



## II-743 - USO DE REDES NEURAS ARTIFICIAIS APLICADA AOS DADOS ANALÍTICOS DO TRATAMENTO DE ESGOTO VISANDO PREVISIBILIDADE

### Allan dos Anjos Pestana<sup>(1)</sup>

Técnico em automação industrial (IFSP). Tecnólogo em automação industrial (FATEC) e Tecnólogo em gestão ambiental (SENAC). Técnico em Sistemas de Saneamento na SABESP-SP. Pós-graduado em indústria 4.0 (SENAI), MBA em gestão de projetos e processos organizacionais (Centro Paula Souza), além de especialista em automação para processos de saneamento (USP). Mestrando profissional em gestão e tecnologia em sistemas produtivos (Centro Paula Souza), além de aluno do MBA em data science e analytics (USP/ESALQ) e Engenharia elétrica (IFSP).

### Rodrigo Alves<sup>(2)</sup>

Pós-graduado em Química Ambiental e Engenharia de Controle de Poluição (2019) pelas Faculdades Oswaldo Cruz. Graduado em Engenharia Química (2014) nas Faculdades Oswaldo Cruz. Graduando em MBA Saneamento Ambiental (FESPSP). Atualmente atua como Gerente da Operação do Sistema ABC na Superintendência de tratamento de esgotos da região metropolitana de São Paulo – SABESP.

### Márcia Ito<sup>(3)</sup>

Médica pela Escola Paulista de Medicina da Universidade Federal de São Paulo. Tecnóloga em Processamento de Dados pela Faculdade de Tecnologia de São Paulo. Mestre e Doutora em Engenharia Elétrica pela Escola Politécnica da Universidade de São Paulo.

### Diogo Alfieri Palma<sup>(4)</sup>

Engenheiro de Produção pela Pontifícia Universidade Católica do Paraná. Mestre em Engenharia de Produção e de Manufatura pela Universidade Estadual de Campinas (UNICAMP). Doutorando em Tecnologia pela UNICAMP.

**Endereço<sup>(1)</sup>:** Rua João Lopes Maciel, 465 – Cidade Nitroquímica – São Paulo - SP - CEP: 08090-040 - Brasil - Tel: +55 (11) 97227-8628 - e-mail: [apestana@sabesp.com.br](mailto:apestana@sabesp.com.br)

**Endereço<sup>(2)</sup>:** Rua Antônio La Giudice, 972 – Jardim Aricanduva – São Paulo - SP - CEP: 03454-000 - Brasil - Tel: +55 (11) 99570-9013 - e-mail: [rodrigoalves@sabesp.com.br](mailto:rodrigoalves@sabesp.com.br)

**Endereço<sup>(3)</sup>:** Rua Bandeirantes, 169 – Bom Retiro – São Paulo – SP – CEP: 04304-010 – Brasil – Tel: +55 (11) 3327-3104 – e-mail: [marcia.ito@cpspos.sp.gov.br](mailto:marcia.ito@cpspos.sp.gov.br)

**Endereço<sup>(4)</sup>:** Rua Paschoal Marmo, 1888 - Jardim Nova Italia, Limeira – SP – CEP: 13484-332 – Brasil - Tel.: +55 (19) 2113-3332 - e-mail: [diogoapalma@gmail.com](mailto:diogoapalma@gmail.com)

## RESUMO

O tratamento de esgoto no Brasil enfrenta desafios diante do novo marco legal do saneamento e da necessidade de lidar com vastos conjuntos de dados. Estações de Tratamento de Esgoto (ETEs) tradicionais, com foco no princípio de lodo ativado, possuem um processo que acelera a decomposição da matéria orgânica, e a Demanda Química de Oxigênio (DQO) é uma análise importante que orienta o tratamento, podendo refletir a eficiência do processo. O estudo propõe a criação de um modelo preditivo para aprimorar o tratamento de esgoto, combinando dados analíticos e técnicas de aprendizado de máquina. Métricas como a Raiz do Erro Quadrático Médio, do inglês *Root Mean Square Error* (RMSE), o Erro Percentual Absoluto Médio, do inglês *Mean Absolute Percentage Error* (MAPE) e coeficiente de determinação ( $R^2$ ) avaliam a qualidade do modelo. Foi realizada a modelagem para DQO no efluente de três ETEs utilizando RNA. Posterior a isso, também foi modelada a amônia afluente em uma dessas Estações com base em parâmetros analíticos das elevatórias a montante, podendo trazer predição dessa análise em até quarenta minutos, com um Erro Médio Absoluto, do inglês *Mean Absolute Error* (MAE), de 2,46; e o  $R^2$  de 0,84. O projeto visa não apenas otimizar processos



tradicionais, mas também introduzir abordagens avançadas, proporcionando eficiência e predição ao setor de saneamento, culminando na aplicação prática em uma ETE na região metropolitana de São Paulo.

**PALAVRAS-CHAVE:** Modelagem; Demanda Química de Oxigênio; Demanda Bioquímica de Oxigênio; Amônia; Saneamento.

## INTRODUÇÃO

O tratamento de esgoto brasileiro tem grande desafio pela frente devido o novo marco legal do saneamento e a quantidade de dados provenientes de seus processos, essa situação representa oportunidade visando eficiência operacional e por consequência maior volume e qualidade do esgoto tratado.

Uma técnica frequentemente empregada para o tratamento de águas residuais é o princípio de lodo ativado. Nesse processo, é introduzido oxigênio no esgoto para permitir que as bactérias existentes decomponham a matéria orgânica presente. Com o decorrer do tempo, os procedimentos de tratamento passaram a requerer tecnologias mais aprimoradas que minimizem os impactos ambientais e os custos ao máximo (Peres et al., 2021).

O procedimento de tratamento mediante lodo ativado essencialmente acelera o processo natural de decomposição da matéria orgânica ao oxigenar as bactérias nas águas residuais. Essa substância orgânica é composta por proteínas (40-60%), carboidratos (25-50%) e lipídios (10%) e serve como fonte de energia para esses microrganismos (Ferreira e Coraiola, 2008). Destaque para a separação das etapas de tratamento (líquida e sólida), essa primeira onde ocorre a depuração da matéria orgânica nos tanques de aeração.

O processo de tratamento, inclui a fase inicial em que as grades removem o material particulado, os tanques de aeração, onde ocorre a depuração da matéria orgânica e a desidratação mecânica do lodo, a fase final do tratamento e envolve a separação do lodo prensado em líquido e sólido.

No tratamento de águas residuais, é comum utilizar várias análises para orientar os processos em cada etapa. Uma das mais importantes é a Demanda Química de Oxigênio (DQO). A DQO é a concentração em massa de oxigênio que é igual à quantidade de um certo reagente químico consumido em seu ensaio, e isso pode estar ligado à carga orgânica presente nas águas residuais (Geerdink, 2017). Em outras palavras, a relação entre DQO de entrada e saída pode fornecer informações sobre a eficácia da ETE, daí a frequente utilização desse ensaio.

Diversas abordagens podem ser utilizadas visando obter direcionamentos através de dados no tratamento de esgoto. Dentre elas é possível citar as Redes Neurais Artificiais (RNA) (Padhma, 2023), Algoritmos Genéticos (AG) (Arismendy, 2021), entre outras. As RNAs podem ser definidas como algoritmos computacionais com capacidade de aproximar o comportamento de determinado processo com precisão através da simulação (Arismendy et al., 2021).

Na arquitetura da rede, os neurônios da camada de entrada são as variáveis explicativas do processo e cada um desses são conectados a cada um dos neurônios da camada oculta, que por fim são conectados ao neurônio da camada de saída, que é a variável dependente. Através de algumas métricas e por meio do algoritmo de retro propagação de Levenberg-Marquardt é possível diminuir a função perda a medida que a rede passa pelos ciclos ou quantidade de épocas definidos (Li e Vanrolleghem, 2022).

Uma vertente das RNAs é a Long Short-Term Memory (LSTM) que surgiu como solução para os problemas com o gradiente das RNAs tradicionais, e funciona através da inclusão de “gates” ou chaves, afim de permitir armazenamento ou não de memórias de longo prazo no algoritmo (Van Houdt, 2020).

O objetivo geral do projeto é a concepção de um modelo preditivo que auxilie no tratamento de esgoto a partir de dados analíticos de seus processos e de técnicas de aprendizado de máquina. Para tal, fora realizada uma bibliometria acerca da aplicação dessas tecnologias no setor de saneamento, uma fundamentação teórica

relacionada aos termos mais relevantes e, finalmente, a aplicação de uma dessas modelagens em ETEs na região metropolitana de São Paulo, entre as finalidades por exemplo, prever a amônia em uma estação.

## MATERIAIS E MÉTODOS

Através da análise dos dados estudados ou treinamento do algoritmo é possível encontrar determinados padrões e, para isso, é fundamental a utilização de algumas métricas que visam identificar a acurácia do modelo utilizado, e poder comparar diferentes modelos. A eq. (1) mostra o cálculo para o Erro Médio Quadrático (do inglês, Mean Square Error) (MSE), usualmente utilizado para determinar a diferença entre o valor predito e esperado em um modelo (Padhma, 2023).

$$MSE = \sum_{i=1}^n \frac{(d-y)^2}{n} \quad (1)$$

onde,  $d$ : é o valor da variável,  $y$ : é o valor estimado pelo modelo, com  $i$  variando de 1 a  $n$ , onde  $n$  é valor de amostras da base de dados modelada.

A eq. (2), a raiz do erro quadrático médio ou Root Mean Square Error (RMSE), análogo ao que ocorre na métrica MSE, porém adicionado a raiz quadrada ao cálculo:

$$RMSE = \sqrt{\sum_{i=1}^n \frac{(d-y)^2}{n}} \quad (2)$$

onde,  $d$ : é o valor da variável,  $y$ : é o valor estimado pelo modelo, com  $i$  variando de 1 a  $n$ , onde  $n$  é valor de amostras da base de dados modelada.

Outra métrica normalmente utilizada é o erro médio absoluto ou Mean Absolute Error (MAE), eq. (3):

$$MAE = \frac{1}{n} \sum_{i=1}^n |d - y| \quad (3)$$

onde,  $d$ : é o valor da variável,  $y$ : é o valor estimado pelo modelo, com  $i$  variando de 1 a  $n$ , onde  $n$  é valor de amostras da base de dados modelada.

Novamente, da mesma forma que ocorre com as métricas MSE e RMSE, o MAPE é uma métrica muito parecida com o MAE, com o acréscimo da divisão, trazendo a porcentagem média da diferença entre o valor predito e real, eq. (4):

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{d-y}{d} \right| \times 100 \quad (4)$$

onde,  $d$ : é o valor da variável,  $y$ : é o valor estimado pelo modelo, com  $i$  variando de 1 a  $n$ , onde  $n$  é valor de amostras da base de dados modelada.

