



**UNIVERSIDADE DE BRASÍLIA**  
**INSTITUTO DE QUÍMICA**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA**

**DISSERTAÇÃO DE MESTRADO**

**Desempenho de uma Língua Eletrônica Impedimétrica com Algoritmo de Aprendizado de Máquina na Análise de Águas Residuais**

Camila Calandriny Rocha Da Costa

Orientador: Prof. Dr. Leonardo Giordano Paterno

Brasília, DF  
(2024)

**CAMILA CALANDRINY ROCHA DA COSTA**

**Desempenho de uma Língua Eletrônica Impedimétrica com Algoritmo de Aprendizado de Máquina na Análise de Águas Residuais**

*Dissertação de metrado apresentado ao Programa de Pós-Graduação em química (PPGQ), IQ-UnB, como requisito para a obtenção de título de Mestre em Química.*

Orientador: Prof. Dr. Leonardo G. Paterno

Brasília, DF

(2024)

## Catálogo na Fonte

C837d

Costa, Camila Calandriny Rocha da, 1990-

Desempenho de uma Língua Eletrônica Impedimétrica com Algoritmo de Aprendizado de Máquina na Análise de Águas Residuais. [manuscrito] / Camila Calandriny Rocha da Costa.

viii, 74 f.: il.

Dissertação (Mestrado em Química) – Instituto de Química, Universidade de Brasília, 2015.

Orientador: Dr. Leonardo G. Paterno

1. Sensores químicos. 2. Espectroscopia de impedância. 3. Filmes nanoestruturados. 4. Qualidade da água. 5. Aprendizado de máquina. I. Título.

CDU: 552.44:543.42(043)

## FOLHA DE APROVAÇÃO

Comunicamos a aprovação da Defesa de Dissertação de Mestrado do (a) aluno (a) **Camila Calandriny Rocha da Costa**, matrícula nº **21/0040963** intitulada “*Desempenho de uma Língua Eletrônica Impedimétrica com Algoritmo de Aprendizado de Máquina na Análise de Águas Residuais*”, apresentada no Instituto de Química (IQ) da Universidade de Brasília (UnB) em 27 de março de 2024.

Prof. Dr. Leonardo Giordano Paterno  
Presidente de Banca (IQ/UnB)

Prof. Dr. José Roberto dos Santos Politi  
Membro Titular (IQ/UnB)

Dra. Flavia Carneiro da Cunha Oliveira  
Membro Titular (Agência Nacional de Águas e Saneamento Básico)

Prof. Dr. Jez Willian Batista Braga  
Membro Suplente (IQ/UnB)

Dra. Michele Avila dos Santos  
Membro Suplente (Polícia Federal)

Em 27 de março de 2024

## AGRADECIMENTOS

Primeiramente gostaria de agradecer a Deus, “Bendito seja o Senhor, Deus de Israel, que, só ele, faz maravilhas. Bendito seja eternamente seu nome glorioso, e que toda a terra se encha de sua glória” (Sl 71, 18-19). Deus fez maravilhas em meu favor e sem sua graça durante todo o caminho, esse trabalho não seria finalizado. Agradeço à Nossa Senhora por sua incansável intercessão por mim e suas mãos maternais em todo o processo.

Ao meu noivo, Lityz, que sempre me apoiou e me incentivou durante todo o tempo. Sua paciência, carinho e motivações foram muito importantes. À minha família: meu pai Rocha, meu irmão Eduardo e minha madrastra Vanda, que se mostraram sempre muito pacientes e sempre estiveram ao meu lado durante todas as minhas decisões. Em especial ao meu pai e à minha mãe (*in memoriam*) que me formaram e me ajudaram a ser a mulher que hoje eu sou e cujos ensinamentos levo com especial carinho no meu coração.

Aos meus colegas e amigos, pelo companheirismo, incentivos e conselhos.

A todos os meus professores que me fizeram me apaixonar pela ciência, em especial ao meu orientador, professor Leonardo Paterno, que sempre acreditou em mim e que com muita paciência e otimismo me orientou no melhor caminho para a obtenção dos resultados. Seu profissionalismo, além de muitas qualidades pessoais me fazem o admirar e a tê-lo muito mais que um orientador, mas um grande amigo e conselheiro de vida.

À CAESB por disponibilizar as amostras e suas características físico-químicas. Sem ela, esse trabalho não existiria. Em particular, gostaria de agradecer aos analistas de sistema de saneamento da CAESB, Sra. Alessandra Momesso e Sr. Fernando Costa Garcia pela disponibilização das amostras, resultados e discussões durante a execução do projeto. Também, ao pessoal do laboratório que com muita paciência separavam as amostras, Stefany, Eric e Luís. Ao Afrânio que, na reta final, tirou muitas dúvidas minhas sobre o processo de tratamento de águas residuais.

