MORGAN, N.; MAIER, H. R.; DANDY, G. C.; NIXON, J. B.; HOLMES, M. Investigation into the relationship between chlorine decay and water distribution parameters using data driven methods. **Mathematical and Computer Modelling**, v. 44, n. 5-6, p. 485-498, set. 2006. DOI: 10.1016/j.mcm.2006.01.007.

GOUDARZI, A. Frequency domain beamforming using neuronal Networks. In: Berlin Beamforming Conference, 9., 2022, Berlin, Germany. **Anais eletrônicos** [...] Berlin, 2022. Disponível em: <<https://elib.dlr.de/186820/1/ELIB-Eintrag-2022-GoudarziA-186820-PaperPublished.pdf>>. Acesso em: 21 ago. 2023.

GUEDES, T. L.; SOUZA, F. H.; GHISI, D. B.; PEREZ, A. B. A.; DALSSASSO, R. L.; SENS, M. L. Aplicação da filtração em margem de rio como alternativa de tratamento de água para comunidades isoladas. **Revista DAE**, v. 67, n. 215, p. 84-94, jan./mar., 2019. DOI: 10.4322/dae.2019.007.

GÜLCÜ, S. An Improved Animal Migration Optimization Algorithm to Train the Feed-Forward Artificial Neural Networks. **Research Article-Computer Engineering and Computer Science**, v. 47, p. 9557–9581, 2022. DOI: 10.1007/s13369-021-06286-z.

HASHIMOTO, T.; STEDINGER, J. R.; LOUCKS, D. P. Reliability, Resiliency, and Vulnerability Criteria For Water Resource System Performance Evaluation. **Water Resources Research**, v. 18, n. 1, p. 14-20, fev. 1982. Disponível em: <<https://core.ac.uk/download/pdf/79433422.pdf>>. Acesso em: 23 jan. 2023.

HASSON, U.; NASTASE, S. A.; GOLDSTEIN, A. Direct Fit to Nature: An Evolutionary Perspective on Biological and Artificial Neural Networks. **Neuron**, v. 105, n. 3, p. 416-434, fev. 2020. DOI: 10.1016/j.neuron.2019.12.002.

HEMDAN, B. A.; EL-TAWEEL, G. E.; GOSWAMI, P.; PANT, D.; SEVDA, S. The role of biofilm in the development and dissemination of ubiquitous pathogens in drinking water

distribution systems: an overview of surveillance, outbreaks, and prevention. **World Journal of Microbiology and Biotechnology**, v. 37, n. 2, p. 1-18, 2021. DOI: 10.1007/s11274-021-03008-3.

HOQUE, Y. M.; HANTUSH, M. M.; GOVINDARAJU, R. S. On the scaling behavior of reliability–resilience–vulnerability indices in agricultural watersheds. **Ecological Indicators**, v. 40, p. 136-146, mai. 2014. DOI: 10.1016/j.ecolind.2014.01.017.

HOQUE, Y. M.; TRIPATHI, S.; HANTUSH, M. M.; GOVINDARAJU, R. S. Watershed reliability, resilience and vulnerability analysis under uncertainty using water quality data. **Journal of Environmental Management**, v. 109, p. 101-112, out. 2012. DOI: 10.1016/j.jenvman.2012.05.010.

HOSSAIN, S.; CHOW, C. W. K.; COZINHEIRO, D.; SAWADE, E.; HEWA, G. A. Review of chloramine decay models in drinking water system. **Environmental Science: Water Research & Technology**, v. 8, p. 926-948, 2022. DOI: 10.1039/D1EW00640A.

HOWE, K. J.; HAND, D. W.; CRITTENDEN, J. C.; TRUSSELL, R. R.; TCHOBANOGLIOUS, G. **Princípios de tratamento de água**. Tradução de Noveritis do Brasil. Revisão técnica de Elvis Carissimi. 1. ed. São Paulo: Cengage, 2016.

IBGE. Instituto Brasileiro de Geografia e Estatística. **Sinopse do Censo Demográfico: 2010**. Rio de Janeiro: IBGE, 2011.

JAFARI, I.; LUO, R.; LIM, F. Y.; HUI, N. S.; JIANGYONG, H. Machine-learning-assisted prediction and optimized kinetic modelling of residual chlorine decay for enhanced water quality management. **Chemosphere**, v. 341, p. 1-12, nov. 2023. DOI: 10.1016/j.chemosphere.2023.140011.

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

JIANG, W.; XU, Y.; CHEN, Z.; ZHANG, N.; ZHOU, J. Fault diagnosis for rolling bearing using a hybrid hierarchical method based on scale-variable dispersion entropy and parametric t-SNE algorithm. **Measurement**, v. 191, p. 1-17, 2022. DOI: 10.1016/j.measurement.2022.110843.

KHEDIKAR, I. P.; TEMBHURKAR, A. R.; DABHEKAR, K. R.; GODBOLEY, B. J. Effect of turbidity on survival of Escherichia coli, faecal coliform and total coliform in grey water by using solar disinfection (SODIS). **Journal of Physics: Conference Series**, v. 1913, p. 1-6, fev., 2021. DOI: 10.1088/1742-6596/1913/1/012068.

KOO, J.; INAKAZU, T.; KOIZUMI, A.; ARAI, Y.; KIM, K.; AHN, J. Application of Artificial Neural Network for Reducing of Chlorine Residual Concentration in Water Distribution Network. **Water Practice and Technology**, v. 3, n. 2, p. 969-978, jun. 2008. DOI: 10.2166/wpt.2008.032.

KOTE, A.; WADKAR, D. Modeling of chlorine and coagulant dose in a water treatment plant by artificial neural. **Engineering, Technology & Applied Science Research**, v. 9, n. 3, p. 4176-4181, 2019. DOI: 10.48084/ETASR.2725.

KURAJICA, L.; UJEVIĆ BOŠNJAK, M.; KINSELA, A.S.; ŠTIGLIĆ, J.; WAITE, T.D.; CAPAK, K.; PAVLIĆ, Z. Effects of changing supply water quality on drinking water distribution networks: Changes in NOM optical properties, disinfection byproduct formation, and Mn deposition and release. **Science of The Total Environment**, v. 762, p. 1-13, mar. 2021. DOI: 10.1016/j.scitotenv.2020.144159.

LECHEVALLIER, E.; STAVRIDOU, E.; GRANELL-RUIZ, R.; KEY, G.; BERCKMOES, E. **FERTINNOWA benchmark report – deliverable 3.3**. CATE, 2018. Disponível em: <<https://www.fertinnowa.com/wp-content/uploads/2018/05/Benchmark-report.pdf>>. Acesso em: 10 out. 2022.

LÉZIART, T.; ROCHERE, P. M. D.; CHESWICK, R.; JARVIS, P. Effect of turbidity on water disinfection by chlorination with the emphasis on humic acids and chalk. **Environmental Technology**, v. 40, n. 13, p. 1734-1743, fev., 2019. DOI: 10.1080/09593330.2019.1585480.