Por fim, o coeficiente de determinação ou  $R^2$ , demonstra a quão ajustada está a equação obtida ao conjunto de dados, eq. (5):

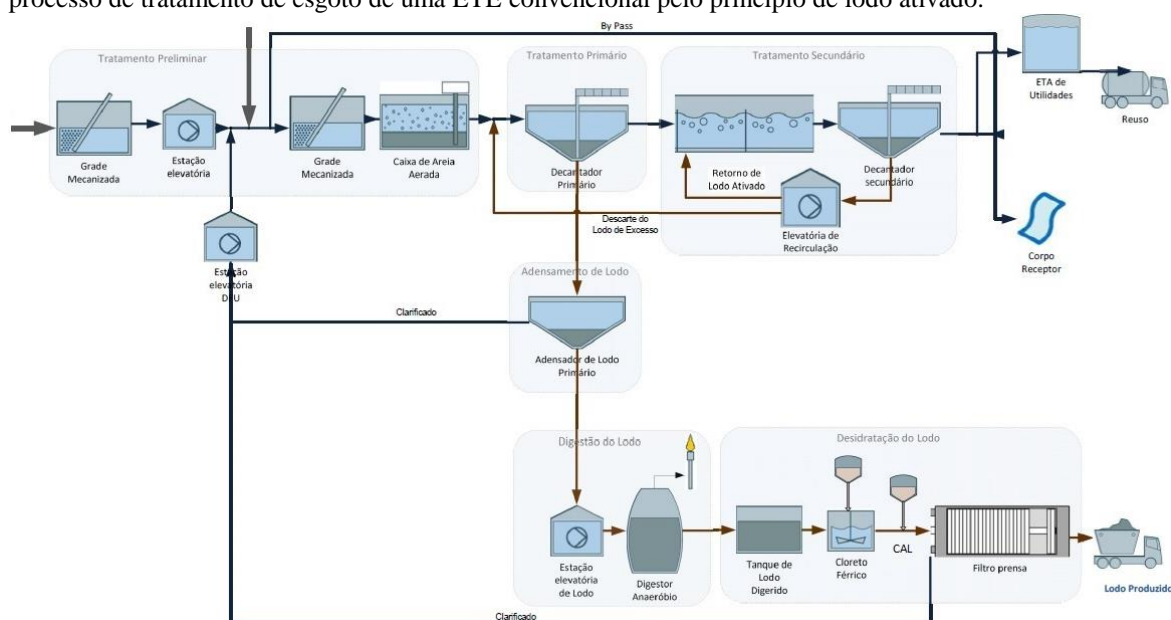


$$R^2 = 1 - \frac{\sum_{i=1}^n (d[i] - \hat{y}[i])^2}{\sum_{i=1}^n d[i] - \hat{y}[i]^2} \quad (5)$$

onde,  $d$ : é o valor da variável,  $y$ : é o valor estimado pelo modelo, com  $i$  variando de 1 a  $n$ , onde  $n$  é valor de amostras da base de dados modelada.

Enquanto nos quatro primeiros, MSE, RMSE, MAE e MAPE, esperam-se valores baixos afim de se atestar a qualidade do modelo, no  $R^2$ , que varia entre 0 e 1, valores que se aproximam de 1, aumentam o poder de predição do modelo.

Antes de explorar as análises detalhadas e a aplicação das RNAs para a predição dos resultados do tratamento de esgoto, é necessário entender as etapas desse processo. O tratamento de esgoto é um procedimento complexo que envolve etapas sequenciais, cada uma desempenhando um papel na remoção de impurezas e na produção de efluentes tratados dentro dos padrões estabelecidos. Neste contexto, na Figura 1, temos o fluxo do processo de tratamento de esgoto de uma ETE convencional pelo princípio de lodo ativado.



**Figura 1. Fluxograma do tratamento - ETE**

Fonte: Dados originais da pesquisa

Com destaque para os produtos finais da Estação nas fases líquidas e sólidas, o efluente tratado e o lodo seco que é encaminhado ao aterro respectivamente. A Tabela 1 traz os parâmetros analisados na Estação que é objeto de estudo, os locais de instalação dessas sondas, faixa de trabalho do instrumento e os identificadores que serão utilizadas para facilitar a modelagem.

**Tabela 1. Parâmetros analíticos no processo de tratamento de esgoto**

| Identificador | Parâmetro           | Local               | Faixa de Trabalho    |
|---------------|---------------------|---------------------|----------------------|
| AIT101        | pH                  | Esgoto desarenado   | 2 a 12               |
| AIT102        | Condutividade       | Esgoto desarenado   | 0-10K $\mu$ S/cm     |
| AIT103        | SST (imersão)       | Esgoto desarenado   | 100 a 500 mg/L       |
| AIT104        | DQO                 | Esgoto desarenado   | 200 a 1.000 mg/L     |
| AIT105        | SST (imersão)       | Efluente Primário   | 20 a 200 mg/L        |
| AIT106        | DQO                 | Efluente final      | 100 a 600 mg/L       |
| LIT301        | Manta de Lodo       | Decantador Prim.    | 0-6metros            |
| LIT302        | Manta de Lodo       | Decantador Prim.    | 0-6metros            |
| AIT109        | SST (imersão)       | Licor misto         | 1.000 a 8.000 mg/L   |
| AIT110        | SST (imersão)       | Lodo de Retorno     | 2.000 a 15.000 mg/L  |
| AIT111        | ORP                 | Reator aeróbio 1    | (-175 a -400 mV)     |
| AIT112        | ORP                 | Reator aeróbio 1    | (-175 a -400 mV)     |
| AIT113        | ORP                 | Reator aeróbio 1    | (-175 a -400 mV)     |
| AIT114        | ORP                 | Reator aeróbio 2    | (-175 a -400 mV)     |
| AIT115        | ORP                 | Reator aeróbio 2    | (-175 a -400 mV)     |
| AIT116        | ORP                 | Reator aeróbio 2    | (-175 a -400 mV)     |
| AIT117        | ORP                 | Reator aeróbio 3    | (-175 a -400 mV)     |
| AIT118        | ORP                 | Reator aeróbio 3    | (-175 a -400 mV)     |
| AIT119        | ORP                 | Reator aeróbio 3    | (-175 a -400 mV)     |
| AIT120        | ORP                 | Reator aeróbio 4    | (-175 a -400 mV)     |
| AIT121        | ORP                 | Reator aeróbio 4    | (-175 a -400 mV)     |
| AIT122        | ORP                 | Reator aeróbio 4    | (-175 a -400 mV)     |
| AIT123        | Oxigênio Dissolvido | Reator aeróbio 1    | 0 a 8 PPM            |
| AIT124        | Oxigênio Dissolvido | Reator aeróbio 1    | 0 a 8 PPM            |
| AIT125        | Oxigênio Dissolvido | Reator aeróbio 1    | 0 a 8 PPM            |
| AIT126        | Oxigênio Dissolvido | Reator aeróbio 2    | 0 a 8 PPM            |
| AIT127        | Oxigênio Dissolvido | Reator aeróbio 2    | 0 a 8 PPM            |
| AIT128        | Oxigênio Dissolvido | Reator aeróbio 2    | 0 a 8 PPM            |
| AIT129        | Oxigênio Dissolvido | Reator aeróbio 3    | 0 a 8 PPM            |
| AIT130        | Oxigênio Dissolvido | Reator aeróbio 3    | 0 a 8 PPM            |
| AIT131        | Oxigênio Dissolvido | Reator aeróbio 3    | 0 a 8 PPM            |
| AIT132        | Oxigênio Dissolvido | Reator aeróbio 4    | 0 a 8 PPM            |
| AIT133        | Oxigênio Dissolvido | Reator aeróbio 4    | 0 a 8 PPM            |
| AIT134        | Oxigênio Dissolvido | Reator aeróbio 4    | 0 a 8 PPM            |
| AIT135        | NH <sub>3</sub>     | Efluente Primário   | 10-150mg/L           |
| AIT136        | NH <sub>3</sub>     | Licor misto         | 0,05-20mg/L          |
| AIT137        | Turbidez            | Efluente tratado    | 0-10 NTU             |
| LIT501        | Manto de Lodo       | Decantador DS1      | 0-6metros            |
| LIT502        | Manto de Lodo       | Decantador DS2      | 0-6metros            |
| LIT503        | Manto de Lodo       | Decantador DS3      | 0-6metros            |
| LIT504        | Manto de Lodo       | Decantador DS4      | 0-6metros            |
| AIT142        | SST (imersão)       | Lodo Aflu. Digestor | 20.000 a 50.000 mg/L |
| AIT143        | SST (imersão)       | Lodo condicionado   | 20.000 a 50.000 mg/L |
| AIT144        | SST (imersão)       | Filtrado FPP 1 a 3  | 1.000 a 10.000 mg/L  |
| AIT145        | NH <sub>3</sub>     | Filtrado FPP 1 a 3  | 10-150mg/L           |
| AIT146        | Condutividade       | Filtrado FPP 1 a 3  | 0-10K $\mu$ S/cm     |
| AIT147        | Cloro (inserção)    | ETA Reuso           | 0 a 20 mg/l          |
| AIT148        | Turbidez            | ETA Reuso           | 0-10 NTU;            |

**Fonte: Dados originais da pesquisa**

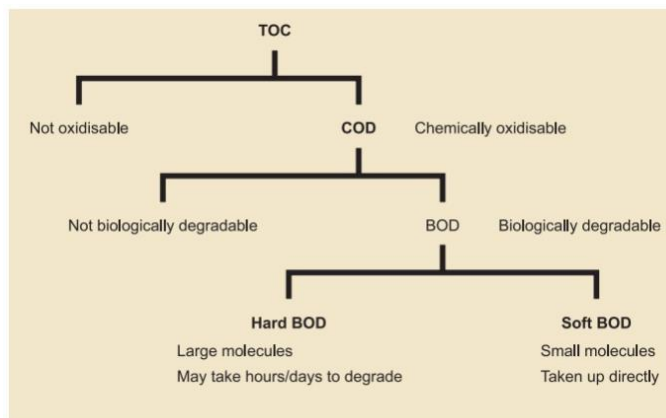
A seguir temos uma descrição dos parâmetros citados e sua relação com o processo de tratamento de esgoto:

- **Demanda Química de Oxigênio**

Para controle de processos em ETEs usualmente são utilizadas medidas para determinação da matéria orgânica existente. O Carbono Orgânico Total, do inglês Total Organic Carbon (TOC), a DQO e a Demanda



Bioquímica de Oxigênio (DBO), do inglês Biological Oxygen Demand (BOD), (Davies, 2005). A Figura 2 nos traz a relação entre esses ensaios.



**Figura 2. Relação entre métricas para determinação da matéria orgânica**

Fonte: Davies (2005).

Conforme observado, a DQO traz o dado do que é quimicamente oxidável na amostra, enquanto a DBO daquilo que é biologicamente degradável.

- **pH**

O pH é o potencial hidrogeniônico de uma solução e é definido como o logaritmo negativo da concentração de íons de hidrogênio, serve para medir o grau de acidez, neutralidade ou alcalinidade, e no tratamento de esgoto normalmente varia entre 6,5 e 7,5 (Jordão, 2011). Para determiná-lo segue eq. (6):

$$pH = \log_{10} \frac{1}{H^+} \quad (6)$$

onde:  $H^+$ : é a concentração do Hidrogênio em íons-grama por Litro de solução.

- **Condutividade**

A condutividade é uma medida utilizada para determinação de capacidade de conduzir corrente elétrica por determinada solução, eq. (7) (APHA, 2023). Normalmente, moléculas de compostos orgânicos não são bons condutores elétricos, por isso a condutividade é um parâmetro utilizado em algumas etapas do tratamento de esgoto.

$$G = \frac{1}{R} \quad (7)$$

onde, G: é a condutância, e R: é a resistência.

- **Sólidos Suspensos Totais (SST)**

Em águas residuais, os Sólidos Suspensos Totais (SST) são compostos por substâncias orgânicas e inorgânicas, tais como partículas de carboidratos, proteínas e lipídios para o primeiro caso e partículas minerais e areia para o segundo. O SST é um parâmetro importante pois pode afetar as características óticas de absorção e dispersão de luz solar (DU et al., 2022).

- **Oxigênio Dissolvido**

O oxigênio dissolvido em águas residuais depende das atividades físicas, químicas e bioquímicas do corpo d'água (APHA, 2023). No tratamento de esgoto, é usualmente medido nos tanques de aeração, sendo de suma importância o controle dessa variável nessa etapa, devido custo com energia elétrica elevado.

