



UNIVERSIDADE FEDERAL DE SANTA CATARINA  
CENTRO TECNOLÓGICO  
DEPARTAMENTO DE ENGENHARIA SANITÁRIA E AMBIENTAL  
CURSO DE ENGENHARIA SANITÁRIA E AMBIENTAL

Rafaela Chiminelli Borth

**CALIBRAÇÃO DE SISTEMAS DE MONITORAMENTO ATMOSFÉRICO DE  
BAIXO CUSTO**

Florianópolis

2024

Rafaela Chiminelli Borth

**CALIBRAÇÃO DE SISTEMAS DE MONITORAMENTO ATMOSFÉRICO DE  
BAIXO CUSTO**

Trabalho de Conclusão de Curso submetido ao curso de Engenharia Sanitária e Ambiental do Centro Tecnológico da Universidade Federal de Santa Catarina como requisito parcial para a obtenção do título de Bacharel em Engenharia Sanitária e Ambiental.

Orientador(a): Prof. Leonardo Hoinaski, Dr.(a)

Florianópolis

2024

Chiminelli Borth, Rafaela  
CALIBRAÇÃO DE SISTEMAS DE MONITORAMENTO ATMOSFÉRICO DE  
BAIXO CUSTO / Rafaela Chiminelli Borth ; orientador,  
Leonardo Hoinaski, 2024.  
87 p.

Trabalho de Conclusão de Curso (graduação) -  
Universidade Federal de Santa Catarina, Centro  
Tecnológico, Graduação em Engenharia Sanitária e  
Ambiental, Florianópolis, 2024.

Inclui referências.

1. Engenharia Sanitária e Ambiental. 2. Poluição do ar.  
3. Monitoramento da qualidade do ar. 4. Monitoramento da  
qualidade do ar de baixo custo. 5. Calibração de sensores  
de monitoramento da qualidade do ar de baixo custo. I.  
Hoinaski, Leonardo. II. Universidade Federal de Santa  
Catarina. Graduação em Engenharia Sanitária e Ambiental.  
III. Título.

Rafaela Chiminelli Borth

**CALIBRAÇÃO DE SISTEMAS DE MONITORAMENTO ATMOSFÉRICO DE  
BAIXO CUSTO**

Este Trabalho de Conclusão de Curso foi julgado adequado para obtenção do título de  
“Bacharel em Engenharia Sanitária e Ambiental” e aprovado em sua forma final pelo Curso  
Engenharia Sanitária e Ambiental

Florianópolis, 04 de agosto de 2024.

Insira neste espaço  
a assinatura

Coordenação do Curso

**Banca examinadora**

Insira neste espaço  
a assinatura

Prof. Leonardo Hoinaski, Dr.(a)  
Orientador(a)

Insira neste espaço  
a assinatura

Prof. Pedro Luiz Borges Chaffe, Dr.(a)

Instituição UFSC

Insira neste espaço  
a assinatura

Fernando Campo Garcia Dr.(a)

Dedico este trabalho de conclusão de curso aos meus pais, Edemir e Mércia, por serem meus grandes exemplos.

## **AGRADECIMENTOS**

Dedico este trabalho de conclusão de curso a todos que fizeram parte da minha trajetória. Aos meus pais, Edemir e Mércia, por fazerem dos meus sonhos os seus, por me permitirem viver essa experiência, pelo amor incondicional e por sempre acreditarem em mim, mesmo nos momentos mais desafiadores. Ao meu irmão, Bernardo, por me mostrar o lado leve da vida. À minha avó, Edília, que sempre me inspirou com sua sabedoria e força, e que será sempre meu exemplo de mulher. À minha tia, Luiza, que nos deixou fisicamente, mas que estará sempre em meu coração. Ao meu namorado Nicolas, por ser meu companheiro nessa jornada e por me inspirar a ser uma pessoa melhor a cada dia. Aos amigos da cidade onde cresci, que estão sempre de portas abertas esperando meu retorno e que me lembram diariamente da minha criança interior. Aos amigos da faculdade, por serem meu lar em Florianópolis e o conforto em meio às dificuldades nesse caminho. Ao meu professor orientador e mentor, Leonardo Hoinaski, pela orientação, paciência, dedicação e por me guiar com sabedoria e competência ao longo deste percurso acadêmico. Aos colegas do LCQAr, por me acolherem de forma tão calorosa e pelo espírito de equipe que foi fundamental para o desenvolvimento deste trabalho. Aos colegas da IRIS, por compartilharem esse sonho comigo. E à Universidade Federal de Santa Catarina, por me dar o privilégio de aprender com os maiores exemplos.

A todos vocês, meu sincero agradecimento.

## RESUMO

A poluição do ar é uma das maiores ameaças ambientais à saúde humana e aos ecossistemas. O monitoramento da qualidade do ar é o primeiro passo para a gestão integrada da poluição atmosférica. Tradicionalmente, o monitoramento da qualidade do ar é realizado por estações fixas de alto custo e operação complexa, limitando seu uso em regiões menos desenvolvidas. O paradigma muda com a incorporação de sensores portáteis e acessíveis, que fornecem dados com alta resolução temporal e espacial. No entanto, a precisão, durabilidade e confiabilidade desses sensores variam conforme a tecnologia de medição, componentes, condições ambientais e métodos de operação. Para contribuir com a confiabilidade dessas medições, este estudo desenvolve uma ferramenta de filtragem e correção de dados, e análise de modelos de calibração para estações de monitoramento atmosférico de baixo custo. As estações utilizadas foram desenvolvidas no Laboratório de Controle da Qualidade do Ar, da Universidade Federal de Santa Catarina, e foram comparadas com uma estação de referência na cidade de Tubarão/SC. Foram aplicadas metodologias de controle de qualidade, detecção de outliers e algoritmos de aprendizado de máquina para calibrar os sensores. A detecção de outliers se mostrou eficiente no controle das flutuações da amostra, enquanto o modelo de regressão multivariada obteve os maiores resultados de  $R^2$ , RMSE e MAE. Os coeficientes de Kendall e Spearman apresentaram melhora significativa, quase dobrando entre as correlações utilizando as amostras brutas e as amostras calibradas. Por fim, esse estudo demonstra que os procedimentos de filtragem, correção e calibração das amostras são importantes e que os métodos apresentados podem fazer parte de uma metodologia generalizada de processamento de dados para sensores de poluição atmosférica de baixo custo.

**Palavras-chave:** Poluição do ar; monitoramento da qualidade do ar; monitoramento da qualidade do ar de baixo custo; calibração de sistemas de monitoramento de baixo custo.

## ABSTRACT

Air pollution is one of the biggest environmental threats to human health and ecosystems. Air quality monitoring is the first step towards integrated air pollution management. Traditionally, air quality monitoring is conducted by high-cost, complex fixed stations, limiting their use in less developed regions. The paradigm shifts with the incorporation of portable, accessible sensors that provide high temporal and spatial resolution data. However, the accuracy, durability, and reliability of these sensors vary based on measurement technology, components, environmental conditions, and operation methods. To enhance the reliability of these measurements, this study develops a data filtering and correction tool, and analyzes calibration models for low-cost air monitoring stations. The stations used were developed at the Air Quality Control Laboratory, Federal University of Santa Catarina, and compared with a reference station in Tubarão, SC. Quality control methodologies, outlier detection, and machine learning algorithms were applied to calibrate the sensors. Outlier detection was efficient in controlling sample fluctuations, while the multivariate regression model achieved the highest  $R^2$ , RMSE, and MAE results. Kendall and Spearman coefficients showed significant improvement, nearly doubling between raw and calibrated sample correlations. Ultimately, this study demonstrates that sample filtering, correction, and calibration procedures are crucial and that the presented methods can be part of a generalized data processing methodology for low-cost air pollution sensors.

**Keywords:** Air pollution; air quality monitoring; low-cost air quality monitoring; low-cost monitoring systems calibration.

## LISTA DE FIGURAS

Figura 1 - Número de estações de monitoramento atmosférico de referência autônomas ativas no Brasil em 2023.....	23
Figura 2 - Esquema de uma célula eletroquímica de dois eletrodos. ....	28
Figura 3 - Componentes da iniciativa CLEAN .....	35
Figura 4 - Estrutura do equipamento de baixo custo CLEAN.....	36
Figura 5 - Equipamento de baixo custo CLEAN instalado na Universidade Federal de Santa Catarina.....	37
Figura 6 - Mapa das estações de monitoramento CLEAN e referência .....	38
Figura 7 - Fluxograma das etapas de filtragem e detecção de falhas nas medições de dispositivos de monitoramento atmosférico de baixo custo .....	41
Figura 8 - Fluxograma da etapa de calibração de dispositivos de monitoramento atmosférico de baixo custo.....	44
Figura 9 - Série temporal do sensor CO-B4 antes e após ser submetido ao processo de controle de qualidade, com dados agrupados em 15 minutos .....	46
Figura 10- Pontos de alteração da linha base do sensor CO-B4 detectados por PELT conforme série temporal, e janelas de dados válidos mantidas após a aplicação do algoritmo, com dados agrupados em 15 minutos.....	48
Figura 11 – Detecção de outliers por quartil e k-vizinhos mais próximos do sensor CO-B4 conforme série temporal de dados, com dados agrupados em 15 minutos.....	49
Figura 12 - Série temporal final do sensor CO-B4 após etapas de filtragem e remoção de falhas, com dados agrupados em médias horárias .....	50
Figura 13 - Histograma dos dados do sensor CO-B4 antes e após as etapas de filtragem e remoção de falhas, com dados agrupados em médias horárias .....	51
Figura 14 - Box-plot mensal das concentrações do sensor CO-B4 antes e depois das etapas de filtragem e remoção de falhas.....	52
Figura 15 - Correlação entre as leituras do sensor CO-B4 e a estação de referência antes da aplicação dos modelos de regressão. Histogramas das concentrações de CO-B4 inicial e de referência. ....	55
Figura 16 - Comportamento das leituras do sensor CO-B4 e dos dados de referência após a aplicação dos dois melhores modelos de regressão .....	57

Figura 17 - Série temporal de concentrações do sensor CO-B4 de entrada do modelo, amostras de CO utilizadas para o teste do melhor modelo de regressão, os valores de CO previstos pelo melhor modelo de regressão e os dados alvo de CO de referência. ....59

## LISTA DE TABELAS

Tabela 1 - Métodos de monitoramento de referência .....	22
Tabela 2 - Total de estações de monitoramento atmosférico de referência no Brasil por poluente e população. ....	24
Tabela 3 - Descrição de características técnicas requeridas de sensores de baixo custo para monitoramento da qualidade do ar. ....	26
Tabela 4 - Intervalos de funcionamento dos sensores eletroquímicos modelo -B4 da Alphasense .....	28
Tabela 5 - Resumo das principais iniciativas de monitoramento atmosférico de baixo custo .	29
Tabela 6 - Sugestões de performance para sensores de baixo custo estabelecidos pela Agência de Proteção Ambiental Americana .....	31
Tabela 7 - Padrões de performance para sensores de baixo custo estabelecidos pela União Europeia.....	31
Tabela 8 - Características dos modelos de regressão aplicados na etapa de calibração .....	34
Tabela 9 - Relação de equipamentos de referência utilizados na estação Vila Moema .....	38
Tabela 10 - Etiquetas de diferenciação dos dados com base nas etapas de pré-processamento .....	42
Tabela 11 - Resultados obtidos após o processo de detecção de anomalias do sensor CO-B4	47
Tabela 12 - Análise estatísticas das etapas de controle de qualidade e detecção de anomalias .	52
Tabela 13 - Análise estatísticas das etapas de controle de qualidade e detecção de anomalias .	53
Tabela 14 - Resultados da calibração do sensor CO-B4 considerando as melhores métricas obtidas.....	54

## LISTA DE ABREVIATURAS E SIGLAS

AIC - Akaike Information Criterion (Critério de Informação de Akaike)

ANN - Artificial Neural Networks (Redes Neurais Artificiais)

BIC - Bayesian Information Criterion (Critério de Informação Bayesiano)

CO - Monóxido de Carbono

CONAMA - Conselho Nacional do Meio Ambiente

IEMA - Instituto de Energia e Meio Ambiente

kNN - k-Nearest Neighbors (k Vizinhos Mais Próximos)

LCs – Low-Cost Sensors (Sensores de baixo custo)

LCQAr - Laboratório de Controle da Qualidade do Ar

MBE - Mean Bias Error (Erro Médio de Tendência)

ML - Machine Learning (Aprendizado de Máquinas)

MLR - Multivariate Linear Regression (Regressão Linear Multivariada)

MP2.5 - Partículas Inaláveis Grossas com Diâmetro 2.5  $\mu\text{m}$

MP10 - Partículas Inaláveis Grossas com Diâmetro 10  $\mu\text{m}$

MSE - Mean Squared Error (Erro Médio Quadrático)

NO<sub>2</sub> - Dióxido de Nitrogênio

NO<sub>x</sub> - Óxidos de Nitrogênio

OMS - Organização Mundial da Saúde

O<sub>3</sub> - Ozônio

$\rho$  - Correlação de Spearman

PELT - Pruned Exact Linear Time (Tempo Linear Exato Poda)

ppb - Partes por Bilhão

ppm - Partes por Milhão

RF - Random Forests (Florestas Aleatórias)

RMSE - Root Mean Squared Error (Raiz do Erro Quadrático Médio)

SO<sub>2</sub> - Dióxido de Enxofre

UE - União Europeia

US EPA - United States Environmental Protection Agency (Agência de Proteção Ambiental Norte-Americana)

## SUMÁRIO

<b>1. INTRODUÇÃO</b> .....	<b>16</b>
<b>1.1 OBJETIVOS</b> .....	<b>17</b>
1.1.1 <b>Objetivo Geral</b> .....	<b>17</b>
1.1.2 <b>Objetivos Específicos</b> .....	<b>18</b>
<b>2. REVISÃO BIBLIOGRÁFICA</b> .....	<b>18</b>
2.1 <b>POLUIÇÃO ATMOSFÉRICA</b> .....	<b>18</b>
2.2 <b>MONITORAMENTO DA QUALIDADE DO AR</b> .....	<b>20</b>
2.2.1 <b>Monitoramento da qualidade do ar de referência</b> .....	<b>21</b>
2.2.2 <b>Monitoramento da qualidade do ar de baixo custo</b> .....	<b>25</b>
2.2.3 <b>Alternativas de monitoramento de baixo custo</b> .....	<b>28</b>
2.2.4 <b>Desafios no monitoramento de baixo custo</b> .....	<b>31</b>
2.2.5 <b>Métodos de correção e calibração no monitoramento de baixo custo</b> .....	<b>32</b>
2.3 <b>A INICIATIVA CLEAN – COLLABORATIVE LOW-COST ENVIRONMENTAL AIR QUALITY NETWORK</b> .....	<b>34</b>
<b>3. MATERIAIS E MÉTODOS</b> .....	<b>37</b>
3.1 <b>DESCRIÇÃO DOS DADOS</b> .....	<b>37</b>
3.2 <b>PROCEDIMENTO DE CALIBRAÇÃO DE SISTEMAS DE MONITORAMENTO DA QUALIDADE DO AR DE BAIXO CUSTO</b> .....	<b>39</b>
3.2.1 <b>Filtragem e detecção de falhas</b> .....	<b>39</b>
3.2.2 <b>Calibração dos dados</b> .....	<b>42</b>
<b>4. RESULTADOS</b> .....	<b>45</b>
4.1 <b>CARACTERIZAÇÃO, FILTRAGEM E DETECÇÃO DE FALHAS</b> .....	<b>45</b>
4.2 <b>CALIBRAÇÃO DOS SENSORES</b> .....	<b>53</b>
<b>5. CONCLUSÃO</b> .....	<b>60</b>
<b>APÊNDICE A – Série Temporal Dos Dados Brutos De Concentração De No2 – CLEAN Com Dados Agrupados Em 15 Minutos</b> .....	<b>68</b>
<b>APÊNDICE B – Série temporal de concentração de NO2 – CLEAN dos dados válidos após etapa de controle de qualidade com dados agrupados em 15 minutos</b> .....	<b>68</b>
<b>APÊNDICE C– Mudanças abruptas geradas por PELT na série temporal de concentração NO2 – CLEAN com dados agrupados em 15 minutos</b> .....	<b>69</b>
<b>APÊNDICE D – Janela de medição após implementa PELT na série temporal de concentração de NO2 – CLEAN com dados agrupados em 15 minutos</b> .....	<b>69</b>

APÊNDICE E – Detecção de outliers por quartil 1% e 99% global e diário das concentrações de NO <sub>2</sub> – CLEAN com dados agrupados em 15 minutos .....	70
APÊNDICE F – Detecção de outliers por kNN das concentrações de NO <sub>2</sub> – CLEAN com dados agrupados em 15 minutos .....	70
APÊNDICE G – Série temporal final das concentrações de NO <sub>2</sub> – CLEAN após etapas de filtragem e detecção de falhas, com dados agrupados em médias horárias .....	71
APÊNDICE H – Série temporal final das concentrações de NO <sub>2</sub> – CLEAN e NO <sub>2</sub> de referência, com dados agrupados em médias horárias .....	71
APÊNDICE I – Correlação e histogramas das concentrações de NO <sub>2</sub> – CLEAN inferida pela calibração, de referência e dos resíduos gerados.....	72
APÊNDICE J – Série temporal de concentrações de NO <sub>2</sub> – CLEAN de entrada do melhor modelo de regressão, amostras de teste, valores previstos e os dados alvo de referência	73
APÊNDICE K – Série temporal dos dados brutos de concentração de O <sub>3</sub> – CLEAN com dados agrupados em 15 minutos .....	73
APÊNDICE L – Série temporal de concentração de O <sub>3</sub> – CLEAN dos dados válidos após etapa de controle de qualidade com dados agrupados em 15 minutos.....	74
APÊNDICE M – Mudanças abruptas geradas por PELT na série temporal de concentração O <sub>3</sub> – CLEAN com dados agrupados em 15 minutos .....	74
APÊNDICE N – Janela de medição após implementa PELT na série temporal de concentração de O <sub>3</sub> – CLEAN com dados agrupados em 15 minutos .....	75
APÊNDICE O – Detecção de outliers por quartil 1% e 99% global e diário das concentrações de O <sub>3</sub> – CLEAN com dados agrupados em 15 minutos.....	75
APÊNDICE P – Detecção de outliers por kNN das concentrações de O <sub>3</sub> – CLEAN com dados agrupados em 15 minutos .....	76
APÊNDICE Q – Série temporal final das concentrações de O <sub>3</sub> – CLEAN após etapas de filtragem e detecção de falhas, com dados agrupados em médias horárias .....	76
APÊNDICE R – Série temporal final das concentrações de O <sub>3</sub> – CLEAN e NO <sub>2</sub> de referência, com dados agrupados em médias horárias .....	77
APÊNDICE S – Correlação e histogramas das concentrações de O <sub>3</sub> – CLEAN inferida pela calibração, de referência e dos resíduos gerados.....	77
APÊNDICE U – Série temporal de concentrações de O <sub>3</sub> – CLEAN de entrada do melhor modelo de regressão, amostras de teste, valores previstos e os dados alvo de referência	78
APÊNDICE V – Série temporal dos dados brutos de concentração de SO <sub>2</sub> – CLEAN com dados agrupados em 15 minutos .....	78

<b>APÊNDICE X – Série temporal de concentração de SO<sub>2</sub> – CLEAN dos dados válidos após etapa de controle de qualidade com dados agrupados em 15 minutos .....</b>	<b>79</b>
<b>APÊNDICE Y – Mudanças abruptas geradas por PELT na série temporal de concentração SO<sub>2</sub> – CLEAN com dados agrupados em 15 minutos .....</b>	<b>79</b>
<b>APÊNDICE Z – Janela de medição após implementa PELT na série temporal de concentração de SO<sub>2</sub> – CLEAN com dados agrupados em 15 minutos .....</b>	<b>80</b>
<b>APÊNDICE AA – Detecção de outliers por quartil 1% e 99% global e diário das concentrações de SO<sub>2</sub> – CLEAN com dados agrupados em 15 minutos.....</b>	<b>80</b>
<b>APÊNDICE BB – Detecção de outliers por kNN das concentrações de SO<sub>2</sub> – CLEAN com dados agrupados em 15 minutos .....</b>	<b>81</b>
<b>APÊNDICE CC – Série temporal final das concentrações de SO<sub>2</sub> – CLEAN após etapas de filtragem e detecção de falhas, com dados agrupados em médias horárias .....</b>	<b>81</b>
<b>APÊNDICE DD – Correlação e histogramas das concentrações de SO<sub>2</sub> – CLEAN inferida pela calibração, de referência e dos resíduos gerados.....</b>	<b>82</b>
<b>APÊNDICE EE – Série temporal de concentrações de SO<sub>2</sub> – CLEAN de entrada do melhor modelo de regressão, amostras de teste, valores previstos e os dados alvo de referência.....</b>	<b>83</b>
<b>APÊNDICE FF – Resultados parciais da calibração dos sensores CO-B4 CLEAN considerando as melhores métricas obtidas em termos de R<sup>2</sup>, RMSE e MAE.....</b>	<b>84</b>
<b>APÊNDICE GG – Resultados parciais da calibração dos sensores SO<sub>2</sub>-B4 CLEAN considerando as melhores métricas obtidas em termos de R<sup>2</sup>, RMSE e MAE.....</b>	<b>85</b>
<b>.....</b>	<b>85</b>
<b>APÊNDICE HH – Resultados parciais da calibração dos sensores NO<sub>2</sub>-B4 CLEAN considerando as melhores métricas obtidas em termos de R<sup>2</sup>, RMSE e MAE.....</b>	<b>86</b>
<b>.....</b>	<b>86</b>
<b>APÊNDICE II – Resultados parciais da calibração dos sensores O<sub>3</sub>-B4 CLEAN considerando as melhores métricas obtidas em termos de R<sup>2</sup>, RMSE e MAE.....</b>	<b>87</b>
<b>.....</b>	<b>87</b>

## 1. INTRODUÇÃO

A poluição do ar é uma das maiores ameaças ambientais à saúde humana e aos ecossistemas, sendo responsável por cerca de 7 milhões de óbitos prematuros anuais, segundo a Organização Mundial da Saúde (WHO, 2021). Além de impactar a saúde, a poluição do ar afeta significativamente o meio ambiente e a economia mundial, contribuindo para fenômenos de alterações climáticas. Neste contexto, primeiro passo para uma gestão integrada da poluição atmosférica é o monitoramento da qualidade do ar (VORMITTAG, 2021).

Tradicionalmente, o monitoramento da qualidade do ar é realizado através de estações fixas de referência, que medem a concentração dos poluentes na atmosfera. Esses sistemas são altamente confiáveis e precisos, a exemplo dos cromatógrafos e espectômetros de massa. No entanto, os altos custos de aquisição, manutenção e operação dessas estações limitam sua utilização, aumentando a lacuna no controle da qualidade do ar em regiões menos desenvolvidas (BORREGO et al., 2016).

O paradigma do monitoramento da qualidade do ar está mudando com a incorporação de sensores portáteis, também chamados de Low-Cost Sensors (LCs), acessíveis e de fácil utilização, que fornecem dados de resolução temporal e espacial significativas (SNYDER et al., 2013). Esses dispositivos oferecem oportunidades para melhorar as alternativas existentes e explorar novas aplicações de monitoramento. A Agência de Proteção Ambiental dos EUA (EPA) identificou que uma das principais utilizações desses sensores é para monitoramento suplementar e informativo não regulamentar. Outras aplicações incluem monitoramento móvel, exposição pessoal e qualidade do ar interior (KILARU et al., 2014). Esses sensores são utilizados para caracterizar a variabilidade espaciotemporal dos poluentes, analisar diferenças nas características de poluição, detectar hotspots, responder a emergências e monitorar tendências de longo prazo.

No entanto, a precisão, durabilidade e confiabilidade dos sensores variam de acordo com a tecnologia de medição, qualidade dos componentes, condições ambientais e métodos de operação (LIANG, 2021). Essas instabilidades são objeto de estudo de diferentes pesquisadores, que buscam identificar e corrigir os erros nas medições. A EPA desenvolveu guias com protocolos de colocação, correção, calibração, avaliação de desempenho e implantação de LCs, a fim de garantir e acompanhar a evolução das medições de baixo custo (DUVALL et al., 2022). Entretanto, as maiores dificuldades no avanço da tecnologia ainda permanecem no desenvolvimento de um modelo de filtragem e correção dos dados, e calibração dos sensores (WANG et al., 2023).

O processo de filtragem e correção prepara os dados para a calibração. Este envolve o processamento dos dados brutos dos sensores para remover ruídos e compensar fatores ambientais que possam distorcer as leituras. A calibração por sua vez, ajusta as saídas dos sensores para corresponder a valores de referência conhecidos, corrigindo imprecisões e vieses inerentes. A calibração estabelece uma linha de base para medições precisas, enquanto a correção de dados refina essas medições, garantindo consistência e confiabilidade. Juntos, eles melhoram a qualidade dos dados coletados, tornando os sensores de baixo custo viáveis para o monitoramento (OTTOSEN; KUMAR, 2019).

Embora frameworks padronizados de calibração, como os desenvolvidos pela EPA, ofereçam diretrizes importantes, sua aplicação ainda é limitada devido às características únicas de cada sistema. A quantidade e o período de coleta de dados, o comportamento dos sensores em diferentes ambientes e a calibração em tempo real são aspectos críticos para aproximar as leituras dos LCs às leituras de referência. Superar esses desafios potencializa o uso de LCs para outras finalidades, como ferramentas complementares para monitoramento regulatório.

Neste sentido, o objetivo do presente trabalho é desenvolver uma ferramenta de filtragem e correção dos dados, e analisar modelos de calibração para estações de monitoramento atmosférico de baixo custo. Para a condução da pesquisa, foram utilizados os LCs desenvolvidos no Laboratório de Controle da Qualidade do Ar, da Universidade Federal de Santa Catarina, em comparação com a estação de referência instalada na cidade de Tubarão/SC.

## 1.1 OBJETIVOS

### 1.1.1 Objetivo Geral

Desenvolver uma metodologia generalista de calibração e correção de dados para sistemas de monitoramento atmosférico de baixo custo, visando melhorar a precisão e a confiabilidade das medições.

### 1.1.2 Objetivos Específicos

**Objetivo específico 1:** Desenvolver e implementar uma ferramenta para filtragem e detecção de falhas nas medições de concentração de poluentes atmosféricos, utilizando sensores de monitoramento de baixo custo.

**Objetivo específico 2:** Avaliar a eficácia de diferentes modelos de calibração de sensores de baixo custo, comparando a precisão e as incertezas dos dados calibrados com as leituras de estações de referência.

## 2. REVISÃO BIBLIOGRÁFICA

### 2.1 POLUIÇÃO ATMOSFÉRICA

A qualidade do ar é medida pelo nível dos poluentes na atmosfera a qual uma população é exposta (ARTAXO, 2021). De acordo com a Resolução CONAMA nº 491 de 19/11/2018, poluente atmosférico é qualquer forma de matéria, em quantidade, concentração, característica ou tempo, que torne o ar nocivo ao bem-estar coletivo. Assim, a determinação da qualidade do ar depende da concentração dos poluentes na atmosfera e o grau de exposição dos receptores (CETESB, 2019).

Existem duas categorias principais de poluentes: os primários e os secundários. Os poluentes primários são aqueles emitidos diretamente pelas fontes de emissão, enquanto os secundários têm sua formação na atmosfera, através de reações químicas entre os compostos nela presentes, como os poluentes primários e componentes naturais (CETESB, 2021).

A poluição pode ocorrer tanto em ambientes internos quanto externos, seja por causas naturais, como queimadas florestais, erupções vulcânicas ou processos biológicos, quanto como resultado de atividades humanas, incluindo queima de combustíveis fósseis, emissões industriais, veiculares e práticas agrícolas (MOSLEY, 2010).

Embora a poluição já fosse um problema nas cidades pré-industriais, a industrialização e o crescimento econômico descontrolado agravaram significativamente os níveis de poluição, especialmente nos países em desenvolvimento (MOSLEY, 2010; BRIMBLECOMBE, 1998). Esses países enfrentam desafios complexos devido às múltiplas fontes de poluentes atmosféricos, tornando a mitigação da poluição do ar uma prioridade urgente. Enquanto os países desenvolvidos começaram a controlar as emissões desde meados

do século XX, as nações em desenvolvimento agora precisam equilibrar crescimento econômico e proteção ambiental (ZHANG et al., 2022).

Entre as consequências da má qualidade do ar, destaca-se a ação do ozônio na atmosfera, que reduz a produtividade agrícola, provoca a ocorrência da chuva ácida, acidifica os corpos hídricos e contribui para as alterações climáticas (ARTAXO, 2021). Além disso, a exposição à poluição do ar é reconhecida como a maior ameaça ambiental à saúde humana, devido à sua significativa contribuição para a carga de doenças. A morbidade e a mortalidade das doenças cardiovasculares e respiratórias não transmissíveis são aumentadas pela poluição do ar, as quais são as principais causas de mortalidade global (WHO, 2021).

A fim de controlar esses impactos, a OMS estabeleceu em 2005 Diretrizes para a Qualidade do Ar (DQA), visando oferecer recomendações quantitativas expressas em termos de concentrações de poluentes. Conforme os efeitos da exposição aos poluentes foram sendo estudados, a OMS ajustou os valores-guia em 2021, tornando os limites de concentração mais rigorosos (WHO, 2021). Apesar dos avanços da última década em relação ao controle das emissões, o relatório recente da OMS apontou que 99% da população mundial ainda respira ar impróprio (WHO, 2022).

No Brasil, os padrões de qualidade do ar são estabelecidos pela Resolução CONAMA nº 491/2018 e são reconhecidos como instrumentos de gestão da qualidade do ar no país. Esses padrões definem o limite de concentração de um poluente na atmosfera para um período específico de exposição (BRASIL, 2018). Os principais poluentes referenciados pela Resolução são material particulado (MP), ozônio ( $O_3$ ), dióxido de nitrogênio ( $NO_2$ ), dióxido de enxofre ( $SO_2$ ) e monóxido de carbono (CO), que devem ter suas concentrações monitoradas pelo órgão ambiental competente.