## RESUMO

A água é crucial para uma variedade de atividades humanas e, portanto, deve ser tratada após o uso para garantir que retorne ao meio ambiente em condições adequadas, evitando danos à saúde humana e/ou desequilíbrios ecológicos. As águas despejadas na rede de esgoto, águas residuais, são direcionadas para estações de tratamento, onde são retirados excessos de resíduos e contaminantes. O controle da eficiência do tratamento é feito por análises laboratoriais de diferentes parâmetros, que por muitas vezes demandam muitos recursos. Assim, é importante a pesquisa e inovações tecnológicas que possam assistir ao processo de controle da qualidade de água. Os sensores químicos são uma tecnologia promissora, sendo dispositivos portáteis capazes de detectar várias substâncias químicas de interesse, o que auxilia no monitoramento em tempo real e *in loco*. Nessa linha, o presente projeto propõe o desenvolvimento e avaliação de um sistema multissensorial tipo língua eletrônica para o auxílio no processo de controle de qualidade de tratamento de água residual. A língua eletrônica a ser desenvolvida é composta de um arranjo de sensores impedimétricos inespecíficos, formados por eletrodos interdigitados modificados com filmes ultrafinos de nanomateriais diversos, um analisador de impedância e um ou mais algoritmos de classificação. As amostras de água residuais foram fornecidas pela Companhia de Saneamento Ambiental do Distrito Federal (CAESB), obtidas de diferentes estações de tratamento de esgoto da região do Distrito Federal. Foram analisadas 212 amostras de 7 estações diferentes, coletadas entre os meses de novembro e dezembro de 2022, em janeiro e de setembro a dezembro de 2023. Verificou-se que a língua eletrônica desenvolvida consegue distinguir as amostras de acordo com o tipo do processo de tratamento operado pela estação e não, necessariamente, pela sua localização geográfica, usando algoritmos classificatórios supervisionados e não supervisionados. A implementação de aprendizado de máquina para quatro parâmetros de qualidade: demanda química de oxigênio (DQO), sólidos suspensos totais (SST), nitrogênio total (NT) e fósforo total (PHOSP), permitiu que a língua eletrônica alcançasse índices de acertos superiores a 50% para a maioria dos parâmetros, chegando a mais de 70% para o caso de NT, mesmo com um banco de dados baixo para a complexidade do projeto. A língua eletrônica desenvolvida mostrou resultados promissores, com leituras estáveis e sensores duráveis para amostras de grande complexidade.

**Palavras-Chave:** Sensores químicos. Espectroscopia de impedância. Filmes nanoestruturados. Qualidade da água. Aprendizado de máquina.

## ABSTRACT

Water is crucial for a variety of human activities and, therefore, must be treated after use to guarantee its return to the environment in suitable conditions, avoiding damage to human health and/or ecological imbalances. Wastewater, discharged into the sewage system, is directed to water treatment stations, where excess residues and contaminants are removed. Control of treatment efficiency is achieved through laboratory analyses of various parameters, which require significant resources. Hence, scientific research and technological innovations are essential to assist in the process of water quality control. Chemical sensors represent a promising technology; these portable devices can detect various chemical substances of interest, facilitating real-time and on-site monitoring. In line with this, the present project proposes the development and evaluation of a multisensorial system, called an electronic tongue, to aid in wastewater treatment quality control. The electronic tongue to be developed will consist of an arrangement of nonspecific impedimetric sensors, comprised of interdigital electrodes modified with ultrathin films of various nanomaterials, an impedance analyzer, and one or more classification algorithms. Wastewater samples were provided by the Companhia de Saneamento Ambiental do Distrito Federal (CAESB) and obtained from different sewage treatment stations in the Distrito Federal region. A total of 212 samples from 7 different stations, collected during November and December of 2022, January of 2023 and from September to December of 2023, were analyzed. It was found that the developed electronic tongue can distinguish samples according to the type of treatment process operated by the station, rather than solely by geographic location, using supervised and unsupervised classification algorithms. By implementing machine learning for four quality parameters – chemical oxygen demand (COD), total suspended solids (TSS), total nitrogen (TN), and total phosphorus (PHOSP) – the electronic tongue achieved a success rate of over 50% for most parameters, exceeding 70% for TN, despite using a small dataset for the complexity of the project. The developed electronic tongue exhibited promising results, with stable readings and durable sensors for samples of high complexity.