LIANG, Y.; NIU, D.; HONG, W. C. Short term load forecasting based on feature extraction and improved general regression neural network model. **Energy**, v. 166, p. 653-663, jan. 2019. DOI: 10.1016/j.energy.2018.10.119.

LIBÂNIO, M. **Fundamentos de qualidade e tratamento de água**. 4. ed. Campinas: Editora Átomo, 2016.

LIBRANTZ, A. F.; SANTOS, F. C. R. dos; DIAS, C. G. Artificial neural networks to control chlorine dosing in a water treatment plant. **Acta Scientiarum. Technology**, v. 40, n. 1, p. 1-9, 2018. DOI: 10.4025/actascitechnol.v40i1.37275.

LINDERMAN, G. C.; STEINERBERGER, S. Clustering with t-SNE, Provably. **SIAM Journal on Mathematics of Data Science**, v. 1, n. 2, p. 313-332, 2019. DOI: 10.1137/18M1216134.

LINDERMAN, G. C.; STEINERBERGER, S. Clustering with t-SNE, Provably. **SIAM Journal on Mathematics of Data Science**, v. 1, n. 2, p. 313-332, 2019. DOI: 10.1137/18M1216134.

LIU, R. A.; MCDONALD, J. A.; SATHASIVAN, A.; KHAN, S. J. Disinfectant residual stability leading to disinfectant decay and by-product formation in drinking water distribution systems: A systematic review. **Water Research**, v. 153, p. 335-348, abr. 2019. DOI: 10.1016/j.watres.2019.01.020.

LOUCKS, D. P. Quantifying trends in system sustainability. **Hydrological Sciences Journal**, v. 42, n. 4, p. 513-530, ago. 1997. DOI: 10.1080/02626669709492051.

MACEDO, R. G.; HERRERO, N. P.; WESSELINK, P.; VERSLUIS, M.; SLUIS, L. V. D. Influence of the dentinal wall on the pH of sodium hypochlorite during root canal irrigation. **Journal of endodontics**, v. 40, n. 7, p. 1005-1008, 2014. DOI: 10.1016/j.joen.2013.12.018.

MAIER, H. R.; DANDY, G. C. Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. **Environmental Modelling & Software**, v. 15, n. 1, p. 101-124, jan. 2000. DOI: 10.1016/S1364-8152(99)00007-9.

MAIER, H. R.; JAIN, A.; DANDY, G. C.; SUDHEER, K. P. Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions. **Environmental Modelling & Software**, v. 25, n. 8, p. 891-909, ago. 2010. DOI: 10.1016/j.envsoft.2010.02.003.

MANUEL, P.; LEITÃO, A. A.; BOAVENTURA, R. A. R. Qualidade da Água para Consumo Humano na Cidade do Uíge (Angola). **Revista Internacional em Língua Portuguesa**, n. 33, p. 75-93, ago., 2018. DOI: 10.31492/2184-2043.RILP2018.33/pp.75-93.

MARCY, M. **Metodologia baseada em redes neurais artificiais para a detecção de danos estruturais**. Tese (Doutorado em Estruturas e Construção Civil) – Faculdade de Tecnologia, Universidade de Brasília, Brasília, 2017.

- MARCY, M.; BRASILIANO, A.; SILVA, G. B. L.; DOZ, G. Locating damages in beams with artificial neural network. **International Journal of Lifecycle Performance Engineering**, v. 1, n. 4, p. 398-413, 2014. DOI: 10.1504/IJLCPE.2014.064110.
- MAY, R. J.; DANDY, G. C.; MAIER, H. R.; NIXON, J. B. Application of partial mutual information variable selection to ANN forecasting of water quality in water distribution systems. **Environmental Modelling & Software**, v. 23, n. 10-11, p. 1289-1299, out./nov. 2008. DOI: 10.1016/j.envsoft.2008.03.008.
- MAZHAR, M. A.; KHAN, N. A.; AHMED, S.; KHAN, A. H.; HUSSAIN, A.; RAHISUDDIN; CHANGANI, F.; YOUSEFI, M.; AHMADI, S.; VAMBOL, V. Chlorination disinfection by-products in municipal drinking water – A review. **Journal of Cleaner Production**, v. 273, p. 123-159, nov. 2020. DOI: 10.1016/j.jclepro.2020.123159.
- MEDEIROS, L. E. L.; RODRIGUES, A. C. L.; BARBOSA, D. L.; FEITOSA, P.; H. C.; COURA, M. A. Avaliação da qualidade do serviço de abastecimento de água em Campina Grande-PB com o uso de indicadores, incluindo o de satisfação dos usuários. **Revista DAE**, v. 68, n. 224, p. 135-152, 2020. DOI: 10.36659/dae.2020.045.
- MELO, C. S. T. **Avaliação da qualidade da água da captação e do sistema de abastecimento de Fagilde, Viseu**. Relatório de estágio (Mestrado em Tecnologias Ambientais) – Escola Superior de Tecnologia e Gestão de Viseu, Instituto Politécnico de Viseu, Viseu, 2018.
- MENTES, A.; GALIATSOTOU, P.; SPYROU, D.; SAMARAS, A.; STOURNARA, P. Hydraulic Simulation and Analysis of an Urban Center's Aqueducts Using Emergency Scenarios for Network Operation: The Case of Thessaloniki City in Greece. **Water**, v. 12, n. 6, p. 1-29, 2020. DOI: 10.3390/w12061627.
- METCALF, L.; EDDY, H. P. **Tratamento de Efluentes e Recuperação de Recursos**. Tradução de Ivanildo Hespanhol e José Carlos Mierzwa. 5. ed. [s. l.]: AMGH Editora LTDA, 2016.
- MICHELAN, D. C. G. S.; BATISTA, I. F.; BATISTA, D. F.; SANTOS, D. G.; MENDONÇA, L. C.; LIMA, D. M. F. Desempenho das etapas de tratamento de água da estação de tratamento de água Poxim. **Scientia cum Industria**, v. 7, n. 3, p. 7-14, 2019. DOI: 10.18226/23185279.v7iss3p7.
- MONTEIRO, L.; CARNEIRO, J.; COVAS, D. I. C. Modelling chlorine wall decay in a full-scale water supply system. **Urban Water Journal**, v. 17, n. 8, p. 754-762, 2020. DOI: 10.1080/1573062X.2020.1804595.
- NASCIMENTO, J. P. B. **Fundamentos Bootcamp Analista de Machine Learning**. 1. ed. Belo Horizonte: IGTI, 2020.
- NOURANI, V.; BAGHANAM, A. H.; ADAMOWSKI, J.; KISI, O. Applications of hybrid wavelet–Artificial Intelligence models in hydrology: A review. **Journal of Hydrology**, v. 514, p. 358-377, jun. 2014. DOI: 10.1016/j.jhydrol.2014.03.057.
- O'REILLY, G.; BEZUIDENHOUT, C. C.; BEZUIDENHOUT, J. J. Artificial neural networks: applications in the drinking water sector. **Water Supply**, v. 18, n. 6, p. 1869-1887, jan. 2018. DOI: 10.2166/ws.2018.016.
- OGATA, I. S.; SANTOS, D. L.; MENESES, I. A.; OLIVEIRA, R.; RODRIGUES, A. C. L. Aplicação do índice de pobreza hídrica na região hidrográfica do alto curso do rio Paraíba, estado da Paraíba, Brasil. In: Congresso Brasileiro de Gestão Ambiental, 7., 2016, Campina Grande. **Anais eletrônicos [...]** Campina Grande: IBEAS, 2016. p. 1-9. Disponível em: <<https://www.ibeas.org.br/congresso/Trabalhos2016/VIII-071.pdf>>. Acesso em: 07 jan. 2024.