O material particulado é constituído por diversas formas de partículas suspensas na atmosfera, como poeiras e fumaças, originadas principalmente de veículos, processos industriais e queima de biomassa. Essas partículas podem se formar a partir de gases emitidos durante atividades de combustão. O tamanho das partículas está diretamente relacionado aos seus efeitos na saúde, sendo as menores mais prejudiciais (CETESB, 2024).

O ozônio é produto secundário de reações entre óxidos de nitrogênio e compostos orgânicos voláteis na presença de luz solar, provenientes da queima incompleta de combustíveis. A exposição prolongada ao ozônio pode causar problemas respiratórios, como asma e redução da função pulmonar (CETESB, 2024).

Dióxido de nitrogênio é originado principalmente da queima de combustíveis fósseis em veículos e processos industriais. Está associado a problemas respiratórios, como bronquite

e asma, além de contribuir para a formação de smog e chuva ácida. Sua exposição prolongada pode resultar em danos ao sistema respiratório e cardiovascular (CETESB, 2024).

O dióxido de enxofre é gerado predominantemente pela queima de combustíveis ricos em enxofre, como óleo diesel, óleo combustível industrial e gasolina. Ele desempenha um papel central na formação da chuva ácida e pode reagir com outros compostos atmosféricos, reduzindo visibilidade (CETESB, 2024).

Monóxido de carbono é um gás sem cor e odor, produzido pela queima incompleta de combustíveis orgânicos, como os combustíveis fósseis e biomassa. Nas cidades, é comumente encontrado em níveis elevados, especialmente nas áreas de tráfego intenso, devido às emissões dos veículos automotores (CETESB, 2024).

O monitoramento da concentração de poluentes direciona a compreensão do quadro da poluição atmosférica em uma determinada região. Esse processo não apenas identifica áreas de intervenção necessária, mas também auxilia na elaboração de planos de ação e no acompanhamento da evolução da qualidade do ar (BRASIL, 2018).

## 2.2 MONITORAMENTO DA QUALIDADE DO AR

O monitoramento da qualidade do ar é o primeiro passo para uma gestão integrada da qualidade do ar, pois auxilia na identificação do problema, embasa medidas regulatórias e estratégias de redução de gastos públicos, como os relacionados à saúde e à remediação ambiental. Ao fornecer dados atualizados, diversos atores, como agentes públicos, entidades privadas, comunidades e centros de pesquisa, são envolvidos em colaboração para enfrentar esse desafio.

Diversos estudos confirmam a importância do monitoramento para a construção de políticas e planos de ação eficientes e ágeis no controle da poluição, entretanto, esse conhecimento se encontra predominantemente difundido em países desenvolvidos (FENG, 2014). Na União Europeia (UE) e na América do Norte, por exemplo, dados de poluição do ar são coletados por estações fixas para aplicação de estratégias de controle, como instalação de zonas de baixa emissão, proibição do uso de combustíveis sólidos em residências e limitação da idade da frota de veículos públicos (TARGINO, 2016). No Brasil, a cobertura de monitoramento territorial é limitada, incompleta ou desatualizada em relação aos poluentes monitorados. A falta de monitoramento prejudica a estruturação dos demais instrumentos necessários à gestão da qualidade do ar (VORMITTAG, 2021).

Ao decidir sobre como desenvolver ou melhorar sua capacidade de monitoramento, os países devem avaliar a facilidade de implementação dentro de critérios como custo, recursos técnicos, requisitos computacionais e energéticos. Por exemplo, caso um país não tenha sistemas de monitoramento estabelecidos, pode-se considerar a instalação de um monitor de referência, e complementar essa medida com sensores de baixo custo (LCs). Esses sensores são conhecidos por seu baixo custo de aquisição e operação, exigindo menos recursos técnicos para calibração. Apresentam também baixo consumo de energia, mas requerem capacidade computacional média para processamento de grandes volumes de dados (EPA, 2017).

### **2.2.1 Monitoramento da qualidade do ar de referência**

Para fins regulatórios, o monitoramento é conduzido por instituições governamentais e privadas por meio de redes autônomas de referência. Essa prática tem dois principais objetivos: avaliar o nível de exposição da população aos poluentes atmosféricos, levando em conta critérios de saúde pública; e acompanhar tendências de médio e longo prazo para avaliar a eficácia dos programas de controle existentes e necessários aprimoramentos (CAMPO, 2023).

As estações de referências monitoram os poluentes de acordo com critérios de desempenho de medição muito rigorosos, para garantir dados de alta precisão e confiabilidade (EPA, 2019). Os dados gerados por essas estações servem, inclusive, como base de comparação para julgar outros métodos de medição. Para garantir a eficácia de uma rede de monitoramento da qualidade do ar, é crucial que as estações tenham objetivos bem definidos, permitindo uma interpretação precisa dos dados coletados. Nesse sentido, critérios simples devem ser observados tanto na escolha de novas localizações para as estações, quanto na descrição das áreas de amostragem já existentes. A determinação da representatividade espacial requer considerações como o tipo de poluente monitorado, a proximidade das fontes de emissão, a intensidade das emissões, bem como as características topográficas, relevo e padrões de transporte dos poluentes (BRASIL, 2020).

O Guia Técnico da Qualidade do Ar, criado pelo Ministério do Meio Ambiente (2020), estabelece equipamentos padrão para coleta dos dados e procedimentos básicos para aquisição, como mostrado na Tabela 1.

Tabela 1 - Métodos de monitoramento de referência

<b>Poluentes</b>	<b>Métodos de medição: Rede Automática</b>	<b>Métodos de medição: Rede Manual</b>
Partículas inaláveis	Radiação Beta	Gravimétrico; Amostrador dicotômico
Dióxido de enxofre	Fluorescência de pulso	Cromatografia iônica; Amostrador passivo
Óxido de nitrogênio	Quimiluminescência	-
Monóxido de carbono	Infravermelho não dispersivo	-
Hidrocarbonetos	Cromatografia grossa; Ionização de chama	-
Ozônio	Ultravioleta	-
Fumaça	-	Refletância da luz
Partículas totais em suspensão	-	Gravimétrico; Amostrador de grandes volumes

Fonte: Adaptado de BRASIL (2020).

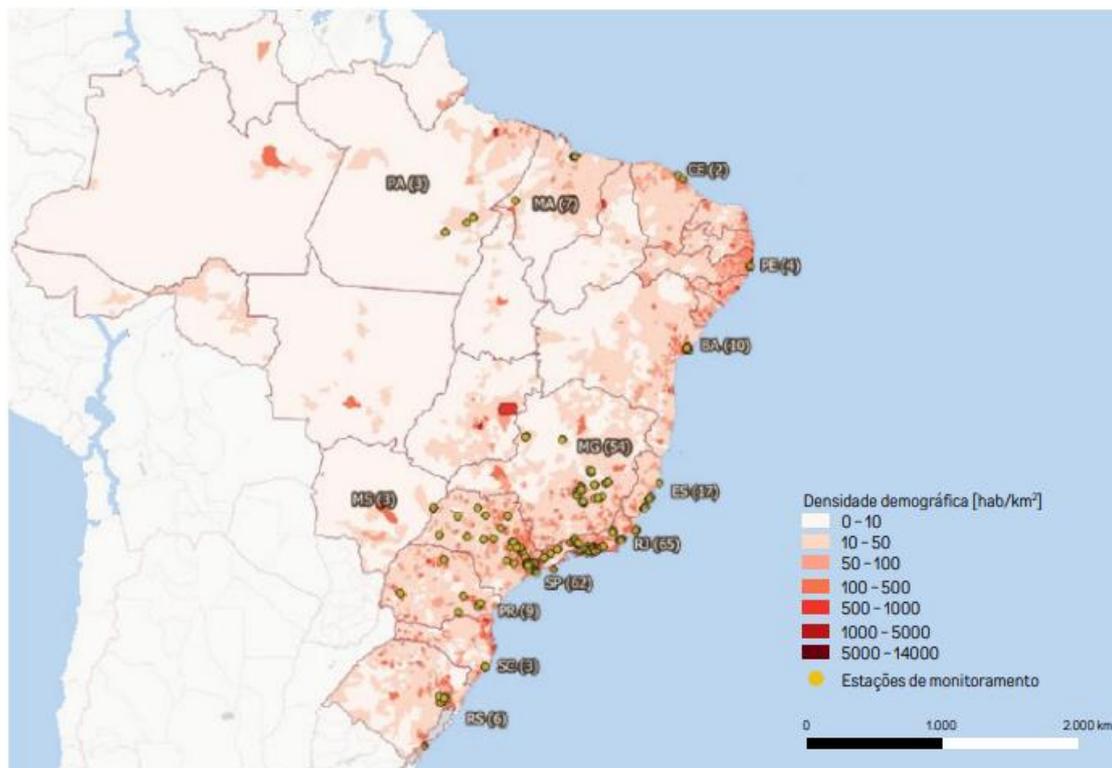
Devido à complexidade envolvida em sua aquisição, instalação, operação e manutenção, essas estações se tornam dispendiosas, o que dificulta sua implantação em larga escala. Como resultado, os dados coletados ficam limitados ao local onde a estação está instalada, o que pode levar a uma percepção distorcida das condições locais. Isso, por sua vez, resulta em análises com alcance limitado e pouca variabilidade espacial (KARAGULIAN et al., 2019). As estimativas de custo das tecnologias variam de \$15.000 a \$100.000, sem considerar despesas adicionais como consumo energético, instalação, operação e manutenção (CONCAS; MINERAUD; LAGERSPETZ; VARJONEN; LIU et al., 2021, apud, CAMPO, 2024). Além disso, é importante destacar que, até o momento, não houve desenvolvimento de tecnologias nacionais nesse campo, o que aumenta a dependência de importações. A exemplo, dados do Ministério do Meio Ambiente apontam a realização de um pregão para aquisição de estações de referência, com valor mínimo de R\$ 350.000 por estação (IEMA, 2024). Portanto, os custos totais da solução tornam-se um obstáculo, limitando o número de estações viáveis financeiramente e comprometendo a resolução e a distribuição espacial das medições.

De acordo com a Agência de Proteção Ambiental Americana (EPA), a escala de distribuição dessas estações pode indicar diferentes tipos de poluentes e seus impactos, tornando a distribuição espacial tão importante quanto a qualidade dos dados gerados (EPA, 2017). Por exemplo, na microescala, com medições de até 100 metros, é possível monitorar fontes de emissões próximas, como escapamentos de veículos individuais ou chaminés industriais. Já na escala urbana, com medições entre 4 a 50 quilômetros, podem ser monitoradas fontes de emissões de larga escala, como o tráfego em toda a cidade e as emissões industriais

de grande porte (EPA, 2017). No recorte brasileiro, o número mínimo de estações de monitoramento sugerido pela EPA e pela União Europeia (EU) não são atendidos, tanto no que se refere à distribuição espacial, quanto ao tamanho da população.

Na Figura 1 são apresentados os números atualizados de estações em operação em cada estado.

Figura 1 - Número de estações de monitoramento atmosférico de referência autônomas ativas no Brasil em 2023.



Fonte: (IEMA, 2024)

A análise considerou apenas estações automáticas, ou seja, que coletam dados em tempo real sem necessidade de um operador exclusivo. As estações manuais, por terem intervalos de coleta de até 6 dias, não são interessantes para formulação de políticas sobre a qualidade do ar. Das quase 250 estações de monitoramento automático operando em todo o Brasil, mais de 80% estão localizadas na Região Sudeste, indicando uma disparidade na distribuição dessas estações pelo país. A região Norte por sua vez, apresenta somente 3 estações de monitoramento, sendo incapaz de suprir a demanda advinda das emissões por queimadas, por exemplo.

A Figura 1 também revela que essas estações estão presentes em apenas 13 dos 26 estados brasileiros, com o Distrito Federal não contando com nenhuma estação automática.

Além disso, apenas sete capitais estaduais possuem esse tipo de monitoramento. Entre os 20 municípios mais populosos do país, apenas nove possuem sistemas automáticos de monitoramento da qualidade do ar.

A Tabela 2 mostra a distribuições das estações por estado e por poluente monitorado. Observa-se a relação entre número de estações e população, um dos critérios analisados na implementação de redes de monitoramento.

Tabela 2 - Total de estações de monitoramento atmosférico de referência no Brasil por poluente e população.

UF	Número de monitores						Total de estações	População (milhões de habitantes)	Estação / Pop. (milhões de habitantes)
	CO	MP10	MP2.5	NO2	O3	SO2			
<b>BA</b>	4	3	2	7	4	10	10	15.0	0.7
<b>CE</b>	2	1	1	2	2	2	2	9.2	0.2
<b>ES</b>	10	13	14	15	6	10	17	4.1	4.1
<b>MA</b>	3	7	3	4	4	3	7	7.2	1.0
<b>MG</b>	15	53	46	18	15	18	54	21.4	2.5
<b>MS</b>	3	1	-	3	3	1	3	2.8	1.1
<b>PA</b>	2	3	3	2	2	2	3	8.8	0.3
<b>PE</b>	4	4	1	4	4	4	4	9.7	0.4
<b>PR</b>	5	5	-	9	5	5	9	11.6	0.8
<b>RJ</b>	30	49	9	38	52	32	65	17.5	3.7
<b>RS</b>	5	6	-	6	6	5	6	11.5	0.5
<b>SC</b>	1	3	1	3	2	-	3	7.3	0.4
<b>SP</b>	18	54	32	40	53	12	62	46.7	1.3

Fonte: Adaptado de IEMA (2024)

A Tabela 2 destaca a predominância de estações no Sudeste, a inadequada cobertura geográfica e a falta de monitores para o material particulado fino MP2.5 em alguns estados. Contudo, vale ressaltar que existe uma inconsistência no registro dos dados, uma vez que na Plataforma da Qualidade do Ar mantida pelo Instituto de Energia e Meio Ambiente (IEMA), o último período de dados registrados foi no ano de 2022 (IEMA, 2024).

O avanço contínuo na pesquisa e no desenvolvimento de outras tecnologias, como os sensores de baixo custo, é crucial para ampliar sua aplicabilidade e precisão no monitoramento da qualidade do ar.

### 2.2.2 Monitoramento da qualidade do ar de baixo custo

O aumento da disponibilidade comercial da tecnologia de micro sensores está impulsionando a rápida adoção LCs para o monitoramento da qualidade do ar (KARAGULIAN et al., 2019). Essa tendência reflete a busca por alternativas mais viáveis de monitoramento, despertando interesse entre pesquisadores, governantes e empreendedores. O uso de LCs não apenas reduziria os custos de monitoramento da poluição do ar, mas também permitiria uma cobertura espacial mais ampla, especialmente em áreas remotas onde o monitoramento tradicional é limitado (SPINELLE et al., 2015). Além disso, a capacidade desses sensores de operar sem intervenção humana constante, facilita seu uso por usuários não especializados, representando um avanço significativo na democratização do monitoramento da qualidade do ar e na disponibilidade de dados em tempo real para embasar decisões.

Para entender melhor a implantação dos sensores, é importante distinguir entre os sensores individuais, produzidos pelos fabricantes originais, e os sistemas de monitoramento de baixo custo, que incluem sensores individuais juntamente com uma variedade de componentes adicionais, como caixa de proteção, sistema de alimentação, hardware eletrônico, software para aquisição de dados, conversão analógica para digital, tratamento de dados e transferência de dados. Do ponto de vista do usuário, os sistemas de monitoramento são sistemas prontos para uso (KARAGULIAN et al., 2019).

Diferentes protocolos foram desenvolvidos por órgãos regulatórios para direcionar o desenvolvimento e aplicação de LCs, como é o caso do *'The Enhanced Air Sensor Guidebook'* (DUVALL et al., 2022) e *'Protocol of evaluation and calibration of low-cost gas sensors for the monitoring of air pollution'* (SPINELLE et al., 2013). Esses protocolos estabelecem diferentes requisitos, incluindo tratamento de dados do sensor, níveis e duração dos testes, sazonalidade dos testes, tempo médio de amostragem do sensor e tipo de medições de referência às quais os dados do sensor são comparados. A depender da implementação, diferentes fatores devem ser considerados no momento de escolha de LCs. Alguns fatores estão descritos na Tabela 3.

Tabela 3 - Descrição de características técnicas requeridas de sensores de baixo custo para monitoramento da qualidade do ar.

<b>Fatores</b>	<b>Descrição</b>
Sensibilidade do sensor	O fator de resposta, ou seja, a capacidade do sensor de detectar com precisão variações na concentração do poluente alvo, deve ser estável ao longo do tempo. Além disso, precisa performar em baixas concentrações de poluentes.
Sensibilidade a condições ambientais	As flutuações de temperatura, pressão e umidade relativa no ambiente podem afetar sua resposta e por isso, devem ser minimizadas.
Sensibilidade cruzada	É comum que os sensores também apresentem flutuações nas medições por serem sensíveis a outros poluentes, além do poluente alvo. Para isso, a sensibilidade cruzada deve ser minimizada.
Tempo de resposta	Os sensores de gases podem exibir tempos de resposta variados diante de mudanças na concentração do poluente alvo. Em contextos que demandam uma alta resolução temporal, é preferível que o sensor responda rapidamente. No entanto, em áreas remotas, essa característica pode ser menos crucial.
Tempo de estabilização	Os sensores de gases podem levar algum tempo para que o sinal se estabilize, o que pode ser especialmente crítico em situações em que se planeja usar o sensor em um dispositivo móvel.
Tempo de vida do sensor	Sensores com maior durabilidade são preferíveis, pois reduzem a necessidade de substituições frequentes e os custos associados. Além disso, um tempo de vida mais longo proporciona estabilidade e consistência nas medições ao longo do tempo.
Limites operacionais de funcionamento	As condições dentro das quais um sensor pode operar de forma confiável devem ser alcançadas. Isso inclui faixas específicas de temperatura, umidade, pressão atmosférica e outros fatores ambientais.
Deriva da linha de base do sensor	É crucial que o sensor mantenha uma resposta estável mesmo na ausência de concentração do poluente alvo ao longo de todo o período de uso.
Consumo energético	A utilização desses sensores em dispositivos móveis, como equipamentos portáteis, implica o uso de baterias para alimentação elétrica. Quanto menor for o consumo de energia do sensor, maior será a autonomia do sistema de medição. Mesmo em dispositivos fixos, a integração de um sensor em um sistema de bateria de menor porte simplifica sua utilização.

<b>Fatores</b>	<b>Descrição</b>
Custos de aquisição, operação e manutenção	O custo de aquisição refere-se ao preço inicial de compra do equipamento. Os custos operacionais abrangem despesas contínuas associadas à operação do sistema. E os custos de manutenção incluem reparos ou calibrações regulares. Dessa forma, busca-se minimizar os custos atrelados.

Fonte: Adaptado de GOMES (2015)

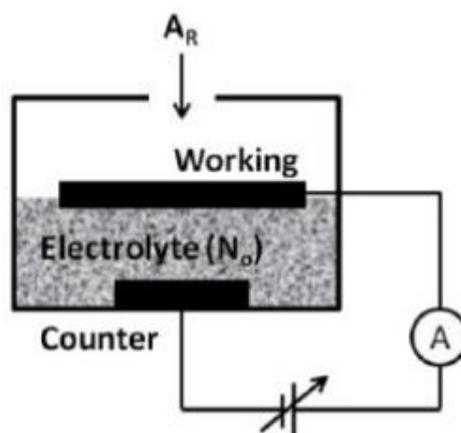
Portanto, como o uso desses sistemas não pode ser generalizado, a escolha de um determinado LCs depende da aplicabilidade e finalidade deste. Nesse sentido, há uma variedade de modelos de sensores disponíveis atualmente, a maioria dos quais opera com base em três princípios distintos: resistivo, ótico e eletroquímico. Este estudo se concentrará na aplicação de sensores eletroquímicos.

#### *2.2.2.1 Sensores de gases eletroquímicos*

Os sensores eletroquímicos operam com base no princípio de transformação de energia química em energia elétrica e vice-versa, e podem ser divididos em três classes principais: amperimétricos, potenciométricos e condutimétricos. Entre esses, os sensores amperimétricos são o foco deste estudo, que operam com eletrólito líquido em temperatura ambiente, sendo ideais para medições de poluentes atmosféricos (GOMES, 2015).

Na célula eletroquímica representada na Figura 2, tem-se dois eletrodos: o eletrodo de trabalho e o contra-eletrodo, imersos em uma solução eletrolítica. Ao aplicar uma determinada tensão elétrica entre esses eletrodos, ocorrem reações eletroquímicas de oxidação-redução nos compostos eletroativos presentes na solução e na superfície do eletrodo de trabalho. Esse desequilíbrio de elétrons, juntamente com a conexão do circuito externo, gera uma corrente elétrica. Desta forma, a corrente elétrica é proporcional à concentração do composto gasoso em contato com o eletrodo (GOMES, 2015).

Figura 2 - Esquema de uma célula eletroquímica de dois eletrodos.



Fonte: (GOMES, 2015)

Na medição de gases, a corrente é gerada quando a célula é exposta a uma atmosfera contendo o poluente. As características combinadas deste sensor o tornam altamente versátil. Com baixo consumo de energia, custo acessível, dimensões reduzidas, seletividade, estabilidade, sensibilidade e rápida resposta, ele oferece alto desempenho analítico a um custo acessível. Existem diferentes compostos gasosos medidos com estes sensores, tais como CO, SO<sub>2</sub>, NO<sub>2</sub> e O<sub>3</sub>.

Cada sensor possui um intervalo de funcionamento, variando entre o valor mínimo e máximo de operação, em relação à medição de poluentes, temperatura, pressão e umidade, como mostrado na Tabela 4.

Tabela 4 - Intervalos de funcionamento dos sensores eletroquímicos modelo -B4 da Alphasense

Parâmetros	CO-B4		O <sub>3</sub> -B4		NO <sub>2</sub> -B4		SO <sub>2</sub> -B4	
	Min	Max	Min	Max	Min	Max	Min	Max
Varição de medição (ppm)	0	1000	0	20	0	20	0	100
Varição de temperatura (C°)	-30	50	-30	40	-30	40	-30	50
Varição de umidade (% RH)	15	90	15	85	15	85	15	90
Varição de pressão (kPa)	80	120	80	120	80	120	80	120

Fonte: Adaptado de ALPHASENSE (s.n.)

### 2.2.3 Alternativas de monitoramento de baixo custo

As iniciativas envolvendo LCs visam, sobretudo, complementar as redes de monitoramento convencionais, estando presentes no meio acadêmico, comercial e no campo da ciência colaborativa. A partir da revisão bibliográfica realizada por Morawska et al. (2018),

a Tabela 5 resume algumas das principais iniciativas de monitoramento de baixo custo desenvolvidas até o presente momento.

Tabela 5 - Resumo das principais iniciativas de monitoramento atmosférico de baixo custo

<b>Projeto</b>	<b>Período</b>	<b>Tipo de Projeto</b>	<b>Resultados Aplicados</b>	<b>Acesso aos Dados</b>
ARC-LP16	2016–2020	Desenvolvimento de Rede	Redes de sensores/monitores de baixo custo em várias cidades	Acesso Público
EuNetAir	2012–2016	Rede de Novas Tecnologias de Sensoreamento	Desenvolvimento e avaliação de novos sensores/monitores	N/A
Citi-Sense	2012–2016	Desenvolvimento de Plataformas Tecnológicas para Monitoramento Distribuído	Teste/monitoramento de sensores/monitores em múltiplos países	Toolbox do Observatório Cidadão (COT)
Citi-Sense-MOB	2013–2015	Medição Móvel de Qualidade do Ar	Avaliação exaustiva de plataformas de baixo custo	N/A
OpenSense	2010–2017	Sensoreamento Baseado em Comunidade	Mapa de poluição do ar com base em plataforma móvel de sensoreamento	Global Sensor Network (GNS)
CONA (Community Observation Networks for Air)	2015 ~	Estabelecimento de Rede de Sensores de Baixo Custo	Monitores desenvolvidos, construção da rede	Relatório para Participantes
PiMi Airbox	2013–2016	Monitoramento de Qualidade do Ar Interno e Mineração de Dados Sensoriais em Grande Escala	Monitores desenvolvidos, testes de rede	N/A
U.S. EPA CAIRSENSE	2013–2016	Avaliação de Desempenho de Longo Prazo de Sensores/Monitores e Rede	Sensores/Monitores testados	N/A
U.S. EPA Village Green	2013~	Construção de Sistemas de Monitoramento Autônomo	Unidades construídas e instaladas em número limitado de locais	Acesso Online
AirVisual	2015 ~	Rede Global de Monitores de Qualidade do Ar	Mapa de locais fixos e aplicativo desenvolvido para todos os usuários	Aplicativo AirVisual e Website
Air Quality Egg	2012 ~	Rede de Sensoreamento de Qualidade do Ar Liderada pela Comunidade	Mapa e função de dados desenvolvidos para todos os usuários	Aplicativo para celular e Website
AirCasting (Monitor AirBeam)	2012 ~	Plataforma para Gravação, Mapeamento, e Compartilhamento de Dados de Saúde e Meio Ambiente Usando o Smartphone	Mapa de dados dos monitores AirBeam e app desenvolvido para todos os usuários	Aplicativo para celular e Website
SMARTCITIZEN	N/A	Plataforma para Gerar Processos Participativos de Pessoas nas Cidades	Mapa de dados dos monitores Smart Citizen e app desenvolvido para todos os usuários	Aplicativo para celular e Website

<b>Projeto</b>	<b>Período</b>	<b>Tipo de Projeto</b>	<b>Resultados Aplicados</b>	<b>Acesso aos Dados</b>
Purple Air	2015 ~	Rede de Monitoramento da Qualidade do Ar Construída em uma Nova Geração de Sensores/Monitores de "Internet das Coisas"	O Mapa do PurpleAir exibe os pontos usando a escala de Índice de Qualidade do Ar (AQI) da U.S. EPA	Registro de Usuário Obrigatório

Fonte: Adaptado de MORAWSKA et al. (2018)

As aplicações dos LCs abrangem monitoramento externo, interno e pessoal, cada uma com requisitos de desempenho específicos. Por exemplo, enquanto os LCs destinados a monitorar poluição relacionada ao tráfego precisam detectar partículas menores, aqueles voltados para poeira de construção só precisam detectar partículas maiores. Em outras palavras, é crucial que os LCs sejam adaptados ao propósito específico em questão. Diante disso, surge a questão de necessidade sobre um "protocolo padrão" para testes, ou se os protocolos devem ser adaptados aos propósitos individuais, com critérios mais flexíveis (MORAWSKA et al., 2018).

Embora diversos estudos tenham avaliado o desempenho de LCs em diferentes condições, os critérios de "suficiência" variam amplamente entre eles. Chu et al. (2020) relataram resultados favoráveis ao aplicarem regressão espacial e interpolação para calibrar e mapear a qualidade do ar regional, com melhor desempenho em ambientes de alta umidade relativa. Por outro lado, Jiao et al. (2016) ao comparar modelos de LCs, encontraram concordância com os equipamentos de referência sem necessidade de ajuste em alguns casos, enquanto outros só alcançaram resultados satisfatórios quando variáveis ambientais foram incorporadas em uma equação de regressão linear múltipla.

Ao testar os LCs da fabricante AQMesh, Castell et al. (2016) observaram que um bom desempenho em laboratório não necessariamente se traduz em bom desempenho em condições reais, mesmo em ambientes com condições ambientais estáveis. Nesse caso, o uso de fatores de conversão mais realistas, adaptados ao local, pode melhorar os resultados.

Além dessas incompatibilidades entre pesquisas, há uma escassez de dados quantitativos para comparar e avaliar a concordância entre os LCs e os dados de referência na maioria das revisões. Esses fatores indicam limitações no monitoramento de baixo custo que devem ser constantemente revisitadas.

## 2.2.4 Desafios no monitoramento de baixo custo

O desafio central do monitoramento de baixo custo é conseguir equiparar a qualidade dos dados medidos pelas estações de referência regulamentadas. Para isso, faz-se necessário atingir um padrão de qualidade das medições, definido para cada tipo de poluente. Governos dos EUA e da UE definiram limites de precisão para guiar o trabalho de pesquisadores e fabricantes em relação a aplicação de LCs, indicados na Tabela 6 e 7 respectivamente.

Tabela 6 - Sugestões de performance para sensores de baixo custo estabelecidos pela Agência de Proteção Ambiental Americana

Tipo de monitoramento	Máximo de incerteza associada a medição				Dados válidos
	SO <sub>2</sub> , CO	MP	O <sub>3</sub>	NO <sub>2</sub>	
Regulatório	10%	10%	7%	15%	>75%
Complementar (indicativo)	20%	20%	20%	20%	>80%
Educacional	50%	50%	50%	50%	>50%
Identificação de hotspots	30%	30%	30%	30%	>75%
Uso pessoal	30%	30%	30%	30%	>80%

Fonte: Adaptado de WILLIAMS et al. (2014)

Tabela 7 - Padrões de performance para sensores de baixo custo estabelecidos pela União Europeia

Tipo de monitoramento	Máximo de incerteza associada a medição			Dados válidos
	SO <sub>2</sub> , NO <sub>x</sub> , CO	MP	O <sub>3</sub>	
Regulatório	15%	25%	15%	>90%
Complementar (indicativo)	25%	50%	30%	>90%

Fonte: Adaptado de UNIAO EUROPEIA (2008)

Diferente dos EUA, onde não há diretrizes específicas para o uso de sensores suplementares no monitoramento regulatório, a diretiva 2008/50/EC da UE permite que medições indicativas complementem, mas não substituam, as medições fixas. No Brasil, ainda não existem padrões de desempenho definidos para esses sensores, uma vez que seu uso não é regulamentado.

A principal barreira para a adoção ampla dos sensores de baixo custo (LCs) é a precisão das medições, que é afetada pela natureza dos sensores e pelo modelo de coleta de dados (CONCAS et al., 2021). O desempenho dos LCs pode ser comprometido por sensibilidade cruzadas entre compostos e por condições climáticas específicas de cada local. Além disso, a

deriva, ou perda de precisão ao longo do tempo, varia conforme a vida útil de cada dispositivo (CAROTENUTO et al., 2023).