**Keywords:** Chemical sensors. Impedimetric spectroscopy. Nanostructured films. Water quality. Machine learning.

## LISTA DE FIGURAS

Figura 1: Esquema ilustrativo do Circuito equivalente utilizado nesse trabalho. ....	17
Figura 2: Ilustração esquemática de um microelétrodo interdigitado. ....	19
Figura 3: Fórmula estrutural dos materiais empregados na fabricação dos sensores. ....	24
Figura 4: Ilustração esquemática da deposição dos filmes pela técnica LbL. ....	26
Figura 5: Visão geral do sistema de medição da língua eletrônica. ....	27
Figura 6: Sistema de medição da língua eletrônica impedimétrica: (a) visão geral e partes indicadas, (b) visão ampliada da cabeça sensora, (c) sensores. ....	27
Figura 7: Fluxograma do tratamento dos dados com o uso da linguagem Python. ....	30
Figura 8: Fluxograma do estudo dos indicadores. ....	31
Figura 9: Mapa de correlação entre sensores e parâmetros ....	33
Figura 10: Mapa de correlação entre sensores. (a) Capacitância. (b) Resistência. ....	35
Figura 11: Gráfico dos pesos de 2 componentes principais, sensores e parâmetros. (a) Capacitância. (b) Resistência. (c) Capacitância e Resistência. ....	38
Figura 12: Gráfico 2D de dois componentes principais dos sensores de todas as amostras para a Capacitância. ....	43
Figura 13: Gráfico 2D de dois componentes principais dos sensores de todas as amostras para a Resistência. ....	44
Figura 14: Gráfico 2D de dois componentes principais dos sensores de todas as amostras para a Capacitância e Resistência. ....	45
Figura 15: Dendrograma dos sensores valores com autoescalados para a Capacitância. ....	47
Figura 16: Dendrograma dos sensores com valores autoescalados para resistência. ....	48
Figura 17: Dendrograma dos sensores com valores autoescalados para a (c) Capacitância e Resistência. ....	49
Figura 18: Matrizes de confusão para a capacitância. Acertos e erros do algoritmo de aprendizado de máquina. (a) DQO, (b) SST, (c) NT e (d) PHOSP. ....	57
Figura 19: Matrizes de confusão para a resistência. Acertos e erros do algoritmo de aprendizado de máquina. (a) DQO, (b) SST, (c) NT e (d) PHOSP. ....	58
Figura 20: Matrizes de confusão para a resistência. Acertos e erros do algoritmo de aprendizado de máquina para os quatro indicadores analisados: (a) DQO, (b) SST, (c) NT e (d) PHOSP. ....	59



Figura 21: PCA 2 componentes para a capacitância dos quatro indicadores analisados: (a) DQO, (b) SST, (c) NT e (d) PHOSP. ....	61
Figura 22: PCA 2 componentes para a resistência dos quatro indicadores analisados: (a) DQO, (b) SST, (c) NT e (d) PHOSP. ....	62
Figura 23: PCA 2 componentes para a capacitância e resistência dos quatro indicadores analisados: (a) DQO, (b) SST, (c) NT e (d) PHOSP. ....	63
Figura 24: Capacitância pela concentração de DQO.....	64
Figura 25: PCA dois componentes para as estações Norte e Sul. (a) Capacitância. (b) Resistência. (c) Capacitância e Resistência. ....	66

## LISTA DE TABELAS

Tabela 1: Descrição dos sensores impedimétricos .....	26
Tabela 2: Variância explicada para capacitância.....	36
Tabela 3: Variância explicada para resistência.....	36
Tabela 4: Variância explicada para capacitância e resistência. ....	36
Tabela 5: Variância explicada para capacitância - Sensores. ....	40
Tabela 6: Variância explicada para resistência - Sensores. ....	40
Tabela 7: Variância explicada para capacitância e resistência - Sensores. ....	40
Tabela 8: Relação de quantidade de cada <i>score</i> . ....	55

## LISTA DE QUADROS

Quadro 1: Formação de 7 grupos para capacitância. Dados autoescalados. ....	50
Quadro 2: Formação de 7 grupos para resistência. Dados autoescalados. ....	51
Quadro 3: Formação de 7 grupos para capacitância e resistência. Dados autoescalados.....	52
Quadro 4: Indicadores e faixa de ótimo.....	54
Quadro 5: Porcentagem de acerto da língua eletrônica para os indicadores. ....	55
Quadro 6: porcentagem de acerto língua eletrônica para os indicadores feita na primeira parte do trabalho. ....	56
Quadro 7: Score estações e quantidades.....	64
Quadro 8: Porcentagem de acerto da língua eletrônica para as estações Sul e Norte. ....	65