- **Oxidation-Reduction Potential (ORP) - Potencial de Oxido-Redução**

As reações de oxidação e redução (redox), mensuram o comportamento de muitos componentes químicos presentes na água no tratamento de esgoto. As reatividades e mobilidade de componentes químicos em sistemas biológicos, como Fe, S, N e C (APHA, 2023).

- **Amônia (NH<sub>3</sub>)**

Baixas concentrações de amônia normalmente estão associadas a água potável, águas superficiais ou efluentes de águas residuais nitrificadas de boa qualificada (APHA, 2023).

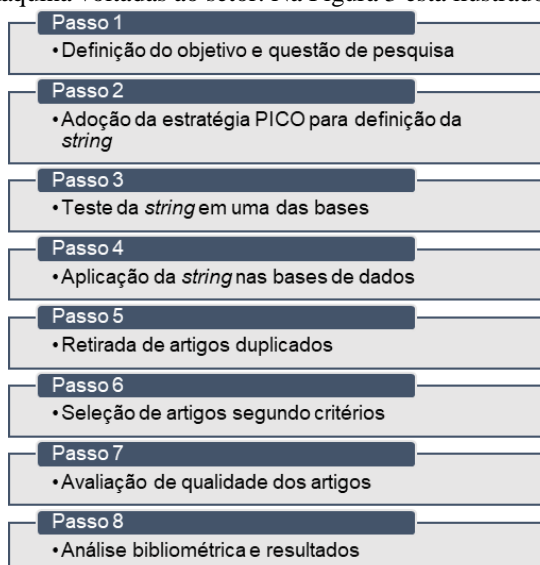
- **Cloro**

A utilização do cloro no abastecimento de água e tratamento de águas residuais tem o objetivo principal de desativar ou destruir microrganismos produtores de doenças (APHA, 2023). Para o fornecimento de água de reuso, por exemplo, é necessária uma concentração mínima do mesmo visando a segurança na utilização da água.

- **Turbidez**

A turbidez é uma propriedade ótica que faz com que a luz seja espalhada e absorvida. A clareza da água é um parâmetro muito utilizado em diversos processos produtivos afim de determinação da qualidade da água utilizada. Nos processos de tratamento de esgoto indica a eficiência dos processos de sedimentação e filtração além de ser um parâmetro legal para disposição do efluente nos córregos. Na água, a turbidez é causada por matéria suspensa e coloidal, como lodo, argila, matéria orgânica e inorgânica (APHA, 2023).

Através de uma bibliometria acerca da transformação digital no tratamento de esgoto e baseado no processo metodológico fundamentado em Kitchenham (2007), (Pestana, 2023) foram obtidas aplicações de diferentes técnicas de aprendizado de máquina voltadas ao setor. Na Figura 3 está ilustrado o método utilizado.

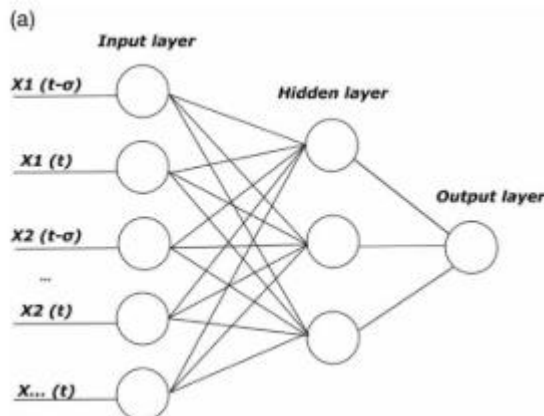


**Figura 3. Procedimento metodológico adotado**

Fonte: Pestana (2023).



A partir desse estudo, sete aplicações de RNA foram encontradas com aplicações no tratamento de esgoto. Na Figura 4 é possível observar um exemplo para arquitetura de RNAs a representação de seus neurônios.



**Figura 4. Exemplo de topologia para Redes Neurais Artificiais**

**Fonte: Li e Vanrolleghem, (2022).**

Na arquitetura da rede, os neurônios da camada de entrada são as variáveis explicativas do processo e cada um desses são conectados a cada um dos neurônios da camada oculta, que após passarem por uma função de ativação, por fim são conectados ao neurônio da camada de saída, que é a variável dependente. A seguir é feita uma síntese das publicações encontradas.

Acerca da função de ativação, elas permitam que as alterações que ocorrem nas variáveis de entrada, causem as mudanças nos neurônios subsequentes e algumas funções usualmente são utilizadas. Como por exemplo: função linear, sigmoide, tangente hiperbólica, ReLu, entre outras. A Equação 8 nos traz a função Sigmoide, que será utilizada nos modelos aqui descritos:

$$\text{Função}(x) = \frac{1}{1 + e^{-x}} \quad (8)$$

onde,  $e$ : é o número de Euler, de valor aproximado = 2,71828.

Arismendy (2021) através do uso de RNAs, LSTM e AG propõe uma redução de DQO em uma ETE. Inicia o trabalho com uma revisão sistemática da literatura e identifica que para a amostra a programação matemática é a técnica de análise prescritiva mais encontrada, representando 39% das publicações. A aplicação ocorreu na China, em uma Estação que analisa 22 variáveis de processo, com uma base de dados de 12/2017 a 07/2020. A rede foi configurada com 18 variáveis nos neurônios de entrada como as variáveis explicativas, e duas camadas ocultas com 2 e 16 neurônios respectivamente, com objetivo de prever a DQO do dia seguinte. Para configuração da rede o autor cria uma matriz de correlação entre todas as variáveis. A intenção é identificar principalmente, altas correlações de análises com parâmetros a manipular. O autor ainda aplica uma árvore de decisão afim de associar o pH em uma das etapas do processo com o nitrogênio, e uma das métricas utilizadas pelo autor para verificação da qualidade do modelo é o MAPE. Por fim a aplicação do AG e da RNA é realizada e o autor demonstra os resultados positivos entre o valor real e predito através de um gráfico em que sobrepõe esses dados.

Shen (2023) em uma base de dados proveniente de uma ETE, propõe um método para preenchimento de dados ausentes de OD através de LSTM, Gated Recurrent Units GRU] uma variante de LSTM, e uma combinação de redes neurais convolucionais. Modelos Florestas aleatórias e o modelo do K-nearest neighbor (KNN) são utilizados como benchmark de modelos para comparação de métricas. As métricas utilizadas foram: o erro





médio absoluto MAE, o RMSE, o MAPE e a acurácia. O método GRU teve o melhor desempenho de preenchimento com uma precisão de superior a 80%.

Li (2022) assim como outros autores, também faz uso de LSTM, nesse caso associado a Influyente Generator (IG) e aprimorado por um AG multiobjeto. As métricas utilizadas foram o MAPE e o KLdivergence que mede a diferença entre duas distribuições de probabilidade. A intenção foi modelar os dados de Sólidos Suspensos Totais (SST), DQO e nitrogênio amoniacal, a partir de dados de chuva, temperatura e vazão. Uma melhoria na série temporal gerada foi obtida através da utilização do processo de passeio aleatório ao LSTM.

Pisa (2020), diferentes dos outros autores utiliza métricas distintas para avaliação do modelo. As métricas: Erro Quadrado Integrado, do inglês Integrated Square Error (ISE) e Erro Absoluto Integrado, do inglês Integrated Absolute Error (IAE) para o modelo implementado. É implementado um Internal Model Controller (IMC) baseado em RNAs também do tipo LSTM. Para eliminação de ruídos foram adotadas duas estratégias a Análise de Componentes Principais (ACP) e um Denoising Autoencoder (DAE). Para comparação dos resultados gerais também foram utilizadas as métricas RMSE e MAPE associados ao coeficiente de determinação  $R^2$ . Toda essa modelagem com intuito de controlar o oxigênio dissolvido em um dos reatores de uma ETE.

Pisa (2023) propõe aplicação de transfer learning de modelos de redes neurais artificiais no tratamento de esgoto. As mesmas métricas de seu trabalho anterior, ISE e IAE são utilizadas para verificar a qualidade da transferência de modelos nos controles de oxigênio dissolvido e nitrogênio.

Padhma (2023) utilizando uma RNA associada a AG, propõe uma modelagem para operar sistemas complexos que tratam compostos orgânicos específicos, como fenol, xilenol e cianeto. Para mensurar a qualidade do modelo, assim como em outros trabalhos, o autor utiliza a métrica RMSE e MAPE.

Sharma (2018) aplica RNAs em conjunto com programação genética. A métrica RMSE é utilizada para determinação da qualidade do modelo. A intenção é modelar o dado de sólidos suspensos totais em uma etapa específica do processo em uma Estação de tratamento.

Foi observado uma técnica em comum entre alguns autores que utilizaram redes neurais artificiais, o LSTM. Enquanto as aplicações, incluem principalmente a predição de variáveis importantes para determinação da eficiência da estação, como a DQO. Já as métricas utilizadas também apresentam concordância entre as publicações, com o MAPE, MAE e RMSE sendo utilizados para verificar a qualidade da modelagem utilizada e determinação do erro entre o valor predito e esperado.

## Resultados e Discussão

Frente aos experimentos propostos espera-se encontrar correlações não só entre pares de dados, mas também a influência de algumas variáveis explicativas em uma variável dependente. No período de 02/2022 e 10/2022 foi criado um banco de dados visando fazer essa análise com intuito de facilitar a parametrização de futuras modelagens.

A partir dos dados analíticos obtidos das instalações de tratamento de esgoto e depois de organizá-los em tabelas, investigamos a viabilidade de identificar associações entre as sequências de informações. O software Excel contém uma função denominada CORREL, destinada a calcular o coeficiente de correlação. A Equação 8 apresentada a fórmula utilizada por essa função:

$$\text{Correl}(X, Y) = \frac{\sum(x - \bar{x})(y - \bar{y})}{\sqrt{\sum(x - \bar{x})^2 \sum(y - \bar{y})^2}} \quad (8)$$



onde,  $x$ : é o valor da matriz 1,  $y$ : é o valor da matriz 2,  $\bar{x}$  é a média das amostras da matriz 1  $\bar{y}$  é a média das amostras da matriz 2.