O custo-benefício é um aspecto central na revolução do monitoramento atmosférico. Se a precisão dos LCs for aceitável, redes mais amplas de LCs poderiam complementar as redes esparsas das agências reguladoras, oferecendo uma cobertura espaço-temporal mais extensa (MEAD et al., 2013, apud CAROTENUTO et al., 2023). Essas redes, operadas por longos períodos, poderiam revelar detalhes importantes sobre a poluição do ar, incluindo tendências temporais, pontos críticos e fontes de poluição, que redes esparsas não conseguiriam detectar (CASTELL et al., 2017).

Embora medidas de controle de qualidade sejam comuns em equipamentos de laboratório e estações robustas, essas não são suficientes para garantir a precisão dos LCs quando instalados em campo. Estudos recentes têm focado em reduzir imprecisões e incertezas nas medições dos LCs através de modelos de calibração e compensação. Modelos de regressão multivariada, tanto paramétricos quanto não paramétricos, têm mostrado resultados promissores (FENG et al., 2014). A eficácia desses modelos está na capacidade de considerar múltiplas variáveis, o que é benéfico para sensores de gases, cujas respostas são influenciadas por sensibilidades cruzadas e condições ambientais como temperatura, pressão e umidade relativa.

### **2.2.5 Métodos de correção e calibração no monitoramento de baixo custo**

De acordo com os protocolos da EPA, a calibração é o procedimento de verificar e ajustar as configurações de um LCs para que as medições sejam comparáveis às medições de referência. Para isso, a etapa de correção dos dados ajusta as medições dos sensores a fim de aumentar sua precisão em relação a esse valor conhecido (DUVALL et al., 2022).

A EPA define duas abordagens gerais para calibrar um LCs. A primeira é fazer uma calibração com padrões, onde é associado um padrão de referência amplamente aceito ao sensor. A segunda é fazer uma comparação com um instrumento de referência por meio da co-localização. A co-localização consiste em instalar o equipamento de baixo custo junto a uma estação de monitoramento de referência para fins de validação e correção (CAMPO, 2024). No presente trabalho foram realizadas análises via co-localização.

Diferentes técnicas de calibração e compensação para LCs vem sendo estudadas, e de forma geral, são divididas em três grupos: modelos paramétricos univariados, modelos paramétricos multivariados e modelos multivariados não-paramétricos. Os modelos

paramétricos univariados consideram apenas uma variável independente para prever a variável dependente, como a regressão linear univariada (ULR). Os modelos paramétricos multivariados consideram múltiplas variáveis independentes para prever a variável dependente, como a regressão multivariada (MLR). Nestes dois modelos, assume-se uma forma específica para a relação entre as variáveis. Já os modelos multivariados não paramétricos, consideram múltiplas variáveis independentes para prever a variável dependente, e não assumem uma forma específica para a relação entre as variáveis. As florestas aleatórias (RF) e redes neurais artificiais (ANN) são exemplos desses modelos.

Os métodos de regressão são populares devido à sua simplicidade e eficiência computacional. A ULR é um ponto de partida comum, mas as MLR são mais utilizadas quando a precisão adicional é necessária. Modelos de aprendizado de máquina (ML) são cada vez mais adotados por sua robustez, mas enfrentam desafios de generalização e eficiência computacional. Além da escolha do melhor modelo, a definição dos coeficientes de calibração pode interferir no resultado da calibração, por serem influenciados por condições ambientais variáveis (LIANG, 2021).

A duração da coleta de dados e da quantidade de amostras válidas afetam na estimativa dos melhores parâmetros ao empregar modelos de regressão. A degradação dos sensores e a variação sazonal do ambiente por sua vez, também podem afetar a precisão do modelo. Logo, para avaliar um modelo de calibração eficiente, deve-se conhecer o tipo de variável, a quantidade de variáveis independentes consideradas e se eles assumem uma forma de relação específica entre si (DESOUZA *et al.*, 2022).

A maioria dos algoritmos de calibração utiliza mais de uma variável no desenvolvimento do modelo, mas os efeitos benéficos dessa inclusão variam. Adicionar variáveis irrelevantes pode reduzir a importância das variáveis essenciais e diminuir o desempenho do modelo, destacando a importância da seleção de variáveis. Por esse motivo, faz-se necessário analisar as particularidades de cada caso (LIANG, 2021).

Dada a variedade de técnicas empregadas, é interessante comparar o desempenho daquelas que já tem sido utilizadas em modelos de calibração. Uma revisão conduzida por Liang (2021) compilou as técnicas empregadas por 24 estudos nos últimos anos, apontando os modelos de MLR como os mais utilizados. Entretanto, com os crescentes estudos em metodologias de calibração híbridas, outras técnicas de ML apresentam bons resultados, como é o caso de Mahajan e Kumar (2020) que reportaram uma melhoria de 10% de  $R^2$  entre os resultados de ULR e ANN.

Sob essas perspectivas, os modelos escolhidos para o presente trabalho são apresentados na Tabela 8.

Tabela 8 - Características dos modelos de regressão aplicados na etapa de calibração

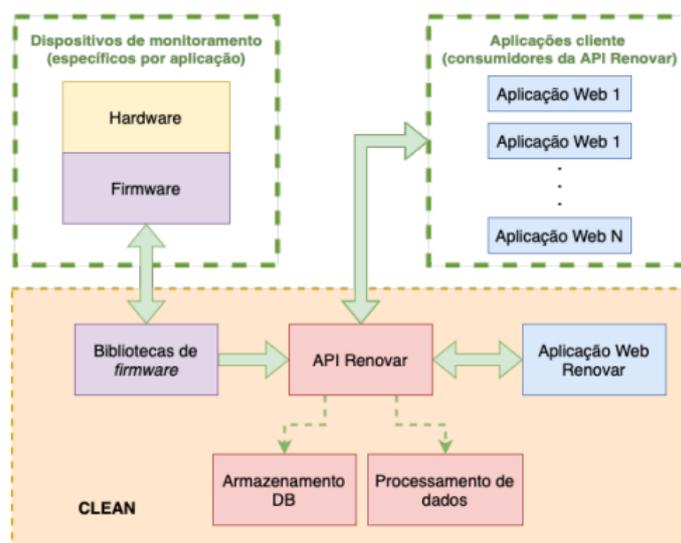
<b>Modelo de regressão</b>	<b>Modelo</b>	<b>Funcionamento</b>	<b>Vantagens</b>
Linear Multivariada	Paramétrico Multivariado	Ajusta uma linha reta (ou hiperplano) que melhor se encaixa aos dados, minimizando a soma dos quadrados dos erros.	Simples e fácil de interpretar; Eficiente para dados lineares.
K-Vizinhos mais Próximos (kNN)	Não-Paramétrico Multivariado	Prediz o valor de uma nova observação com base na média dos valores das k observações mais próximas no espaço de entrada.	Fácil de entender e implementar; Não assume nenhuma forma paramétrica dos dados.
Florestas Aleatórias	Não-Paramétrico Multivariado	Constrói uma "floresta" de árvores de decisão durante o treino e usa a média das previsões das árvores para fazer a predição.	Pode modelar relações complexas e não lineares; Resistente a overfitting (over ajuste); Lida bem com grandes conjuntos de dados e muitas features.
Perceptron Multicamadas (MLP)	Não-Paramétrico Multivariado	Usa uma rede neural com múltiplas camadas (camada de entrada, camadas ocultas e camada de saída) para modelar relações complexas nos dados.	Capaz de modelar relações altamente complexas e não lineares; Flexível e pode aprender representações complexas dos dados.

Fonte: Elaborado pela autora (2024)

### 2.3 A INICIATIVA CLEAN – COLLABORATIVE LOW-COST ENVIRONMENTAL AIR QUALITY NETWORK

A iniciativa CLEAN (Collaborative Low-cost Environmental Air-quality Network) tem como objetivo promover e facilitar o desenvolvimento de monitores de qualidade do ar de baixo custo por meio de uma plataforma colaborativa (CAMPO, 2023). A iniciativa é composta por quatro elementos principais: i) dispositivos de hardware, ii) firmware reutilizável, iii) guias e documentação para a reprodução do hardware e adesão à rede, e iv) a Plataforma Web Renovar para visualização e acesso remoto de dados em tempo real. Essas componentes são ilustradas graficamente na Figura 3.

Figura 3 - Componentes da iniciativa CLEAN

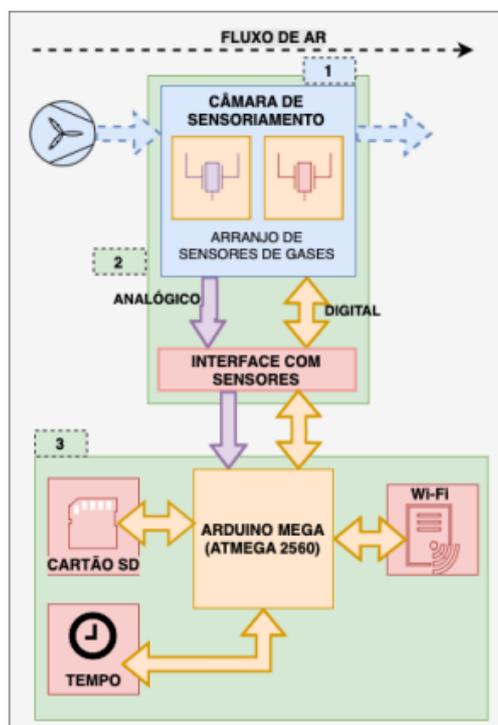


Fonte: CAMPO (2023)

O hardware e firmware são os dispositivos específicos para a etapa de coleta de dados. O hardware representa os dispositivos físicos responsáveis por esta fase, enquanto o firmware, que é o software integrado nesses dispositivos, controla suas operações. O firmware estabelece a conexão entre esses dispositivos e a API Renovar, uma Interface de Programação de Aplicações especializada. A Renovar atua como um hub central para a troca eficiente de dados entre os dispositivos de monitoramento e as aplicações dos usuários. Recebendo dados diretamente dos dispositivos, ela os processa e armazena de maneira organizada em um banco de dados centralizado. As informações coletadas são então acessadas e visualizadas através de uma aplicação web, que interage com a API para recuperar os dados necessários. Esta abordagem colaborativa permite que outras aplicações também acessem ou contribuam com dados para a plataforma Renovar, ampliando sua utilidade e impacto.

Como objeto de estudo, foram analisados os dados do equipamento CLEAN, com estrutura representada na Figura 4.

Figura 4 - Estrutura do equipamento de baixo custo CLEAN



Fonte: CAMPO (2023)

Os dispositivos são compostos por três componentes principais: 1) transporte de gás, 2) sensoriamento e 3) microcontrolador. No estágio de transporte de gás, o ar ambiente é capturado pelos sensores, os quais geram um sinal analógico proporcional à concentração do gás (CAMPO, 2023). O microcontrolador, integrado a uma placa Arduino Mega, recebe as respostas dos sensores e as converte em dados de concentração de gás. O hardware também registra a data, hora e a localização de cada medição. Essas informações são armazenadas em um cartão micro SD e transmitidas para um servidor web hospedado na Superintendência de Tecnologia da Informação e Comunicação da Universidade, utilizando o aplicativo Renovar Web. O dispositivo instalado é mostrado na Figura 5.

Figura 5 - Equipamento de baixo custo CLEAN instalado na Universidade Federal de Santa Catarina



Fonte: Elaborado pela autora (2024)

A placa eletrônica que abriga os sensores é montada de forma fixa na câmara de medição, com o objetivo de proteger os componentes eletrônicos e minimizar interferências externas. A alimentação do dispositivo é fornecida pela rede elétrica, embora haja a opção de utilizar baterias em situações de falta de energia ou quando o acesso à rede elétrica é limitado. Essa flexibilidade permite que o dispositivo mantenha seu funcionamento mesmo em condições adversas.

### 3. MATERIAIS E MÉTODOS

#### 3.1 DESCRIÇÃO DOS DADOS

Para o estudo em questão, foram utilizados dados de concentração de uma das estações CLEAN instaladas em campo, juntamente com dados de concentração de referência. Os dados foram coletados através do banco de dados Renovar, local de armazenamento das informações provenientes dos equipamentos CLEAN. Estes dispositivos são projetados para coletar em um período de amostragem de 15 minutos. A estação de monitoramento de referência registra dados diários, agrupados de forma horária. O período total analisado compreende os meses de setembro de 2023 a maio de 2024, considerando um período mínimo de coleta de três meses, conforme sugerido pela literatura (SPINELLE et al., 2013).

O dispositivo CLEAN avaliado foi instalado na cidade de Tubarão, no estado de Santa Catarina. A estação de referência é operada pela empresa Diamante Geração de Energia Ltda., igualmente localizada no município de Tubarão, no bairro Vila Moema (IMA/SC, 2023). Os dados da estação de referência foram obtidos através da base de dados pública do Instituto do Meio Ambiente de Santa Catarina. A Figura 6 mostra o mapa das estações.

Figura 6 - Mapa das estações de monitoramento CLEAN e referência



Fonte: IMA - SC (2023)

Na Tabela 9 estão dispostos os modelos de equipamento de referência da estação Vila Moema, no município de Tubarão - SC.

Tabela 9 - Relação de equipamentos de referência utilizados na estação Vila Moema

Equipamento	Poluentes	Princípio de operação	Fabricante
Monitor APNA-370	NO, NO <sub>2</sub> e NO <sub>x</sub>	Quimioluminescência	Horiba
Monitor APOA-370	O <sub>3</sub>	Adsorção ultravioleta	Horiba
Monitor APSA-370	SO <sub>2</sub>	Fluorescência ultravioleta	Horiba
Monitor APMA-370	CO	Modulação cruzada infravermelha sem dispersão	Horiba
Monitor BAM 1020	MP <sub>2,5</sub> , MP <sub>10</sub>	Atenuação de raios beta	Met One

Fonte: CAMPO (2024)

Para desenvolver a ferramenta de calibração, foram utilizados os dados dos poluentes monóxido de carbono (CO), dióxido de nitrogênio (NO<sub>2</sub>), dióxido de enxofre (SO<sub>2</sub>) e ozônio (O<sub>3</sub>), por serem alguns dos gases de referência para indicar a qualidade do ar, conforme definido pela Resolução nº 491/2018 do CONAMA. Na seção de resultados, serão abordados os resultados obtidos para o poluente CO, devido a maior presença de dados válidos no período. Dado os resultados similares observados entre os demais poluentes, os resultados estão dispostos nos apêndices.

## 3.2 PROCEDIMENTO DE CALIBRAÇÃO DE SISTEMAS DE MONITORAMENTO DA QUALIDADE DO AR DE BAIXO CUSTO

O procedimento de calibração foi separado em duas etapas: filtragem e detecção de falhas, e calibração dos dados. Para a realização de todo o processo foi utilizada a linguagem de programação Python, com o uso da ferramenta Anaconda *Spyder*.

### 3.2.1 Filtragem e detecção de falhas

A filtragem e detecção de falhas nos dados visa reduzir ruído, identificar anomalias e valores atípicos. Nessa etapa, os dados coletados foram filtrados, pré-processados e reamostrados em médias horárias antes de efetuar as rotinas de correção das leituras com base nas medições da estação de referência. Procedimentos descritos por Concas et al. (2021), Ottosen e Kumar (2019) e pelo Guia Técnico para Monitoramento e Avaliação da Qualidade do Ar (MMA, 2019) foram usados nessa etapa.

Esta etapa foi dividida em dois processos, sendo o primeiro denominado controle de qualidade, e que compreende os seguintes processos:

- Reamostragem dos dados para período de 15 minutos: As medições precisam ser reamostradas para uma resolução temporal adequada. Nesse caso, buscou-se reduzir o impacto das flutuações ao longo do tempo a partir do cálculo de médias em 15 minutos. Este processo garante um melhor desempenho dos métodos matemáticos de identificação de anomalias na segunda etapa.
- Remoção do período de estabilização: O período inicial após a instalação está sujeito a flutuações na medição, logo, os primeiros 7 dias foram desconsiderados a fim de evitar distorções nas análises (CAMPO, 2024).

- Remoção de valores fora dos limites de medição: Os valores abaixo da resolução do sensor e acima do valor máximo do sensor foram removidos a fim de remover falhas de medição.

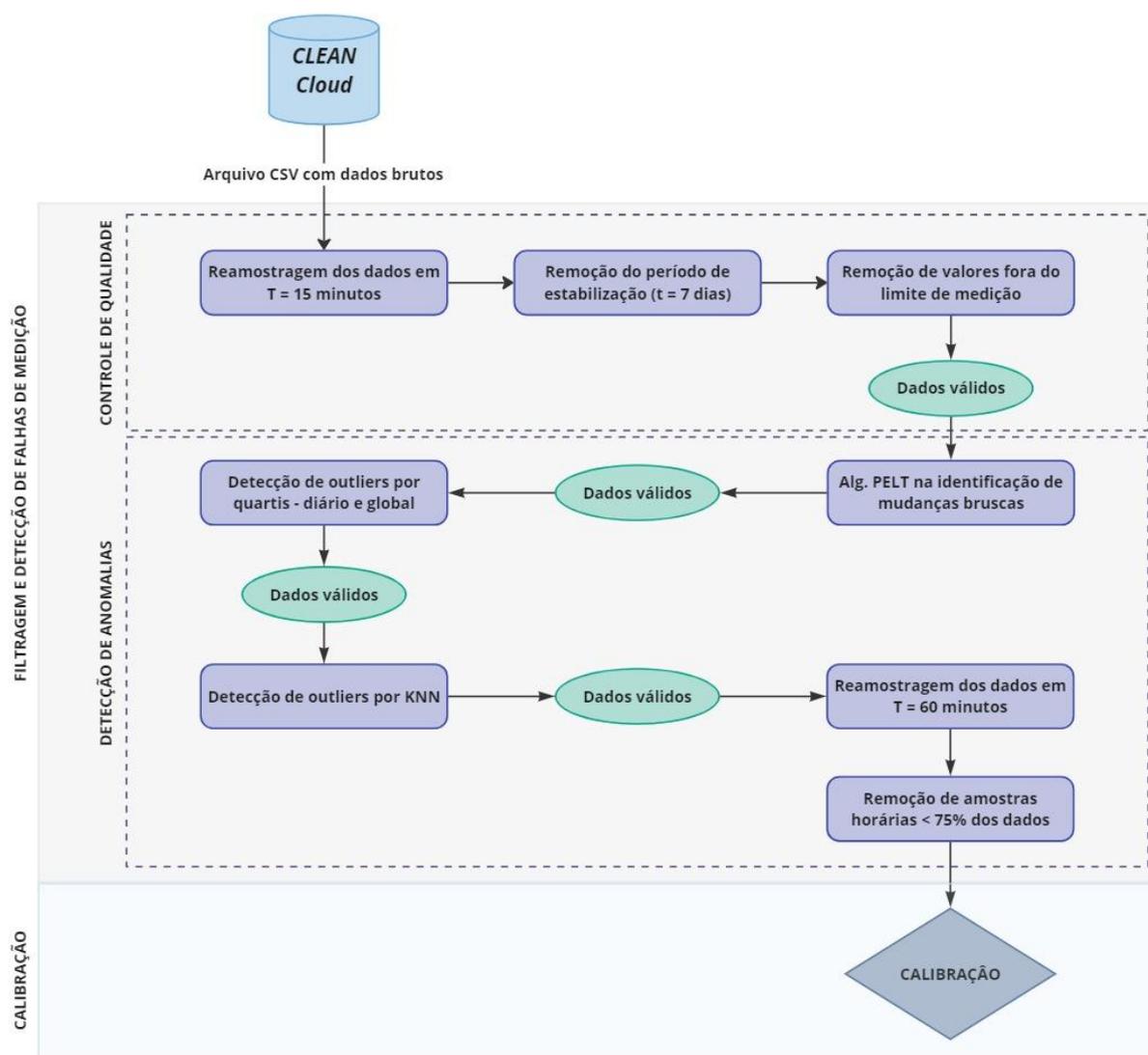
A segunda etapa, denominada de detecção de anomalias, segue o princípio dos seguintes processos:

- Análise de valores com alteração na linha de base: Em termos de séries temporais, a linha de base é o nível esperado de leitura quando não há eventos ou anomalias significativas. Para isso, foi aplicado o algoritmo PELT (Killick; Fearnhead; Eckley, 2012, apud CAMPO, 2024). Este método foi utilizado por Ottosen e Kumar (2019) para identificar mudanças abruptas na média e/ou na variância das séries temporais de sensores de gases. Para melhor desempenho do algoritmo, foi aplicada uma análise em grid para definir as constantes de entrada do modelo. Os parâmetros utilizados na análise em grid incluíram o valor de penalização, a largura da janela de análise e os limites de detecção de mudanças, ajustados conforme as variações nos dados de concentração de poluentes. Dados fora dos intervalos de melhor desempenho foram removidos.
- Identificação e remoção de outliers por quartis: A remoção dos percentis 1% e 99% dos dados foi realizada para mitigar a influência de valores extremos que podem introduzir viés e distorções na análise. Os outliers foram calculados globalmente (utilizando todos os dados do intervalo) e diariamente (analisando somente os intervalos de 24 horas). Foram submetidos para esta etapa, somente os dados válidos após a aplicação do método PELT.
- Identificação e remoção de outliers pelo método k-NN (*K-Nearest Neighbors* ou k-vizinhos mais próximos): O método calcula a distância entre um ponto e seu  $n$ -ésimo vizinho, e essa distância é utilizada como a pontuação de outlier. Quanto maior a distância, mais provável é que o ponto seja um outlier, já que ele está significativamente afastado de seus vizinhos imediatos. Com base na literatura e nas características dos sensores, foi adotado um intervalo máximo de 2 dias de medição para o cálculo das distâncias (OTTOSEN; KUMAR, 2019). Foram submetidos para esta etapa, somente os dados válidos após a identificação de outliers por quartis.

- Reamostragem dos dados para período de 1 hora: Após a filtragem das amostras, para fim de comparação e calibração junto aos dados de referência, foi aplicado uma reamostragem nos dados para um período de 1 hora.
- Validação de dados por intervalo temporal: Para garantir a integridade dos dados em médias horárias, foi verificado se cada intervalo de 1 hora continha pelo menos 75% dos dados de medições de 15 minutos. Este passo é sugerido pelos protocolos de calibração da EPA (DUVALL et al., 2022). Logo, apenas intervalos com, no mínimo,  $\frac{3}{4}$  de dados no cálculo das médias, foram considerados válidos.

As etapas desta fase são sumarizadas na Figura 7.

Figura 7 - Fluxograma das etapas de filtragem e detecção de falhas nas medições de dispositivos de monitoramento atmosférico de baixo custo



Fonte: Elaborado pela autora (2024)

Para cada estágio, foram utilizadas etiquetas para identificar os dados, garantindo que apenas amostras válidas parciais avancem para a próxima etapa de verificação. A Tabela 10 apresenta um resumo das etiquetas criadas.

Tabela 10 - Etiquetas de diferenciação dos dados com base nas etapas de pré-processamento

<b>Etiquetas</b>	<b>Tipo de dados</b>
VALID	Dados válidos em todas as etapas
MISSING	Dados faltantes
STAB	Dados gerados durante o período de estabilização de 7 dias. A fabricante recomenda, no mínimo, 2 dias de estabilização para sensores eletroquímicos
LOWSAMPLE	Intervalos horários com menos de 3 medidas de 15 minutos, ou seja, menos de 75% de representatividade
UPLIMIT	Dados acima do limite superior definido pelo fabricante dos sensores. Esse limite varia conforme o poluente analisado
LOWLIMIT	Dados abaixo do limite inferior definido pelo fabricante dos sensores. Esse limite varia conforme o poluente analisado
CHANGEPOINT	Indica quando houve mudança brusca na série de dados com base no algoritmo PELT
GQTL99	Dados pertencentes ao quartil 99% global no histograma dos dados
GQTL01	Dados pertencentes ao quartil 1% global no histograma dos dados
DQTL99	Dados pertencentes ao quartil 99% diário no histograma dos dados
DQTL01	Dados pertencentes ao quartil 1% diário no histograma dos dados
KNNOUT	Outliers detectados pelo método k-NN

Fonte: Elaborado pela autora (2024)

### 3.2.2 Calibração dos dados

A etapa final do procedimento implica na avaliação de modelos matemáticos que melhor convertam a concentração de poluentes medida pelos sensores CLEAN, na concentração de referência. A partir da literatura e das características dos sensores CLEAN, as técnicas utilizadas nesta etapa foram: regressão linear multivariada, regressão por k-vizinhos

mais próximos (kNN), regressão por florestas aleatórias e regressão com perceptron multicamadas (MLP).

Para identificar o modelo que melhor explicasse os dados, foi realizada uma busca em grid, combinando diferentes parâmetros e variáveis de entrada. Os parâmetros variam conforme os modelos: na regressão linear multivariada, não foram especificados parâmetros adicionais. Na regressão kNN, o parâmetro utilizado foi o número de vizinhos. Na regressão por florestas aleatórias, os parâmetros incluíram o número de estimadores, a profundidade máxima, o número mínimo de amostras para dividir um nó, e o número mínimo de amostras por folha. Na regressão MLP, foram usados os tamanhos das camadas ocultas e a regularização alpha. Já nas variáveis, foram considerados: os poluentes medidos pelo equipamento (CO, NO<sub>2</sub>, SO<sub>2</sub>, MP e O<sub>3</sub>) e as variáveis ambientais, também medidas pelo equipamento (temperatura externa, interna, pressão externa, interna e humidade). As medidas internas fazem referência àquelas medidas por sensores dentro da câmara de medição dos poluentes, enquanto as externas, são medidas externamente ao equipamento.

Os modelos foram avaliados utilizando validação cruzada com  $k = 3$ . Essa validação cruzada permite que os dados sejam separados em etapas de treino e teste, garantindo que os subconjuntos sejam testados somente uma vez, evitando viés.

Cada modelo foi ajustado com diferentes combinações de parâmetros específicas para cada técnica, também por meio da busca em grid. Por exemplo, para a MLP, foram testadas diferentes estruturas de rede e valores de regularização, enquanto para o KNN, foram testados diferentes números de vizinhos.

Durante a busca em grid, todas as combinações possíveis de variáveis de entrada foram geradas. Essa busca é uma técnica utilizada para encontrar os melhores parâmetros de um modelo de aprendizado de máquina. O processo envolve definir um conjunto de valores possíveis para cada parâmetro e, em seguida, testar todas as combinações desses valores. Cada combinação é avaliada usando validação cruzada para medir o desempenho do modelo. O objetivo é identificar a combinação de parâmetros que proporciona o melhor desempenho do modelo em termos de métricas de avaliação. Garantiu-se que cada combinação incluísse as medidas do sensor do poluente em questão, e cada sensor representasse apenas uma entrada.

Cada combinação submetida aos modelos de regressão foram avaliadas pelas seguintes métricas: coeficiente de determinação ( $R^2$ ), erro médio quadrático (RMSE), erro absoluto médio (MAE), critério de informação de Akaike (AIC) e critério de informação Bayesiano (BIC). O valor de  $R^2$  mede a proporção da variabilidade nos dados que é explicada pelo modelo,

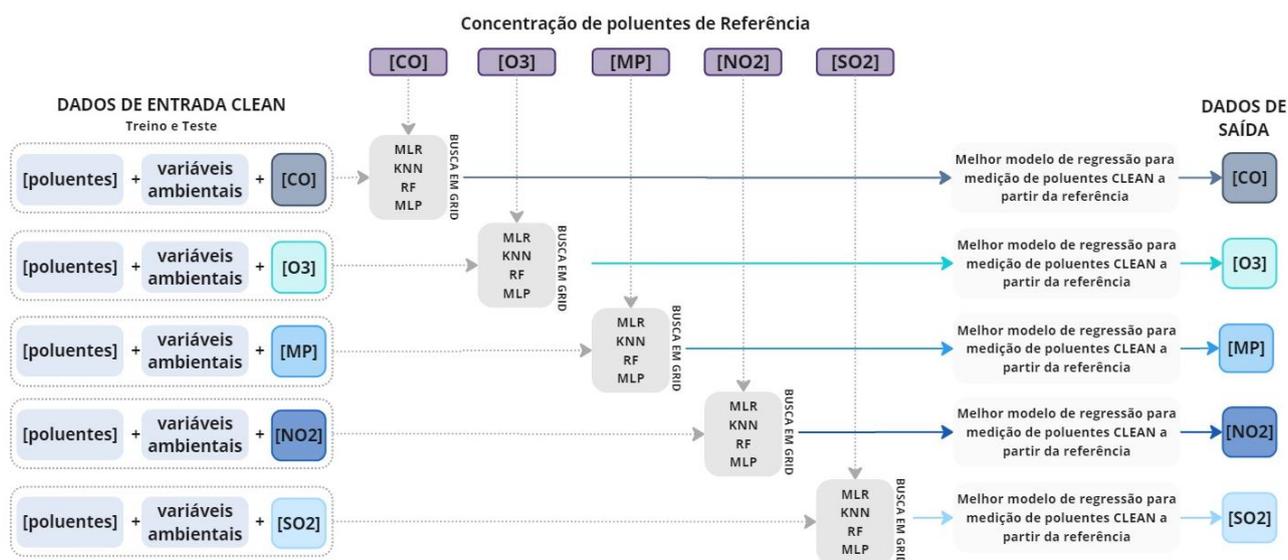
variando de 0 a 1; quanto mais próximo de 1, melhor o modelo é capaz de explicar os dados, indicando uma maior precisão na correção das leituras de concentração de poluentes.

O RMSE mede a raiz da média dos quadrados dos erros entre os valores previstos pelo modelo e os valores reais observados. Quanto menor o valor do RMSE, melhor o modelo é capaz de prever os valores reais de concentração. O MAE, por sua vez, calcula a média das diferenças absolutas entre os valores previstos pelo modelo e os valores reais observados.