## LISTA DE ABREVIATURAS, SIGLAS E FÓRMULAS

**AGL** – Águas Lindas Final

**ALA** – Alagado Final

**BRAZ** – Brazlândia Final

**CAESB** - Companhia de Saneamento Ambiental do Distrito Federal

**CP** – Capacitância

**DQO** – Demanda Química de oxigênio

**ETE** – Estação de Tratamento de Esgoto

**Ftc** – ftalocianina

**Ftc-Cu** – ftalocianina de cobre (II)

**Ftc-M** – ftalocianina metalada ou metaloftalocianina

**Ftc-Ts-Cu** – ftalocianina de cobre (II) tetrassulfonada

**Ftc-Ts-Ni** – ftalocianina de níquel (II) tetrassulfonada

**GO** - óxido de grafeno

**HAuCl<sub>4</sub>** – ácido cloroáurico

**HPEI** – poli(etilenimina) hiperamificada

**LbL** – técnica de automontagem camada-por-camada (do inglês, *layer-by-layer*)

**Mag** – suspensão coloidal de magnetita

**MEL** – Melchior Final

**NORBIO** – Norte Biológico

**NORFIN** – Norte Final

**NT** – Nitrogênio Total

**NTC** – Nanotubos de carbono

**PANI** - polianilina

**PDAC** – hidrocloreto de poli(dialil dimetilamônio)

**PHOSP** – Fósforo Total

**PSS** – poli(4 – estireno-sulfonato de sódio)

**RFI** – Riacho Fundo reator I Final

**RFII** – Riacho Fundo reator II Final

**RFIII** – Riacho Fundo reator III Final

**rGO** - óxido de grafeno reduzido

**RS** - Resistência

**SST** – Sólidos Suspensos Totais

**SULBIO** – Sul Biológico

**SULFIN** – Sul Final

## SUMÁRIO

1	INTRODUÇÃO .....	10
2	REVISÃO BIBLIOGRÁFICA.....	12
2.1	Saneamento Básico.....	12
2.2	Métodos de monitoramento da qualidade.....	13
2.3	Língua Eletrônica .....	15
2.3.1	Impedância Elétrica .....	17
2.4	Sensores químicos .....	18
2.5	Análise de Componentes Principais (PCA) e de Agrupamento Hierárquico de Clusters (HCA) .....	19
2.6	Coeficiente de Correlação de Pearson .....	20
2.7	Aprendizado de Máquina.....	21
2.7.1	Regressão Logística .....	22
3	OBJETIVOS.....	23
4	MATERIAIS E MÉTODOS .....	24
4.1	Materiais .....	24
4.2	Metodologia.....	25
4.2.1	Fabricação dos sensores impedimétricos.....	25
4.2.2	Sistema de medidas da língua eletrônica e procedimento de medição .....	26
4.2.3	Coleta e preparo das amostras CAESB .....	28
4.2.4	Programação, tratamento dos dados e aprendizado de máquina .....	29
5	RESULTADOS E DISCUSSÃO .....	32
5.1	Tratamento dos dados .....	32
5.2	Mapa de Calor .....	32
5.3	Análise dos dados por PCA .....	35
5.4	Análise dos sensores - PCA.....	39
5.4.1	Dendrograma e agrupamento.....	45
5.5	Estudo dos indicadores e aprendizado de máquina .....	53
5.6	Estudo das Estações Norte e Sul.....	64
6	CONCLUSÕES.....	68
7	PERSPECTIVAS .....	70
	REFERÊNCIAS .....	71
	ANEXOS .....	78
	ANEXO A - Tabela dos dados coletados. ....	78

ANEXO B - Valores de resistência e capacitância de cada sensor.....	79
ANEXO C- Leituras de capacitância da água para verificação do reset. ....	79
ANEXO D - Gráfico de barras de capacitância e resistência para amostra da ETE Norte Final do dia 14/12/2022 .....	80
ANEXO E - Quadros de variância (Capacitância, Resistência, Capacitância e Resistência). .....	81
ANEXO F - Gráficos 2D dos componentes principais dos sensores e parâmetros. ....	83
ANEXO G - Quadros de variância para os sensores .....	85
ANEXO H - Análise dos parâmetros físicos. ....	86

## 1 INTRODUÇÃO

O tratamento de água é fundamental para a manutenção e desenvolvimento de uma população. A água é vital para o ser humano e todos devem ter acesso a água de qualidade pois é um direito (NAÇÕES UNIDAS, [2023]). A água, além da ingestão para o funcionamento correto do organismo, também é utilizada em diversas atividades humanas, principalmente para a irrigação de lavouras, abastecimento público para uso doméstico, atividades industriais, geração de energia, extração mineral, aquicultura, navegação turismo e lazer (AGÊNCIA..., [2023b]). Devido à sua extrema importância, é essencial que esse recurso seja bem cuidado, tratado e preservado.