A Tabela 2 traz a correlação entre os parâmetros e os seus subsequentes, com a cor verde para as maiores correlações positivas, enquanto o azul é utilizado para as correlações negativas:

**Tabela 2. Parâmetros analíticos no processo de tratamento de esgoto**

|                       | pH | Condutividade | SST (imersão) | DQO   | SST (imersão) | DQO   | Manto de lodo (média) | SST (imersão) | SST (imersão) | ORP (Médio) - Entrada | ORP (Médio) - Saída | Oxigênio Dissolvido (Médio) - Entrada | Oxigênio Dissolvido (Médio) - Saída | NH3   | NH3   | Turbidez | Manto de lodo (Médio) | SST (imersão) |
|-----------------------|----|---------------|---------------|-------|---------------|-------|-----------------------|---------------|---------------|-----------------------|---------------------|---------------------------------------|-------------------------------------|-------|-------|----------|-----------------------|---------------|
| pH                    |    | 0,58          | 0,34          | 0,17  | 0,21          | -0,40 | -0,32                 | 0,07          | -0,07         | 0,69                  | 0,90                | 0,08                                  | 0,33                                | 0,27  | 0,49  | -0,04    |                       |               |
| Condutividade         |    |               | -0,12         | 0,48  | 0,03          | -0,41 | -0,18                 | -0,17         | -0,17         | 0,72                  | 0,81                | 0,04                                  | -0,16                               | 0,07  | 0,45  | 0,05     |                       |               |
| SST (imersão)         |    |               |               | -0,20 | 0,24          | -0,49 | -0,08                 | 0,17          | 0,02          | 0,51                  | 0,79                | 0,24                                  | 0,26                                | 0,14  | 0,12  | -0,36    |                       |               |
| DQO                   |    |               |               |       | -0,02         | 0,32  | -0,14                 | -0,10         | -0,01         | 0,74                  | 0,81                | 0,15                                  | -0,10                               | -0,09 | -0,03 | 0,23     |                       |               |
| SST (imersão)         |    |               |               |       |               |       |                       |               |               |                       |                     |                                       |                                     |       |       |          |                       |               |
| DQO                   |    |               |               |       |               |       | -0,21                 | -0,12         | -0,08         | -0,14                 | 0,42                | 0,66                                  | 0,29                                | 0,26  | -0,01 | 0,10     | -0,14                 |               |
| Manto de lodo (média) |    |               |               |       |               |       |                       | 0,25          | -0,15         | -0,04                 | -0,66               | 0,55                                  | 0,02                                | 0,17  | -0,33 | -0,16    | 0,28                  |               |
| SST (imersão)         |    |               |               |       |               |       |                       |               |               |                       |                     |                                       |                                     |       |       |          |                       |               |
| SST (imersão)         |    |               |               |       |               |       |                       |               | 0,42          | 0,38                  | -0,40               | -0,53                                 | -0,19                               | -0,09 | 0,13  | -0,21    | -0,23                 |               |
| ORP (Médio) - Entrada |    |               |               |       |               |       |                       |               |               | 0,91                  | -0,43               | -0,20                                 | -0,11                               | -0,35 | 0,17  | -0,04    | 0,03                  |               |
| ORP (Médio) - Saída   |    |               |               |       |               |       |                       |               |               |                       | -0,40               | -0,28                                 | -0,12                               | -0,42 | 0,08  | -0,04    | 0,15                  |               |
| OD (Médio) - Entrada  |    |               |               |       |               |       |                       |               |               |                       |                     | 0,67                                  | 0,66                                | -0,05 | 0,40  | 0,57     |                       |               |
| OD (Médio) - Saída    |    |               |               |       |               |       |                       |               |               |                       |                     |                                       | 0,29                                | 0,05  | 0,73  | 0,86     |                       |               |
| NH3                   |    |               |               |       |               |       |                       |               |               |                       |                     |                                       |                                     | 0,34  | -0,17 | -0,03    | -0,13                 |               |
| NH3                   |    |               |               |       |               |       |                       |               |               |                       |                     |                                       |                                     |       | 0,00  | 0,01     | -0,43                 |               |
| Turbidez              |    |               |               |       |               |       |                       |               |               |                       |                     |                                       |                                     |       |       | 0,21     | 0,02                  |               |
| Manto de lodo (Médio) |    |               |               |       |               |       |                       |               |               |                       |                     |                                       |                                     |       |       |          | -0,02                 |               |
| SST (imersão)         |    |               |               |       |               |       |                       |               |               |                       |                     |                                       |                                     |       |       |          |                       |               |

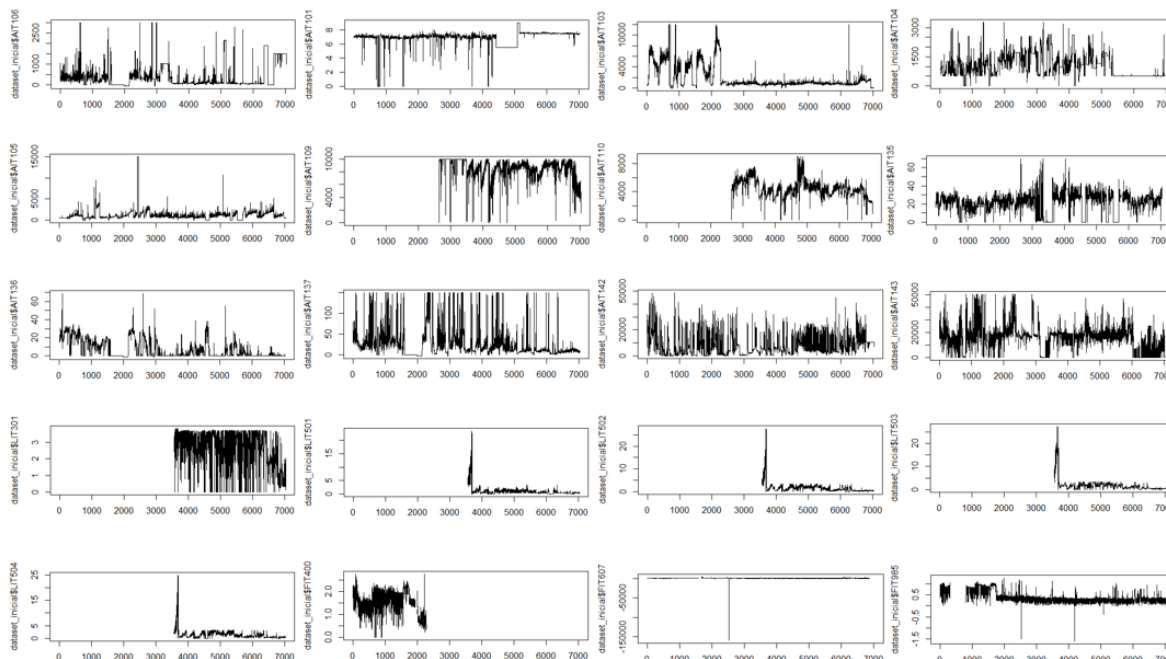
**Fonte: Resultados originais da pesquisa**

Foi possível observar que algumas variáveis se correlacionam diretamente em alguns momentos e inversamente em relação a outros parâmetros. Isso se mostrará importante no momento de parametrizar a rede em passos futuros.

Nesse passo foram agrupados dados analíticos do ano de 2023 juntamente com a vazão parcial de entrada, vazão de saída e os dados de pluviometria, afim de realizar uma modelagem com RNA. Os dados de vazão afluente total não foi utilizado devido a indisponibilidade dessa medição para o período. O mesmo acontece para os dados de OD e ORP.

Após ter os dados em uma mesma planilha, foi observado que devido a periodicidade de registro diferente para cada variável, seria necessário obter uma média horária, com objetivo de se ter uma leitura válida para cada variável em cada amostra.

Outro fato observado é que para o período completo do ano de 2023, particularidades da planta em questão, impediam que todas as análises tivessem seus registros sem falhas durante todo o período, como observado na Figura 5, inclusive com a presença de muitos outliers.



**Figura 5. Registro para o ano de 2023**  
**Fonte: Dados originais da pesquisa**

Por tanto, decidiu-se por limitar a base de dados que seria analisada ao período entre jun. de 2023 até a primeira quinzena de set. de 2023, onde a maior parte dos dados eram válidos. Na Tabela 3, temos a caracterização desses dados.

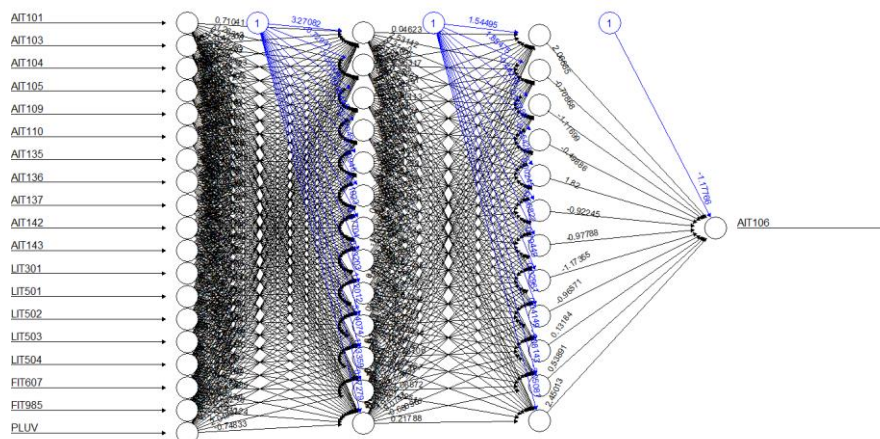
**Tabela 3. Caracterização dos dados**

| Identificador | Min. :  | 1st Qu.: | Median : | Mean :   | 3rd Qu.: | Max. :   | Desv. Pad: |
|---------------|---------|----------|----------|----------|----------|----------|------------|
| AIT101        | 0,00    | 5,50     | 6,93     | 6,64     | 7,36     | 9,00     | 423,65     |
| AIT103        | 0,00    | 699,40   | 798,50   | 835,70   | 935,50   | 2832,8   | 1,02       |
| AIT104        | 0,00    | 900,00   | 1125,90  | 1163,10  | 1460,00  | 3174,70  | 212,02     |
| AIT105        | 0,00    | 631,50   | 849,50   | 885,50   | 1176,10  | 3314,80  | 477,01     |
| AIT106        | 0,00    | 58,00    | 71,00    | 203,00   | 112,60   | 2794,00  | 478,73     |
| AIT109        | 1039,00 | 7944,00  | 8638,00  | 8329,00  | 9161,00  | 10016,00 | 1357,88    |
| AIT110        | 1127,00 | 3614,00  | 4230,00  | 4401,00  | 4812,00  | 9018,00  | 1217,42    |
| AIT135        | 0,25    | 24,12    | 27,25    | 27,69    | 32,00    | 70,00    | 8,09       |
| AIT136        | 0,00    | 0,00     | 0,50     | 4,98     | 7,50     | 55,25    | 8,18       |
| AIT137        | 0,00    | 9,75     | 14,00    | 24,20    | 25,00    | 150,00   | 29,60      |
| AIT142        | 0,00    | 3541,00  | 5218,00  | 6913,00  | 7808,00  | 34705,00 | 5795,38    |
| AIT143        | 0,00    | 15982,00 | 18275,00 | 18113,00 | 19818,00 | 47483,00 | 5080,27    |
| LIT301        | 0,00    | 2,38     | 2,95     | 2,73     | 3,55     | 3,83     | 1,01       |
| LIT501        | 0,00    | 0,38     | 0,89     | 0,93     | 1,42     | 3,25     | 0,67       |
| LIT502        | 0,00    | 0,73     | 1,77     | 1,62     | 2,33     | 3,47     | 0,98       |
| LIT503        | 0,00    | 1,00     | 1,82     | 1,68     | 2,42     | 3,26     | 0,99       |
| LIT504        | 0,00    | 0,44     | 1,54     | 1,47     | 2,30     | 3,26     | 1,04       |
| FIT607        | 0,00    | 748,30   | 953,30   | 926,20   | 1103,30  | 2020,00  | 230,51     |
| FIT985        | 0,01    | 0,16     | 0,25     | 0,24     | 0,32     | 1,59     | 0,11       |
| PLUV          | 0,00    | 0,00     | 0,00     | 0,03     | 0,00     | 3,60     | 0,20       |

**Fonte: Resultados originais da pesquisa**

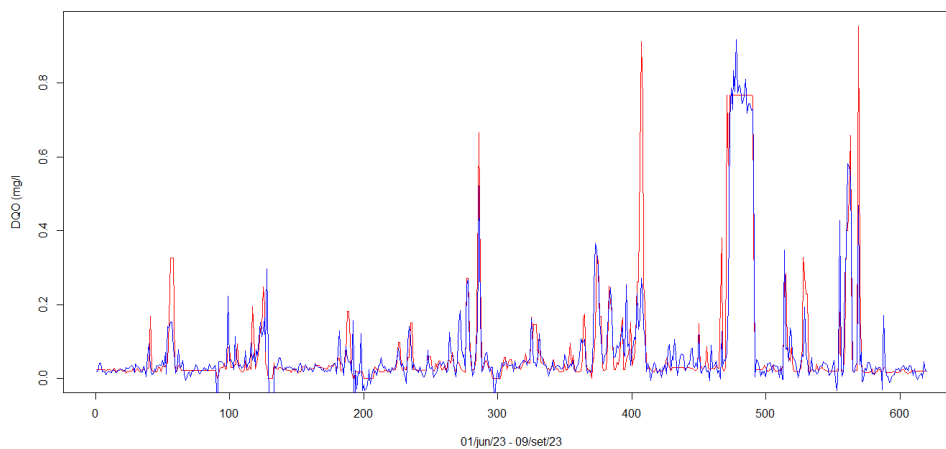


Para a base de dados escolhida, foi determinada uma separação de 70% dos dados para o treinamento da rede e 30% para o teste e a arquitetura de neurônios na camada escondida que obteve os melhores resultados foi (13,12) com a função de ativação logística, e é possível visualiza-la na Figura 6.