Os valores de AIC e BIC são utilizados para avaliar a qualidade dos modelos de predição de concentração de poluentes, penalizando a complexidade do modelo. O AIC estima a qualidade de um modelo em relação a outros modelos, favorecendo modelos mais simples que explicam bem os dados de concentração. O BIC, similar ao AIC, também penaliza modelos mais complexos, mas de maneira mais rigorosa. Modelos com menores valores de AIC e BIC são considerados melhores, pois indicam um equilíbrio ideal entre a complexidade do modelo e seu ajuste aos dados, crucial para a correção das leituras de concentração de poluentes.

Modelos com maior  $R^2$  e menores RMSE e MAE foram considerados os de melhor desempenho. Esse processo matematicamente robusto e detalhado auxilia no entendimento do comportamento dos poluentes medidos pelos equipamentos CLEAN, com os demais poluentes no ar, variáveis ambientais e dados de referência. A Figura 8 apresenta um fluxograma da metodologia aplicada.

Figura 8 - Fluxograma da etapa de calibração de dispositivos de monitoramento atmosférico de baixo custo



Fonte: Elaborado pela autora (2024)

## 4. RESULTADOS

As etapas do procedimento de filtragem, detecção de falhas e calibração foram aplicadas para os sensores de poluentes do equipamento CLEAN, a fim de avaliar o desempenho dos métodos escolhidos e possíveis correções. Neste momento serão discutidos os resultados para o sensor CO-B4. Os demais resultados estão disponíveis nos apêndices.

### 4.1 CARACTERIZAÇÃO, FILTRAGEM E DETECÇÃO DE FALHAS

O sensor CO-B4 instalado no equipamento CLEAN registrou dentro do período de setembro de 2023 a maio de 2024, 17.852 dados de concentração. A Tabela 11 apresenta o número de amostras e o desempenho de cada etapa dos processos de controle de qualidade. Para ambas as análises, foram usados os dados agrupados em 15 minutos, para posterior reamostragem de 1 hora.

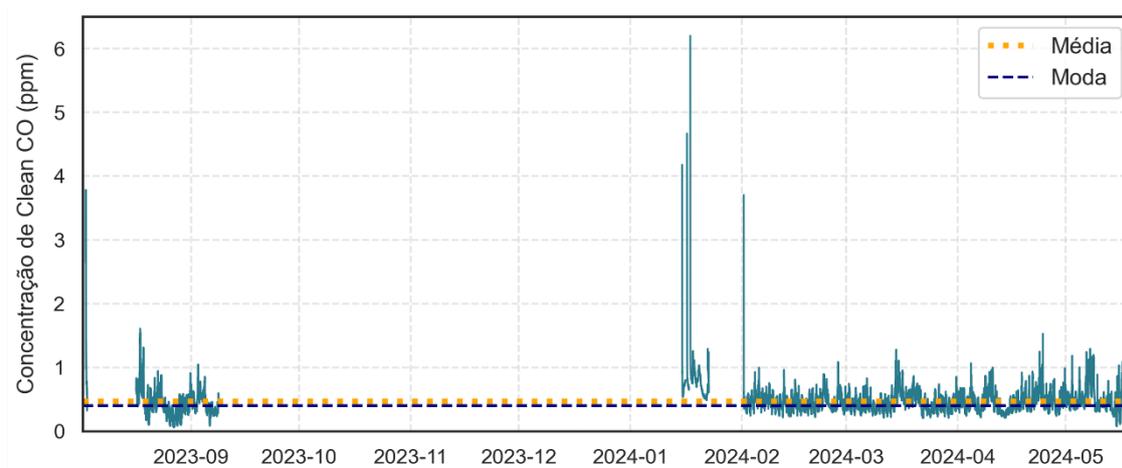
Tabela 11 - Resultados obtidos após o processo de controle de qualidade do sensor CO-B4

Etapa	Etiquetas	Nº amostras	% amostras
	<b>Total</b>	<b>17852</b>	<b>100%</b>
Caracterização por limites de medição	STAB	46	0.3%
	MISSING	4591	26%
	LOWLIMIT	-	-
	<i>Amostras válidas parciais</i>	<i>13215</i>	<i>74%</i>

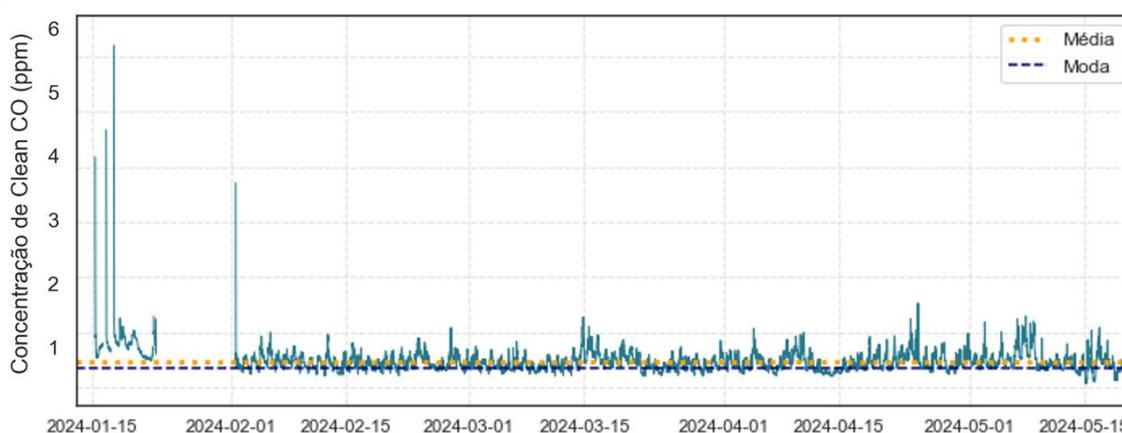
Fonte: Elaborado pela autora (2024)

Nota-se que após o processo de controle de qualidade, 74% dos dados totais são classificados como válidos. O comportamento da amostra antes e depois de ser submetida ao processo é demonstrado na Figura 9a e 9b.

Figura 9 - Série temporal do sensor CO-B4 antes e após ser submetido ao processo de controle de qualidade, com dados agrupados em 15 minutos



a) Série temporal do sensor CO-B4 anterior ao processo de controle de qualidade, com indicação dos valores de média e moda do intervalo. Os dados são agrupados em média de 15 minutos.



b) Série temporal do sensor CO-B4 posterior ao processo de controle de qualidade, com indicação dos valores de média e moda do intervalo. Os dados são agrupados em média de 15 minutos.

Fonte: Elaborado pela autora (2024)

Os dados coletados em 2023 foram retirados da amostra devido os longos períodos sem registro, em decorrência, principalmente, da falta de conexão à internet na estação de monitoramento. Contudo, um período mínimo de 3 meses de dados ainda foi mantido.

No processo de detecção de anomalias, a aplicação do método PELT resultou em 81% das amostras válidas, como mostrado na Tabela 12. Nota-se que somente a fase de controle de qualidade não é capaz de identificar todos os comportamentos discrepantes na amostra devido aos dados removidos ao longo do processo de detecção de anomalias.

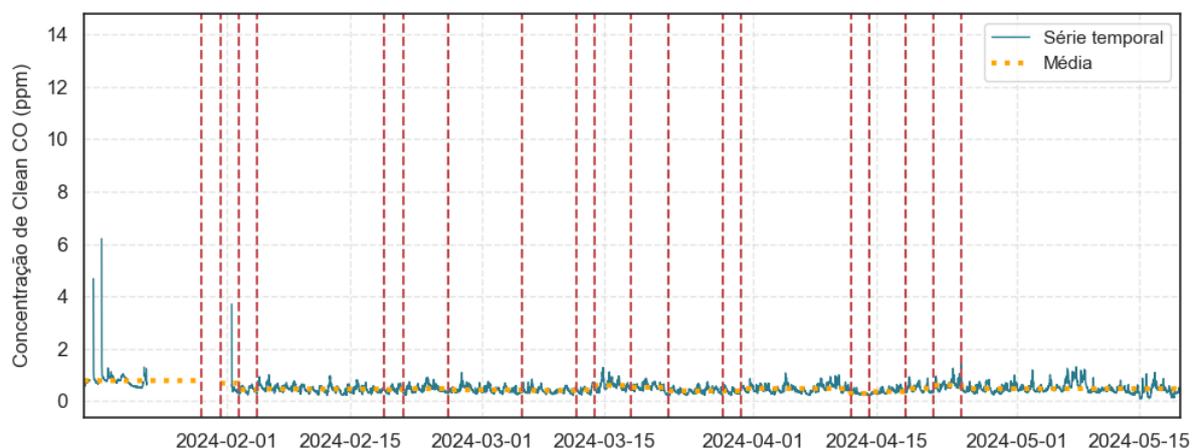
Tabela 12 - Resultados obtidos após o processo de detecção de anomalias do sensor CO-B4

<b>Etapa</b>	<b>Etiquetas</b>	<b>Nº amostras</b>	<b>% amostras</b>
	<b>Total - Intervalo definido por PELT*</b>	<b>10709</b>	<b>81%</b>
Correção por mudanças de base	LOWLIMIT	-	-
	STAB	-	-
	MISSING	327	3%
	<i>Amostras válidas parciais</i>	<i>10382</i>	
Correção por outlier por quartil	GQTL01	104	1%
	GQTL99	104	1%
	DQTL99	46	0.4%
	DQTL01	28	0.3%
	<i>Amostras válidas parciais</i>	<i>10100</i>	
Correção por outlier por KNN	KNN_OUT	262	3%
	<b>Amostras válidas finais</b>	<b>9838</b>	<b>92%</b>

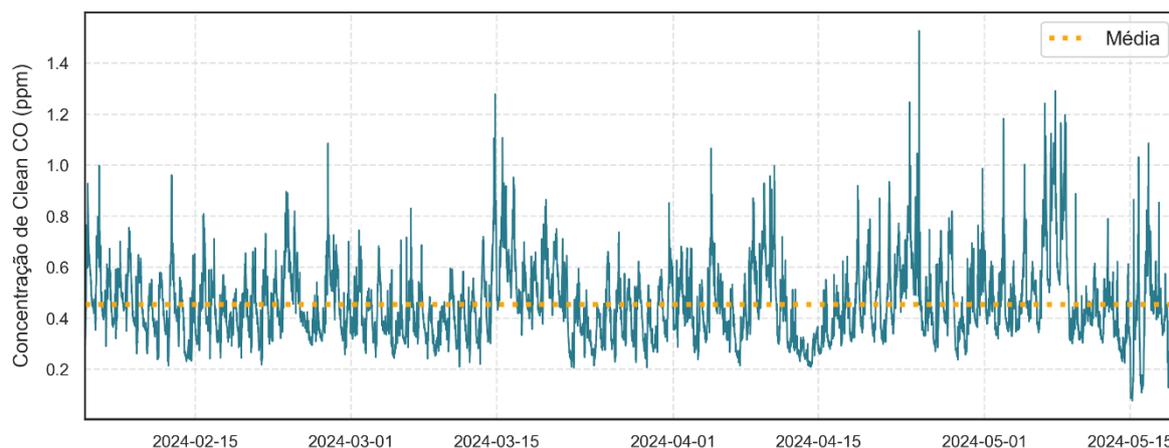
Fonte: Elaborado pela autora (2024)

Após submeter a amostras às duas etapas, 55% dos dados totais foram definidos como válidos para seguir com o processo de calibração, o que corresponde à 92% da amostra válida definida por PELT. A Figura 10a e 10b apresenta as janelas de segmentação na série temporal, e o resultado após a eliminação de janelas com alteração brusca de linha de base. Para dados agrupados em médias de 15 minutos, foram identificados 22 pontos de mudança da linha de base.

Figura 10- Pontos de alteração da linha base do sensor CO-B4 detectados por PELT conforme série temporal, e janelas de dados válidos mantidas após a aplicação do algoritmo, com dados agrupados em 15 minutos



a) Série temporal do sensor CO-B4 com pontos de alteração da linha de base definidos através do algoritmo PELT, com dados agrupados em média de 15 minutos. Os traços verticais em vermelho indicam os pontos onde houve mudanças significativas na linha base. A média foi calculada com base no intervalo entre duas mudanças bruscas seguidas.

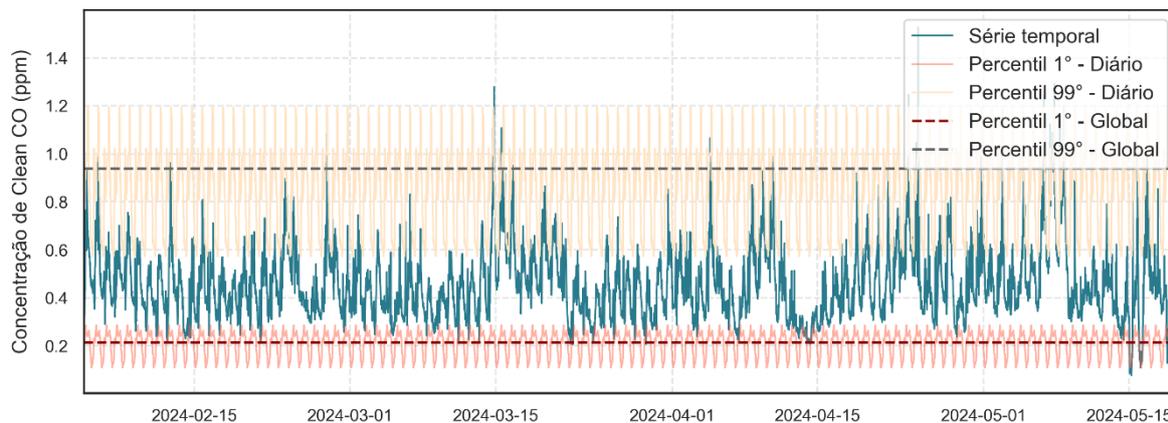


b) Série temporal do sensor CO-B4 com amostra considerada válida após a aplicação do algoritmo PELT, com dados agrupados em média de 15 minutos. Os intervalos escolhidos para serem mantidos no processo foram definidos com base na média horária.

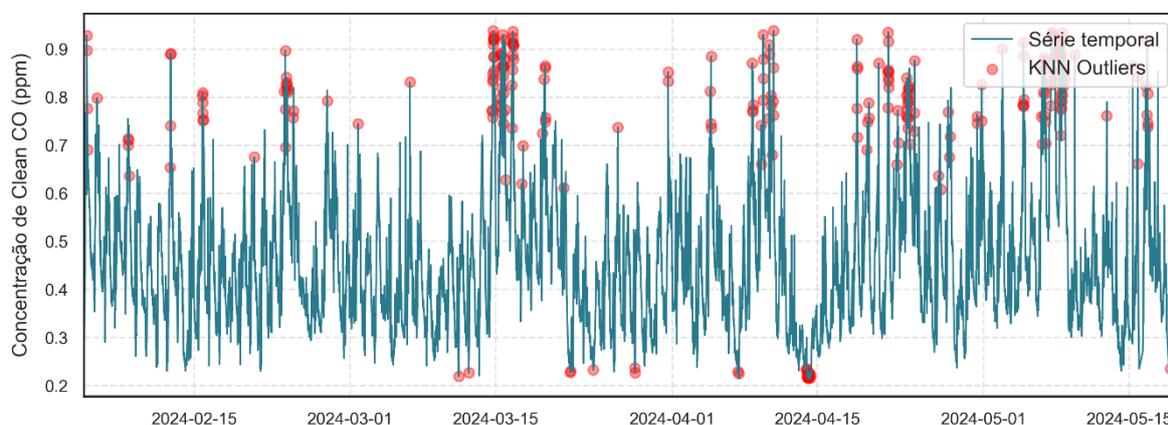
Fonte: Elaborado pela autora (2024)

O intervalo válido da amostra seguiu para a detecção de outliers por quartis. Para serem classificados como válidos, os dados deveriam estar entre as faixas de quartis global e diário de 1% e 99%. Após esse processo, houve ainda um refinamento dos dados pelo método de *k*-vizinhos mais próximos. Os resultados são apresentados na Figura 11a e 11b.

Figura 11 – Detecção de outliers por quartil e k-vizinhos mais próximos do sensor CO-B4 conforme série temporal de dados, com dados agrupados em 15 minutos



a) Série temporal com detecção de outliers por quartis global e diário do sensor CO-B4, com dados agrupados em 15 minutos. A série conta com os dados válidos após a detecção de mudanças bruscas por PELT.

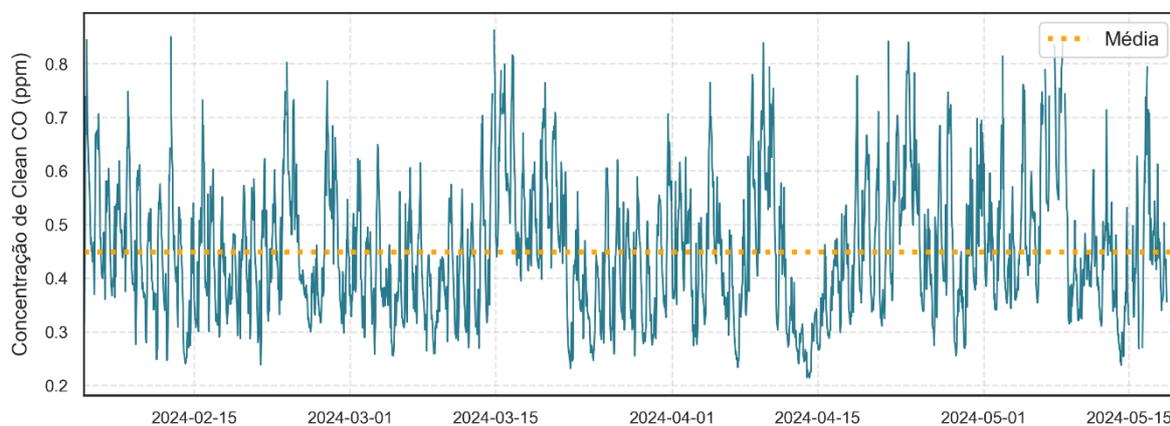


b) Série temporal com detecção de outliers pelo método de k-vizinhos mais próximos do sensor CO-B4, com dados agrupados em 15 minutos. A série conta com os dados validados após a detecção de outliers por quartis.

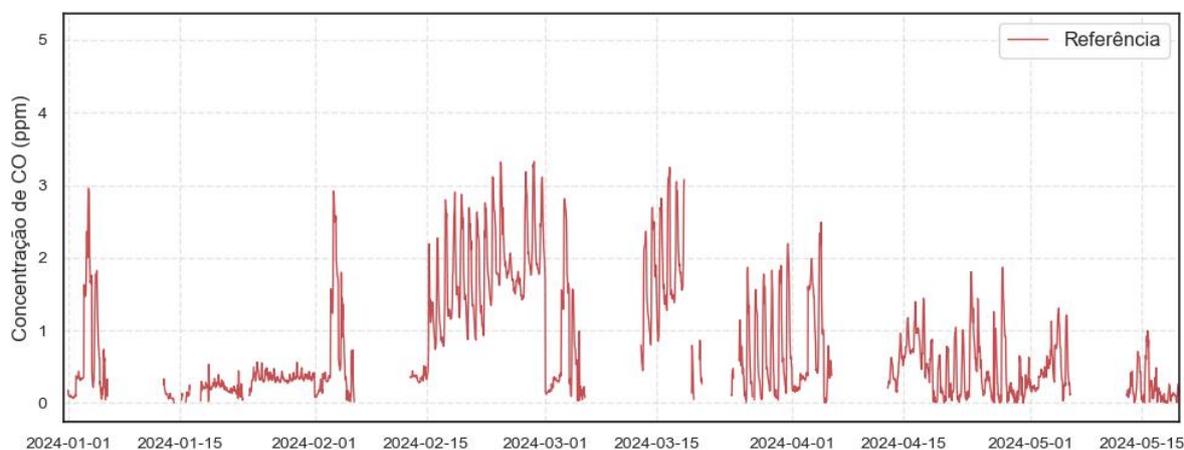
Fonte: Elaborado pela autora (2024)

Após a correção dos dados, foi realizada uma reamostragem de 1 hora para possibilitar a comparação com os dados de referência, como mostra a Figura 12a, 12b e 12c. Essa reamostragem possibilita a etapa de calibração, além de suavizar as flutuações das leituras. Nota-se que o sensor CO-B4 não acompanha as grandes flutuações das medições de referência, subestimando ou superestimando os valores mais discrepantes.

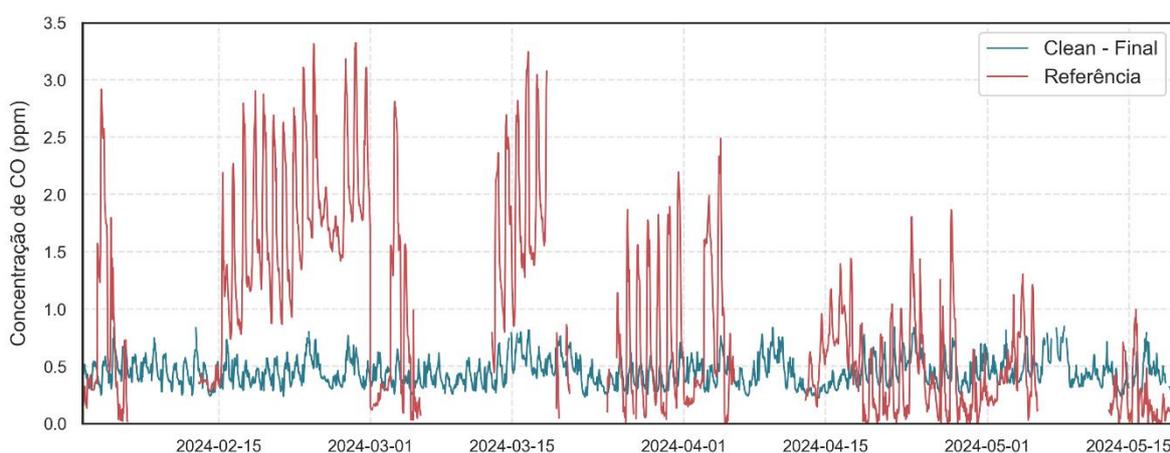
Figura 12 - Série temporal final do sensor CO-B4 após etapas de filtragem e remoção de falhas, com dados agrupados em médias horárias



a) Série temporal final do sensor CO-B4 após etapas de filtragem e detecção de falhas, com dados agrupados em médias horárias.



b) Série temporal do poluente CO medido pela estação de referência.

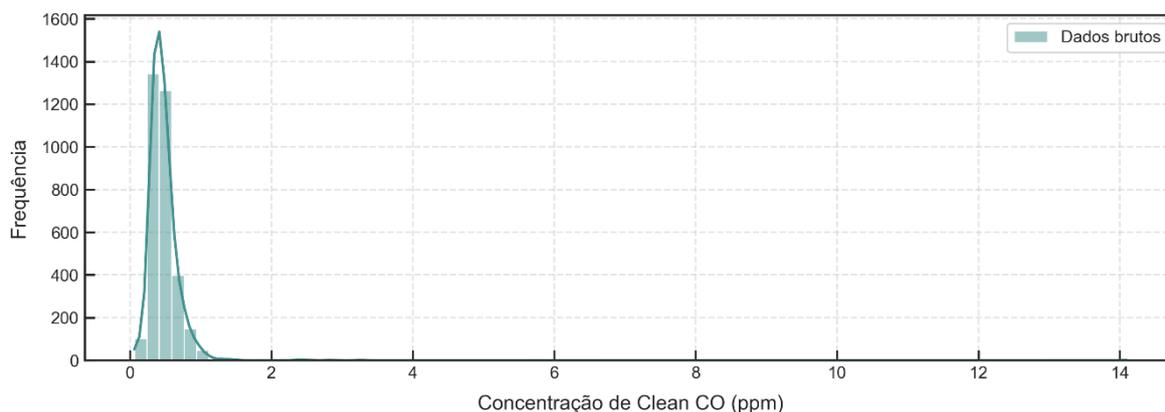


c) Série temporal final do sensor CO-B4 em relação à série temporal das concentrações de CO medidas pela estação de referência, após etapas de filtragem e detecção de falhas, com dados agrupados em médias horárias.

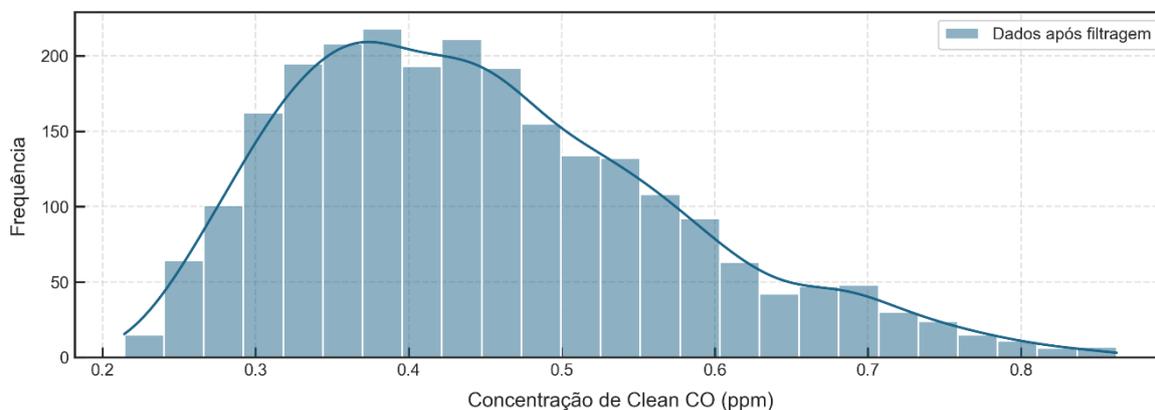
Fonte: Elaborado pela autora (2024)

A partir da análise do histograma da Figura 13a e 13b, observa-se que os dados brutos apresentaram um comportamento log-normal, um padrão para dados de concentração. Após os procedimentos de filtragem e remoção de falhas, os dados passam a se assemelhar a um comportamento normal. Essa característica possibilita a aplicação de modelos de calibração.

Figura 13 - Histograma dos dados do sensor CO-B4 antes e após as etapas de filtragem e remoção de falhas, com dados agrupados em médias horárias



a) Histograma do sensor CO-B4 anteriormente às etapas de filtragem e remoção de falhas, com dados agrupados em médias horárias. O histograma apresenta esse comportamento devido a presença de outliers extremos na amostra.

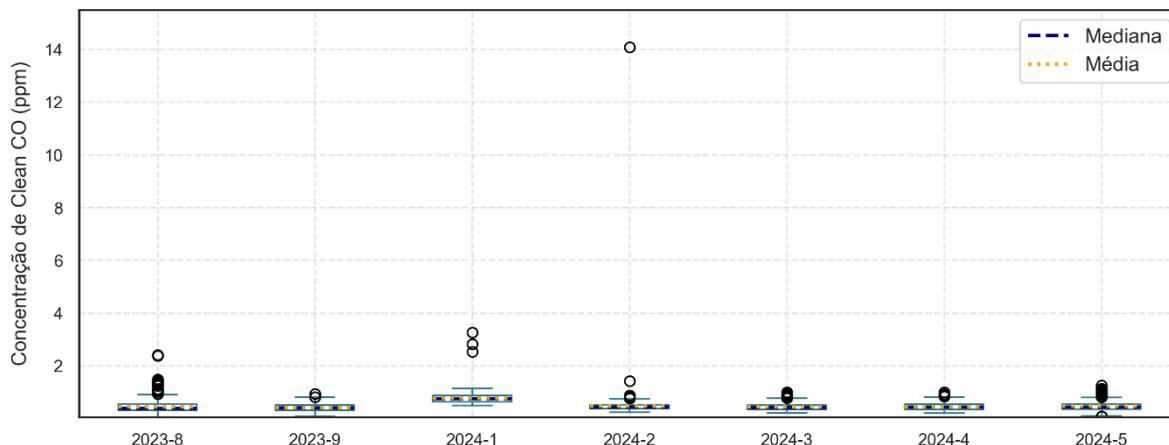


b) Histograma do sensor CO-B4 posteriormente às etapas de filtragem e remoção de falhas, com dados agrupados em médias horárias. Com a remoção dos dados extremos, o comportamento dos dados é suavizado.

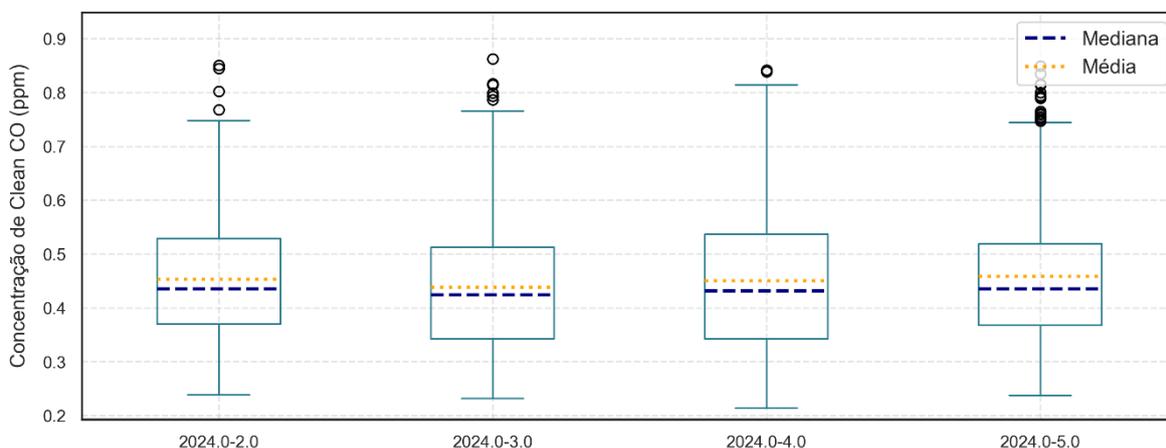
Fonte: Elaborado pela autora (2024)

A fim de comparar o comportamento estatístico da amostra anterior e posteriormente aos processos de filtragem e remoção de falhas, também foram realizadas análises de box-plot dos dados, como mostra a Figura 14a e 14b. Nota-se que nos primeiros meses existem pontos discrepantes, o que também explica a característica do histograma inicial.