- OLIVEIRA, L. **Decaimento da concentração de cloro residual livre nas redes de abastecimento de água**. Dissertação (Mestrado em Engenharia Ambiental e Sanitária) – Escola de Engenharia Civil e Ambiental, Universidade Federal de Goiás, Goiânia, 2018.
- OLIVEIRA, L.; REIS, D. R.; HOFFMANN, N. K. S. A. Decay process of free residual chlorine concentration affected by travel time in water distribution systems. **Rev. Ambient. Água**, v. 17, n. 3, p. 1-14, Taubaté, 2020. DOI: 10.4136/ambi-agua.2830.
- OLIVEIRA, L.; REIS, D. R.; HOFFMANN, N. K. S. A. Influência da temperatura e da concentração inicial no decaimento do cloro. **Revista DAE**, v. 69, n. 228, p. 117-131, jan./mar. 2021. DOI: 10.36659/dae.2021.009.
- OLIVEIRA, M. L. **Desenvolvimento de método para avaliação de desempenho de sistemas de abastecimento de água: aplicação ao caso da rede DF e entorno**. Dissertação (Mestrado em Tecnologia Ambiental e Recursos Hídricos) – Faculdade de Tecnologia, Universidade de Brasília, Brasília, 2016.
- OLIVEIRA, P. P. **Proposição de tecnologias de tratamento de água aplicáveis em comunidades rurais e tradicionais**. Dissertação (Mestrado em Engenharia Ambiental e Sanitária) – Escola de Engenharia Ambiental e Sanitária, Universidade Federal de Goiás, Goiânia, 2021.
- ONYUTHA, C. A hydrological model skill score and revised R-squared. **Hydrology Research**, v. 53, n. 1, p. 51-64, 2021. DOI: 10.2166/nh.2021.071.
- ONYUTHA, C. From R-squared to coefficient of model accuracy for assessing "goodness-of-fits". **Geoscientific Model Development Discuss**, 2020. DOI: 10.5194/gmd-2020-51.
- ONYUTHA, C. Multiple Statistical Model Ensemble Predictions of Residual Chlorine in Drinking Water: Applications of Various Deep Learning and Machine Learning Algorithms. **Journal of Environmental and Public Health**, v. 2022, p. 1-15, 2022. DOI: 10.1155/2022/7104752.
- ONYUTHA, C.; KWIO-TAMALE, J. C. Modelling chlorine residuals in drinking water: a review. **International Journal of Environmental Science and Technology**, v. 19, n. 1, p. 1-18, 2022. DOI: 10.1007/s13762-022-03924-3.
- OYEBODE, O.; STRETCH, D. Neural network modeling of hydrological systems: A review of implementation techniques. **Natural Resource Modeling**, v. 32, n. 1, p. 1-14, set. 2018. DOI: 10.1111/nrm.12189.
- PENG, X.; LI, Q.; JING, J. CNGAT: A Graph Neural Network Model for Radar Quantitative Precipitation Estimation. **IEEE Transactions on Geoscience and Remote Sensing**, v. 60, p. 1-14, 2021. DOI: 10.1109/TGRS.2021.3120218.
- PEREIRA, M. G.; SOUZA, A. R.; SILVA, S. L. O.; BRITO, M. R. Qualidade da água para consumo humano e doenças diarreicas agudas no estado do Tocantins. **Revista de Engenharia e Tecnologia**, v. 13, n. 2, p. 259-273, jun., 2021. Disponível em: <<https://revistas.uepg.br/index.php/ret/article/view/18081>>. Acesso em: 15 out. 2022.
- PIRES, W. D. F. **Gestão de riscos em sistemas de abastecimento de água - uma abordagem**. Dissertação (Mestrado em Engenharia do Ambiente) – Faculdade de Engenharia, Universidade do Porto, Porto, 2020.
- QUITANA, G.; MOLINOS-SENANTE, M.; CHAMORRO, A. Resilience of critical infrastructure to natural hazards: A review focused on drinking water systems. **International Journal of Disaster Risk Reduction**, v. 48, p. 1-13, set. 2020. DOI: 10.1016/j.ijdr.2020.101575.

REN, K.; HUANG, S.; HUANG, Q.; WANG, H.; LENG, G.; FANG, W.; LI, P. Assessing the reliability, resilience and vulnerability of water supply system under multiple uncertain sources. **Journal of Cleaner Production**, v. 252, p. 1-14, abr. 2020. DOI: 10.1016/j.jclepro.2019.119806.

RISSO, S. S. O.; ASANO, P. T. L.; SUBTIL, E. L.; BRAMBILA, M. C. V. Indicadores de desempenho estatísticos para reservatórios: uma revisão de literatura. *In*: OLIVEIRA, R. J (org.). **Recursos Hídricos: gestão, planejamento e técnicas em pesquisa**. 1. ed. [S.l.]: Editora Científica Digital, 2021. p. 112-139. DOI: 10.37885/210504613.

RODRIGUES, D. G. **Sistema alternativo para desinfecção da água por pasteurização solar para pequenas comunidades**. Tese (Doutorado em Engenharia Agrícola) – Faculdade de Engenharia Agrícola, Universidade Estadual de Campinas, Campinas, 2018.

RODRIGUES, M. F. da S.; SCALIZE, P. S. Decaimento de cloro residual livre em águas distribuídas em redes de abastecimento. **Brazilian Journal of Development**, v. 5, n. 9, p. 16366-16375, set., 2019. DOI:10.34117/bjdv5n9-187.

RUSTAM, F.; ISHAQ, A.; KOKAB, S. T.; DIEZ, I. T.; MAZÓN, J. L. V.; RODRÍGUEZ, C. L.; ASHRAF, I. An Artificial Neural Network Model for Water Quality and Water Consumption Prediction. **Water**, v. 14, n. 21, p. 1-23, 2022. DOI: 10.3390/w14213359.

SANDOVAL-SOLIS, S.; MCKINNEY, D. C.; LOUCKS, D. P. Sustainability Index for Water Resources Planning and Management. **Journal of Water Resource Planning Management**, v. 137, n. 5, p. 380-391, 2011. DOI: 10.1061/(ASCE)WR.1943-5452.0000134.