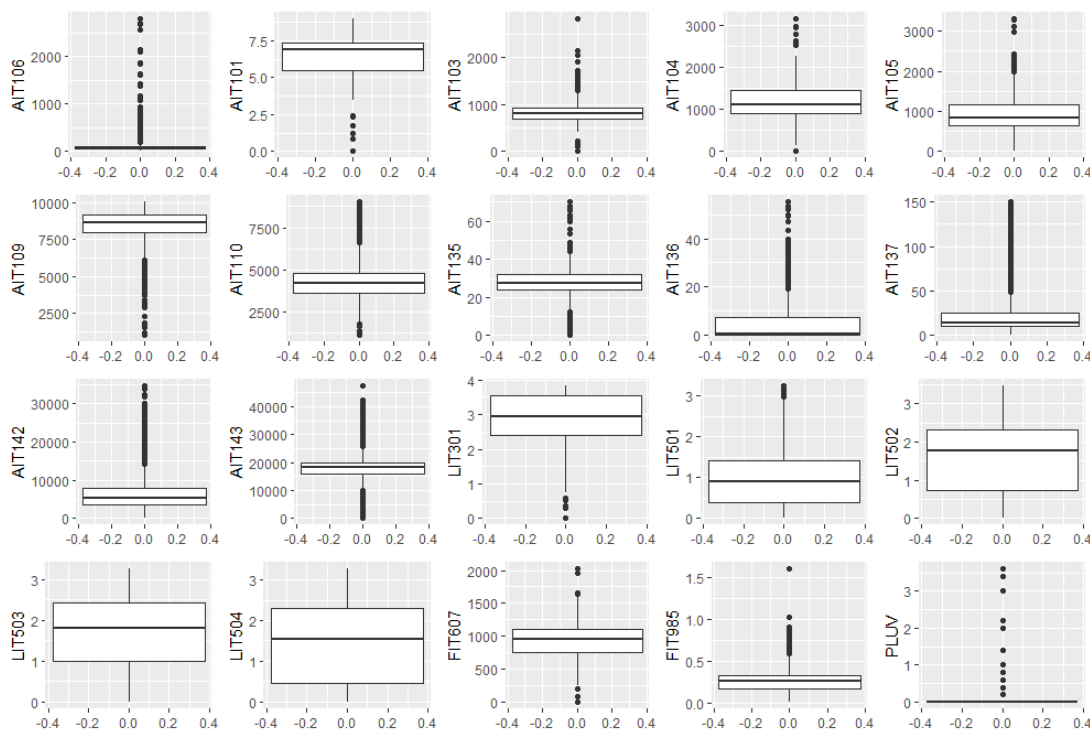
**Figura 6. RNA obtida através do software RStudio**  
**Fonte: Resultados originais da pesquisa**

Por fim, é possível visualizar o gráfico com os dados da DQO de saída da Estação predito pela RNA com a DQO real, separada para o teste da rede, conforme Figura 7. Em azul temos o dado previsto e em vermelho o dado real.



**Figura 7. DQO modelada vs. real**  
**Fonte: Resultados originais da pesquisa**

É possível observar que existe convergência entre o dado modelado e real, porém com as métricas utilizadas para verificar a acurácia do modelo ainda altas. Nesse ponto e a partir do uso dos gráficos box-plot foi possível observar a presença de outliers na base de dados utilizada, conforme ilustra a Figura 8.



**Figura 8. Box-plot dos dados utilizados**  
**Fonte: Resultados originais da pesquisa**

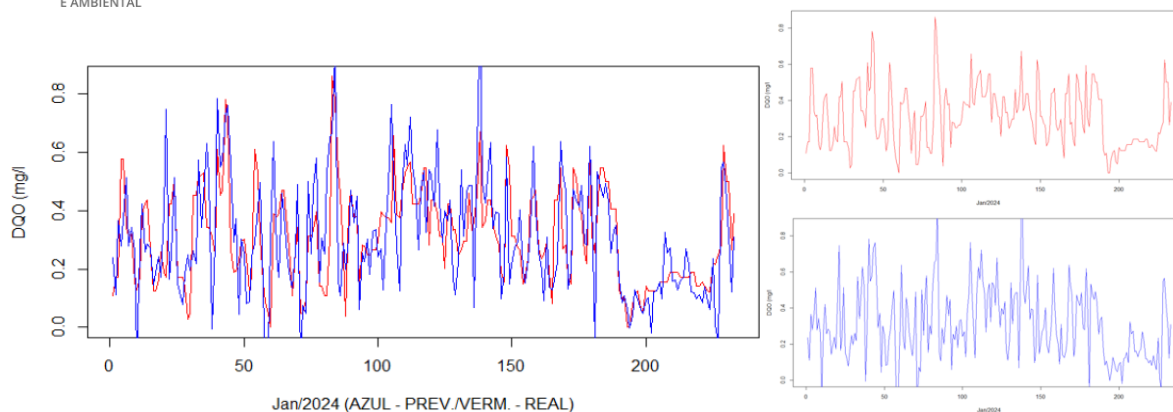
Na Tabela 4 temos a comparação dos resultados obtidos por diferentes técnicas e em diferentes bases, com e sem a retirada de outliers.

**Tabela 4. Comparação dos resultados**

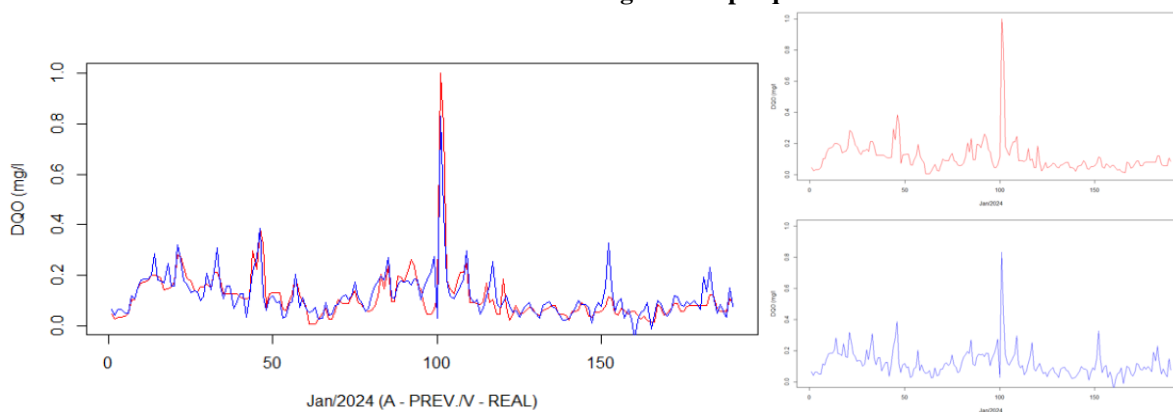
|                   | 01/06/23 - 10/09/23     | jun/23 |
|-------------------|-------------------------|--------|
|                   | MSE                     |        |
| Árvore de decisão | 88907                   | 5448   |
| RNA               | 43601                   | 9728   |
|                   | (MSE) Retirado Outliers |        |
| Árvore de decisão | 3047                    | 3138   |
| RNA               | 2199                    | 2163   |

**Fonte: Resultados originais da pesquisa**

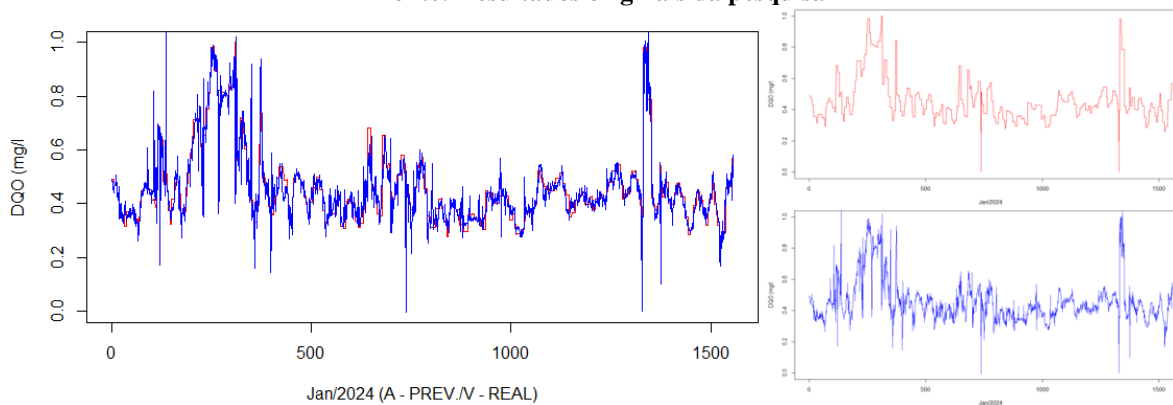
Foram então inclusos os dados mais recentes de OD e ORP, para o período de janeiro/2024, com a atualização da base de dados dos outros parâmetros para esse período na ETE São Miguel, e também realizada a modelagem para outras duas ETES, são elas ETE ABC e Suzano. Observou-se que seja pela inclusão desses dados, retirada de outliers, ou mesmo o amadurecimento do contrato e das análises realizadas, contribuíram para melhor performance da modelagem, conforme Figuras 9, 10 e 11.



**Figura 9. DQO efluente modelada vs. real – ETE São Miguel**  
**Fonte: Resultados originais da pesquisa**



**Figura 10. DQO efluente modelada vs. real – ETE ABC**  
**Fonte: Resultados originais da pesquisa**



**Figura 11. DQO efluente modelada vs. real – ETE Suzano**  
**Fonte: Resultados originais da pesquisa**

As métricas utilizadas para verificação do desempenho desses modelos corroboram para a existência de uma relação causal entre o dado de DQO na saída da Estação com os parâmetros intermediários analisados na planta. Outro fato observado é que quanto maior a quantidade de dados disponíveis para criação da rede e seu treino, existe uma tendência de melhor desempenho, conforme pode ser observado nas Tabelas 5 e 6.



**Tabela 5. Métricas de desempenho da RNA para dados efluentes**

|                | ETE SMP | ETE ABC | ETE SUZ |
|----------------|---------|---------|---------|
| MSE            | 92,67   | 286,07  | 36,41   |
| RMSE           | 9,62    | 16,91   | 6,03    |
| MAE            | 7,18    | 10,7    | 3,36    |
| MAPE           | 53,06   | 19,37   |         |
| R <sup>2</sup> | 0,461   | 0,718   | 0,808   |

Fonte: Resultados originais da pesquisa

**Tabela 6. Caracterização da base de dados**

|                           | ETE SMP       | ETE ABC       | ETE SUZ    |
|---------------------------|---------------|---------------|------------|
| Observações               | 777           | 632           | 5189       |
| Variáveis                 | 41            | 42            | 29         |
| Periodicidade do registro | Média horária | Média horária | 10 minutos |
| Treino/Teste              | 544/233       | 442/190       | 3632/1557  |

Fonte: Resultados originais da pesquisa

Na Tabela 7 é possível observar limites e pontos de ajustes estabelecidos para cada parâmetro obtidos em conjunto com um especialista em tratamento de esgoto, visando nortear etapas futuras do trabalho.