Figura 14 - Box-plot mensal das concentrações do sensor CO-B4 antes e depois das etapas de filtragem e remoção de falhas



a) Box-plot mensal das concentrações do sensor CO-B4 anterior às etapas de filtragem e remoção de falhas. Devido à presença de outliers extremos, há um achatamento do gráfico.



b) Box-plot mensal das concentrações do sensor CO-B4 posterior às etapas de filtragem e remoção de falhas

Na Tabela 13 são encontradas as estatísticas das amostras do sensor CO-B4 ao longo das etapas de filtragem e detecção de falhas para dados agrupados em médias de 15 minutos.

Tabela 13 - Análise estatísticas das etapas de controle de qualidade e detecção de anomalias

Estatística	Dados brutos	Dados após controle de qualidade	Dados após detecção de anomalias
Amostras	13261.0	13215.0	9838.00
Média	0.47	0.48	0.44
Desvio Padrão	0.23	0.23	0.12
Mínimo	0.06	0.08	0.21
25%	0.35	0.36	0.35
50%	0.44	0.44	0.43
75%	0.55	0.55	0.52
Máximo	14.10	14.10	0.87

Fonte: Elaborado pela autora (2024)

Já a Tabela 14 apresenta as estatísticas das amostras do sensor CO-B4 ao longo das etapas para dados agrupados em médias horárias.

Tabela 14 - Análise estatísticas das etapas de controle de qualidade e detecção de anomalias

<b>Estatística</b>	<b>Dados brutos</b>	<b>Dados após controle de qualidade</b>	<b>Dados após detecção de anomalias</b>
Amostras	3323.0	2758.0	2540.0
Média	0.47	0.48	0.45
Desvio Padrão	0.31	0.18	0.12
Mínimo	0.06	0.09	0.21
25%	0.35	0.36	0.36
50%	0.44	0.44	0.43
75%	0.55	0.55	0.53
Máximo	14.10	3.26	0.85

Fonte: Elaborado pela autora (2024)

É possível observar que não há alterações significativas no valor de média entre as etapas, mas que a remoção de anomalias influencia nos valores máximos e mínimos do intervalo.

## 4.2 CALIBRAÇÃO DOS SENSORES

Após a filtragem e remoção de falhas, os dados válidos seguiram para a calibração. As variáveis de entrada incluíram: concentração de CO e O<sub>3</sub>, temperatura interna, pressão interna, temperatura externa e umidade externa, todas medidas pelo dispositivo CLEAN. As variáveis internas são dos sensores alocados dentro da câmara de medição do dispositivo. Outras concentrações de poluentes não foram consideradas devido à falta de dados no período de medição de CO. A concentração de CO medida pela estação de referência foi a variável alvo para a calibração.

Os melhores resultados da busca em grid, em termos de  $R^2$ , RMSE e MAE, para os modelos de regressão são descritos na Tabela 15.

Tabela 15 - Resultados da calibração do sensor CO-B4 considerando as melhores métricas obtidas em termos de  $R^2$ , RMSE e MAE.

Modelo	Variáveis de entrada	R2	RMSE	MAE	AIC	BIC
Regressão multivariada	CO, Temp. Interna, Temp. Externa,	0.191	0.530	0.428	813.426	799.986
Regressão multivariada	CO, O3, Temp. Interna, Temp. Externa,	0.188	0.531	0.428	812.490	794.569
Regressão multivariada	CO, Temp. Interna, Umidade Interna, Temp. Externa,	0.172	0.534	0.432	818.472	800.551
Regressão multivariada	CO, O3, Temp. Interna, Umidade Interna, Temp. Externa,	0.171	0.534	0.432	817.062	794.661
Regressão multivariada	CO, O3, Temp. Interna, Press. Interna, Temp. Externa,	0.151	0.539	0.434	814.390	791.989
Regressão multivariada	CO, Temp. Interna, Press. Interna, Temp. Externa,	0.149	0.540	0.434	815.132	797.212
Regressão multivariada	CO, O3, Temp. Interna, Umidade Interna, Press. Interna, Temp. Externa,	0.113	0.546	0.441	819.777	792.896
Regressão multivariada	CO, Temp. Interna, Umidade Interna, Press. Interna, Temp. Externa,	0.109	0.547	0.441	821.041	798.641
Regressão MLP	CO, O3, Temp. Interna, Temp. Externa,	0.102	0.554	0.447	875.623	857.703
Regressão MLP	CO, O3, Temp. Interna, Umidade Interna, Temp. Externa,	0.090	0.554	0.451	849.532	827.132
Florestas Aleatórias	CO, Temp. Interna, Temp. Externa,	-0.005	0.573	0.467	885.365	871.925
Florestas Aleatórias	CO, O3, Temp. Interna, Umidade Interna, Temp. Externa,	-0.008	0.573	0.464	901.882	879.482
Florestas Aleatórias	CO, Temp. Interna, Umidade Interna, Temp. Externa,	-0.016	0.574	0.468	905.168	887.248
Florestas Aleatórias	CO, O3, Temp. Interna, Temp. Externa,	-0.019	0.576	0.466	888.878	870.958
Regressão multivariada	CO, O3, Temp. Interna, Umidade Interna,	-0.051	0.585	0.471	750.495	732.574

Fonte: Elaborado pela autora (2024)

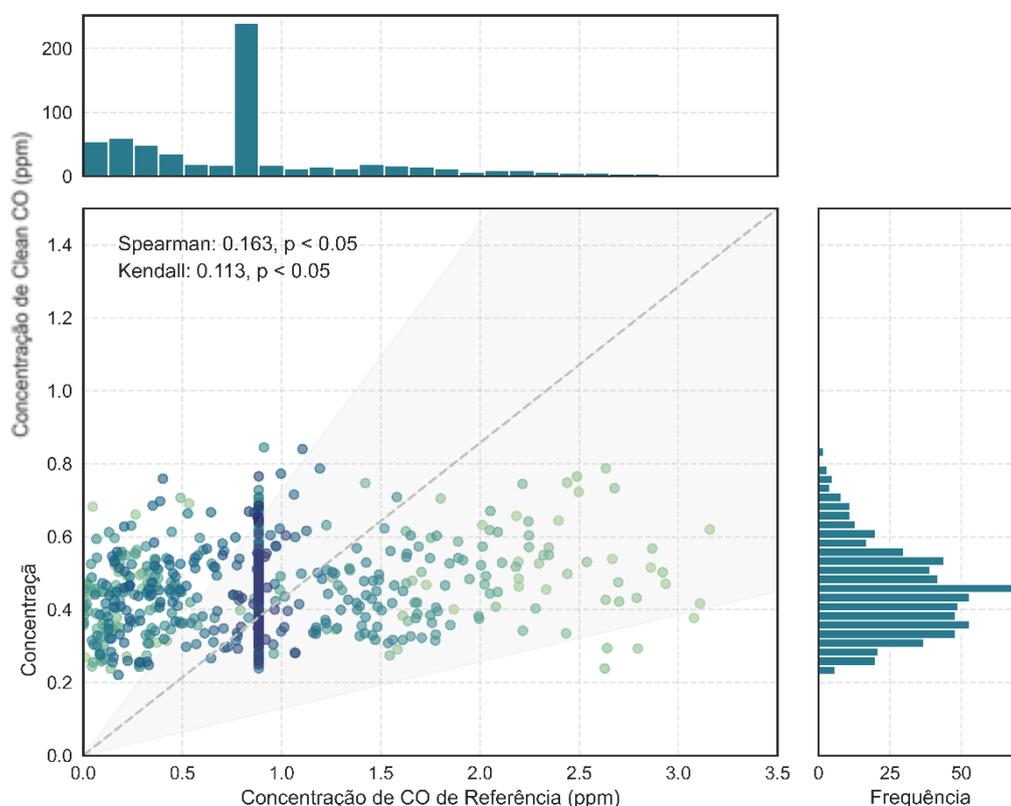
O modelo com melhor desempenho em relação ao  $R^2$ , RMSE e MAE foi a regressão multivariada. Isso pode ser explicado pela característica do próprio poluente CO, uma vez que em grandes concentrações os modelos multivariados possuem melhor desempenho. Os modelos floresta aleatória, k-vizinhos mais próximos e regressão MLP apresentaram desempenho inferior devido a possíveis problemas como overfitting, configuração inadequada de hiperparâmetros e alta variabilidade no desempenho.

Os dados preditos foram submetidos aos testes de correlação de Spearman e Kendall, a fim de comparar o desempenho do modelo na estimativa dos dados. Os testes foram realizados com base nos dois melhores modelos gerados pela calibração: o modelo de regressão

multivariada utilizando como parâmetros as variáveis de CO, temperatura interna e temperatura externa; e o modelo de regressão multivariada utilizando como parâmetros as variáveis de CO, O<sub>3</sub>, temperatura interna e temperatura externa.

Primeiramente, os dados de teste de concentração de CO, anterior a aplicação do modelo, foram correlacionados com os dados de concentração provenientes da estação de referência, como mostrado na Figura 15. Os coeficientes de Spearman e Kendall foram de 0.163 e 0.113 respectivamente, implicando em uma baixa correlação positiva.

Figura 15 - Correlação entre as leituras do sensor CO-B4 e a estação de referência antes da aplicação dos modelos de regressão. Histogramas das concentrações de CO-B4 inicial e de referência.

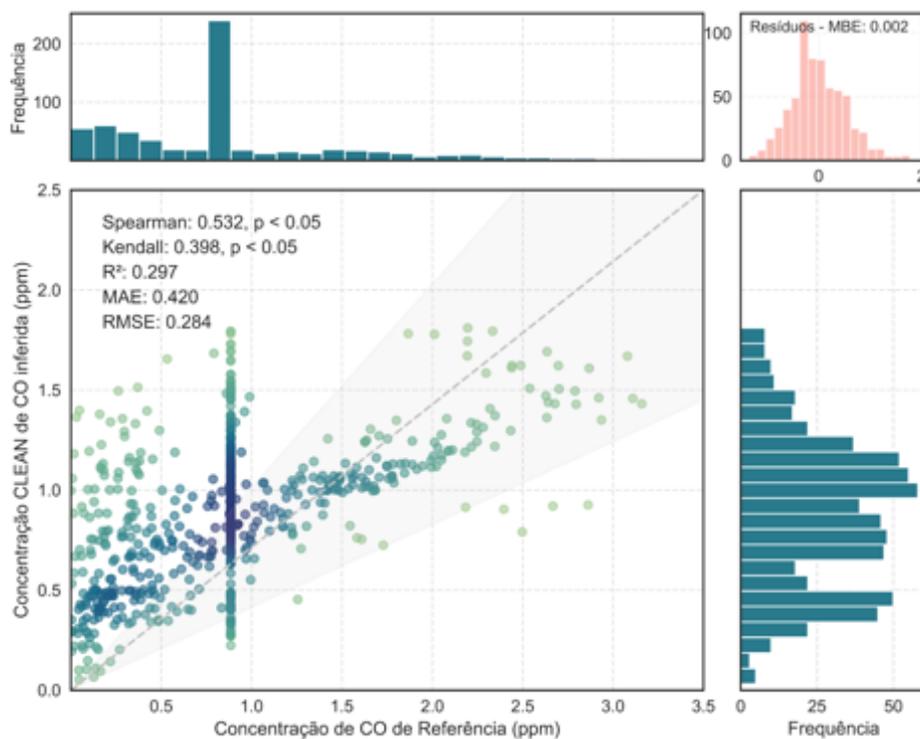


Fonte: Elaborado pela autora (2024)

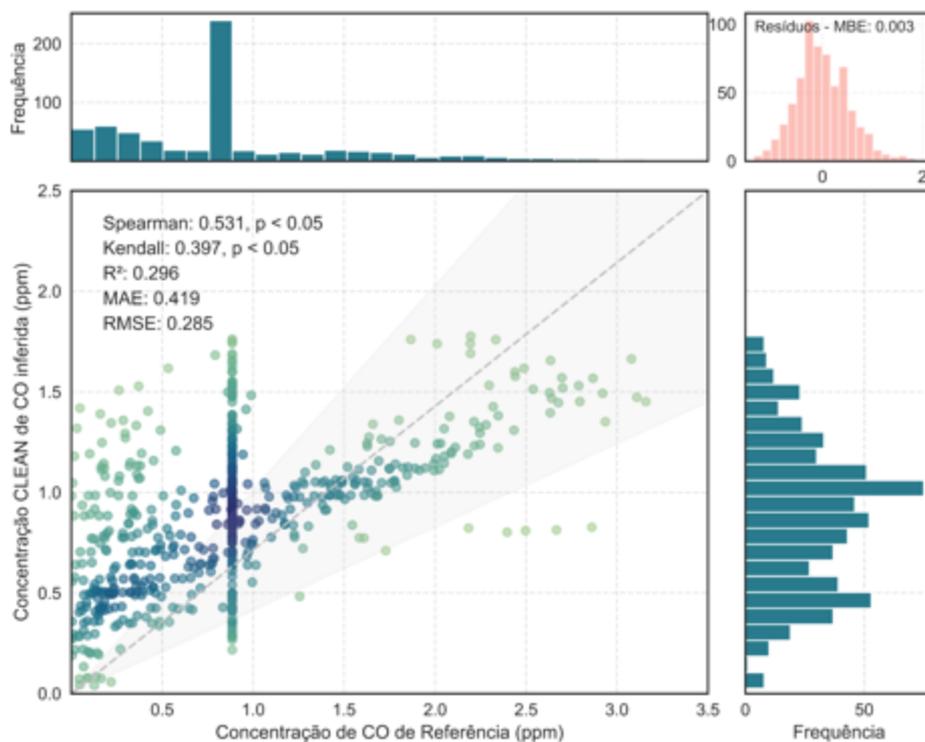
A Figura 16a e 16b apresenta os resultados de correlação após a aplicação do modelo de regressão com os dois melhores resultados para as métricas de qualidade. No primeiro caso, obtiveram-se os valores de 0.532 para o coeficiente de correlação de Spearman, e 0.398 para o coeficiente de Kendall, indicando uma leve correlação positiva. O valor  $p < 0.05$  sugere que essa correlação é estatisticamente significativa. O coeficiente de determinação,  $R^2$ , é 0.297, indicando que aproximadamente 29.7% da variabilidade na concentração de CO inferida pelo sensor pode ser explicada pela concentração de CO de referência. O erro absoluto médio indica

que, em média, a diferença absoluta entre os valores previstos pelo sensor e os valores de referência é 0.420 ppm. Já o RMSE de 0.284 pode ser considerado baixo, dadas as características das medições. Nota-se que após a adição da concentração de O<sub>3</sub> nas variáveis de entrada, as métricas não sofrem mudanças significativas.

Figura 16 - Comportamento das leituras do sensor CO-B4 e dos dados de referência após a aplicação dos dois melhores modelo de regressão



a) Correlação entre os dados de CO-B4 previsto pelo melhor modelo de calibração e os dados de CO de referência. Os histogramas fazem referência às concentrações de CO-B4 inferidas, de referência e dos resíduos após a calibração.



b) Correlação entre os dados de CO-B4 previsto pelo segundo melhor modelo de calibração e os dados de CO de referência. Os histogramas fazem referência às concentrações de CO-B4 inferidas, de referência e dos resíduos após a calibração.

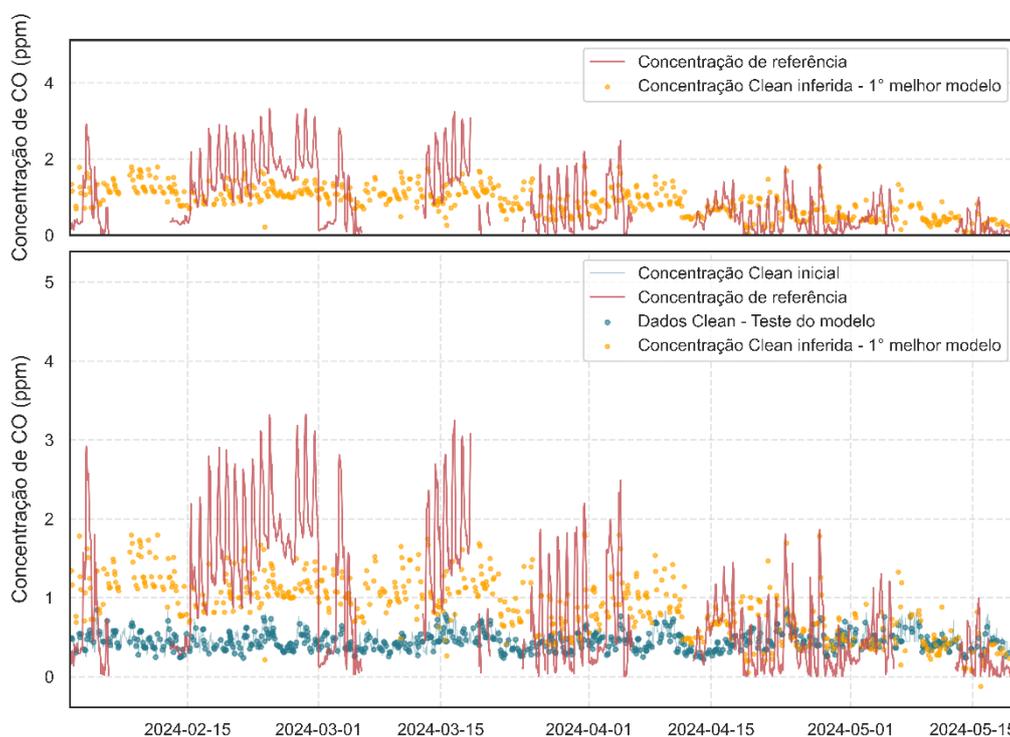
Fonte: Elaborado pela autora (2024)

Após a aplicação da regressão multivariada, nota-se uma melhora significativa nos coeficientes de correlação, o que implica no funcionamento razoável do modelo. O  $R^2$  relativamente baixo sugere que o modelo de regressão linear explica apenas uma parte modesta da variação nos dados do sensor, principalmente em altas concentrações (acima de 1ppm). O MAE e o RMSE indicam um certo grau de precisão, mas ainda com margem para melhorias, principalmente devido ao perfil das leituras do sensor. Em relação aos resíduos, apesar de o erro médio ser próximo de zero, há pontos que podem indicar uma subestimação ou superestimação pelo modelo.

O gráfico de dispersão revela um comportamento peculiar, com uma concentração de pontos alinhados verticalmente em torno de aproximadamente 1.0 ppm de CO, conforme medido pela estação de referência. Este padrão indica que, para várias medições de CO pela estação de referência nessa faixa, o sensor CLEAN infere uma ampla gama de concentrações de CO, resultando em uma dispersão vertical dos pontos. Esse comportamento foi observado também em análises de outros sensores e pode estar relacionado à imputação da média nos dados para preencher valores ausentes, contribuindo para a variabilidade observada.

Por fim, a série de dados inicial, submetida ao processo de calibração, os dados de teste do modelo, os dados alvo de concentração e os valores preditos, são representados pela Figura 17. Observa-se que as concentrações de CO preditas seguem a tendência geral dos dados de referência, embora com uma menor amplitude de variação, sugerindo que o modelo está capturando a tendência geral, mas não as variações extremas.

Figura 17 - Série temporal de concentrações do sensor CO-B4 de entrada do modelo, amostras de CO utilizadas para o teste do melhor modelo de regressão, os valores de CO previstos pelo melhor modelo de regressão e os dados alvo de CO de referência.



Legenda: os dados de concentração CLEAN antes do processo de calibração são demonstrados pela serie em azul; dentre esses dados, a amostra utilizada para teste do melhor modelo é representada pelo pontilhado azul; após a aplicação do melhor modelo de calibração, os dados de teste são utilizados para inferir a concentração final, representada pelo pontilhado laranja.

Fonte: Elaborado pela autora (2024)

Ao comparar o comportamento entre os valores de CO-B4 usados no teste do modelo e os valores de CO-B4 preditos em relação aos valores de CO alvo, infere-se que o modelo conseguiu reduzir a discrepância inicial. No entanto, as predições não capturam completamente as flutuações mais extremas dos dados de referência, possivelmente devido à combinação entre a complexidade dos dados de entrada, limitações do modelo de regressão multivariada e características dos dados de referência.

As métricas de desempenho apresentaram contradições com os valores de AIC e BIC, que não foram os mais baixos dentro do intervalo. Embora um alto valor de  $R^2$  indique que o modelo explica bem a variabilidade dos dados, valores elevados de AIC e BIC podem sinalizar um problema de overfitting. Isso ocorre porque, apesar do bom ajuste aparente, o modelo pode estar capturando variações aleatórias e não generalizando bem para novos dados.

## 5. CONCLUSAO

O presente trabalho teve como objetivo o desenvolvimento de uma ferramenta de calibração de sistemas de monitoramento atmosférico de baixo custo, incluindo procedimentos de filtragem e detecção de falhas, além da análise de modelos de calibração. Dada a urgente necessidade de novas alternativas de monitoramento da qualidade do ar, complementares às soluções tradicionais, este estudo busca contribuir significativamente para essa área.

Os sistemas de monitoramento atmosférico de baixo custo surgem como alternativas viáveis aos métodos tradicionais de referência, sendo mais acessíveis, de fácil operação e manutenção. Estes sensores podem ampliar a cobertura de monitoramento e fornecer uma visão mais detalhada da variabilidade espacial e temporal dos poluentes atmosféricos. Contudo, a precisão, durabilidade e confiabilidade desses sensores variam conforme a tecnologia de medição, qualidade dos componentes, condições ambientais e métodos de operação.

A ferramenta de filtragem e detecção de falhas desenvolvida demonstrou ser eficaz na preparação dos dados para os modelos de calibração. No entanto, é necessário aplicar a metodologia a um número maior de amostras, com maior tempo de coleta e menos instabilidade, para garantir que essa etapa filtre apenas os dados necessários, sem superestimar ou subestimar os valores. Das leituras coletadas pelo dispositivo CLEAN instalado em campo, observou-se que as principais causas de dados inválidos foram os dados fora dos limites de medição e os períodos sem dados coletados devido a quedas de energia, falhas na comunicação com os sensores e falhas do equipamento.

A aplicação de algoritmos de aprendizado de máquina e métodos de regressão, como a regressão multivariada, mostrou-se eficaz na calibração dos sensores, apesar de não capturar completamente as flutuações extremas dos dados de referência. Com a metodologia proposta de correção de dados, foi possível agregar valor às leituras, permitindo a correção de medições ruidosas e com dados faltantes. Observou-se que o modelo de regressão linear multivariada apresentou os melhores resultados para as métricas de  $R^2$ , RMSE e MAE, com aumento significativo nos coeficientes de Spearman e Kendall após a aplicação do modelo. No entanto, os modelos como floresta aleatória, k-vizinhos mais próximos e regressão MLP apresentaram desempenho inferior devido a problemas de overfitting, inadequação dos hiperparâmetros e características das concentrações.

Apesar dos resultados promissores, melhorias adicionais são necessárias. Recomenda-se ajustes nos hiperparâmetros dos modelos, inclusão de variáveis adicionais e experimentação com outros modelos de calibração, como é o caso de modelos híbridos, para aprimorar a

precisão das medições. O teste da metodologia em outros dispositivos de baixo custo também é necessário, dado o interesse em entender o poder de generalização da proposta. Além disso, o aprimoramento da aquisição de sinal do equipamento CLEAN, visando reduzir o ruído e estabilizar o sinal de linha base, é fundamental para aumentar a confiabilidade dos sensores.

A implementação dessas metodologias pode potencialmente transformar a maneira como a qualidade do ar é monitorada, especialmente em regiões com recursos limitados. A infraestrutura dos dispositivos deve ser melhorada para garantir robustez e autonomia, reduzindo a dependência de recursos externos e minimizando falhas operacionais.

Em suma, este trabalho estabelece uma base para o uso dos dispositivos CLEAN no monitoramento da qualidade do ar, oferecendo uma solução inicial para os problemas de incerteza dos dados. As melhorias sugeridas e os estudos futuros poderão consolidar ainda mais a metodologia e, conseqüentemente, as tecnologias de baixo custo, tornando-a uma ferramenta poderosa no avanço do monitoramento de baixo custo.

## REFERÊNCIAS

ALPHASENSE. **Alphasense Application Note AAN 104 How Electrochemical Gas Sensors Work**. [S.l.], 2013. P. 1–4. Disponível em: [https://www.alphasense.com/WEB1213/wp-content/uploads/2013/07/AAN\\_104.pdf](https://www.alphasense.com/WEB1213/wp-content/uploads/2013/07/AAN_104.pdf). .

ALPHASENSE. **Alphasense Application Note AAN 110 ENVIRONMENTAL CHANGES: TEMPERATURE, PRESSURE, HUMIDITY**. [S.l.], 2013. P. 1–6. Disponível em: [https://www.alphasense.com/WEB1213/wp-content/uploads/2013/07/AAN\\_110.pdf](https://www.alphasense.com/WEB1213/wp-content/uploads/2013/07/AAN_110.pdf)

ARROYO, Patricia *et al.* Low-Cost Air Quality Measurement System Based on Electrochemical and MP Sensors with Cloud Connection. **Sensors**, [S.L.], v. 21, n. 18, p. 6228, 16 set. 2021. MDPI AG. <http://dx.doi.org/10.3390/s21186228>

ARTAXO, Paulo. **O Estado da Qualidade do Ar no Brasil**. Working Paper WRI Brasil, p. 32, 2021 Tradução. . Disponível em: <https://wribrasil.org.br/sites/default/files/wri-o-estado-da-qualidade-do-ar-no-brasil.pdf>. Acesso em: 08 mar. 2024.

BORREGO, C. *et al.* Assessment of air quality microsensors versus reference methods: the eunetair joint exercise. **Atmospheric Environment**, [S.L.], v. 147, p. 246-263, dez. 2016. Elsevier BV. <http://dx.doi.org/10.1016/j.atmosenv.2016.09.050>.

BORREGO, C. et al. Assessment of air quality microsensors versus reference methods: The EuNetAir Joint Exercise – Part II. **Atmospheric environment (Oxford, England: 1994)**, v. 193, p. 127–142, 2018.

BRASIL. **Lei nº 6.938, de 31 de agosto de 1981**. Dispõe sobre a Política Nacional do Meio Ambiente. Disponível em: [https://www.planalto.gov.br/ccivil\\_03/leis/16938.htm](https://www.planalto.gov.br/ccivil_03/leis/16938.htm). Acesso em: 29 de setembro de 2023.

**BRASIL**. Ministério do meio ambiente. Guia técnico para o monitoramento e avaliação da qualidade do ar. Brasília, DF: MMA, 2020. Disponível em: <https://www.mma.gov.br/agenda-ambiental-urbana/qualidade-do-ar.html>.

**BRASIL**. Ministério da Saúde. Poluição atmosférica. Série A. Normas e Manuais Técnicos. Brasília: Ministério da Saúde, 2006. Disponível em: [https://bvsm.sau.gov.br/bvs/publicacoes/poluicao\\_atmosferica\\_SUS\\_saude\\_ambiental.pdf](https://bvsm.sau.gov.br/bvs/publicacoes/poluicao_atmosferica_SUS_saude_ambiental.pdf). Acesso em: 29 de setembro de 2023.

BRIMBLECOMBE, P. History of urban air pollution. Em: **Urban Air Pollution — European Aspects**. Dordrecht: Springer Netherlands, 1998. p. 7–20.

CAMPO, Fernando; FRANCO, Davide; SANTOS, Felipe de Campos; BLANCO-RODRÍGUEZ, Andy; GARCIA-RAMIREZ, Alejandro Rafael; RATÃO, Gabriel; HOINASKI,

Leonardo. CLEAN - Collaborative low-cost environmental and air-quality network. **Environmental Modelling & Software**, [S.L.], v. 163, p. 105664, maio 2023. Elsevier BV. <http://dx.doi.org/10.1016/j.envsoft.2023.105664>.

CASTELL, N. et al. Can commercial low-cost sensor platforms contribute to air quality monitoring and exposure estimates? **Environment international**, v. 99, p. 293–302, 2017.

CAROTENUTO, F. et al. Low-cost air quality monitoring networks for long-term field campaigns: A review. **Meteorological applications**, v. 30, n. 6, 2023.

CETESB (org). **Poluentes Atmosféricos**. Disponível em: <https://cetesb.sp.gov.br/ar/poluentes/>. Acesso em: 12 mar. 2024.

CHU, H.-J.; ALI, M. Z.; HE, Y.-C. Spatial calibration and PM<sub>2.5</sub> mapping of low-cost air quality sensors. **Scientific reports**, v. 10, n. 1, 2020.

CONAMA. Resolução nº 491, de 19 de novembro de 2018. Dispõe sobre os padrões da qualidade do ar.

CONCAS, F. et al. Low-cost outdoor air quality monitoring and sensor calibration: A survey and critical analysis. **ACM transactions on sensor networks**, v. 17, n. 2, p. 1–44, 2021.

CRILLEY, Leigh R. *et al.* Evaluation of a low-cost optical particle counter (Alphasense OPC-N2) for ambient air monitoring. **Atmospheric Measurement Techniques**, [S.L.], v. 11, n. 2, p. 709–720, 7 fev. 2018. Copernicus GmbH. <http://dx.doi.org/10.5194/amt-11-709-2018>.

COST ACTION TD1105 SECOND SCIENTIFIC MEETING. **Next generation air monitoring: An overview of US EPA activities**. Disponível em: <[http://www.eunetair.it/cost/meetings/Cambridge/02-PRESENTATIONS/01\\_PLENARY/T02\\_TD1105\\_Queens\\_Plenary\\_Watkins.pdf](http://www.eunetair.it/cost/meetings/Cambridge/02-PRESENTATIONS/01_PLENARY/T02_TD1105_Queens_Plenary_Watkins.pdf)>. Acesso em: 27 jun. 2024.

DESOUZA, P. et al. Calibrating networks of low-cost air quality sensors. **Atmospheric measurement techniques**, v. 15, n. 21, p. 6309–6328, 2022.

DE VITO, S. et al. Adaptive machine learning strategies for network calibration of IoT smart air quality monitoring devices. **Pattern recognition letters**, v. 136, p. 264–271, 2020.

DE VITO, S. et al. Calibrating chemical multisensory devices for real world applications: An in-depth comparison of quantitative machine learning approaches. **Sensors and actuators. B, Chemical**, v. 255, p. 1191–1210, 2018.