SANTANA, A. B. C.; LOPES, A. F.; MENDES, A. P.; YAMAGUCHI, K. K. L. Análise de dados do Sistema de Informação de Vigilância da Qualidade da Água para Consumo Humano (Sisagua) no estado do Amazonas, 2016-2020. **Vigilância Sanitária em Debate**, v. 9, n. 4, p. 25-34, 2021. DOI: 10.22239/2317-269X.01936.

SANTI, M. D.; KAHN, U. T.; ARNOLD, M.; FESSELET, J. F.; ALI, S. I. Forecasting point-of-consumption chlorine residual in refugee settlements using ensembles of artificial neural networks. **Clean Water**, v. 4, n. 35, p. 1-16, jun. 2021. DOI: 10.1038/s41545-021-00125-2.

SANTI, M. **Forecasting chlorine residual for water safety using Artificial neural networks ensembles in Humanitarian water systems**. Dissertação (Mestrado em Ciências Aplicadas), York University, 2021.

SANTI, M.; ALI, S. I.; ARNOLD, M.; FESSELET, J. F.; HYVÄRINEN, A. M. J.; TAYLOR, D.; KHAN, U. T. Modelling point-of-consumption residual chlorine in humanitarian response: Can cost-sensitive learning improve probabilistic forecasts? **Plos Water**, v. 1, n. 9, p. 1-30, set. 2022. DOI: 10.1371/journal.pwat.0000040.

SANTOS, C. T. **Proposição de alternativa para a desinfecção de água no Campus Central da UFRN**. Trabalho de Conclusão de Curso (Bacharelado em Engenharia Ambiental) – Centro de Tecnologia, Universidade Federal do rio Grande do Norte, Natal, 2020.

SANTOS, G. R.; SANTANA, A. S. **Gestão comunitária da água: Soluções e dificuldades do saneamento rural no Brasil**. Brasília: Instituto de Pesquisa Econômica Aplicada, 2020.

SANTOS, L. E. A. **Modelação do decaimento do cloro residual numa água superficial**. Dissertação (Mestrado em Tecnologia Ambiental) – Escola Superior Agrária de Bragança, Instituto Politécnico de Bragança, Bragança, 2018.

SETSHEDI, K. J.; MUTINGWENDE, N.; NGQUALA, N. P. The Use of Artificial Neural Networks to Predict the Physicochemical Characteristics of Water Quality in Three District Municipalities, Eastern Cape Province, South Africa. **International Journal of**

Environmental Research and Public Health, v. 18, n. 10, 2021. DOI: 10.3390/ijerph18105248.

SEYMOUR, I.; O’SULLIVAN, B.; LOVERA, P.; ROHAN, J. F.; O’RIORDAN, A. Electrochemical detection of free-chlorine in Water samples facilitated by in-situ pH control using interdigitated microelectrodes. **Sensors and Actuators B: Chemical**, v. 325, p. 1-9, dez. 2020. DOI: 10.1016/j.snb.2020.128774.

SHHIPI-SHRESTHA, G.; MIAN, H. R.; MOHAMMADIUN, S.; RODRIGUEZ, M.; HEWAGE, K.; SADIQ, R. Digital water: artificial intelligence and soft computing applications for drinking water quality assessment. **Clean Technologies and Environmental Policy**, v. 25, p. 1409–1438, nov. 2023. DOI: 10.1007/s10098-023-02477-4.

SHI, X.; CLARK, G. G.; HUANG, C.; NGUYEN, T. H.; YUAN, B. Chlorine decay and disinfection by-products formation during chlorination of biofilms formed with simulated drinking water containing corrosion inhibitors. **Science of The Total Environment**, v. 815, abr. 2020. DOI: 10.1016/j.scitotenv.2021.152763.

SHORR, A. S. **Tratamento de Águas e Efluentes**. 1. ed. Rio de Janeiro: Freitas Bastos, 2022.

SILVA FILHO, V. G.; SANTOS, V. S.; GONÇALVES, S. N. O.; COSTA NETO, F. A.; BRITO, V. C. Avaliação da sustentabilidade do uso dos recursos hídricos no sistema Curema –Mãe d’Água. **Concilium**, v. 23, n.7, p. 156-172, 2, p. 279-290, mai. 2023. DOI: 10.53660/CLM-1192-23E09.

SILVA, A. H. **Relação entre parâmetros de qualidade da água e a concentração de cloro residual livre em um setor de abastecimento de água em Ponta Grossa - PR**. Dissertação (Mestrado em Engenharia Sanitária e Ambiental) – Programa de Pós-Graduação em Engenharia Sanitária e Ambiental, Universidade Estadual de Ponta Grossa, Ponta Grossa, 2020.

SILVA, A. S. C.; MORAES, L. R. S. Direito ao saneamento básico em localidades rurais: estudo no oeste do estado da Bahia. **Revista Eletrônica de Gestão e Tecnologias Ambientais**, v. 10, n. 1, p. 5-23, 2022. DOI: 10.9771/gesta.v0i1.47769.

SILVA, G. A. B. da; MEIRA, C. M. B. S.; SANTANA, C. F. D. de; COURA, M. de A.; OLIVEIRA, R. de; NASCIMENTO, R. S. do; SANTOS, W. B. Simulação do decaimento de cloro residual livre em reservatórios de distribuição de água. **Revista DAE**, v. 67, n. 218, p. 92-103, jul./set., 2019. DOI: 10.4322/dae.2019.036.

SILVA, G. L.; AURELIANO, J. T.; LUCENA, S. V. O. Proposição de um índice de qualidade de água bruta para abastecimento público. **REGA**, v. 9, n. 1, p. 17-24, jan./jun. 2012. Disponível em: <https://abrh.s3.sa-east-1.amazonaws.com/Sumarios/79/331e4cbc56baf4d2af833a342d58cee5_c610687167375878c127dcecb49b5bc2.pdf>. Acesso em: 07 jan. 2024.

SINZ, F. H.; PITKOW, X.; REIMER, J.; BETHGE, M.; TOLIAS, A. S. Engineering a Less Artificial Intelligence. **Neuron**, v. 103, n. 6, p. 967-979, set. 2019. DOI: 10.1016/j.neuron.2019.08.034.

SISAGUA. Sistema de Informação de Vigilância da Qualidade da Água para Consumo Humano. **Dados de vigilância de qualidade da água para consumo humano: banco de dados**. Disponível em: <<https://dados.gov.br/dados/busca?termo=sisagua>>. Acesso em: 16 set. 2023.

SISAR. Sistema Integrado de Saneamento Rural. **Treinamento de Operadores 2018: Sisar BBJ – Bacia do Baixo e Médio Jaguaribe**. Russas: SISAR, 2018.

SNIS. Sistema Nacional de Informações sobre Saneamento. **SNIS – Série Histórica**. 2022. Disponível em: <<http://app4.mdr.gov.br/serieHistorica/>>. Acesso em: 02 ago. 2023.

SOARES, A. F. S.; FRANCO, R.; ASSIS, J. M. G. Análise preliminar da nova portaria de potabilidade da água (PRT GM/MS nº 888/2021). In: Congresso Brasileiro de Gestão Ambiental, 12., 2021, Salvador. **Anais eletrônicos** [...] Salvador: IBEAS, 2021. p. 1-10. Disponível em: <<https://www.ibeas.org.br/congresso/Trabalhos2021/V-002.pdf>>. Acesso em: 08 out. 2022.