A água, tanto para consumo quanto para retornar ao ecossistema, é submetida a diferentes análises para verificar se atende aos padrões e normas técnicas determinadas pelo Ministério da Saúde (BRASIL, 2011), isso depois de passar por diversos e variados processos e análises, tanto laboratoriais quanto em escala industrial. Muitos demandam equipamento, tempo, soluções, reagentes diversos e operadores especializados. São, em geral, demoradas e de custo considerável.

Um único dispositivo capaz de definir se determinada amostra de água cumpre os padrões estabelecidos e ainda com uma rápida resposta seria uma ótima solução para o controle da qualidade da água. A língua eletrônica é um arranjo de sensores químicos inespecíficos que apresenta essa capacidade e está se mostrando muito eficaz na detecção de múltiplos componentes. Seu manuseio simplificado, rápida resposta, baixo custo, fácil integração com outros sistemas, além de uma alta sensibilidade para sistemas complexos (SHIMIZU; BRAUNGER; RIUL, 2019) faz com que possa ser aplicado para o sistema de tratamento de água. Esses sensores são produzidos por nanopartículas de materiais já bem conhecidos e com boa sensibilidade, além de que a deposição desses materiais para a formação dos sensores não exige uma técnica sofisticada e pode ser realizada de maneira simples como a técnica de deposição camada por camada ou LbL (PATERNO, L.G., SOLER, 2013).

O arranjo de sensores da língua eletrônica gera uma quantidade considerável de dados não-lineares. Dessa forma, é essencial o uso de métodos computacionais que processarão esses dados e reconhecerão os padrões entre as amostras. O tratamento desses dados é feito por aprendizado de máquina, que ainda permite que o sistema possa se ajustar e se calibrar sozinho, à medida que novas amostras são inseridas e novos dados são gerados (RIUL JÚNIOR *et al.*, 2010), aumentando ainda mais o desempenho desse dispositivo.



As estações de tratamento de água residual (ETEs) removem as cargas de poluentes por meio de processos físicos, químicos ou biológicos, devolvendo ao meio ambiente a água final tratada conforme os padrões exigidos pela legislação ambiental. As ETEs do Distrito Federal (DF) são administradas pela Companhia de Saneamento do Distrito Federal (CAESB) e contam com laboratórios para realizar análises de determinados parâmetros em pontos específicos do processo e certificam se a água pode ser devolvida à natureza sem causar danos ambiental e à saúde humana. (CAESB, Análises e resultados). Contudo, muitas dessas análises levam um tempo considerável para retornar uma resposta e o fluxo que passa pelo processo é muito alto. Então, a língua eletrônica seria um dispositivo que aumentaria a eficiência de análise e controle do efluente e ajudaria ainda mais o processo da ETE.

Assim, esse projeto de mestrado propõe o desenvolvimento de um sistema multissensorial tipo língua eletrônica e a avaliação de seu desempenho no controle de qualidade de tratamento de água residual por meio de dois pontos: se a língua eletrônica é capaz de diferenciar as amostras de água de acordo com a localidade das estações de tratamento e se, utilizando um algoritmo de aprendizado supervisionado, a língua eletrônica consegue classificar os parâmetros de amostras coletadas. A língua eletrônica desenvolvida nesse trabalho é composta por um arranjo de sensores impedimétricos inespecíficos, formados por eletrodos interdigitados modificados com filmes ultrafinos de nanomateriais diversos; um analisador de impedância; e um ou mais métodos analíticos e lógicos para o reconhecimento de padrões e classificação das amostras. As amostras de água foram fornecidas pela CAESB, obtidas de diferentes estações de tratamento de esgoto da cidade de Brasília-DF e regiões administrativas.

## **2 REVISÃO BIBLIOGRÁFICA**

### **2.1 Saneamento Básico**

A água é essencial para a manutenção e desenvolvimento da sociedade. É o principal recurso e ter acesso a água de qualidade é um direito humano. A água é utilizada nas mais diferentes atividades humanas e cuidar para que esse recurso esteja sempre disponível é imprescindível. Para isso, o saneamento básico tem que ser eficiente e de forma que consiga tratar, se não todo o volume de água, mas a maior parte para que não haja prejuízo tanto à utilização pela sociedade, como ao ecossistema.

O saneamento básico consiste na prestação de serviços e infraestrutura adequadas para que a população tenha água de qualidade nas casas e empresas, destinação segura para o esgoto, destinação segura para os resíduos sólidos (aterro sanitário), cidade limpa e sem alagamentos, abastecimento de água, drenagem da água de chuva por calçadas drenantes, limpeza urbana e coleta de lixo (destinação de lixo), coleta e tratamento de esgoto. De acordo com a Agência Nacional de Águas e Saneamento Básico, em 2020, 93,4% da população urbana foi atendida com rede de abastecimento de água e 63,2% da população urbana foi atendida com rede de coleta de esgotos sanitários. Apenas 50,8% do esgoto gerado no país era tratado e 84,1% da população brasileira era abastecida com água (AGÊNCIA..., 2023).