**Tabela 7. Set-points esperados para parâmetros analíticos de acordo com especialista**

| Parâmetro           | Local             | Limites ou set-point desejado? |
|---------------------|-------------------|--------------------------------|
| pH                  | Esgoto desarenado | 6,5 a 7,5                      |
| Condutividade       | Esgoto desarenado | 450 a 1.600 µS/cm              |
| SST (imersão)       | Esgoto desarenado | 120 a 280 mg/L                 |
| DQO                 | Esgoto desarenado | 400 a 500 mg/L                 |
| SST (imersão)       | Efluente Primário | 100 a 180 mg/L                 |
| DQO                 | Efluente Final    | 24 a 60 mg/L                   |
| Manta de Lodo       | Decantador Prim.  | ≤ 1,5 m                        |
| Manta de Lodo       | Decantador Prim.  | ≤ 1,5 m                        |
| SST (imersão)       | Licor misto       | 1.500 a 3.500 mg/L             |
| SST (imersão)       | Lodo de Retorno   | 3.000 a 5.000 mg/L             |
| ORP                 | Reator aeróbio 1  | + 80 a + 180 mV                |
| ORP                 | Reator aeróbio 1  | + 30 a + 100 mV                |
| ORP                 | Reator aeróbio 1  | + 100 a + 300 mV               |
| ORP                 | Reator aeróbio 2  | + 80 a + 180 mV                |
| ORP                 | Reator aeróbio 2  | + 30 a + 100 mV                |
| ORP                 | Reator aeróbio 2  | + 100 a + 300 mV               |
| ORP                 | Reator aeróbio 3  | + 80 a + 180 mV                |
| ORP                 | Reator aeróbio 3  | + 30 a + 100 mV                |
| ORP                 | Reator aeróbio 3  | + 100 a + 300 mV               |
| ORP                 | Reator aeróbio 4  | + 80 a + 180 mV                |
| ORP                 | Reator aeróbio 4  | + 30 a + 100 mV                |
| ORP                 | Reator aeróbio 4  | + 100 a + 300 mV               |
| Oxigênio Dissolvido | Reator aeróbio 1  | 1 a 3 mg/L                     |
| Oxigênio Dissolvido | Reator aeróbio 1  | 0,1 a 2 mg/L                   |
| Oxigênio Dissolvido | Reator aeróbio 1  | 1 a 3 mg/L                     |
| Oxigênio Dissolvido | Reator aeróbio 2  | 1 a 3 mg/L                     |
| Oxigênio Dissolvido | Reator aeróbio 2  | 0,1 a 2 mg/L                   |
| Oxigênio Dissolvido | Reator aeróbio 2  | 1 a 3 mg/L                     |
| Oxigênio Dissolvido | Reator aeróbio 3  | 1 a 3 mg/L                     |
| Oxigênio Dissolvido | Reator aeróbio 3  | 0,1 a 2 mg/L                   |
| Oxigênio Dissolvido | Reator aeróbio 3  | 1 a 3 mg/L                     |
| Oxigênio Dissolvido | Reator aeróbio 4  | 1 a 3 mg/L                     |
| Oxigênio Dissolvido | Reator aeróbio 4  | 0,1 a 2 mg/L                   |



|                            |                     |                      |
|----------------------------|---------------------|----------------------|
| <b>Oxigênio Dissolvido</b> | Reator aeróbio 4    | 1 a 3 mg/L           |
| <b>NH3</b>                 | Efluente Primário   | Até 30 mg/L          |
| <b>NH3</b>                 | Licor misto         | Até 25 mg/L          |
| <b>Turbidez</b>            | Efluente tratado    | Até 10 NTU           |
| <b>Manto de Lodo</b>       | Decantador DS1      | ≤ 50 cm              |
| <b>Manto de Lodo</b>       | Decantador DS2      | ≤ 50 cm              |
| <b>Manto de Lodo</b>       | Decantador DS3      | ≤ 50 cm              |
| <b>Manto de Lodo</b>       | Decantador DS4      | ≤ 50 cm              |
| <b>SST (imersão)</b>       | Lodo Aflu. Digestor | 23.000 a 38.000 mg/L |
| <b>SST (imersão)</b>       | Lodo condicionado   | 23.000 a 35.000 mg/L |
| <b>SST (imersão)</b>       | Filtrado FPP 1 a 3  | Até 220 mg/L         |
| <b>NH3</b>                 | Filtrado FPP 1 a 3  | Até 350 mg/L         |
| <b>Condutividade</b>       | Filtrado FPP 1 a 3  | 200 a 1.500 µS/cm    |
| <b>Cloro (inserção)</b>    | ETA Reuso           | >1 e <10 mg/L        |
| <b>Turbidez</b>            | ETA Reuso           | Até 10 NTU           |

Fonte: Resultados originais da pesquisa

Nessa etapa, modelou-se o dado de DBO, que ocorre com frequência bem menor do que a DQO e traria resultados expressivos para equipe operacional, conforme ilustra a Figura 12.

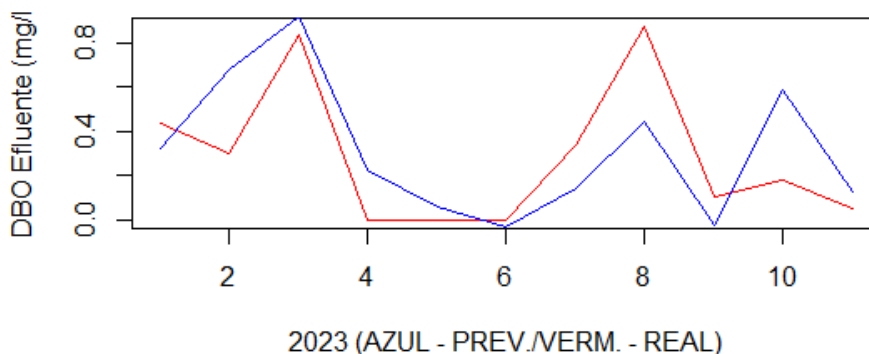


Figura 12. DBO modelada vs. DBO real – 2023

Fonte: Resultados originais da pesquisa

Nesse momento, foi realizada a modelagem agora não só considerando as análises realizadas na mesma Estação que a variável de interesse, mas também levando em conta valores em elevatórias de esgoto antes da planta. Por exemplo, com interesse de identificar cargas tóxicas na ETE ABC, através da análise de Amônia, condutividade e DQO afluente, foi realizada a modelagem com os dados de pH e condutividade nas elevatórias Guamiranga, Mooca e Ribeirão Pires, somados aos dados de vazão e nível de alguns desses locais. Os resultados podem ser vistos na Figuras 13, 14 e 15.

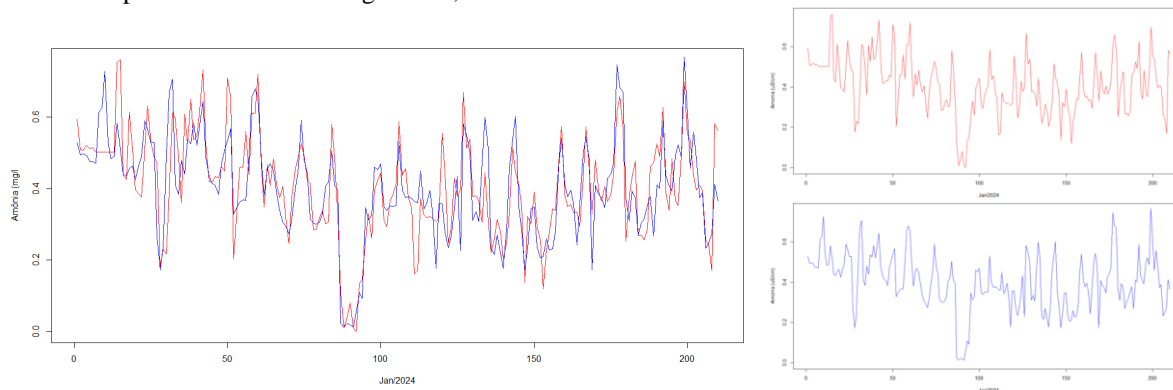
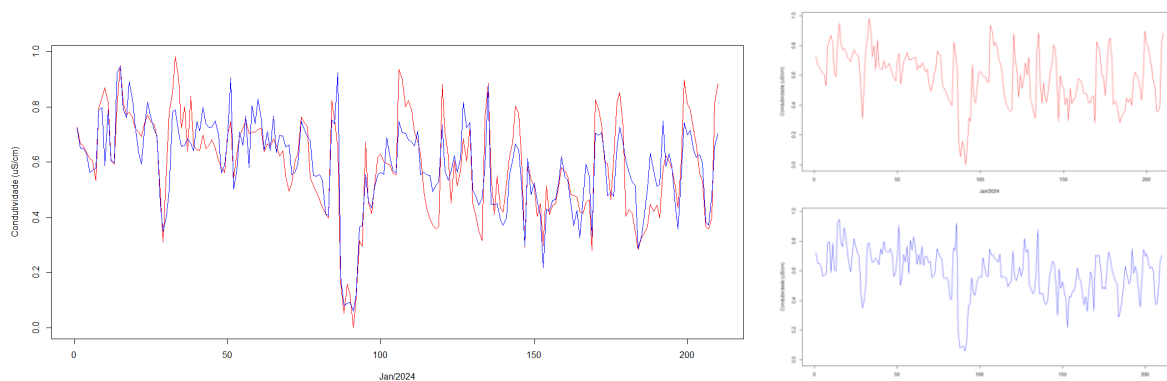


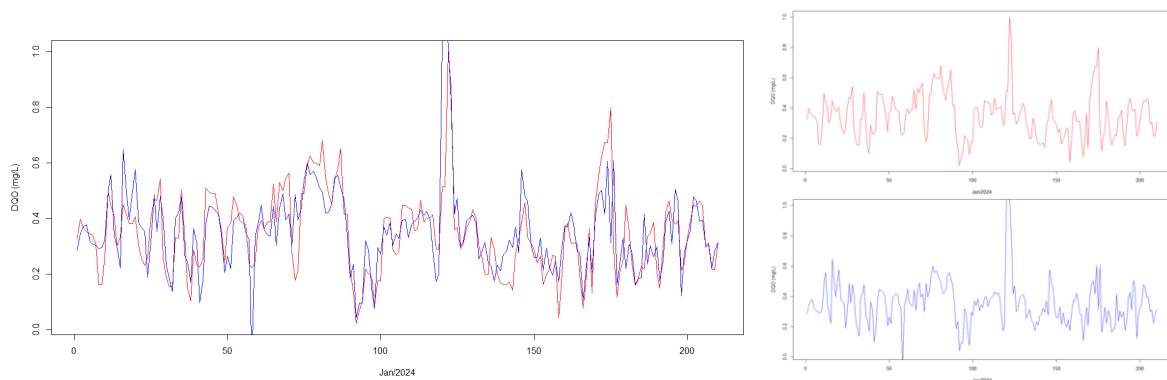
Figura 13. Amônia afluente modelada vs. real – Jan./2024

Fonte: Resultados originais da pesquisa



**Figura 14. Condutividade afluente modelada vs. real – Jan./2024**

Fonte: Resultados originais da pesquisa



**Figura 15. DQO afluente modelada vs. real – Jan./2024**

Fonte: Resultados originais da pesquisa

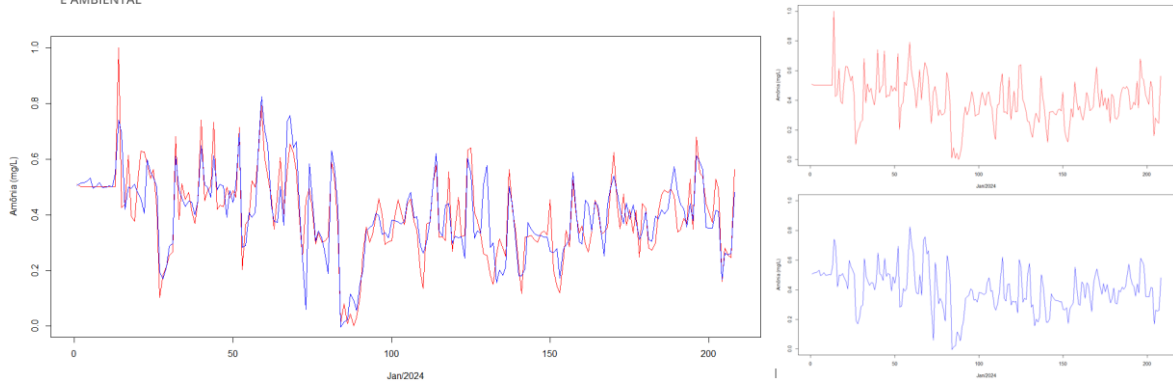
As métricas utilizadas novamente mostram a correlação entre esses parâmetros, sobretudo para análise de Amônia, conforme Tabela 8.