SOARES, S. S.; ARRUDA, P. N.; LOBÓN, G. S.; SCALIZE, P. S. L. Avaliação de métodos para determinação de cloro residual livre em águas de abastecimento público. **Semina: Ciências Exatas e Tecnológicas**, v. 37, n. 1, p. 119-130, jan./jun. 2016. DOI: 10.5433/1679-0375.2016v37n1p119.

SOUSA, A. C. A. O que esperar do novo marco do saneamento? **Perspectivas**, v. 36, n. 12, p. 1-4, dez. 2020. DOI: 10.1590/0102-311X00224020.

SOYUPAK, S.; KILIC, H.; KARADIREK, I. E.; MUHAMMETOGLU, H. On the usage of artificial neural networks in chlorine control applications for water distribution networks with high quality water. **Journal of Water Supply: Research and Technology-Aqua**, v. 60, n. 1, p. 51-60, fev. 2011. DOI: 10.2166/aqua.2011.086.

STEFÁN, S.; ERDÉLYI, N.; IZSÁK, B.; ZÁRAY, G.; VARGHA, M. Formation of chlorination by-products in drinking water treatment plants using breakpoint chlorination. **Microchemical Journal**, v. 149, p. 104-108, set. 2019. DOI: 10.1016/j.microc.2019.104008.

SUNDAS, K.; CÂNDIDO, M.; GHAFAR, M. S.; RASHEED, S.; WASEEM, A.; IQBAL, M. M.; NIAZI, M. B. K.; ZAFAR, M. I. Occurrence, influencing factors, toxicity, regulations, and abatement approaches for disinfection by-products in chlorinated drinking water: A comprehensive review. **Environmental Pollution**, v. 281, p. 1-15, jul. 2021. DOI: 10.1016/j.envpol.2021.116950.

TABARI, H.; TALAEE, P. H. Reconstruction of river water quality missing data using artificial neural networks. **Water Quality Research Journal**, v. 50, n. 4, p. 326-335, nov. 2015. DOI: 10.2166/wqrjc.2015.044.

TINELLI, S.; JURAN, I. Artificial intelligence-based monitoring system of water quality parameters for early detection of non-specific bio-contamination in water distribution systems. **Water Supply**, v. 19, n. 6, p. 1785-1792, abr. 2019. DOI: 10.2166/ws.2019.057.

TIRUNEH, A. T.; DEBESSAI, T. Y.; BWEMBYA, G. C.; NKAMBULE, S. J.; ZWANE, L. Variable Chlorine Decay Rate Modeling of the Matsapha Town Water Network Using EPANET Program. **Journal of Water Resource and Protection**, v. 11, n. 1, p. 37-52, 2019. DOI: 10.4236/jwarp.2019.111003.

TONEV, R.; DIMOVA, G. Investigation of chlorine wall decay in an old, decommissioned metallic pipe using a pipe section reactor. **Water Supply**, v. 20, n. 3, p. 963-962, mai. 2020. DOI: 10.2166/ws.2020.017.

VAN DER MAATEN, L.; HINTON, G. Visualizing Data using t-SNE. **Journal of Machine Learning Research**, v. 9, n. 86, p. 2579--2605, 2008. Disponível em: <<http://jmlr.org/papers/v9/vandermaaten08a.html>>. Acesso em: 16 ago. 2023.

VARGAS, T. F. **Qualidade das águas subterrâneas de soluções alternativas individuais da cidade de Porto Velho (RO): estudo cinético do cloro e no desenvolvimento de biofilmes**. Tese (Doutorado em Biodiversidade e Biotecnologia) – Programa de Pós-Graduação em

Biodiversidade e Biotecnologia da Rede Bionorte, Fundação Universidade Federal de Rondônia, Porto Velho, 2019.

VENTURA, K. S.; SARDINHA A. S.; REZENDE, D. Plano de segurança em sistemas de abastecimento de água: limitações e tendências no meio urbano. **International Workshop for Innovation in Safe Drinking Water**, n. 1, p. 45-50, 2022. DOI: 10.20396/iwisdw.n1.2022.4797.

WADKAR, D.; KOTE, A. Prediction of residual chlorine in a water treatment plant using generalized regression neural network. **International Journal of Civil Engineering and Technology**, v. 8, n. 8, p. 1264-1270, ago. 2017. Disponível em: <https://www.researchgate.net/publication/319929916_Prediction_of_residual_chlorine_in_a_water_treatment_plant_using_generalized_regression_neural_network>. Acesso em: 23 jan. 2023.

WANG, A. Q.; LIN, Y. L.; XU, B.; HU, C. Y.; GAO, Z. C.; LIU, Z.; CAO, T. C.; GAO, N. Y. Factors affecting the water odor caused by chloramines during drinking water disinfection. **Science of The Total Environment**, v. 639, n. 15, p. 687-694, out. 2018. DOI: 10.1016/j.scitotenv.2018.05.188.

WANG, Y. H.; WU, Y. H.; DU, Y.; LI, Q.; CONG, Y.; HUO, Z. Y.; CHEN, Z.; YANK, H. W.; LIU, S. M.; HU, H. Y. Quantifying chlorine-reactive substances to establish a chlorine decay model of reclaimed water using chemical chlorine demands. **Chemical Engineering Journal**, v. 356, p. 791-798, jan. 2019. DOI: 10.1016/j.cej.2018.09.091.

WU, W.; DANDY, G. C.; MAIER, H. R. Application of artificial neural networks to forecasting water quality in a chloraminated water distribution system. **Water Quality Research Journal**, v. 50, n. 4, p. 326-335, nov. 2015. DOI: 10.2166/wqrjc.2015.044.

WU, W.; DANDY, G. C.; MAIER, H. R. Application of artificial neural networks to forecasting water quality in a chloraminated water distribution system. In: International Congress on Modelling and Simulation, 19., 2011, Perth, Australia. **Anais eletrônicos [...]** Perth: MSSANZ, 2011. p. 1112-1118. Disponível em: <<https://www.mssanz.org.au/modsim2011/C1/wu.pdf>>. Acesso em: 23 jan. 2023.

XIAO, C.; HONG, S.; HUANG, W. Optimizing graph layout by t-SNE perplexity estimation. **International Journal of Data Science and Analytics**, v. 15, p. 159-171, 2023. DOI: 10.1007/s41060-022-00348-7.

ZHANG, Q.; ZHENG, F.; KAPELAN, Z.; SAVIC, D.; ELE, G.; YIYI, M. Assessing the global resilience of water quality sensor placement strategies within water distribution systems. **Water Research**, v. 172, p. 1-14, abr. 2020. DOI: 10.1016/j.watres.2020.115527.