DUVALL, Andrea Clements And Rachelle *et al.* **The Enhanced Air Sensor Guidebook**. Us: Epa, 2022. 195 p.

EPA. Environmental Protection Agency. **Low-Cost Air Pollution Monitors and Indoor Air Quality**. Disponível em: <https://www.epa.gov/indoor-air-quality-iaq/low-cost-air-pollution-monitors-and-indoor-air-quality#Intro>. Acesso em: 27 de outubro de 2023.

EPA. Environmental Protection Agency. **Quality Assurance Handbook for Air Pollution Measurement Systems: ambient air quality monitoring program**. Us: Air Quality Assessment Division, 2017. 346 p.

FENG, X. et al. Artificial neural networks forecasting of PM<sub>2.5</sub> pollution using air mass trajectory based geographic model and wavelet transformation. **Atmospheric environment (Oxford, England: 1994)**, v. 107, p. 118–128, 2015.

GARCÍA, Fernando Campo. **Desenvolvimento de uma rede colaborativa de monitoramento de baixo custo da qualidade do ar**. 2024. 422 f. Tese (Doutorado) - Curso de Engenharia Ambiental, Universidade Federal de Santa Catarina, Florianópolis, 2024.

GIORDANO, Michael R.; MALINGS, Carl; PANDIS, Spyros N.; PRESTO, Albert A.; MCNEILL, V.F.; WESTERVELT, Daniel M.; BEEKMANN, Matthias; SUBRAMANIAN, R.. From low-cost sensors to high-quality data: a summary of challenges and best practices for effectively calibrating low-cost particulate matter mass sensors. **Journal Of Aerosol Science**, [S.L.], v. 158, p. 105833, nov. 2021. Elsevier BV. <http://dx.doi.org/10.1016/j.jaerosci.2021.105833>.

GOMES, Paulo Alexandre Marques Vieira. **Utilização de sensores de baixo custo na medição de monóxido de carbono no ar ambiente**. 2015. 133 f. Dissertação (Mestrado) - Curso de Engenharia de Instrumentação e Metrologia, Instituto Politécnico do Porto, Cidade do Porto, 2015.

IDREES, Zeba; ZHENG, Lirong. Low cost air pollution monitoring systems: a review of protocols and enabling technologies. **Journal Of Industrial Information Integration**, [S.L.], v. 17, p. 100123, mar. 2020. Elsevier BV. <http://dx.doi.org/10.1016/j.jii.2019.100123>.

IEMA. Instituto de Energia e Meio Ambiente. **Plataforma da Qualidade do Ar**. Disponível em: <https://energiaeambiente.org.br/qualidadedoar/>. Acesso em: 05 maio 2024.

IMA/SC. **Rede de Monitoramento - SC**. [S.l.: s.n.], 2023. Disponível em: <https://www.ima.sc.gov.br/index.php/qualidade-ambiental/menu-qualidadear/rede-de-monitoramento#:~:text=Localizada%20na%20Rua%20Dr.,Vereadores%20da%20Cidade%20de%20Tubar%20C3%A3o.&text=Localizada%20na%20Rua%20Ant%C3%B4nia%20Goularte,370%20da%20Cidade%20de%20Tubar%20C3%A3o>.

ISS - INSTITUTO DE SAÚDE E SUSTENTABILIDADE. Monitoramento da Qualidade do Ar no Brasil. 2014. Disponível em:

<https://www.saudeesustentabilidade.org.br/publicacao/pesquisa-monitoramento-da-qualidade-do-ar-no-brasil/>. Acesso em: 4 de outubro 2023.

JIAO, W. et al. Community Air Sensor Network (CAIRSENSE) project: evaluation of low-cost sensor performance in a suburban environment in the southeastern United States. **Atmospheric measurement techniques**, v. 9, n. 11, p. 5281–5292, 2016.

KADRI, A. et al. **Wireless sensor network for real-time air pollution monitoring**. 2013 1st International Conference on Communications, Signal Processing, and their Applications (ICCSPA). **Anais...IEEE**, 2013.

KARAGULIAN, F. et al. Review of the performance of low-cost sensors for air quality monitoring. **Atmosphere**, v. 10, n. 9, p. 506, 2019.

KILARU, Ron Williams And Vasu *et al.* **Air Sensor Guidebook**. Us: Epa, 2014. 73 p.

KUMAR, P. et al. The rise of low-cost sensing for managing air pollution in cities. **Environment international**, v. 75, p. 199–205, 2015.

Landrigan, P. J. et al. (2018). The Lancet Commissions: The Lancet Commission on pollution and health. **The Lancet**, 391(10119), 462-512. [https://doi.org/10.1016/S0140-6736\(17\)32345-0](https://doi.org/10.1016/S0140-6736(17)32345-0)

LEWIS, Alastair C.; LEE, James D.; EDWARDS, Peter M.; SHAW, Marvin D.; EVANS, Mat J.; MOLLER, Sarah J.; SMITH, Katie R.; BUCKLEY, Jack W.; ELLIS, Matthew; GILLOT, Stefan R.. Evaluating the performance of low cost chemical sensors for air pollution research. **Faraday Discussions**, [S.L.], v. 189, p. 85-103, nov. 2016. Royal Society of Chemistry (RSC). <http://dx.doi.org/10.1039/c5fd00201j>.

LIANG, Lu. Calibrating low-cost sensors for ambient air monitoring: techniques, trends, and challenges. **Environmental Research**, [S.L.], v. 197, p. 111163, jun. 2021. Elsevier BV. <http://dx.doi.org/10.1016/j.envres.2021.111163>.

MAHAJAN, Sachit *et al.* Evaluation of low-cost sensors for quantitative personal exposure monitoring. **Sustainable Cities And Society**, [S.L.], v. 57, p. 102076, jun. 2020. Elsevier BV. <http://dx.doi.org/10.1016/j.scs.2020.102076>.

MARGARITIS, Dimitris; KERAMYDAS, Christos; PAPACHRISTOS, Ioannis; LAMBROPOULOU, Dimitra. Calibration of Low-cost Gas Sensors for Air Quality Monitoring. **Aerosol And Air Quality Research**, [S.L.], v. 21, n. 11, p. 210073, 2021. Taiwan Association for Aerosol Research. <http://dx.doi.org/10.4209/aaqr.210073>.

MEAD, M. I. et al. The use of electrochemical sensors for monitoring urban air quality in low-cost, high-density networks. **Atmospheric environment (Oxford, England: 1994)**, v. 70, p. 186–203, 2013.

MORAWSKA, Lidia *et al.* Applications of low-cost sensing technologies for air quality monitoring and exposure assessment: how far have they gone?. **Environment International**, [S.L.], v. 116, p. 286-299, jul. 2018. Elsevier BV. <http://dx.doi.org/10.1016/j.envint.2018.04.018>.

MOSLEY, S. Environmental history of air pollution and protection. Em: **Environmental History**. Cham: Springer International Publishing, 2014. p. 143–169.

NGO, N.s.; KOKOYO, S.; KLOPP, J.. Why participation matters for air quality studies: risk perceptions, understandings of air pollution and mobilization in a poor neighborhood in nairobi, kenya. **Public Health**, [S.L.], v. 142, p. 177-185, jan. 2017. Elsevier BV. <http://dx.doi.org/10.1016/j.puhe.2015.07.014>.

OTTOSEN, T.-B.; KUMAR, P. Outlier detection and gap filling methodologies for low-cost air quality measurements. **Environmental science. Processes & impacts**, v. 21, n. 4, p. 701–713, 2019.

RIBEIRO, A. K. C.; GALVÃO, E. S.; ALBUQUERQUE, T. T. DE A. Air quality characterization and trend analysis in a Brazilian industrialized Metropolitan Area in the period from 1995 to 2022. **Atmosphere**, v. 14, n. 12, p. 1792, 2023.

**WHO**. World Health Organization. Protecting health through ambient air quality management: A resource package for the WHO European Region. 2023. Disponível em: <https://iris.who.int/bitstream/handle/10665/366687/WHO-EURO-2023-6898-46664-67857-eng.pdf?sequence=1&isAllowed=y>. Acesso em: 01 de outubro de 2023.

**WHO**. World Health Organization. Air Quality Guidelines - Update 2021. Copenhagen, Dinamarca: Escritório Regional da OMS para a Europa; 2021. Disponível em: <https://iris.who.int/bitstream/handle/10665/345329/9789240034228-eng.pdf?sequence=1>. Acesso em: 01 de outubro de 2023.

**WHO**. World Health Organization. Overview of methods to assess population exposure to ambient air pollution. **Geneva**: [S.N.], 2023.

SCHNEIDER, P. et al. Mapping urban air quality in near real-time using observations from low-cost sensors and model information. **Environment international**, v. 106, p. 234–247, 2017.

SINGH, D. et al. Sensors and systems for air quality assessment monitoring and management: A review. **Journal of environmental management**, v. 289, n. 112510, p. 112510, 2021.

SNYDER, E. G. et al. The Changing Paradigm of Air Pollution Monitoring. **Environmental Science & Technology**, v. 47, n. 20, p. 11369–11377, 3 out. 2013.

SPINELLE, L. et al. Field calibration of a cluster of low-cost available sensors for air quality monitoring. Part A: Ozone and nitrogen dioxide. **Sensors and actuators. B, Chemical**, v. 215, p. 249–257, 2015.

TARGINO, Admir Créso *et al.* Hotspots of black carbon and PM<sub>2.5</sub> in an urban area and relationships to traffic characteristics. **Environmental Pollution**, [S.L.], v. 218, p. 475-486, nov. 2016. Elsevier BV. <http://dx.doi.org/10.1016/j.envpol.2016.07.027>.

TESSUM, Christopher W. *et al.* Inequity in consumption of goods and services adds to racial–ethnic disparities in air pollution exposure. **Proceedings Of The National Academy Of Sciences**, [S.L.], v. 116, n. 13, p. 6001-6006, 11 mar. 2019. Proceedings of the National Academy of Sciences. <http://dx.doi.org/10.1073/pnas.1818859116>.

UNIAO EUROPEIA. **Diretiva 2008/50/EC** do Parlamento Europeu e do Conselho, de 21 de maio de 2008, sobre a qualidade do ar ambiente e um ar mais limpo para a Europa. Jornal Oficial da União Europeia, 11 jun. 2008. Disponível em: <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:152:0001:0044:EN:PDF>.

VORMITTAG, Evangelina da Motta P. A. de Araújo; CIRQUEIRA, Samirys Sara Rodrigues; WICHER NETO, Hélio; SALDIVA, Paulo Hilário N.. Análise do monitoramento da qualidade do ar no Brasil. **Estudos Avançados**, [S.L.], v. 35, n. 102, p. 7-30, ago. 2021. FapUNIFESP (SciELO). <http://dx.doi.org/10.1590/s0103-4014.2021.35102.002>.

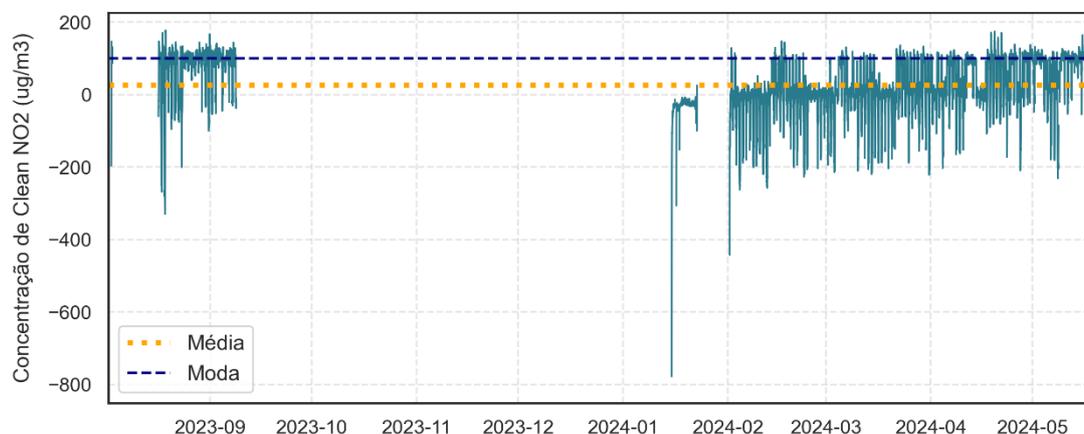
VELASCO, Román Pérez *et al.* Update of the WHO global air quality guidelines: systematic reviews ∴ an introduction. **Environment International**, [S.L.], v. 170, p. 107556, dez. 2022. Elsevier BV. <http://dx.doi.org/10.1016/j.envint.2022.107556>.

WILLIAMS, Ron; KILARU, Vasu; SNYDER, Emily; KAUFMAN, Amanda; DYE, Timothy; RUTTER, Andrew; RUSSELL, Ashley; HAFNER, Hilary. **Air Sensor Guidebook**. Research Triangle Park, NC: U.S. Environmental Protection Agency, Office of Research and Development, National Exposure Research Laboratory, 2014. Disponível em: [www.epa.gov/ord](http://www.epa.gov/ord).

WANG, A. et al. Leveraging machine learning algorithms to advance low-cost air sensor calibration in stationary and mobile settings. **Atmospheric environment (Oxford, England: 1994)**, v. 301, n. 119692, p. 119692, 2023.

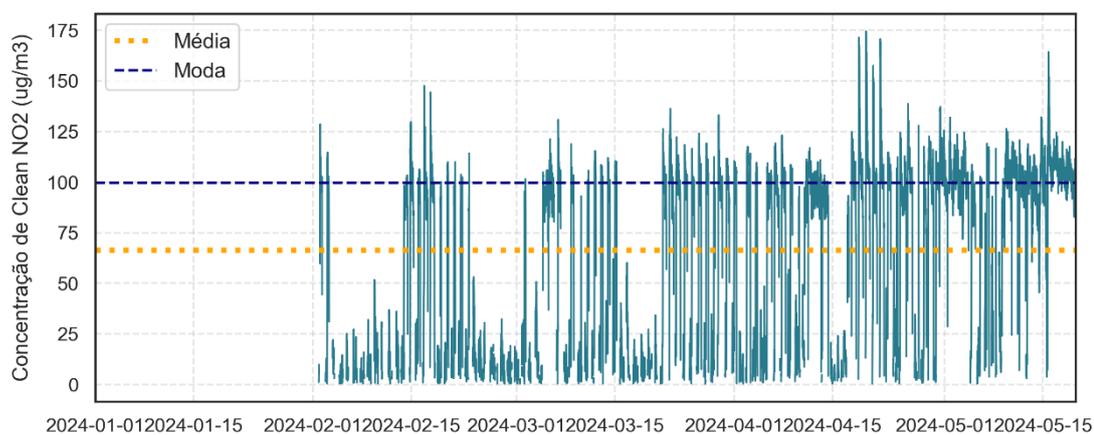
ZHANG, Xiaoyan *et al.* Linking urbanization and air quality together: a review and a perspective on the future sustainable urban development. **Journal Of CLEANer Production**, [S.L.], v. 346, p. 130988, abr. 2022. Elsevier BV. <http://dx.doi.org/10.1016/j.jclepro.2022.130988>.

## APÊNDICE A – SÉRIE TEMPORAL DOS DADOS BRUTOS DE CONCENTRAÇÃO DE NO<sub>2</sub> – CLEAN COM DADOS AGRUPADOS EM 15 MINUTOS



Fonte: Elaborado pela autora (2024)

## APÊNDICE B – SÉRIE TEMPORAL DE CONCENTRAÇÃO DE NO<sub>2</sub> – CLEAN DOS DADOS VÁLIDOS APÓS ETAPA DE CONTROLE DE QUALIDADE COM DADOS AGRUPADOS EM 15 MINUTOS



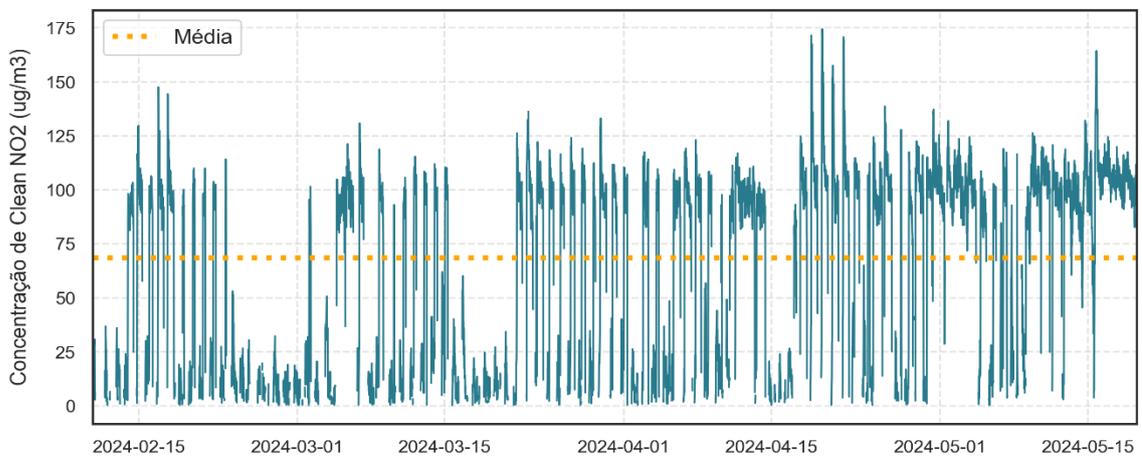
Fonte: Elaborado pela autora (2024)

**APÊNDICE C– MUDANÇAS ABRUPTAS GERADAS POR PELT NA SÉRIE TEMPORAL DE CONCENTRAÇÃO NO<sub>2</sub> – CLEAN COM DADOS AGRUPADOS EM 15 MINUTOS**



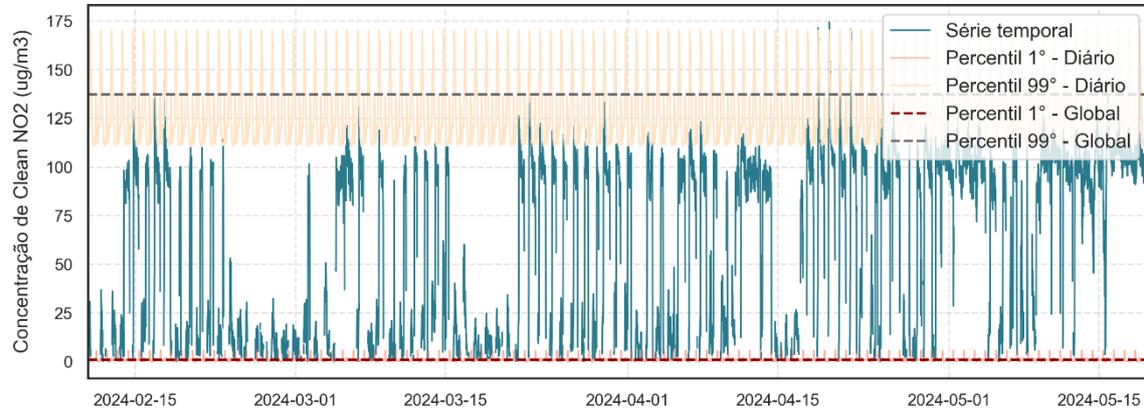
Fonte: Elaborado pela autora (2024)

**APÊNDICE D – Janela de medição após implementa PELT na série temporal de concentração de NO<sub>2</sub> – CLEAN com dados agrupados em 15 minutos**



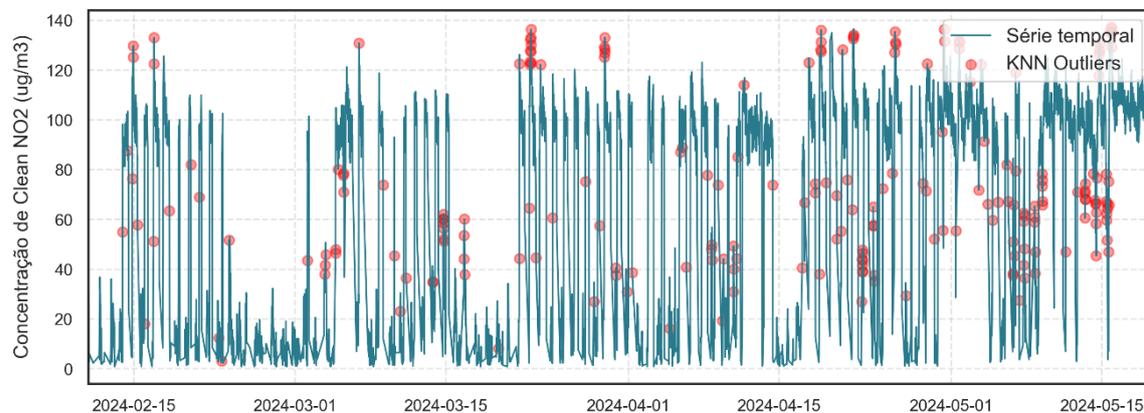
Fonte: Elaborado pela autora (2024)

## APÊNDICE E – DETECÇÃO DE OUTLIERS POR QUARTIL 1% E 99% GLOBAL E DIÁRIO DAS CONCENTRAÇÕES DE NO<sub>2</sub> – CLEAN COM DADOS AGRUPADOS EM 15 MINUTOS



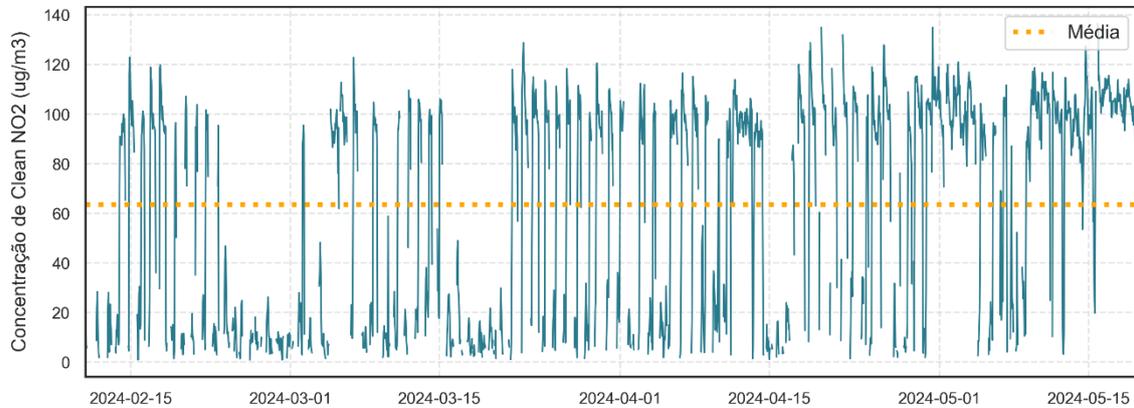
Fonte: Elaborado pela autora (2024)

## APÊNDICE F – DETECÇÃO DE OUTLIERS POR KNN DAS CONCENTRAÇÕES DE NO<sub>2</sub> – CLEAN COM DADOS AGRUPADOS EM 15 MINUTOS



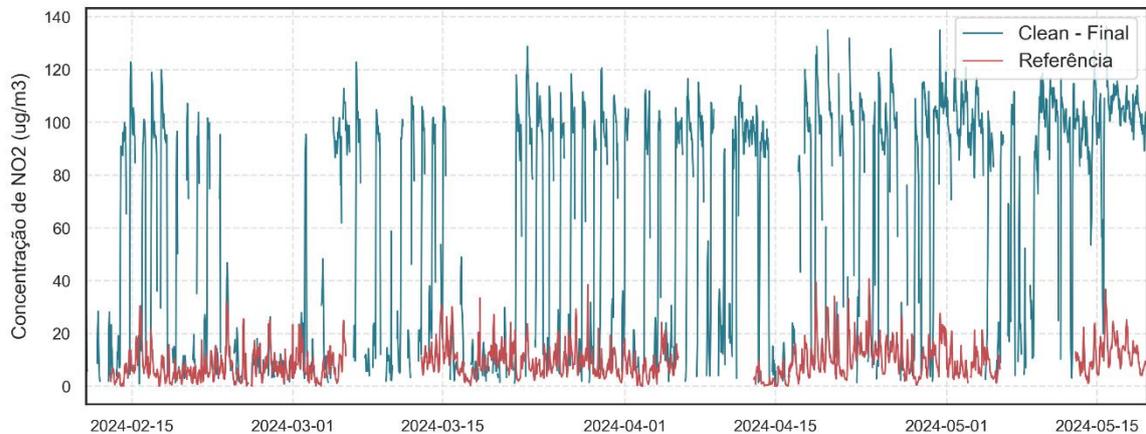
Fonte: Elaborado pela autora (2024)

**APÊNDICE G – SÉRIE TEMPORAL FINAL DAS CONCENTRAÇÕES DE NO<sub>2</sub> – CLEAN APÓS ETAPAS DE FILTRAGEM E DETECÇÃO DE FALHAS, COM DADOS AGRUPADOS EM MÉDIAS HORÁRIAS**



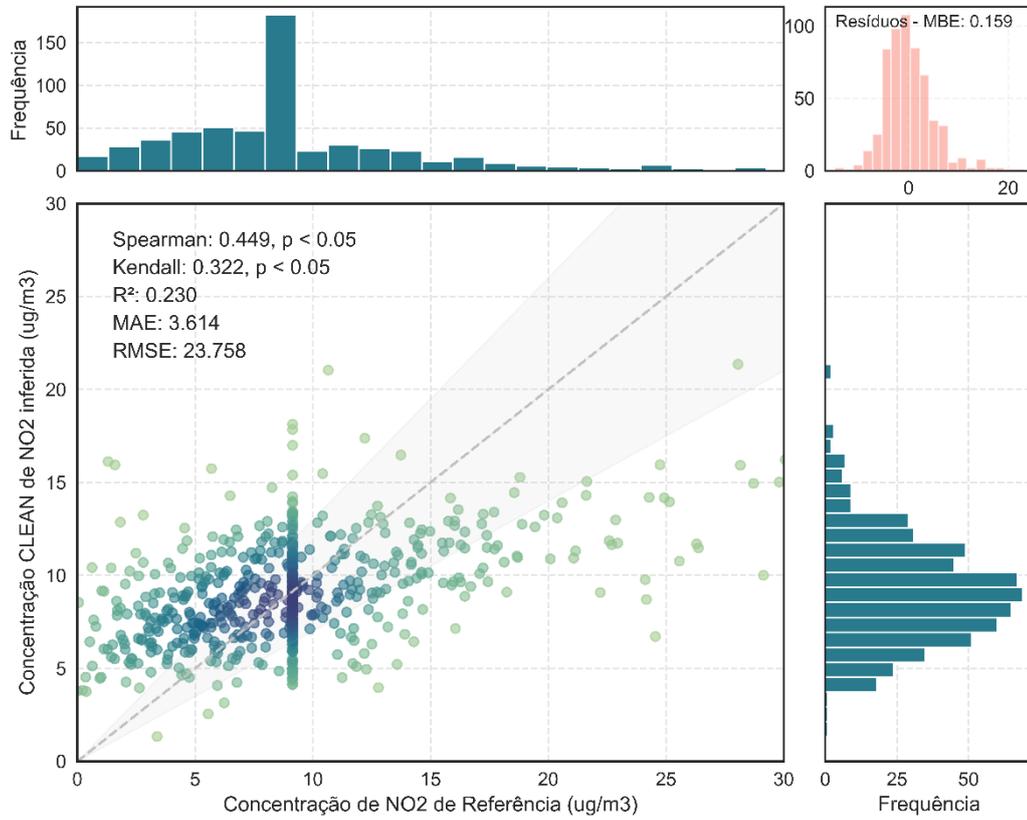
Fonte: Elaborado pela autora (2024)

**APÊNDICE H – SÉRIE TEMPORAL FINAL DAS CONCENTRAÇÕES DE NO<sub>2</sub> – CLEAN E NO<sub>2</sub> DE REFERÊNCIA, COM DADOS AGRUPADOS EM MÉDIAS HORÁRIAS**



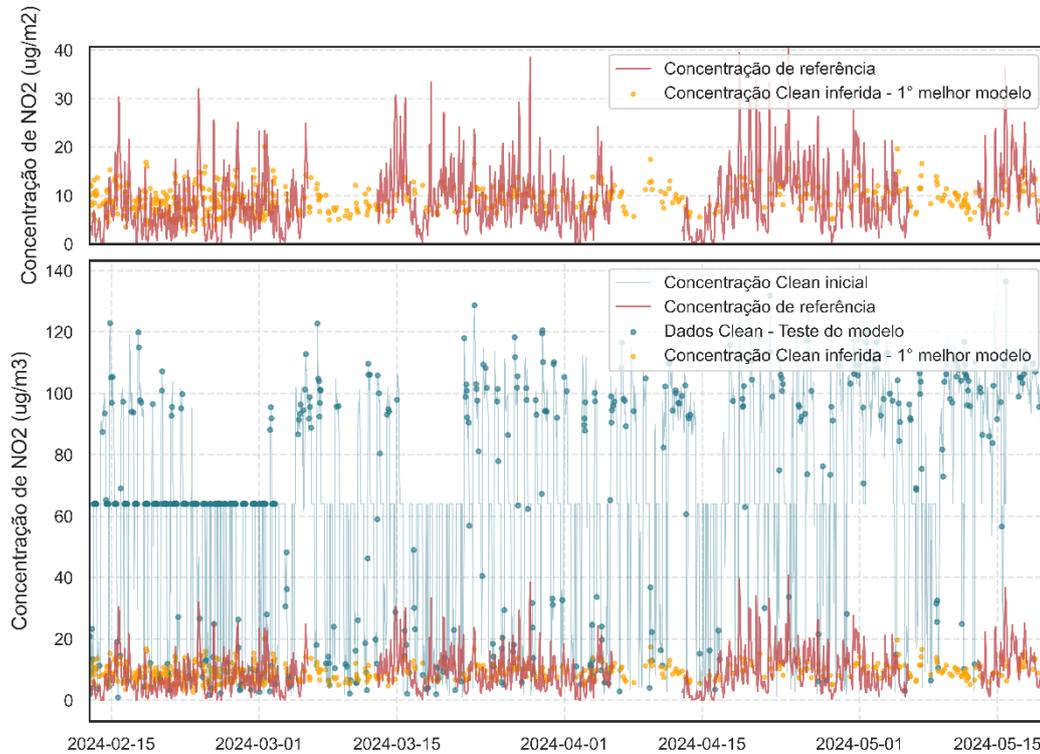
Fonte: Elaborado pela autora (2024)