**Tabela 8. Métricas de desempenho da RNA para dados afluentes**

|                      | Amônia | Condutividade | DQO     |
|----------------------|--------|---------------|---------|
| <b>MSE</b>           | 20,77  | 4202,91       | 7041,09 |
| <b>RMSE</b>          | 4,55   | 64,82         | 83,91   |
| <b>MAE</b>           | 3,49   | 50,86         | 57,75   |
| <b>MAPE</b>          | 14,86  | 8,37          | 16,55   |
| <b>R<sup>2</sup></b> | 0,64   | 0,73          | 0,54    |

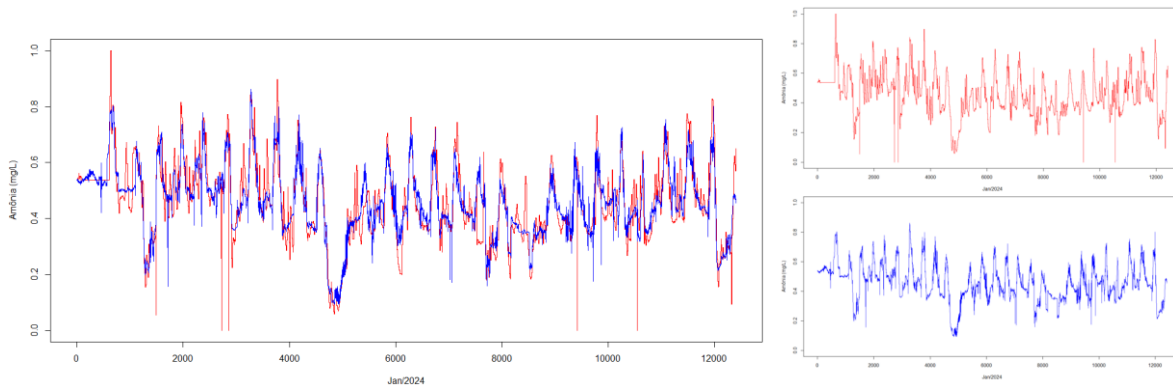
Fonte: Resultados originais da pesquisa

A partir desse ponto foi considerado também o tempo de deslocamento do esgoto entre esses locais, a fim de poder realizar a modelagem com as variáveis explicativas deslocadas em seu registro horário. Isso aconteceria por exemplo, considerando os dados de Guamiranga e Mooca desse instante, de Ribeirão Pires de 7h42m atrás e da variável de interesse considerando 42m a frente, devido à distância e velocidade do fluído na rede calculados pelo especialista local. Essas novas características apresentadas poderiam trazer um caráter preditivo a modelagem e ao ser colocado em produção, estimar o valor de Amônia na Estação com 40 minutos de antecedência. Seguem os resultados nas Figuras 16, 17 e 18.



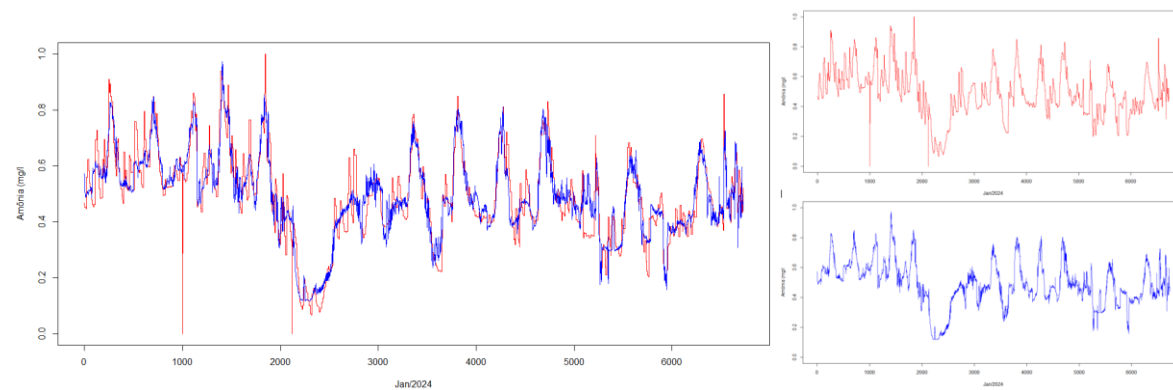
**Figura 16. Amônia af luente modelada vs. real – com previsibilidade de 1 hora – Jan./2024 - Média horária**

**Fonte: Resultados originais da pesquisa**



**Figura 17. Amônia af luente modelada vs. real – com previsibilidade de 1 hora – Jan./2024 – (Dados por minuto) – 41345 reg**

**Fonte: Resultados originais da pesquisa**



**Figura 18. Amônia af luente modelada vs. real – previsibilidade de 40 min. - Jan/2024 – (Dados por minuto) – Sem outliers - 22422 reg**

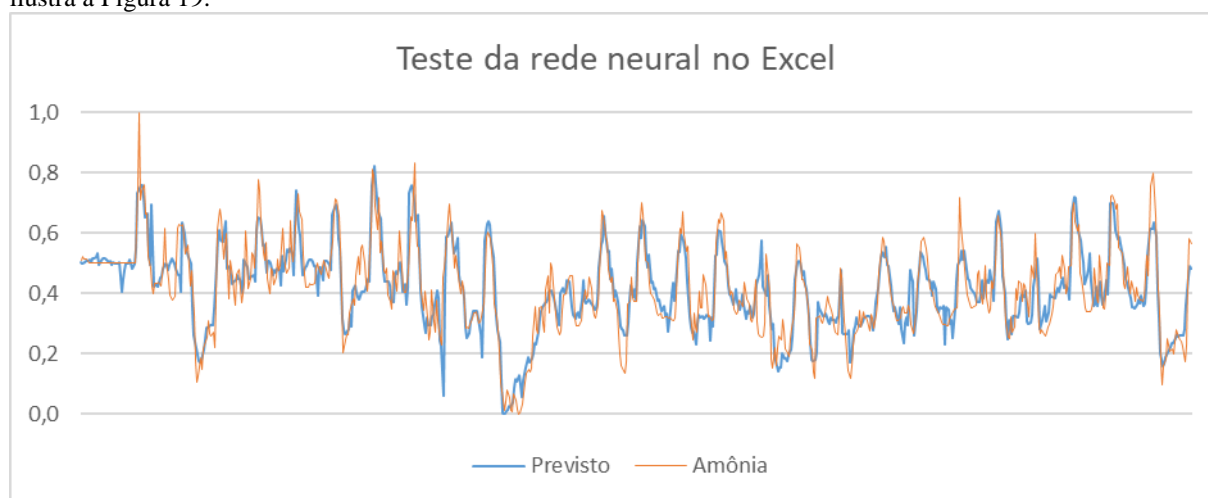
**Fonte: Resultados originais da pesquisa**

**Tabela 9. Métricas de desempenho da RNA para dados afluentes com previsibilidade**

|                      | Amônia (Média horária) | Amônia (Dados por minuto) –<br>41345 reg | Amônia (Dados por minuto) –<br>Sem outliers - 22422 reg |
|----------------------|------------------------|--|---|
| <b>MSE</b>           | 18,36                  | 14,40                                    | 10,16   |
| <b>RMSE</b>          | 4,28                   | 3,79                                     | 3,18  |
| <b>MAE</b>           | 3,07                   | 2,78                                     | 2,46  |
| <b>MAPE</b>          | 14,13                  |  |   |
| <b>R<sup>2</sup></b> | 0,70                   | 0,76                                     | 0,84  |

**Fonte: Resultados originais da pesquisa**

Após obtidos os pesos e vieses da RNA foi possível realizar testes com modelo já treinado no Excel visando identificar possíveis divergências entre os valores calculados e reais, e o mesmo se mostrou assertivo conforme ilustra a Figura 19.



**Figura 19. Testes da utilização da RNA treinada no Excel**

**Fonte: Resultados originais da pesquisa**

Identificado que o modelo respondeu bem após os cálculos matemáticos foi possível então iniciar os testes e implantação no Sistema Supervisório Elipse E3, utilizado na ETE objeto de estudo, conforme é possível observar na Figura 20.



### Testes com Redes Neurais Artificiais (RNA) no Elipse E3

| NEURÔNIO 1 |            | NEURÔNIO 2 |             | NEURÔNIO 3 |            | NEURÔNIO 4 |             | NEURÔNIO 5 |            | NEURÔNIO 6 |            |
|------------|------------|------------|-------------|------------|------------|------------|-------------|------------|------------|------------|------------|
| BIAS       | -0,9273869 | BIAS       | -3,72792447 | BIAS       | -1,5770182 | BIAS       | -0,7262368  | BIAS       | -1,454445  | BIAS       | -5,29076   |
| w11        | 0,9200479  | w21        | -1,07467879 | w31        | 5,8113062  | w41        | 6,664697    | w51        | -3,8667377 | w61        | -2,553677  |
| w12        | 4,0974762  | w22        | -6,32044144 | w32        | 6,0877358  | w42        | 16,0353445  | w52        | 31,8543794 | w62        | 15,385867  |
| w13        | 0,3719638  | w23        | -0,67296048 | w33        | 4,5294133  | w43        | 0,203535    | w53        | -3,6148446 | w63        | 4,436913   |
| w14        | 1,0491108  | w24        | -0,34823274 | w34        | 5,4145986  | w44        | 1,2340347   | w54        | -5,4081072 | w64        | -23,974697 |
| w15        | -1,5498669 | w25        | -0,14361207 | w35        | -1,071306  | w45        | -14,3529228 | w55        | -0,3016867 | w65        | -8,91391   |
| w16        | 1,1275334  | w26        | -0,48772559 | w36        | -1,3747883 | w46        | -2,8903923  | w56        | 2,9679507  | w66        | 3,536966   |
| w17        | 2,4567398  | w27        | -1,06336451 | w37        | -2,2092813 | w47        | -1,4685416  | w57        | 4,7060548  | w67        | 9,328852   |
| w18        | 2,2162638  | w28        | 3,8940417   | w38        | -1,3164816 | w48        | 2,1688343   | w58        | 1,0412048  | w68        | 6,076109   |
| w19        | 2,5899357  | w29        | 1,02291404  | w39        | 0,2742533  | w49        | 2,1792473   | w59        | 5,0886751  | w69        | 2,322738   |
| w110       | 0,2136627  | w210       | -0,06005336 | w310       | 0,2444713  | w410       | 1,6379652   | w510       | -2,00866   | w610       | -3,974607  |