ZHANG, X.; CHEN, Z.; SHEN, J.; ZHAO, S.; KANG, J.; CHU, W.; ZHOU, Y.; WANG, B. Formation and interdependence of disinfection byproducts during chlorination of natural organic matter in a conventional drinking water treatment plant. **Chemosphere**, v. 242, p. 127-227, mar. 2020. DOI: 10.1016/j.chemosphere.2019.125227.

ZHAO, Y.; YANG, Y. J.; SHAO, Y.; NEAL, J.; ZHANG, T. The dependence of chlorine decay and DBP formation kinetics on pipe flow properties in drinking water distribution. **Water Research**, v. 141, p. 32-45, set. 2018. DOI: 10.1016/j.watres.2018.04.048.

ZHOU, J.; CHU, F.; LI, X.; MA, H.; XIAO, F.; SUN, L. Water Quality Prediction Approach Based on t-SNE and SA-BiLSTM. In: 22th International Conference on High Performance Computing and Communications (HPCC), 18th International Conference on Smart City and 6th International Conference on Data Science and Systems, HPCC-SmartCity-DSS, 2020,

Cuvu, Fiji. **Anais eletrônicos** [...] Cuvu: IEEE, 2020. DOI: 10.1109/HPCC-SmartCity-DSS50907.2020.00092.

APÊNDICES

APÊNDICE A: Pesos e vieses para a rede neural artificial de melhor desempenho treinada com o erro médio quadrático (EMQ).

Camada de entrada (5 variáveis) → Camada oculta 1 (30 neurônios)	
Pesos	<p>1: -0,49197796; 0,20870557; 0,3888996; 0,31224418; -0,4538584; 0,2568301; 0,24529773; 0,33705136; -0,10309071; -0,45902938; 0,4165272; 0,45955768; -0,36802617; -0,2506269; -0,3882372; 0,44727257; -0,276222; 0,27997; -0,0744359; 0,4820206; 0,08677031; -0,33380765; -0,35608515; -0,20813228; -0,18990701; -0,26757878; -0,9574824; -0,01837804; -0,40182894; 0,15972435</p> <p>2: 0,02108528; 0,11499155; -0,34791398; -0,2741375; -0,2763209; 0,36040953; 0,09532482; 0,36290255; -0,09816024; -0,09600132; -0,2351324; -0,5322064; 0,02542178; -0,8140465; -0,13461645; -0,06746987; -0,6003504; 0,07574358; 1,4932109; -0,15294574; -0,6463322; 0,11858758; 0,4026383; -0,2769794; 0,3590386; 0,52632177; -0,10235415; -0,2910703; 0,04665104; -0,11662529</p> <p>3: -0,17391837; -0,20399308; 0,12444665; -0,15962031; -0,22661196; 0,1653212; -0,29428682; -0,10670307; -0,5237657; -0,01020324; 0,05272544; 0,16630983; 0,14601225; 0,6100164; 0,19122806; -0,01238199; -0,31981504; -0,83754915; -0,09382813; 0,15257922; -0,78247255; -0,0515027; 0,15069169; -0,37953097; 1,0669314; 0,28289482; -0,04478293; -0,12079348; 0,19425486; 0,5312512</p> <p>4: -0,20116454; -0,37310675; 0,2749119; -1,2976967; 0,19318086; 0,17282183; -0,53820145; -0,63168603; 0,5610563; -0,10991406; 0,05436293; -1,0653677; 0,94608605; -0,49802482; -0,3869254; 0,09120293; 0,1390135; 0,40279138; -0,07202432; -0,11671309; -0,8262036; 0,04278668; 0,6645242; -0,23502286; -0,25215507; 0,8242849; 0,13260725; -0,11590751; -0,03961056; 0,39349127</p> <p>5: -0,29995507; -0,16429526; 0,2382829; 0,47277066; -0,13579679; -0,18122888; -0,38632584; -0,41342962; 0,06196128; -0,10781801; -0,8009885; 0,08554607; 0,03700253; 0,13049515; -0,32278123; -0,12271334; 0,2422245; -0,42250085; -0,18741271; 0,32944053; 0,45337677; -0,07482678; -0,14734145; 0,6143659; 0,04418358; 0,54287934; -0,20616455; -0,3481602; 0,33997333; 0,9099608</p>
Vieses	<p>0,56925184; -0,34252506; -0,32820788; 0,18849461; 0,21599443; 0,04569776; -0,23665535; -0,5682958; 0,30624402; -0,02129236; 0,24678104; 0,05843692; 0,59991664; -0,05281118; 0,22094378; -0,04384895; 0,36803988; -0,42166036; 0,27372625; -0,20078006; 0,2672819; 0,0481835; 0,45280212; 0,14556691; -0,49909097; -0,11744541; -0,78659904; 0,02550635; -0,00410832; 0,11629186</p>
Camada oculta 1 (30 neurônios) → Camada oculta 2 (8 neurônios)	
Pesos	<p>N1₁: -0,17604618; 0,413923; 0,03740185; -0,14770955; -0,6239395; -0,19710977; 0,5328317; -0,5732666</p> <p>N2₁: -0,01848093; 0,12546293; -0,2151085; -0,05657623; 0,4067339; 0,4240317; -0,57953763; 0,36402723</p> <p>N3₁: -0,5272637; -0,0966213; -0,2405272; -0,09185582; 0,10778435; 0,5017036; -0,63459504; -0,39138922</p> <p>N4₁: -0,24772249; -0,20087183; 0,39737573; 0,04273885; 0,22114426; -0,33810318; 0,34544536; -0,3282591</p> <p>N5₁: 0,40283394; 0,16451517; -0,26987702; 0,14080796; 0,26146832; 0,2495621; -0,19378148; 0,3940032</p> <p>N6₁: 0,41314453; 0,3034978; 0,35328606; -0,43101215; -0,07191415; 0,456406; -0,02000049; -0,0911938</p> <p>N7₁: 0,12817389; 0,21280943; -0,15689293; 0,13885188; 0,14761537; 0,491301; -0,6292173; 0,2608221</p> <p>N8₁: 0,8192169; -0,3633481; -0,28211573; -0,05373108; 0,1713437; -0,00645556; -0,70809525; 0,149567</p> <p>N9₁: 0,09357604; -0,07695668; 0,241495; -0,02562534; -0,31066456; -0,09648203; 0,5108953; 0,10888083</p> <p>N10₁: -0,42857134; -0,36306748; 0,26725376; 0,10088196; -0,12354222; -0,45324412; -0,25287247; -0,09625369</p> <p>N11₁: -0,2203911; 0,33067915; 0,09839595; 0,15753569; 0,30642965; 0,00130262; 0,36351117; -0,5890008</p>