## APÊNDICE I – CORRELAÇÃO E HISTOGRAMAS DAS CONCENTRAÇÕES DE NO<sub>2</sub> – CLEAN INFERIDA PELA CALIBRAÇÃO, DE REFERÊNCIA E DOS RESÍDUOS GERADOS



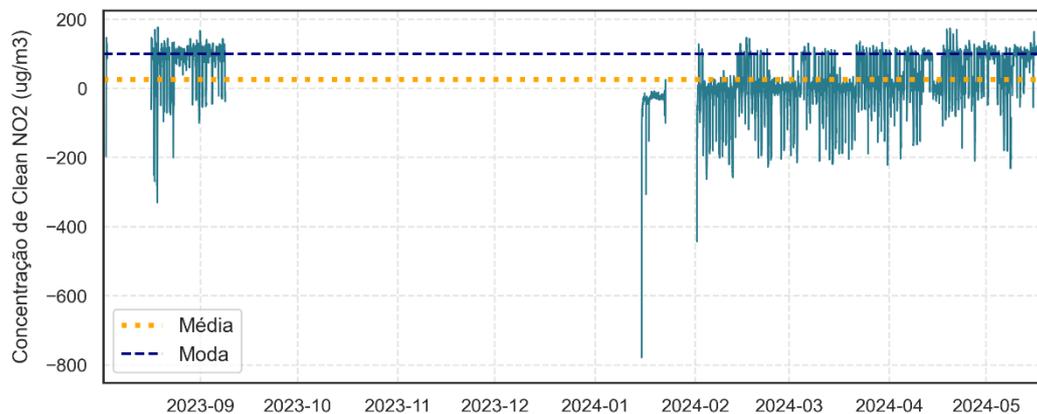
Fonte: Elaborado pela autora (2024)

**APÊNDICE J – SÉRIE TEMPORAL DE CONCENTRAÇÕES DE NO<sub>2</sub> – CLEAN DE ENTRADA DO MELHOR MODELO DE REGRESSÃO, AMOSTRAS DE TESTE, VALORES PREVISTOS E OS DADOS ALVO DE REFERÊNCIA**



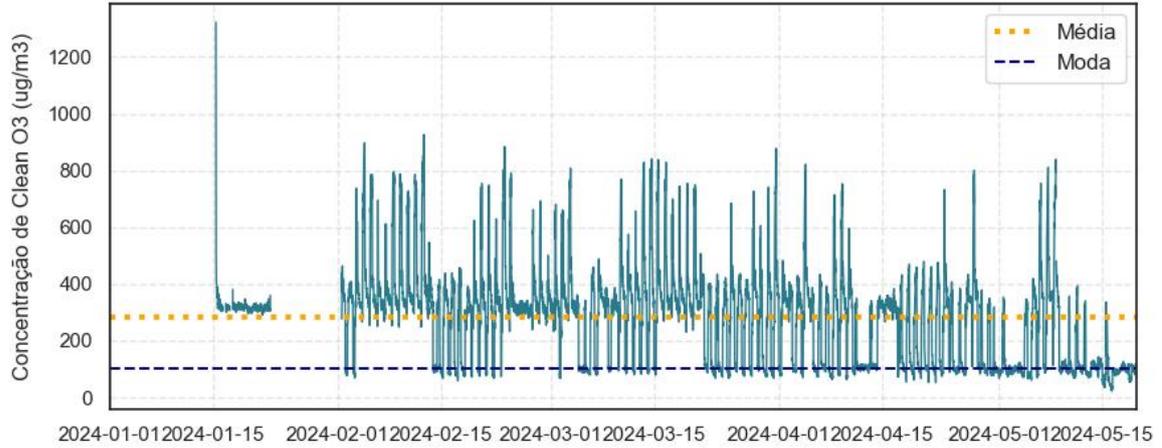
Fonte: Elaborado pela autora (2024)

**APÊNDICE K – SÉRIE TEMPORAL DOS DADOS BRUTOS DE CONCENTRAÇÃO DE O<sub>3</sub> – CLEAN COM DADOS AGRUPADOS EM 15 MINUTOS**



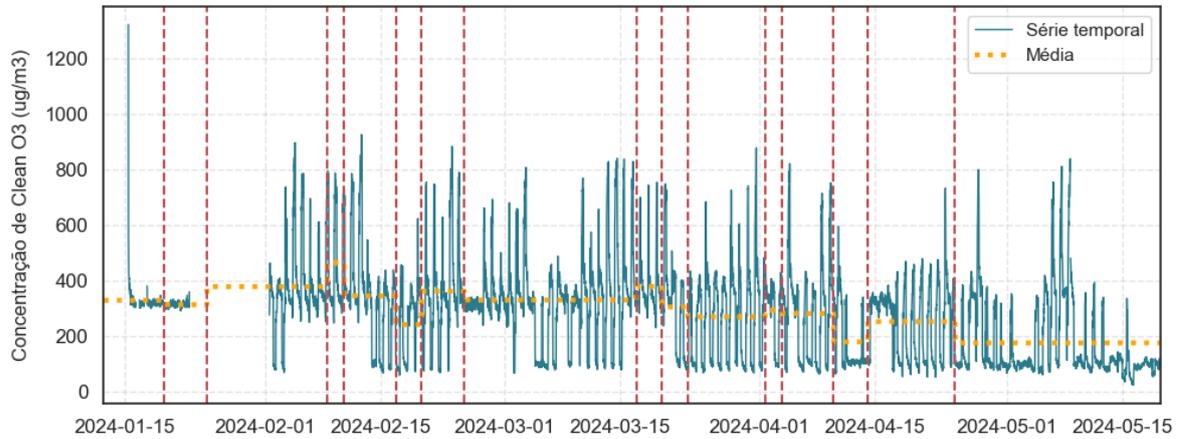
Fonte: Elaborado pela autora (2024)

**APÊNDICE L – SÉRIE TEMPORAL DE CONCENTRAÇÃO DE O3 – CLEAN DOS DADOS VÁLIDOS APÓS ETAPA DE CONTROLE DE QUALIDADE COM DADOS AGRUPADOS EM 15 MINUTOS**



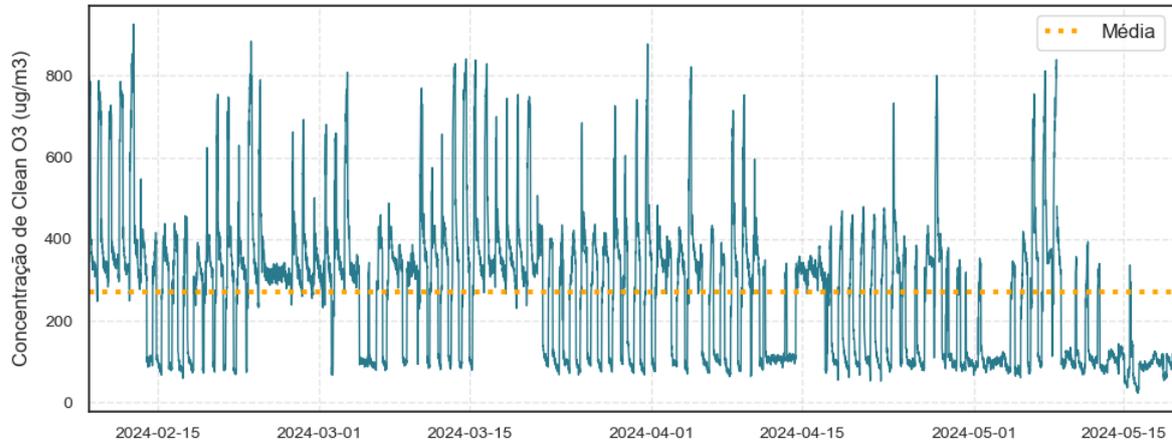
Fonte: Elaborado pela autora (2024)

**APÊNDICE M – MUDANÇAS ABRUPTAS GERADAS POR PELT NA SÉRIE TEMPORAL DE CONCENTRAÇÃO O3 – CLEAN COM DADOS AGRUPADOS EM 15 MINUTOS**



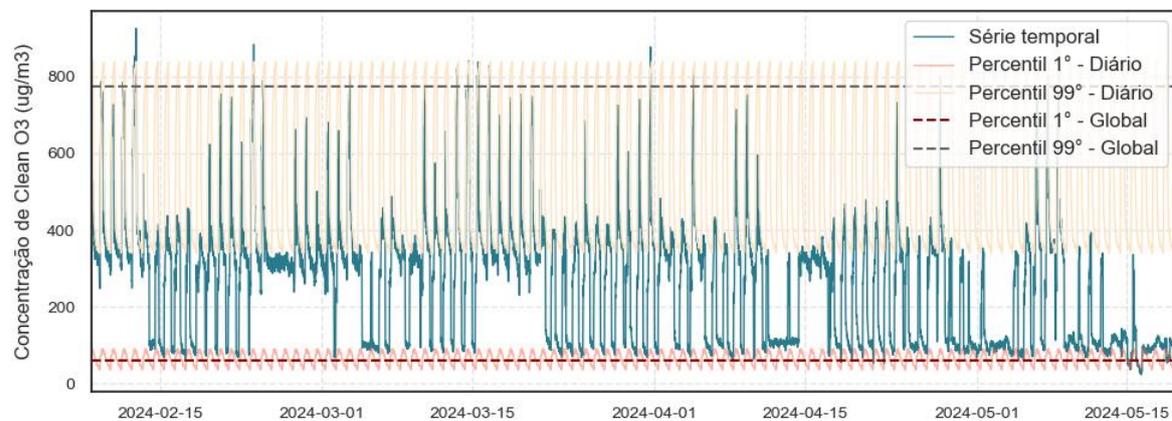
Fonte: Elaborado pela autora (2024)

**APÊNDICE N – JANELA DE MEDIÇÃO APÓS IMPLEMENTA PELT NA SÉRIE TEMPORAL DE CONCENTRAÇÃO DE O<sub>3</sub> – CLEAN COM DADOS AGRUPADOS EM 15 MINUTOS**



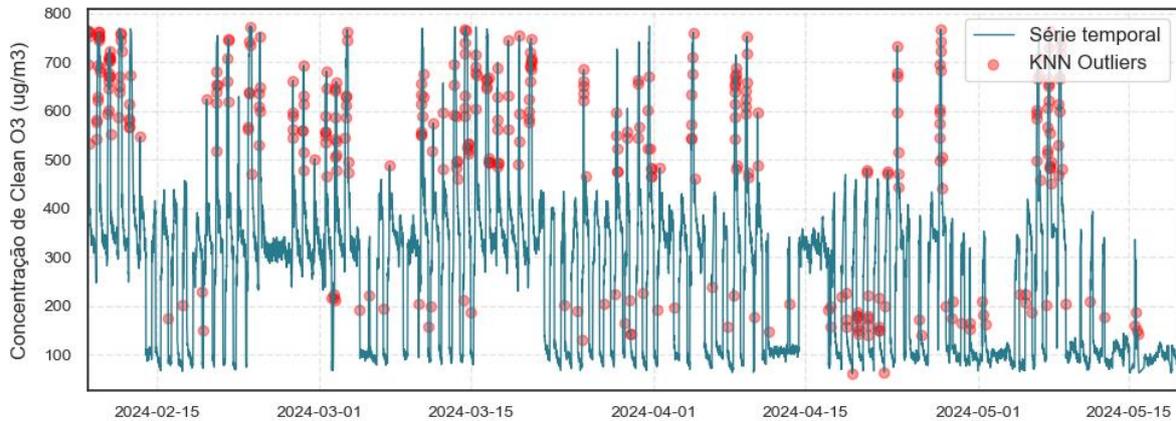
Fonte: Elaborado pela autora (2024)

**APÊNDICE O – DETECÇÃO DE OUTLIERS POR QUARTIL 1% E 99% GLOBAL E DIÁRIO DAS CONCENTRAÇÕES DE O<sub>3</sub>– CLEAN COM DADOS AGRUPADOS EM 15 MINUTOS**



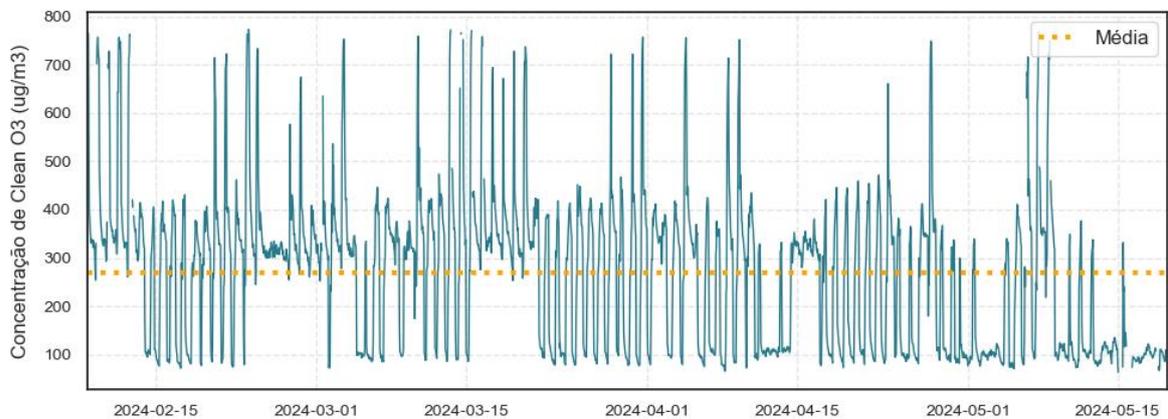
Fonte: Elaborado pela autora (2024)

## APÊNDICE P – DETECÇÃO DE OUTLIERS POR KNN DAS CONCENTRAÇÕES DE O<sub>3</sub> – CLEAN COM DADOS AGRUPADOS EM 15 MINUTOS



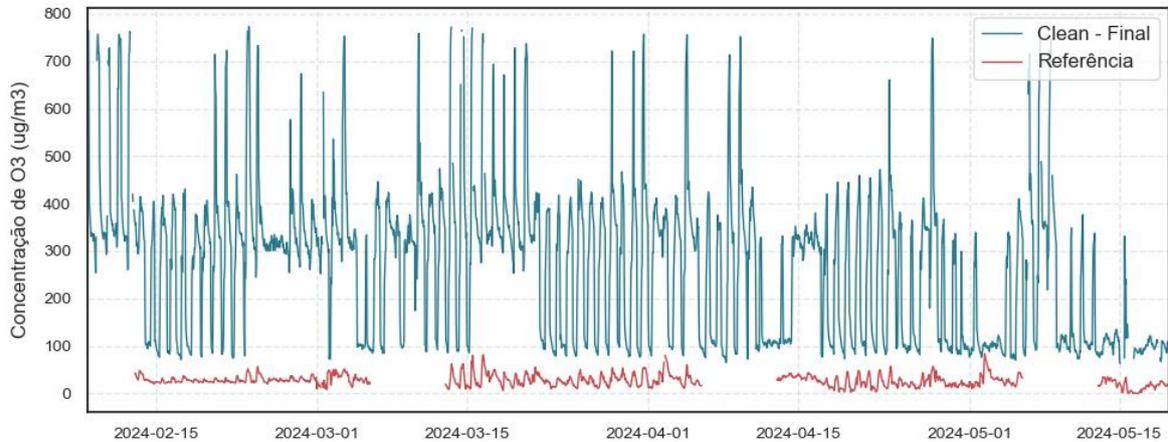
Fonte: Elaborado pela autora (2024)

## APÊNDICE Q – SÉRIE TEMPORAL FINAL DAS CONCENTRAÇÕES DE O<sub>3</sub> – CLEAN APÓS ETAPAS DE FILTRAGEM E DETECÇÃO DE FALHAS, COM DADOS AGRUPADOS EM MÉDIAS HORÁRIAS



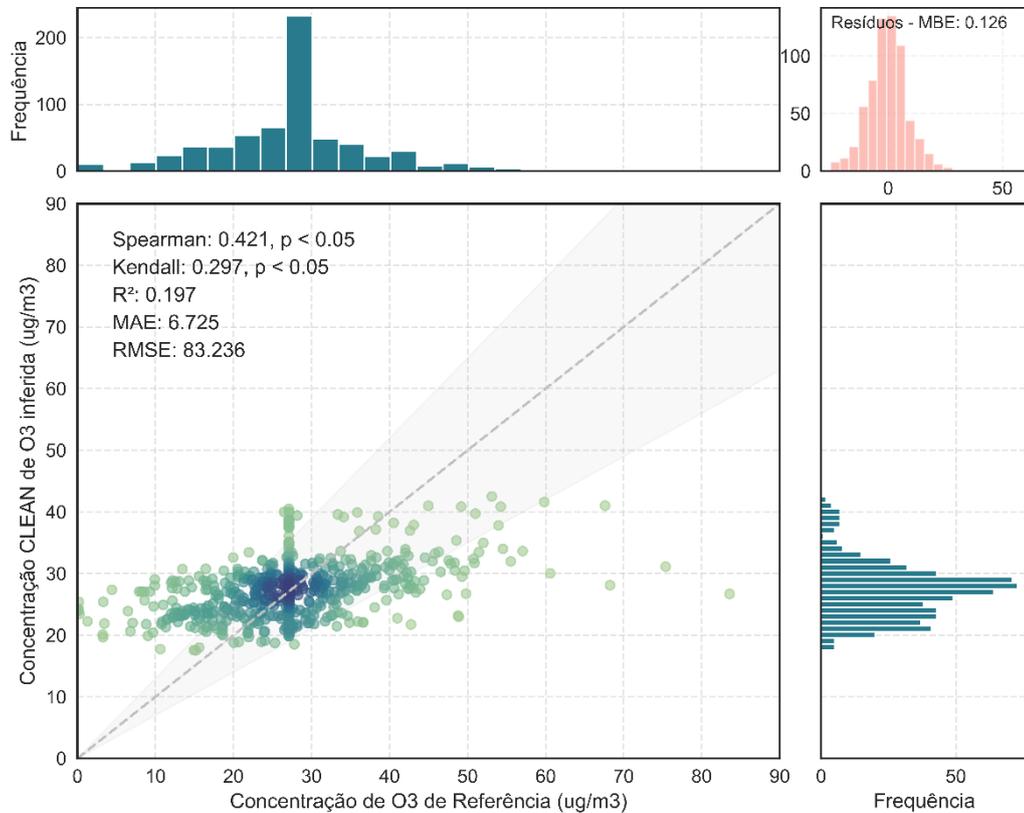
Fonte: Elaborado pela autora (2024)

## APÊNDICE R – SÉRIE TEMPORAL FINAL DAS CONCENTRAÇÕES DE O<sub>3</sub> – CLEAN E NO<sub>2</sub> DE REFERÊNCIA, COM DADOS AGRUPADOS EM MÉDIAS HORÁRIAS



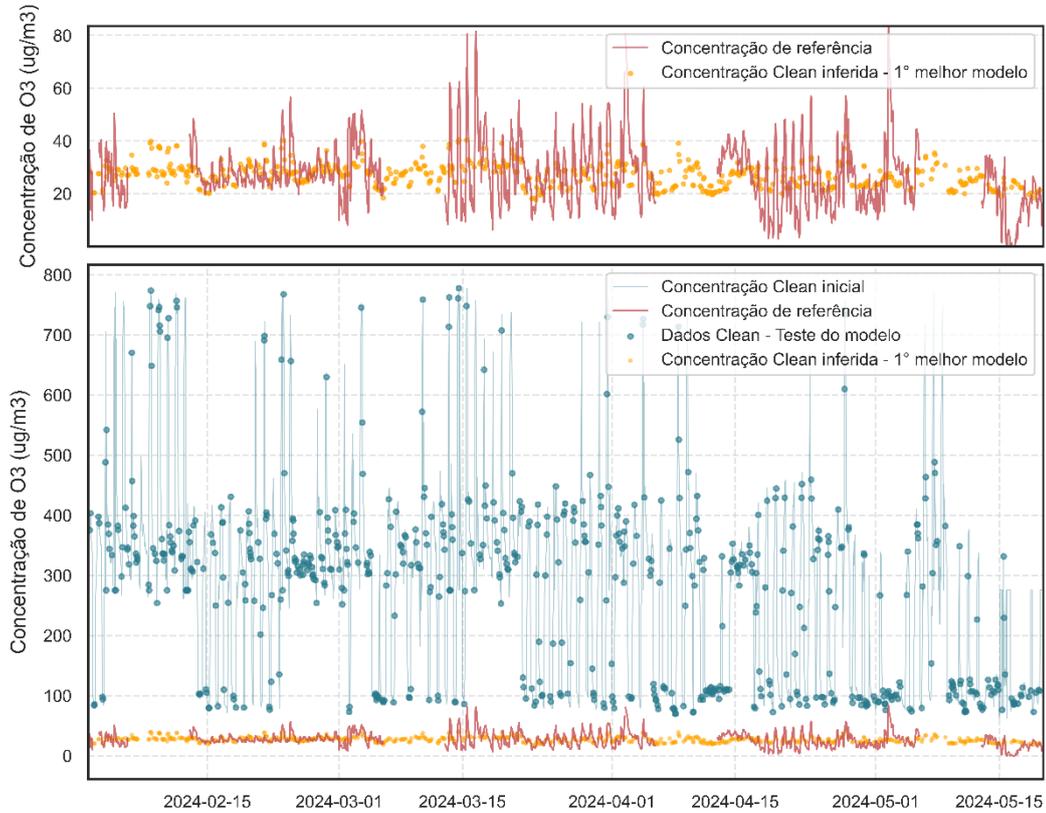
Fonte: Elaborado pela autora (2024)

## APÊNDICE S – CORRELAÇÃO E HISTOGRAMAS DAS CONCENTRAÇÕES DE O<sub>3</sub> – CLEAN INFERIDA PELA CALIBRAÇÃO, DE REFERÊNCIA E DOS RESÍDUOS GERADOS



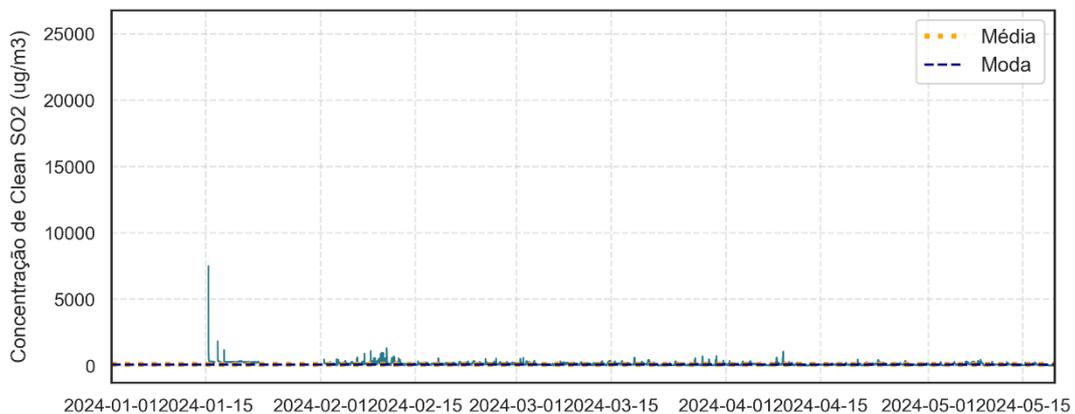
Fonte: Elaborado pela autora (2024)

**APÊNDICE U – SÉRIE TEMPORAL DE CONCENTRAÇÕES DE O<sub>3</sub> – CLEAN DE ENTRADA DO MELHOR MODELO DE REGRESSÃO, AMOSTRAS DE TESTE, VALORES PREVISTOS E OS DADOS ALVO DE REFERÊNCIA**



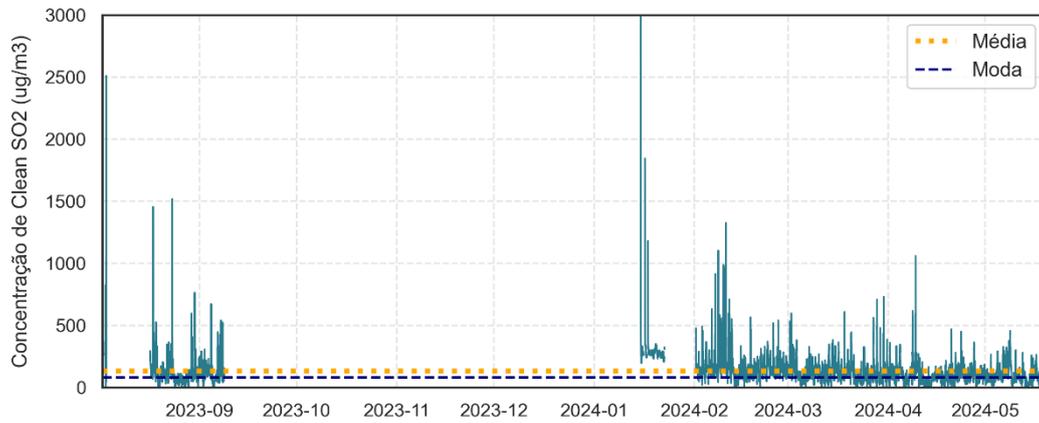
Fonte: Elaborado pela autora (2024)

**APÊNDICE V – SÉRIE TEMPORAL DOS DADOS BRUTOS DE CONCENTRAÇÃO DE SO<sub>2</sub> – CLEAN COM DADOS AGRUPADOS EM 15 MINUTOS**



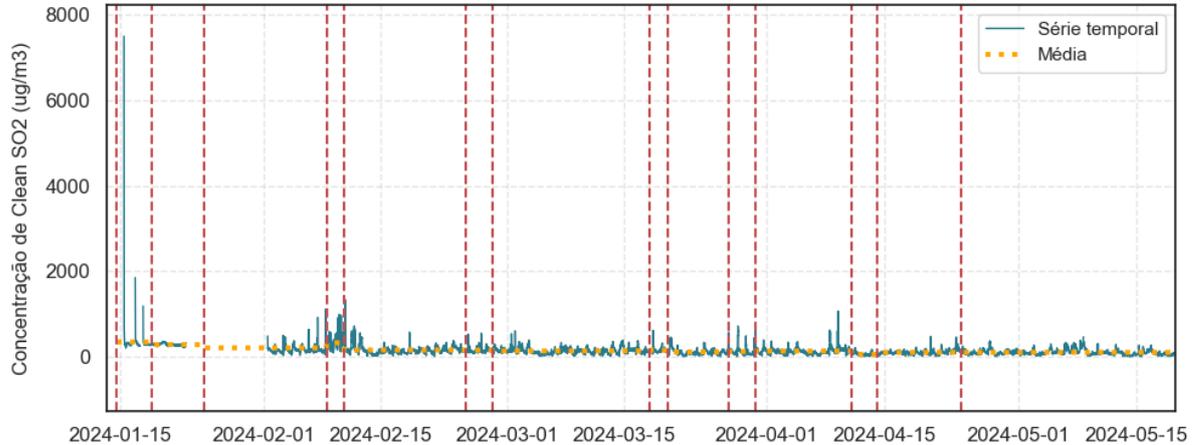
Fonte: Elaborado pela autora (2024)

## APÊNDICE X – SÉRIE TEMPORAL DE CONCENTRAÇÃO DE SO<sub>2</sub> – CLEAN DOS DADOS VÁLIDOS APÓS ETAPA DE CONTROLE DE QUALIDADE COM DADOS AGRUPADOS EM 15 MINUTOS



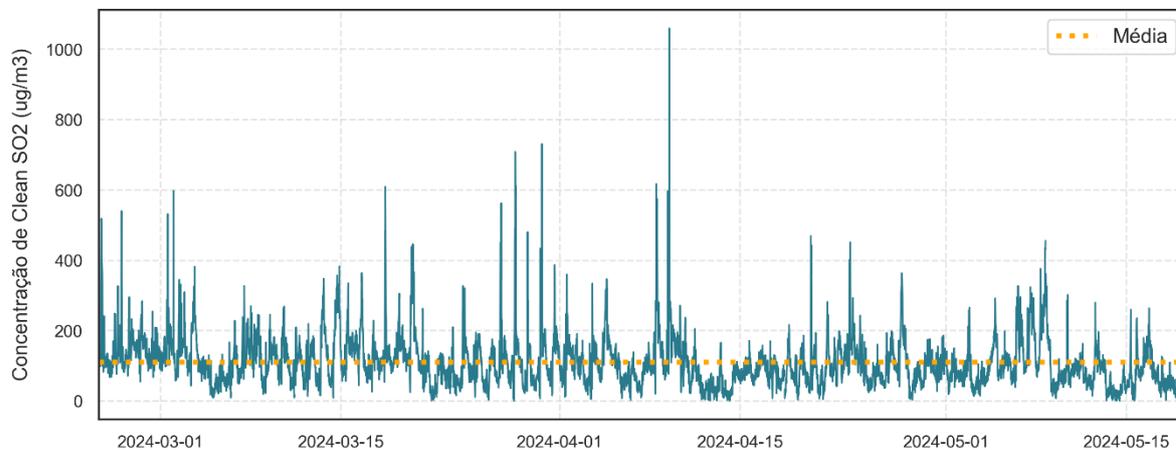
Fonte: Elaborado pela autora (2024)

## APÊNDICE Y – MUDANÇAS ABRUPTAS GERADAS POR PELT NA SÉRIE TEMPORAL DE CONCENTRAÇÃO SO<sub>2</sub> – CLEAN COM DADOS AGRUPADOS EM 15 MINUTOS