O Brasil é um país com grande recurso hídrico. Porém, o Distrito Federal possui diversas nascentes que fluem para distintas regiões hidrográficas e, por isso, é um território com baixa disponibilidade hídrica, embora seja uma região com grande volume populacional (CODEPLAN, 2020). Dessa forma, essa região necessita de um saneamento básico muito eficiente, que envolva desde o tratamento, a distribuição e o consumo consciente.

No DF, o sistema de abastecimento de água é operado pela CAESB. São 17 ETEs espalhadas pela região (CAESB, 2023). O monitoramento das captações utilizadas pela CAESB é realizado sistematicamente de acordo com as resoluções nº 357/2005 do Conselho Nacional do Meio Ambiente (CONAMA). O Índice de Qualidade da Água (IQA) é utilizado para avaliar as condições bacteriológicas e físico-químicas de um corpo d'água (CODEPLAN, 2020).

O controle operacional e monitoramento da eficiência do processo de tratamento são realizados em laboratórios que executam análises físico-químicas e microbiológicas. O tratamento adequado permite a retirada de impurezas. A ETE remove cargas de poluentes por

meio de processos físicos, químicos e biológico em conformidade com os padrões exigidos pela legislação ambiental.

O tratamento de esgoto doméstico é dividido em 4 partes: tratamento preliminar, que consiste na utilização de grades para retenção de resíduos maiores; tratamento primário, que consiste na sedimentação dos sólidos em suspensão que se acumulam no fundo do decantador; tratamento secundário, onde os microrganismos se alimentam de matéria orgânica convertendo-a em gás carbônico e água e tratamento terciário, onde são removidos poluentes específicos como fósforo e nitrogênio. O tratamento terciário só está presente em 9 das 17 estações de tratamento do DF. Das estações presentes nesse trabalho, apenas as ETEs Norte, Sul e Alagado contam com o tratamento terciário (CAESB, 2023).

## **2.2 Métodos de monitoramento da qualidade**

O CONAMA, por meio de suas legislações, estabelece condições padrões de lançamentos de efluentes. A legislação entende a importância da água bem como a necessidade do desenvolvimento sustentável. Além disso, considera os termos de convenção de Estocolmo que trata dos Poluentes Orgânicos Persistentes (POPs). A resolução nº357/2005 dispõe sobre a classificação e diretrizes ambientais para o enquadramento dos corpos de água superficiais. Para isso, essa resolução contém a relação de todos os parâmetros que devem ser analisados e qual o padrão que esses parâmetros devem seguir para que a água atenda aos padrões de qualidade exigidos (BRASIL, 2005).

O monitoramento da água envolve análises físico-químicas e biológicas. A água não pode promover efeitos tóxicos crônicos a organismos, materiais flutuantes (inclusive espumas não naturais), óleos e graxas, substâncias que comuniquem gosto ou odor, corantes artificiais, resíduos sólidos e coliformes termotolerantes (BRASIL, 2005). A qualidade da água abrange ainda o nível máximo de certos componentes como fósforo total, manganês, nitrogênio, nitrogênio amoniacal total, sulfato total, nitrito, nitrato etc. Também verifica parâmetros como cor, turbidez, demanda bioquímica de oxigênio (DBO), demanda química de oxigênio (DQO), pH, sólidos suspensos totais (SST), sólidos sedimentáveis, fósforo total (PHOSP), alcalinidade, nitrogênio total (NT), entre outros. Nesse trabalho, o foco foram os parâmetros de demanda química de oxigênio, sólidos suspensos totais, nitrato, nitrogênio de amônia, nitrogênio total e fósforo total.

O parâmetro de DQO indica a presença de matéria orgânica na água. Ela é a principal responsável pelo problema de poluição das águas, pois reduz a concentração de oxigênio dissolvido. Essa diminuição ocorre devido a atividade respiratória das bactérias (BRASIL, 2006).

No meio aquático, o nitrogênio pode ser encontrado em diversas formas: nitrogênio molecular ( $N_2$ ), nitrogênio orgânico, íon amônio ( $NH_4^+$ ), íon nitrito ( $NO_2^-$ ) e íon nitrato ( $NO_3^-$ ). A participação das bactérias no ciclo do nitrogênio é bastante intensa, tanto no processo de nitrificação (oxidação bacteriana do amônio a nitrito e depois nitrato) quanto no de desnitrificação (redução bacteriana do nitrato ao gás nitrogênio). O nitrogênio é um dos nutrientes mais importantes para o crescimento de algas e macrófitas (plantas aquáticas superiores), sendo facilmente assimilável nas formas de amônio e nitrato. Em condições alcalinas, ocorre o predomínio da amônia livre, que é bastante tóxica a vários organismos aquáticos. O nitrogênio é amplamente encontrado na natureza. No entanto, o excesso de nitrogênio em corpos d'água é de origem antropogênica, como consequência do despejo doméstico, industrial e de criatórios de animais, bem como fertilizantes (BRASIL, 2006).