Camada oclusa 1 (30 neurônios) → Camada oclusa 2 (8 neurônios)	
Pesos	N12₁ : -0,07659561; 0,21512946; -0,65199155; 0,46395925; 0,06262057; 0,10755087; -0,02636802; -0,24917616 N13₁ : -0,09661005; 0,33619604; 0,6861534; 0,29680482; 0,2161639; -0,5293045; 0,03789771; -0,4519908 N14₁ : 0,14365868; -0,265502; -0,43455455; 0,7516261; -0,16372897; 0,17425698; -0,44318995; -0,30150557 N15₁ : -0,36984947; -0,27622548; 0,37232637; 0,10080165; -0,04615113; -0,3932268; 0,26026347; -0,2943537 N16₁ : 0,05230672; 0,49975386; 0,08288816; 0,4533154; 0,30115122; 0,5799825; -0,08260377; 0,25463334 N17₁ : 0,00695622; -0,13290806; 0,11118538; 0,5471141; -0,335531; -0,2679305; -0,04747767; -0,10714545 N18₁ : 0,28953013; -0,31743884; -0,00436917; -0,40513715; 0,20279507; 0,63085204; 0,66723007; 0,36534163 N19₁ : -0,5471393; 0,01172948; -0,04851305; -0,53072214; -0,610937; -0,2344969; 0,3948894; -0,5624971 N20₁ : 0,1118257; 0,3414547; -0,52649295; -0,21137951; 0,381804; -0,16192521; -0,03685857; 0,62791324 N21₁ : -0,07565012; 0,12542826; 0,04743688; 0,14520733; 0,15317841; 0,39344382; -0,4615913; -0,5690035 N22₁ : -0,06560064; -0,32847092; -0,09081153; 0,19810885; -0,2979693; -0,09251747; -0,16693893; -0,35799566 N23₁ : 0,2997735; 0,1760881; 0,6391011; 0,16324095; -0,3025279; 0,07142127; -0,02763123; -0,06370275 N24₁ : -0,19234495; 0,25530288; 0,18886682; 0,14223476; 0,37092555; -0,08820906; 0,24249986; -0,18041565 N25₁ : 0,52433103; -0,17881358; -0,03256804; 0,5118921; 0,6673045; -0,08100536; -0,00856793; -0,16135333 N26₁ : -0,01572615; -0,17527282; 0,14316022; 0,5391296; 0,35330912; -0,00884664; 0,20975342; -0,1837688 N27₁ : 0,0010819; -0,44030932; 0,14607558; -0,3255995; 0,10873906; 0,1803098; -0,42205313; 0,4198752 N28₁ : -0,04462598; 0,01208768; 0,09560836; -0,16220796; -0,32392123; -0,42603627; -0,16557027; 0,09140411 N29₁ : -0,03560105 -0,44163156; 0,25253016 -0,06124268 -0,19927539 -0,17597781; -0,0521454 0,42754325 N30₁ : -0,6327723 0,15012187 0,23970975 -0,41051212 -0,21052094 -0,0178667; -0,5975408 -0,19782671
Vieses	-0,04853654; 0,31439948; 0,10738815; 0,14344437; -0,0741998; 0,00501257; 0,23889521; -0,48000103
Camada oclusa 2 (8 neurônios) → Camada oclusa 3 (4 neurônios)	
Pesos	N1₂ : -0,34292555; -0,17336638; 1,0682638; -0,05930331 N2₂ : 0,22938383; 0,37851956; 0,62419516; 0,41139975 N3₂ : 0,14978717; -0,5900489; 0,6142566; 0,08063142 N4₂ : -0,1501803; 0,2246914; -0,74185294; 0,7923082 N5₂ : -0,85489607; 0,27809432; 0,517029; -0,44624317 N6₂ : -0,42157796; 0,19971181; -0,48122895; 0,35391596 N7₂ : 0,28646556; 0,18712766; 0,09703597; 0,7607676 N8₂ : -0,6229776; -0,752471; 0,9268338; -0,21275735
Vieses	-0,09944738 0,01233378 0,23717654 0,18118995
Camada oclusa 3 (4 neurônios) → Camada de saída (1 neurônio)	
Pesos	N1₃ : -0,6949428 / N2₃ : 1,1138028 / N3₃ : 0,62426865 / N4₃ : 1,1676431
Vieses	0,07074422

Onde:

1	CRL no reservatório	2	Turbidez no reservatório	3	Temperatura do dia	4	pH no reservatório	5	Comprimento da tubulação
----------	---------------------	----------	--------------------------	----------	--------------------	----------	--------------------	----------	--------------------------

N_{xy} → N é o neurônio; x é o número do neurônio e y é o número da camada onde o neurônio se encontra

APÊNDICE B: Pesos e vieses para a rede neural artificial de melhor desempenho treinada com o erro médio quadrático ponderado (EMQP).

Camada de entrada (5 variáveis) → Camada oculta 1 (30 neurônios)	
Pesos	1: -0,63013875; -0,23646994; 0,25896636; 0,07625376; -0,04278184; -0,0329534; -0,09626534; -0,09153128; 0,08341369; 0,07417943; -0,20882247; 0,53120553; -0,15348269; 0,45187658; 0,11850304; 0,69975793; -0,20371608; -0,60716647; -0,3032769; 0,42856628; 0,22451825; -0,37225837; -0,11196519; -0,83608824; 0,32720992; 0,00736395; -0,18559156; -0,65525186; 0,02501607; 1,0171556
	2: -0,49450517; -0,07515923; 0,44864616; -0,2014114; 0,9931524; -0,4894108; -0,23707378; -0,21804339; 0,0732648; 0,10994989; 2,2613058; 0,02539463; 0,23824684; -0,3031104; 0,00883323; 0,08224653; 0,17377989; 0,79792964; 0,63342804; 0,10627123; 0,03894636; 0,05018288; -0,18734844; 0,09441938; 0,54399884; 0,0163841; 0,04374334; 0,13144575; 0,22772239; -0,7349084
	3: 0,99591273; -0,03619906; 0,48800167; -0,96851915; -0,55716926; 0,01987914; 0,48207954; -0,8452762; -0,10872447; -0,15249966; -0,02127318; 0,21890996; -1,0362331; -0,26441073; 0,7028071; 0,01763098; 0,07535988; 0,76219106; 0,04046876; -0,38285396; 0,23490421; -0,05610623; -0,36612895; -0,21123196; -0,42143217; -0,4901869; 0,30536288; 0,00421031; -0,0341978; 0,6853168
	4: 0,96315646; 0,36986792; -0,5891819; -0,0641882; -0,15958767; -0,14123863; 0,11556572; -1,1233903; -0,09605028; -1,0221823; -0,08417952; 0,25984272; 0,7245327; 0,47827378; -0,22316948; 0,37355697; -1,0665632; -0,456253; 0,11831778; 0,0270247; 0,17829916; 0,2904604; -0,07341804; 0,46465117; -0,5672474; -0,1520322; 0,11952472; -0,01077537; -0,25050348; -0,12300517
	5: -1,1520854; -0,28150895; -1,3217276; -1,6770692; 1,2193937; -0,3546027; 0,5123468; 0,46757635; -0,215398; -0,52752537; 0,3164615; 0,04690802; 0,51324207; -0,58327; 0,20706728; 0,8345443; 0,54582846; 0,5627738; 0,9442214; -0,16622953; -0,13186528; -0,11354954; -0,18384787; 0,0521876; -0,31108692; -0,33138195; 0,07423371; 0,247834; 0,25279695; -0,5704607
Vieses	-0,45967132; 0,33734718; -0,016821967; 0,34564024; -0,22777414; -0,00064662547; 0,59573442; -0,27214381; -0,050117552; 0,015654616; 0,78650016; 0,015386452; 0,59116435; -0,26459095; -0,69599187; -0,24345271; -0,16479258; 0,39144534; 0,12608662; -0,68161666; -0,52155024; 0,24302390; 0,61745828; -0,054681137; -0,20661029; 0,075444616; 0,21082506; 0,16655390; 0,001430190; 0,090334579
Camada oculta 1 (30 neurônios) → Camada oculta 2 (8 neurônios)	
Pesos	N1₁: -0,01520352; 0,05048892; 0,12251062; 0,22000133; 0,3729557; -0,2446102; 0,33264762; 0,754601
	N2₁: -0,28878343; -0,3831144; 0,1043498; -0,90380293; 0,11082426; -0,04650551; -0,44139475; 0,22897382
	N3₁: -0,5756407; -0,00399856; 0,8246563; 0,7568762; 0,06856316; 0,45155543; 0,5410698; 0,61928415
	N4₁: 0,3285539; -0,24660222; 0,42189798; -0,10041634; 0,34723923; -0,10154448; -0,4763011; 0,6912137
	N5₁: 0,07607924; -0,11369332; -0,36639977; 0,38903317; 0,29171574; -0,41393313; -0,51077956; 0,26949143
	N6₁: 0,22992341; -0,3760161; 0,23294711; -0,3041461; 0,20153752; 0,09970588; 0,2650502; 0,04419677
	N7₁: -0,3398945; 0,07331103; -0,2054637; -0,00274185; -0,52063936; 0,01624546; -0,17759004; -0,34230664
	N8₁: 0,08634069; 0,16956459; -0,5690082; 0,8816174; 0,23277187; 0,53507054; -0,13834895; -0,11325219
	N9₁: 0,36826676; 0,00749776; 0,39772624; -0,1972552; -0,524543; 0,0126652; 0,08799327; 0,0015266
	N10₁: 0,23614001; 0,16529368; 0,33260244; 0,16239505; -0,11581685; 0,1130193; -0,40154344; 0,66711205
	N11₁: -0,30552498; 0,1904482; 0,37636694; -0,4419589; 0,11065645; 1,0150734; -0,99236184; 0,73901737
	N12₁: 0,2556975; -0,4063274; -0,28454733; -0,11794153; -0,6230733; -0,20389758; -0,11211902; 0,2014087