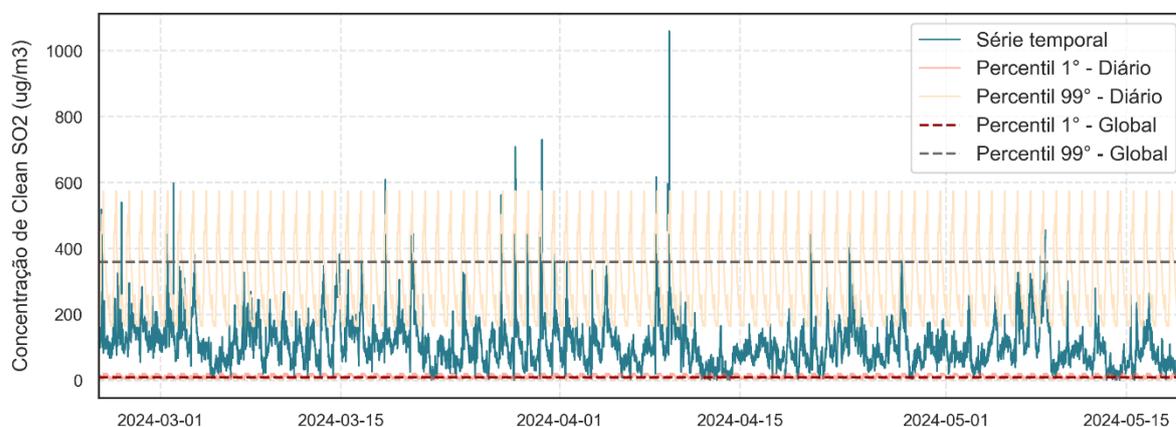
Fonte: Elaborado pela autora (2024)

**APÊNDICE Z – JANELA DE MEDIÇÃO APÓS IMPLEMENTA PELT NA SÉRIE TEMPORAL DE CONCENTRAÇÃO DE SO<sub>2</sub> – CLEAN COM DADOS AGRUPADOS EM 15 MINUTOS**



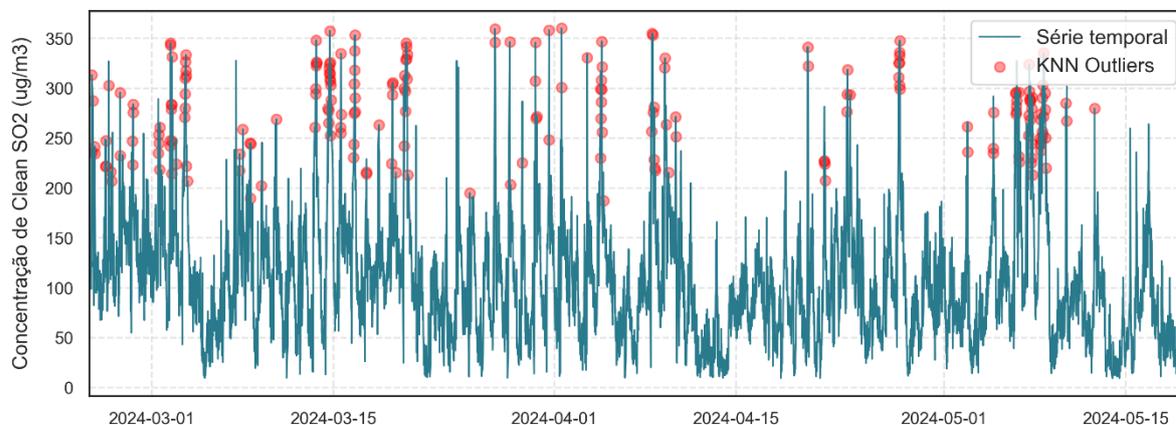
Fonte: Elaborado pela autora (2024)

**APÊNDICE AA – DETECÇÃO DE OUTLIERS POR QUARTIL 1% E 99% GLOBAL E DIÁRIO DAS CONCENTRAÇÕES DE SO<sub>2</sub> – CLEAN COM DADOS AGRUPADOS EM 15 MINUTOS**



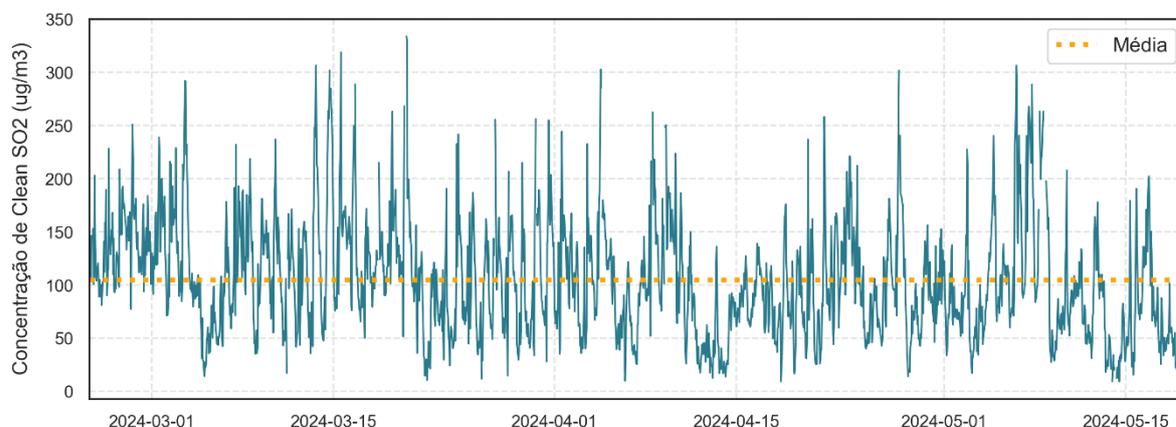
Fonte: Elaborado pela autora (2024)

## APÊNDICE BB – DETECÇÃO DE OUTLIERS POR KNN DAS CONCENTRAÇÕES DE SO<sub>2</sub> – CLEAN COM DADOS AGRUPADOS EM 15 MINUTOS



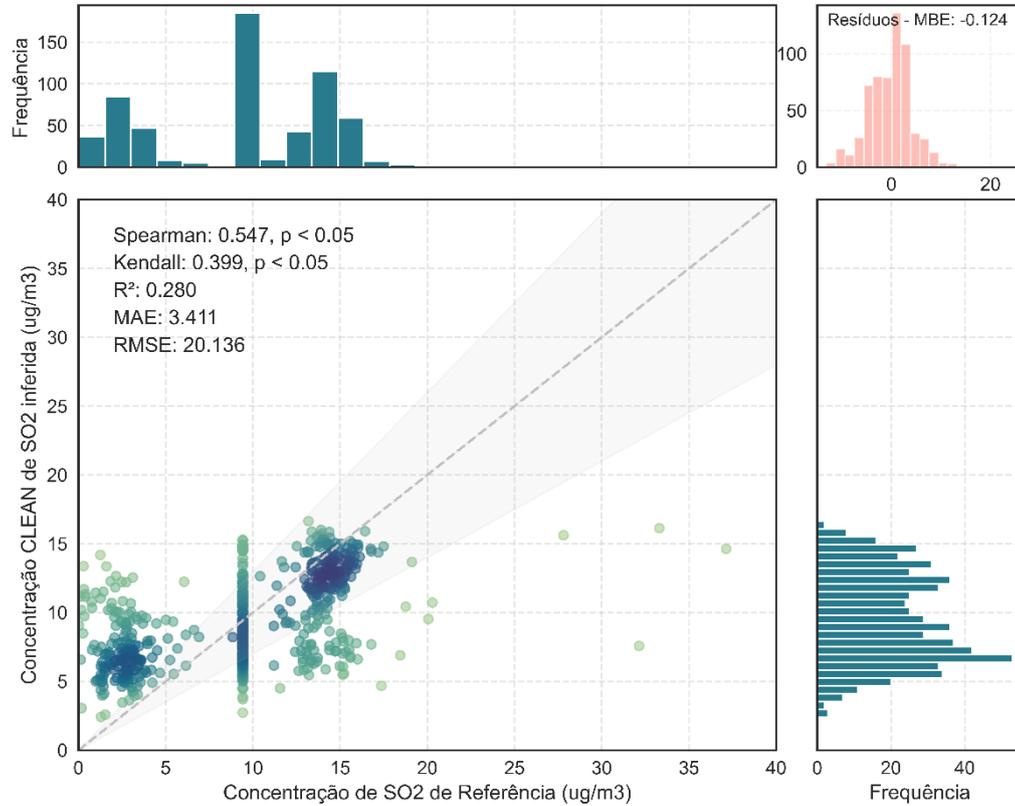
Fonte: Elaborado pela autora (2024)

## APÊNDICE CC – SÉRIE TEMPORAL FINAL DAS CONCENTRAÇÕES DE SO<sub>2</sub> – CLEAN APÓS ETAPAS DE FILTRAGEM E DETECÇÃO DE FALHAS, COM DADOS AGRUPADOS EM MÉDIAS HORÁRIAS



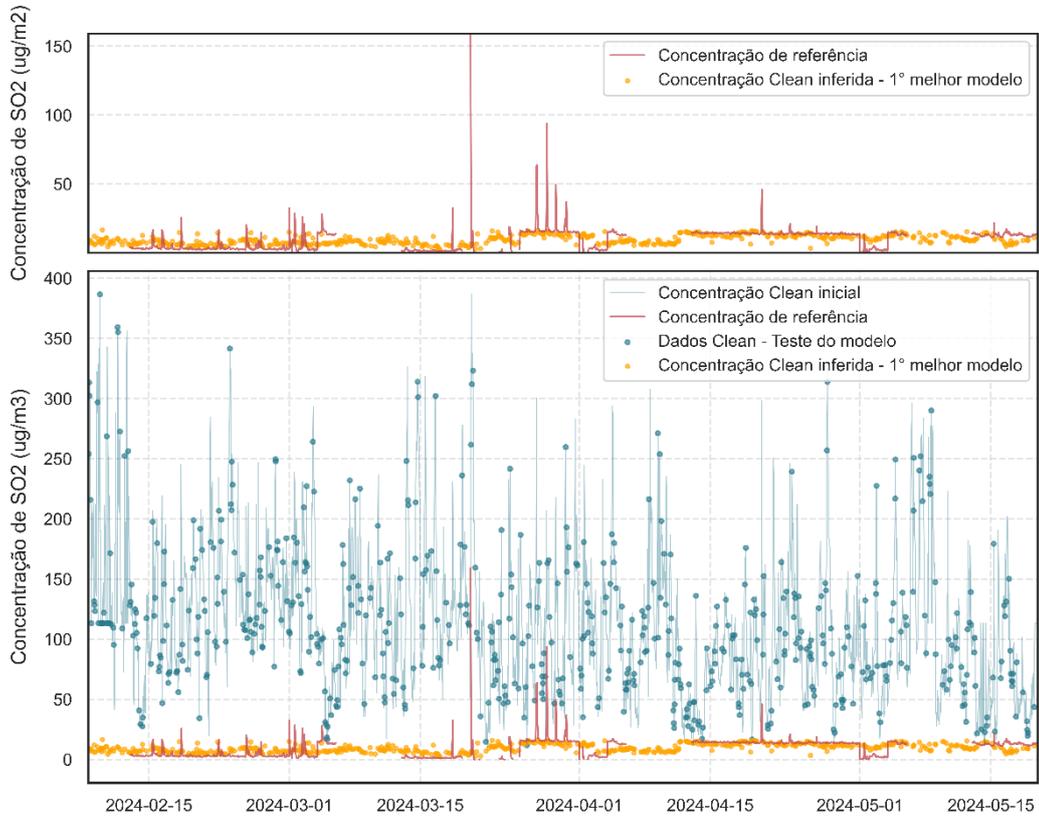
Fonte: Elaborado pela autora (2024)

## APÊNDICE DD – CORRELAÇÃO E HISTOGRAMAS DAS CONCENTRAÇÕES DE SO<sub>2</sub> – CLEAN INFERIDA PELA CALIBRAÇÃO, DE REFERÊNCIA E DOS RESÍDUOS GERADOS



Fonte: Elaborado pela autora (2024)

**APÊNDICE EE – SÉRIE TEMPORAL DE CONCENTRAÇÕES DE SO<sub>2</sub> – CLEAN DE ENTRADA DO MELHOR MODELO DE REGRESSÃO, AMOSTRAS DE TESTE, VALORES PREVISTOS E OS DADOS ALVO DE REFERÊNCIA**



Fonte: Elaborado pela autora (2024)

**APÊNDICE FF – RESULTADOS PARCIAIS DA CALIBRAÇÃO DOS SENSORES CO-B4 CLEAN CONSIDERANDO AS MELHORES MÉTRICAS OBTIDAS EM TERMOS DE R<sup>2</sup>, RMSE E MAE**

Item	Modelo	Variáveis de entrada	R2	RMSE	MAE	AIC	BIC
1 °	Regressão multivariada	CO, Temp. Interna, Temp. Externa,	0.191	0.530	0.428	813.426	799.986
2 °	Regressão multivariada	CO, O3, Temp. Interna, Temp. Externa,	0.188	0.531	0.428	812.490	794.569
3 °	Regressão multivariada	CO, Temp. Interna, Umidade Interna, Temp. Externa,	0.172	0.534	0.432	818.472	800.551
4 °	Regressão multivariada	CO, O3, Temp. Interna, Umidade Interna, Temp. Externa,	0.171	0.534	0.432	817.062	794.661
5 °	Regressão multivariada	CO, O3, Temp. Interna, Press. Interna, Temp. Externa,	0.151	0.539	0.434	814.390	791.989
6 °	Regressão multivariada	CO, Temp. Interna, Press. Interna, Temp. Externa,	0.149	0.540	0.434	815.132	797.212
7 °	Regressão multivariada	CO, O3, Temp. Interna, Umidade Interna, Press. Interna, Temp. Externa,	0.113	0.546	0.441	819.777	792.896
8 °	Regressão multivariada	CO, Temp. Interna, Umidade Interna, Press. Interna, Temp. Externa,	0.109	0.547	0.441	821.041	798.641
9 °	Regressão MLP	CO, O3, Temp. Interna, Temp. Externa,	0.102	0.554	0.447	875.623	857.703
10 °	Regressão MLP	CO, O3, Temp. Interna, Umidade Interna, Temp. Externa,	0.090	0.554	0.451	849.532	827.132
11 °	Florestas Aleatórias	CO, Temp. Interna, Temp. Externa,	-0.005	0.573	0.467	885.365	871.925
12 °	Florestas Aleatórias	CO, O3, Temp. Interna, Umidade Interna, Temp. Externa,	-0.008	0.573	0.464	901.882	879.482
13 °	Florestas Aleatórias	CO, Temp. Interna, Umidade Interna, Temp. Externa,	-0.016	0.574	0.468	905.168	887.248
14 °	Florestas Aleatórias	CO, O3, Temp. Interna, Temp. Externa,	-0.019	0.576	0.466	888.878	870.958
15 °	Regressão multivariada	CO, O3, Temp. Interna, Umidade Interna,	-0.051	0.585	0.471	750.495	732.574
16 °	Regressão multivariada	CO, O3, Temp. Interna,	-0.061	0.587	0.473	750.064	736.624
17 °	Regressão multivariada	CO, Temp. Interna, Umidade Interna,	-0.071	0.589	0.471	746.947	733.506
18 °	Regressão multivariada	CO, Temp. Interna,	-0.075	0.590	0.472	746.649	737.689
19 °	Regressão MLP	CO, Umidade Interna, Temp. Externa,	-0.096	0.591	0.473	879.926	866.486
20 °	Regressão MLP	CO, Temp. Interna, Umidade Interna, Temp. Externa,	-0.106	0.593	0.480	855.955	838.034
21 °	k-Vizinhos mais próximos	CO, O3, Temp. Interna, Umidade Interna, Temp. Externa,	-0.110	0.600	0.480	909.826	887.426
22 °	Florestas Aleatórias	CO, Temp. Externa,	-0.119	0.593	0.471	830.339	821.379
23 °	k-Vizinhos mais próximos	CO, O3, Temp. Interna, Umidade Interna, Press. Interna, Temp. Externa,	-0.120	0.596	0.473	1005.882	979.002
24 °	Florestas Aleatórias	CO, Temp. Interna, Umidade Interna, Press. Interna, Temp. Externa,	-0.123	0.598	0.471	1078.300	1055.900
25 °	k-Vizinhos mais próximos	CO, O3, Temp. Interna, Temp. Externa,	-0.125	0.601	0.480	888.007	870.087
26 °	Regressão multivariada	CO, O3, Temp. Externa,	-0.127	0.600	0.474	744.405	730.965
27 °	Florestas Aleatórias	CO, O3, Temp. Interna, Umidade Interna, Press. Interna, Temp. Externa,	-0.140	0.604	0.478	1094.468	1067.588
28 °	Florestas Aleatórias	CO, Temp. Interna, Umidade Interna,	-0.143	0.604	0.491	827.711	814.270
29 °	Florestas Aleatórias	CO, O3, Temp. Externa,	-0.147	0.600	0.477	838.922	825.482
30 °	Regressão multivariada	CO, O3, Temp. Interna, Press. Interna,	-0.148	0.601	0.484	748.802	730.882

**APÊNDICE GG – RESULTADOS PARCIAIS DA CALIBRAÇÃO DOS SENSORES  
SO2O-B4 CLEAN CONSIDERANDO AS MELHORES MÉTRICAS OBTIDAS EM  
TERMOS DE R<sup>2</sup>, RMSE E MAE**

Item	Modelo	Variáveis de entrada	R2	RMSE	MAE	AIC	BIC
1 °	Regressão multivariada	SO2, Temp. Interna, Umidade Interna, Temp. Externa,	-0.164	5.902	3.842	1839.538	1857.192
2 °	Regressão multivariada	SO2, O3, Temp. Interna, Umidade Interna, Temp. Externa,	-0.187	5.943	3.870	1841.050	1863.117
3 °	Regressão multivariada	SO2, Umidade Interna, Temp. Externa,	-0.191	5.935	3.901	1848.232	1861.473
4 °	Regressão multivariada	SO2, CO, Temp. Interna, Umidade Interna, Temp. Externa,	-0.197	5.964	3.917	1844.493	1866.561
5 °	Regressão multivariada	SO2, O3, Umidade Interna, Temp. Externa,	-0.202	5.955	3.906	1847.861	1865.515
6 °	Regressão multivariada	SO2, CO, O3, Temp. Interna, Umidade Interna, Temp. Externa,	-0.212	5.991	3.929	1845.955	1872.436
7 °	Regressão multivariada	SO2, CO, Umidade Interna, Temp. Externa,	-0.233	6.019	4.004	1852.327	1869.981
8 °	Regressão multivariada	SO2, CO, O3, Umidade Interna, Temp. Externa,	-0.234	6.019	3.983	1852.013	1874.080
9 °	Regressão MLP	SO2, O3, Umidade Interna, Temp. Externa,	-0.261	6.055	4.110	1733.953	1751.607
10 °	Regressão multivariada	SO2, Temp. Interna, Temp. Externa,	-0.273	6.116	4.039	1886.888	1900.129
11 °	Regressão MLP	SO2, Temp. Interna, Umidade Interna, Temp. Externa,	-0.291	6.071	4.014	1736.277	1753.931
12 °	Regressão multivariada	SO2, CO, Temp. Interna, Temp. Externa,	-0.305	6.184	4.121	1894.197	1911.851
13 °	Regressão multivariada	SO2, O3, Temp. Interna, Temp. Externa,	-0.308	6.169	4.089	1886.335	1903.989
14 °	Regressão multivariada	SO2, Temp. Externa,	-0.309	6.171	4.099	1897.254	1906.081
15 °	k-Vizinhos mais próximos	SO2, O3, Temp. Externa,	-0.321	6.252	4.141	1886.818	1900.059
16 °	Regressão multivariada	SO2, CO, O3, Temp. Interna, Temp. Externa,	-0.324	6.212	4.124	1893.062	1915.129
17 °	Regressão MLP	SO2, Temp. Externa,	-0.330	6.151	3.988	1855.880	1864.706
18 °	Regressão multivariada	SO2, O3, Temp. Externa,	-0.343	6.217	4.116	1899.553	1912.794
19 °	k-Vizinhos mais próximos	SO2, CO, O3, Temp. Interna, Umidade Interna, Temp. Externa,	-0.359	6.246	4.026	1690.511	1716.992
20 °	Regressão MLP	SO2, Umidade Interna, Temp. Externa,	-0.366	6.187	4.234	1785.492	1798.732
21 °	Regressão multivariada	SO2, CO, Temp. Externa,	-0.372	6.298	4.255	1904.134	1917.375
22 °	k-Vizinhos mais próximos	SO2, Temp. Interna, Umidade Interna, Temp. Externa,	-0.373	6.247	4.073	1754.117	1771.771
23 °	Regressão multivariada	SO2, CO, O3, Temp. Externa,	-0.381	6.306	4.227	1906.344	1923.998
24 °	k-Vizinhos mais próximos	SO2, CO, Temp. Interna, Umidade Interna, Temp. Externa,	-0.389	6.277	4.071	1711.476	1733.543
25 °	k-Vizinhos mais próximos	SO2, O3, Temp. Interna, Umidade Interna, Temp. Externa,	-0.390	6.286	4.083	1707.416	1729.484
26 °	Regressão MLP	SO2, CO, Umidade Interna, Temp. Externa,	-0.392	6.248	4.239	1750.601	1768.255
27 °	Florestas Aleatórias	SO2, CO, Temp. Externa,	-0.400	6.340	4.125	1814.611	1827.851
28 °	k-Vizinhos mais próximos	SO2, O3, Temp. Interna, Temp. Externa,	-0.412	6.398	4.226	1870.301	1887.955
29 °	k-Vizinhos mais próximos	SO2, CO, Umidade Interna, Temp. Externa,	-0.424	6.334	4.059	1740.545	1758.199
30 °	k-Vizinhos mais próximos	SO2, CO, O3, Umidade Interna, Temp. Externa,	-0.433	6.380	4.125	1697.321	1719.389

**APÊNDICE HH – RESULTADOS PARCIAIS DA CALIBRAÇÃO DOS SENSORES  
NO2-B4 CLEAN CONSIDERANDO AS MELHORES MÉTRICAS OBTIDAS EM  
TERMOS DE R<sup>2</sup>, RMSE E MAE**

Item	Modelo	Variáveis de entrada	R2	RMSE	MAE	AIC	BIC
1 °	Regressão multilinear	NO2, SO2, CO, O3, Umidade Interna,	0.154	4.953	3.678	1860.063	1881.913
2 °	Regressão multilinear	NO2, SO2, CO, O3, Umidade Interna, Temp. Externa,	0.151	4.964	3.685	1861.380	1887.599
3 °	Regressão multilinear	NO2, CO, O3, Umidade Interna,	0.141	4.991	3.698	1858.899	1876.378
4 °	Regressão multilinear	NO2, SO2, CO, O3, Temp. Interna, Umidade Interna,	0.137	5.009	3.711	1861.159	1887.378
5 °	Regressão multilinear	NO2, SO2, CO, Temp. Interna, Umidade Interna,	0.135	5.012	3.724	1857.478	1879.328
6 °	Regressão multilinear	NO2, CO, O3, Umidade Interna, Temp. Externa,	0.133	5.015	3.713	1861.628	1883.478
7 °	Florestas Aleatórias	NO2, CO, Temp. Interna,	0.131	5.022	3.611	1819.378	1832.487
8 °	Florestas Aleatórias	NO2, SO2, CO, O3, Temp. Interna, Umidade Interna,	0.130	5.008	3.726	1711.782	1738.001
9 °	Regressão multilinear	NO2, SO2, CO, O3, Temp. Interna, Umidade Interna, Temp. Externa,	0.129	5.034	3.730	1862.080	1892.670
10 °	Regressão MLP	NO2, CO, Temp. Interna,	0.125	5.036	3.633	1786.801	1799.911
11 °	Regressão multilinear	NO2, CO, Temp. Interna, Umidade Interna,	0.125	5.042	3.739	1857.657	1875.137
12 °	Florestas Aleatórias	NO2, SO2, CO, O3, Temp. Interna,	0.124	5.035	3.734	1767.508	1789.358
13 °	Regressão multilinear	NO2, CO, Umidade Interna,	0.122	5.051	3.760	1863.256	1876.366
14 °	Regressão multilinear	NO2, SO2, CO, Temp. Interna, Umidade Interna, Temp. Externa,	0.121	5.057	3.764	1858.303	1884.522
15 °	Regressão multilinear	NO2, CO, O3, Temp. Interna, Umidade Interna,	0.121	5.053	3.731	1860.361	1882.210
16 °	Florestas Aleatórias	NO2, CO, O3,	0.114	5.066	3.639	1826.006	1839.116
17 °	Regressão multilinear	NO2, SO2, CO, Umidade Interna,	0.113	5.081	3.788	1865.237	1882.717
18 °	Regressão multilinear	NO2, CO, Umidade Interna, Temp. Externa,	0.113	5.080	3.783	1866.707	1884.186
19 °	Regressão multilinear	NO2, SO2, CO, O3,	0.111	5.081	3.735	1891.185	1908.665
20 °	k-Vizinhos mais próximos	NO2, SO2, CO, O3,	0.109	5.073	3.755	1827.338	1844.818
21 °	Florestas Aleatórias	NO2, SO2, CO, O3, Umidade Interna,	0.109	5.065	3.786	1726.077	1747.927
22 °	Regressão multilinear	NO2, CO, O3, Temp. Interna, Umidade Interna, Temp. Externa,	0.108	5.096	3.764	1863.116	1889.335
23 °	Regressão multilinear	Externa,	0.107	5.096	3.781	1860.085	1881.934
24 °	Florestas Aleatórias	NO2, CO,	0.106	5.096	3.660	1867.733	1876.473
25 °	Regressão MLP	NO2, CO,	0.106	5.083	3.666	1893.798	1902.538
26 °	Florestas Aleatórias	NO2, SO2, CO, O3, Umidade Interna, Temp. Externa,	0.106	5.062	3.716	1700.914	1727.134
27 °	Regressão multilinear	NO2, SO2, CO, Temp. Interna,	0.103	5.101	3.743	1881.993	1899.473
28 °	Regressão multilinear	NO2, SO2, CO, O3, Temp. Interna,	0.103	5.103	3.747	1887.138	1908.987
29 °	Regressão multilinear	NO2, SO2, CO, O3, Temp. Externa,	0.103	5.103	3.750	1893.721	1915.571
30 °	Regressão multilinear	NO2, CO, O3,	0.103	5.104	3.741	1888.968	1902.077

**APÊNDICE II – RESULTADOS PARCIAIS DA CALIBRAÇÃO DOS SENSORES O3-B4 CLEAN CONSIDERANDO AS MELHORES MÉTRICAS OBTIDAS EM TERMOS DE R<sup>2</sup>, RMSE E MAE**

Item	Modelo	Variáveis de entrada	R2	RMSE	MAE	AIC	BIC
1 °	Regressão multilinear	O3, Umidade Interna,	0.067	9.403	7.029	2900.173	2909.133
2 °	Regressão multilinear	O3, Humidade Externa, Umidade Interna,	0.050	9.454	7.022	2888.936	2902.376
3 °	Florestas Aleatórias	Externa,	0.048	9.339	6.828	2473.867	2496.267
4 °	Regressão multilinear	O3, Temp. Interna, Umidade Interna,	0.047	9.480	7.125	2893.525	2906.966
5 °	Regressão multilinear	O3, Umidade Interna, Temp. Externa,	0.046	9.474	7.031	2893.575	2907.015
6 °	Regressão multilinear	O3, CO, Umidade Interna,	0.041	9.524	7.091	2900.117	2913.557
7 °	Regressão multilinear	Externa,	0.040	9.486	7.041	2889.982	2907.903
8 °	Regressão multilinear	O3, Temp. Interna, Umidade Interna, Temp. Externa,	0.035	9.544	7.132	2890.098	2908.019
9 °	Regressão multilinear	O3,	0.034	9.544	6.982	2905.321	2909.801
10 °	Florestas Aleatórias	O3, CO, Temp. Interna, Humidade Externa, Umidade Interna, Temp. Externa,	0.032	9.410	6.898	2460.031	2486.911
11 °	Regressão multilinear	O3, CO, Umidade Interna, Temp. Externa,	0.026	9.581	7.091	2892.757	2910.677
12 °	Regressão multilinear	O3, CO, Temp. Interna, Umidade Interna,	0.024	9.589	7.201	2893.886	2911.806
13 °	Florestas Aleatórias	Externa,	0.024	9.491	6.932	2499.498	2521.898
14 °	Regressão multilinear	O3, Humidade Externa,	0.021	9.575	6.912	2895.530	2904.490
15 °	Regressão multilinear	O3, Temp. Externa,	0.019	9.586	6.936	2894.508	2903.468
16 °	Regressão MLP	O3,	0.017	9.529	6.962	2859.860	2864.340
17 °	Regressão multilinear	O3, CO, Humidade Externa, Umidade Interna,	0.016	9.627	7.138	2890.714	2908.635
18 °	k-Vizinhos mais próximos	O3, CO, Temp. Interna, Humidade Externa, Temp. Externa,	0.015	9.629	6.966	2717.072	2739.472
19 °	Regressão multilinear	Externa,	0.015	9.651	7.208	2889.595	2911.995
20 °	Regressão multilinear	O3, Temp. Interna, Temp. Externa,	0.013	9.621	7.067	2889.730	2903.171
21 °	Regressão multilinear	O3, Temp. Interna,	0.013	9.619	7.114	2895.989	2904.949
22 °	Regressão multilinear	O3, Humidade Externa, Temp. Externa,	0.011	9.609	6.918	2892.706	2906.146
23 °	Regressão multilinear	Externa,	0.010	9.646	7.147	2890.726	2913.126
24 °	Regressão multilinear	O3, CO,	0.010	9.674	7.084	2907.394	2916.354
25 °	Florestas Aleatórias	O3, Umidade Interna,	0.007	9.641	7.227	2798.327	2807.287
26 °	Florestas Aleatórias	Interna,	0.007	9.542	7.128	2512.562	2534.963
27 °	Regressão multilinear	Interna,	0.006	9.698	7.280	2886.632	2904.552
28 °	Florestas Aleatórias	O3, CO, Umidade Interna, Temp. Externa,	0.005	9.566	6.993	2512.584	2530.504
29 °	Florestas Aleatórias	O3, CO, Temp. Interna,	0.004	9.589	6.883	2745.722	2759.162
30 °	Regressão multilinear	O3, Temp. Interna, Humidade Externa, Umidade Interna, Temp. Externa,	0.003	9.708	7.280	2888.127	2910.527