| NEURÔNIO 1 |             | NEURÔNIO 2 |            | NEURÔNIO 3 |            | SAÍDA |            | ENTRADAS |             |
|------------|-------------|------------|------------|------------|------------|-------|------------|----------|-------------|
| BIAS       | -0,5466747  | BIAS       | 0,6205481  | BIAS       | 0,4208446  | BIAS  | -1,7504422 | AIT1     | 0           |
| w71        | 0,769088100 | w81        | -2,5728524 | w91        | 0,5534176  | wS1   | -0,5208933 | AIT2     | 0,546579505 |
| w72        | -2,8984635  | w82        | 4,0457874  | w92        | 13,5236412 | wS2   | 1,4108676  | AIT3     | 7,23078E-06 |
| w73        | -0,4667597  | w83        | 1,6445564  | w93        | 2,0910044  | wS3   | 1,495028   | AIT4     | 0,193508691 |
| w74        | -6,6582836  | w84        | -0,8975276 | w94        | -2,419711  |       |            | AIT5     | 0,613390405 |
| w75        | -2,8708379  | w85        | 0,8088837  | w95        | 2,7552995  |       |            | AIT6     | 0,995397393 |
| w76        | 5,5608949   | w86        | 2,2141014  | w96        | -3,1061401 |       |            | AIT7     | 0,061608825 |
|            |             |            |            |            |            |       |            | AIT8     | 4,98496E-05 |
|            |             |            |            |            |            |       |            | AIT9     | 0,096822595 |
|            |             |            |            |            |            |       |            | AIT10    | 0,072345733 |

|        |          |
|--------|----------|
| Euler  | 2,718282 |
| Min    | 4,33     |
| MaxMin | 51,47    |

|    |             |
|----|-------------|
| N1 | 0,89134792  |
| N2 | 0,000411976 |
| N3 | 0,663893152 |
| N4 | 0,04051403  |
| N5 | 0,99999998  |
| N6 | 0,019358586 |

Neuronio5 = 30,1481844554534

TESTAR RNA

Figura 20. Testes da utilização da RNA treinada no supervisor Elipse E3  
Fonte: Resultados originais da pesquisa

Por fim, o cálculo pode ser realizado de maneira automática, sendo necessário somente a troca das variáveis de entrada, que nesse exemplo estão sendo simuladas, por dados reais das Estações Elevatórias, que hoje chegam no Supervisor (Figura 21).



### Testes com Redes Neurais Artificiais (RNA) no Elipse E3

TESTAR RNA

Amônia prevista para daqui 1h com base em pH, condutividade, nível e vazão de Ribeirão Pires, Guairanga, Mooca e TAM3

18,44

Amônia prevista para daqui 1h = 18,3948836475225

ENTRADAS

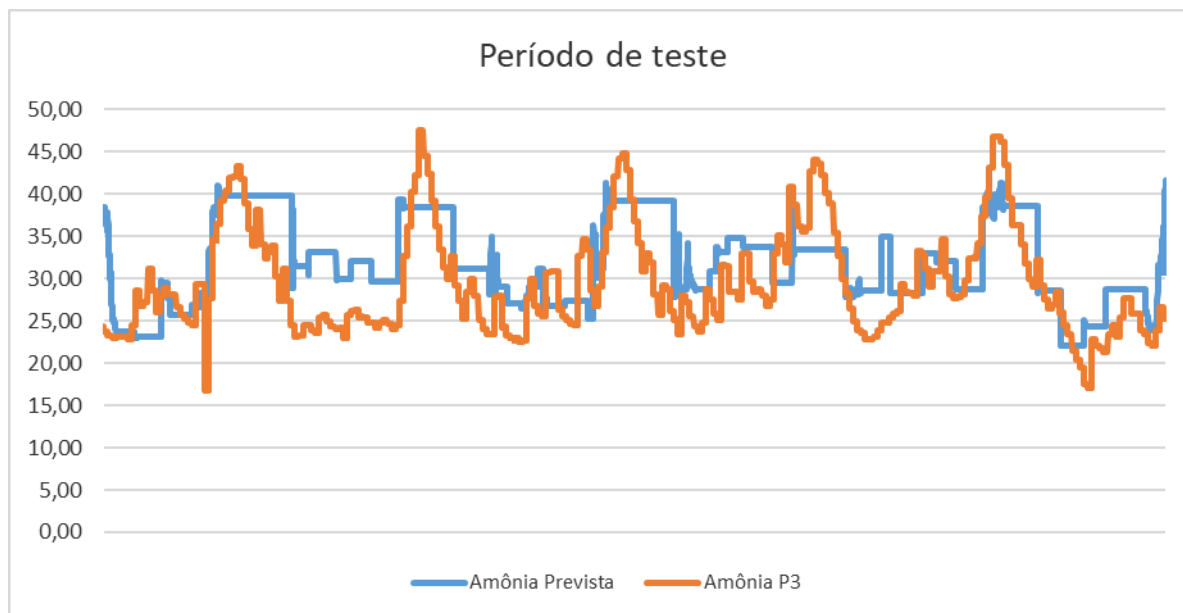
|                         |       |
|-------------------------|-------|
| Condutividade Ribeirão  | 16,92 |
| pH Ribeirão             | 2,87  |
| Condutividade Mooca     | 3,38  |
| Nível Mooca             | 3,05  |
| pH Mooca                | 5,01  |
| Nível TAM3              | 2,53  |
| Vazão TAM3              | 0,61  |
| Condutividade Guairanga | 4,89  |
| pH Guairanga            | 5,01  |
| Nível Guairanga         | 1,01  |

Liga/Desliga Simulação

Figura 21. Testes da utilização da RNA treinada no supervisor Elipse E3 com pesos e vieses no script  
Fonte: Resultados originais da pesquisa



Para o período de experimentação na primeira semana de abril de 2024, o valor de amônia previsto e modelado observado nos registros possuem correlação nos momentos de pico, contudo o resultado sugere que novas bases podem ser utilizadas, com mais dados e tratamento de outliers.



## CONCLUSÕES

Este estudo propôs uma abordagem inovadora para o tratamento de esgoto, de forma a criar um modelo preditivo para aprimorar os processos de tratamento. A combinação de dados analíticos e técnicas de aprendizado de máquina oferece uma nova maneira para melhorar a eficiência e a precisão do processo de tratamento. Métricas como RMSE, MAPE e  $R^2$  são utilizadas para avaliar a qualidade do modelo, o que permite insights valiosos sobre sua capacidade preditiva.

Uma revisão sistemática da literatura destaca o uso de RNAs, com ênfase para o modelo LSTM, como uma solução para a predição no tratamento de esgoto. A partir dessa revisão, o estudo avança para a modelagem da DQO no efluente de três ETEs, através de RNAs. Além disso, a modelagem da amônia afluente em uma dessas estações, com base em pH e condutividade em elevatórias a montante, demonstra a capacidade do modelo de fornecer previsões a equipe operacional, além de poder antecipar um dado mais complexo e caro a partir de duas análises relativamente mais baratas e simples. Por fim, a RNA treinada foi implementada em script no software supervisorio Elipse E3.

O projeto não apenas buscou otimizar os processos tradicionais de tratamento de esgoto, mas também introduziu abordagens avançadas, visando oferecer eficiência e capacidade de previsão ao setor de saneamento. A aplicação prática desses modelos na região metropolitana de São Paulo representa um passo significativo em direção a uma gestão mais inteligente e eficaz dos recursos hídricos no país, rumo a universalização dos serviços.

## REFERÊNCIAS BIBLIOGRÁFICAS

1. APHA American Public Health Association, American Water Works Association, Water Environment Federation, 2510 Conductivity, 2023.
2. APHA American Public Health Association, American Water Works Association, Water Environment Federation, 4500 Cl Chlorine, 2023.



SIMPÓSIO LUSO-BRASILEIRO  
DE ENGENHARIA SANITÁRIA  
E AMBIENTAL



3. APHA American Public Health Association, American Water Works Association, Water Environment Federation, 4500-NH3 Nitrogen (Ammonia), 2023.
4. Arismendy, L.; Cárdenas, C.; Gomes, D.; Maturana, A.; Mejía, R.; Quintero, C. G. A Prescriptive Intelligent System for an Industrial Wastewater Treatment Process: Analyzing pH as a First Approach. Sustainability (Basel, Switzerland), v. 13, n. 8, p. 4311, 2021.
5. Davies, P.S; The biological basis of wastewater treatment. Strathkelvin Instruments Ltd. 2005.
6. Ferreira, D.; Coraiola, M. EFICIÊNCIA DO LODO ATIVADO EM FLUXO CONTÍNUO PARA TRATAMENTO DE ESGOTO. Revista acadêmica: ciências agrárias e ambientais, v. 6, n. 2, p. 259, 2008
7. Geerdink, R. B.; Sebastiaan Van Den Hurk, R.; Epema, O. J. Chemical oxygen demand: Historical perspectives and future challenges. Analytica chimica acta, v. 961, p. 1–11, 2017.
8. Jordão, E. P.; Pessoa, C. A. Tratamento de esgotos domésticos. 7. ed. Rio de Janeiro: ABES, 2014.
9. Kitchenham, B; Charters, S. Guidelines for performing Systematic Literature Reviews in Software Engineering. Technical Report EBSE 2007-001, Keele University and Durham University Joint Report, 2007.
10. Li, F.; Vanrolleghem, P. A. An influent generator for WRRF design and operation based on arecurrent neural network with multi-objective optimization using a genetic algorithm. WaterScience and TechnologyIWA Publishing, 2022.
11. Padhma Priya P; Jenit A; Pearlin D DJ; Sharma NK. Predictive Biodegradation of Multiple Toxic Pollutants in Bioreactors TreatingReal Wastewater using ANN and GP. IOP conference series. Earth and environmental science, v.1130, n. 1, p. 12040, 2023.
12. Peres, J. et al. Modelagem matemática para sistema de tratamento de esgoto doméstico com remoção simultânea de matéria orgânica e nitrogênio. Engenharia Sanitaria e Ambiental, v. 26, n. 4, p. 765–773, 2021.
13. Pestana, A. A. Análise bibliométrica sobre a transformação digital no tratamento de esgoto. XVIII Simpósio Dos Programas De Mestrado Profissional Unidade De Pós-Graduação, Extensão E Pesquisa. Centro Paula Souza, 2023.
14. Pisa, I. et al. Denoising Autoencoders and LSTM-Based Artificial Neural Networks Data Processing for Its Application to Internal Model Control in Industrial Environments—The Wastewater Treatment Plant Control Case. Sensors (Basel, Switzerland), v. 20, n. 13, p. 3743, 2020.
15. Pisa, I. et al. Transfer Learning in wastewater treatment plants control: Measuring the transfer suitability. Journal of process control, v. 124, p. 36–53, 2023.
16. Resolução Nº 430, De 13 De Maio De 2011. Dispõe sobre as condições e padrões de lançamento de efluentes, complementa e altera a Resolução no 357, de 17 de março de 2005, do Conselho Nacional do Meio Ambiente-CONAMA. Disponível em: [http://conama.mma.gov.br/?option=com\\_sisconama&task=arquivo.download&id=627](http://conama.mma.gov.br/?option=com_sisconama&task=arquivo.download&id=627). Acesso em: 12 ago. 2023.
17. Sharma, N.; Chandrasekar, S.; Sundar, K. Suitability of ANN and GP for Predicting Soak Pit Tank Efficiency under Limited Data Conditions. MATEC Web of Conferences, v. 203, p. 3001, 2018.
18. Shen, Y. et al. An artificial neural network-based data filling approach for smart operation of digital wastewater treatment plants. Environmental research, v. 224, p. 115549–115549, 2023.
19. Van Houdt, G.; Mosquera, C.; Nápoles, G. A review on the long short-term memory model. The Artificial intelligence review, v. 53, n. 8, p. 5929–5955, 2020.