Camada oculta 1 (30 neurônios) → Camada oculta 2 (8 neurônios)	
Pesos	N13₁ : 0,4491738; -0,68165725; -0,3219088; -0,34473; -0,08785754; 0,2509251; -0,84269655; -0,318755 N14₁ : 0,07037049; 0,50476533; 0,8403708; 0,13534361; -0,47904533; -0,49438438; 0,01889; 0,03499744 N15₁ : 0,50167364; 0,2956463; -0,1411221; 0,13559991; -0,15119058; -0,4804073; 0,7432106; 0,4368107 N16₁ : 0,3995931; -0,64406943; -0,3316814; -0,21562234; 0,05117732; -0,74670917; 0,20675741; 0,5294183 N17₁ : 0,13109319; 0,18032281; -0,16064101; 0,38928768; 0,24967077; -0,12016513; -0,39093664; 0,2504215 N18₁ : 0,36994955; -0,12074632; -0,9944436; -0,6588423; -0,0338318; 0,4077281; -0,18260488; 0,5348489 N19₁ : 0,2903661; -0,1850485; -1,4821935; 0,06776235; -0,22946164; 0,33066687; 0,04889598; -0,1547808 N20₁ : 0,08721076; 0,09256309; 0,7051303; 0,93891394; 0,21001521; -0,50326365; 0,56451005; -0,03733835 N21₁ : -0,25572824; 0,20321402; 0,09670378; 0,32381693; 0,19251622; -0,3390157; 0,42878473; 0,3475096 N22₁ : -0,3378225; -0,14531684; 0,14813416; -0,63056564; 0,19504939; 0,12081739; -0,07420623; 0,16742706 N23₁ : -0,08049673; -0,26803783; 0,17506139; -0,34527144; -0,3488635; 0,37700334; -0,6800617; -0,31855366 N24₁ : 0,00257851; 0,08466291; 0,49455285; -0,27714002; 0,5509662; 0,44984525; 0,5248525; 0,271632 N25₁ : 0,24000365; 0,34575006; -0,0621122; -0,01147007; -0,17557335; 0,45312038; 0,5427089; -0,4203097 N26₁ : 0,07010657; -0,01367597; 0,813; 0,28817698; -0,21223368; -0,09061216; -0,513014; -0,20805117 N27₁ : -0,39899233; -0,06935377; -0,21408941; -0,23722324; 0,43733668; 0,07414493; -0,04714029; 0,133533 N28₁ : -0,0422786; 0,37264034; -0,00816695; -0,48334345; 0,1825818; 0,45029753; 0,12480341; -0,3113078 N29₁ : -0,24454997; 0,57418865; -0,03722951; 0,3983615; -0,40996346; -0,0836992; -0,25320858; -0,04965453 N30₁ : 0,42253402; -0,15834408; -0,01889281; 0,32705554; -0,02852097; 0,51850826; -0,41422164; 0,4707511
Vieses	0,16192059; -0,13347742; -0,0709751; -0,60005677; -0,53464216; 0,11782957; -0,37632802; -0,12302422
Camada oculta 2 (8 neurônios) → Camada oculta 3 (4 neurônios)	
Pesos	N1₂ : 0,52572685; 0,6657698; -0,21782726; -0,18059917 N2₂ : -0,49601576; -0,33261275; 0,6802243; 0,15855856 N3₂ : -0,87680775; -1,0522572; -0,76856774; -0,17783812 N4₂ : -0,9194224; 0,11601332; -1,2518225; -0,4330253 N5₂ : -0,727394; 0,72153294; -0,29305753; 0,698936 N6₂ : 0,5346793; -0,7543013; 1,1607413; 0,049013 N7₂ : -1,2301952; 0,24012153; -0,7354199; 0,55643046 N8₂ : 0,6480365; 0,8200537; -0,7029679; -0,06629908
Vieses	0,32595828; -0,20248523; 0,01546451; -0,14565334
Camada oculta 3 (4 neurônios) → Camada de saída (1 neurônio)	
Pesos	N1₃ : 1,0062989 / N2₃ : 0,6961042 / N3₃ : -0,5651 / N4₃ : -1,1372969
Vieses	0,24492513

Onde:

1	CRL no reservatório	2	Turbidez no reservatório	3	Temperatura do dia	4	pH no reservatório	5	Comprimento da tubulação
----------	---------------------	----------	--------------------------	----------	--------------------	----------	--------------------	----------	--------------------------

N_{xy} → N é o neurônio; x é o número do neurônio e y é o número da camada onde o neurônio se encontra