O fósforo é um nutriente muito importante para o crescimento de plantas aquáticas. Quando há o crescimento excessivo dessas plantas, se dá o fenômeno de eutrofização, e isso prejudica o uso da água, além de impactar o ecossistema aquático. O despejo de águas residuais contendo detergentes, pesticidas e fertilizantes é uma das formas em que o fósforo aumenta a sua concentração nas águas (BRASIL, 2006).

Os SSTs podem ser definidos como partículas que podem ser retidas por processos de filtração (BRASIL, 2006). São responsáveis pelo transporte de matéria orgânica e diversos contaminantes. Muitos metais se associam às partículas, sendo os SSTs, então, um vetor de transporte. Além disso, os sólidos suspensos são responsáveis pelo aumento de turbidez da água, o que causa um efeito também estético, além de diminuir a penetração de luz no corpo d'água (MONITORAMENTO..., 2017).

As análises de água residual são conduzidas por meio de diversos métodos e equipamentos. São feitas análises gravimétricas, manométricas, cromatografia iônica e nefelometria (medição da luz que atravessa a solução) e análises com a utilização de pHmetros, espectrofotômetros de infravermelho e UV-VIS. Nos laboratórios da CAESB, a DQO é determinada por espectrofotometria no UV-VIS, assim como nitrogênio de amônia e nitrogênio total. Os íons nitrito e nitrato são analisados por cromatografia iônica e SST é determinado por gravimetria.

Algumas dessas análises demoram horas para serem concluídas e outras utilizam equipamentos muito sofisticados e caros, como no caso dos cromatógrafos. Essas análises obedecem a métodos padronizados internacionalmente.

Nas estações de tratamento da CAESB, a água tratada é coletada e com as análises é verificado se atende aos padrões exigidos. As análises são realizadas em laboratórios dentro das estações.

### **2.3 Língua Eletrônica**

O paladar é um dos 4 sentidos e é utilizado para identificar sabores para assim, avaliar alimentos ou possíveis substâncias tóxicas ao organismo que podem causar algum mal. A língua é um órgão composto de músculo e um arranjo de sensores, chamados de papilas gustativas. Essas enviam sinais ao cérebro que identifica, resumidamente, os sabores: salgado, doce, amargo e azedo. Com o uso constante desses sensores e associando ao cérebro informações sobre os alimentos, os seres humanos são capazes de montar um banco de dados para identificar alimentos e os classificar quanto a bom ou ruim (TOKO, 1996; RIUL, A *et al.*, 2002).

A língua eletrônica é uma ferramenta analítica que mimetiza o sistema gustativo biológico. Ela é composta por uma matriz de sensores químicos, um equipamento de interrogação, p. ex. analisador de impedância, e um algoritmo de aprendizado de máquina. Os sensores são, geralmente, inespecíficos, mas com sensibilidade cruzada, pois todos respondem com menor ou maior sensibilidade à composição global do líquido, e não obrigatoriamente a uma substância em específico. Isso resulta num padrão de reconhecimento chamado de seletividade global (RIUL *et al.*, 2002). As amostras analisadas com a língua eletrônica são líquidas, desde soluções aquosas com concentrações conhecidas de analitos conhecidos até líquidos mais complexos, com muitos componentes dos quais se sabe pouco ou nada da concentração. Os sensores podem ser fabricados com diferentes materiais sensíveis, cuja escolha dependerá do método de transdução. Sua utilização abrange diversas tarefas relacionadas ao paladar como a avaliação de vinhos, sucos, café (RIUL JÚNIOR *et al.*, 2010), mascaramento de sabor em medicamentos (MACHADO *et al.*, 2018) e até a detecção de substâncias tóxicas e pesticidas (CHRISTINELLI *et al.*, 2021; COSTA *et al.*, 2022).

Há diversos modelos de língua eletrônicas relatados na literatura e até mesmo em versões comerciais (ASTREE, 2023), cujas principais diferenças residem no método de transdução e tipos de sensores. Historicamente, são relatados três modelos principais:

potenciométricas (TOKO, 1996; VLASOV, Y, LEGIN, A, 1998), voltamétricas (WINQUIST; WIDE; LUNDSTRÖM, 1997; WINQUIST; KRANTZ-RÜLCKER; WIDE; LUNDSTRÖM, 1998) e impedimétricas (RIUL, A *et al.*, 2002). Em 2021 foi publicado um livro que descreve em detalhes a construção, o funcionamento e as aplicações dos principais tipos de língua eletrônica (SHIMIZU, BRAUNGER, RIUL, 2021).

A língua eletrônica impedimétrica (LEI) utilizada nesse trabalho tem como uma grande vantagem que o material sensório não precisa ser eletroativo e não há a necessidade de eletrodos de referência, como no caso das línguas potenciométricas e voltamétricas (RIUL JÚNIOR *et al.*, 2010). Ela se baseia na resposta de diferentes sensores impedimétricos e a impedância elétrica no sensor imerso em um líquido se deve às diferentes impedâncias combinadas nas interfaces eletrodo-filme, filme-solução e eletrodo-solução. Portanto, a impedância depende da amostra analisada, bem como da frequência do potencial aplicado (RIUL JÚNIOR *et al.*, 2013). Além disso, não exigem eletrodos de referência ou unidades de detecção contendo materiais eletroativos ou espécies ativas na amostra (CHRISTINELLI *et al.*, 2021). Contudo, há um imenso volume de dados gerado por cada análise o que torna o tratamento dos dados coletados para a interpretação da resposta uma tarefa complexa. As línguas eletrônicas, então, utilizam métodos computacionais e estatísticos mais complexos para o tratamento desses dados, como aprendizado de máquina (RIUL JÚNIOR *et al.*, 2010).

O aprendizado de máquina também pode ser utilizado sem exigir a calibração, uma vez que o banco de dados e a programação inerente faz com que o sistema não necessite de referência. Além disso, com o aprendizado de máquina, o instrumento pode ser capaz de identificar tanto qualitativamente, no sentido de identificar a presença ou não de um analito de interesse na amostra, como também sua quantidade (CHRISTINELLI *et al.*, 2021).

Nos últimos anos, centenas de trabalhos utilizando a língua eletrônica vêm sendo publicados. A simplicidade de manuseio, baixo custo de produção devido ao uso de materiais simples e a fabricação de sensores não exigir equipamentos ou tecnologia complexa, além de uma alta sensibilidade e a não destruição da amostra de interesse, torna a língua eletrônica uma ferramenta muito eficaz e útil. Além disso, tem aplicação e uso em diversos cenários, desde laboratórios, como em campo, e até dentro de processos.

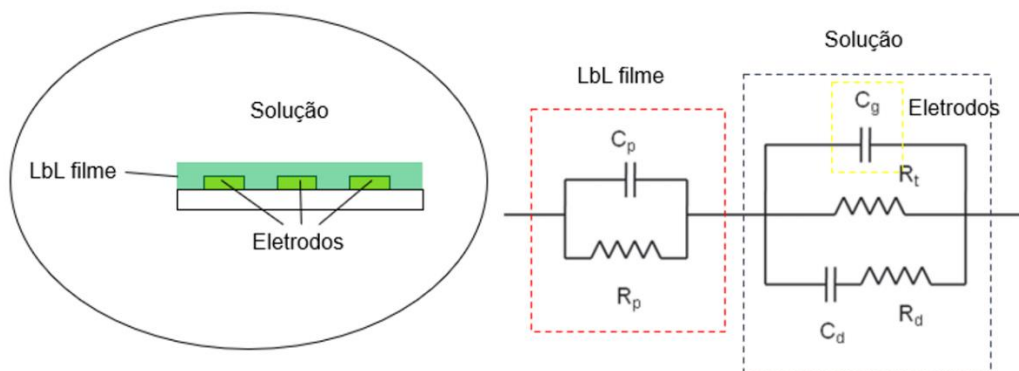
### 2.3.1 Impedância Elétrica

A impedância é uma medida de oposição à corrente elétrica alternada por parte do circuito ou do sistema. Se assemelha à resistência, tanto que sua medida é em ohms. A impedância é composta pela resistência, onde parte da energia é gasta na forma de calor (efeito Joule) e pela reatância capacitiva e indutiva (NILSON; RIEDEL, 2015). Apesar de não ser um fasor, ela é representada por um número complexo, conforme estabelecido pela equação 1, onde a parte real é a resistência e a parte imaginária a reatância.

$$Z = Z_{Re} + jZ_{Im} \quad (1)$$

Circuitos elétricos equivalentes são geralmente utilizados para analisar os dados de impedância coletados, uma vez que a compreensão dos fenômenos físicos e a interpretação dos dados da análise não são triviais. Esse circuito equivalente geralmente é composto por elementos de circuitos básicos, como resistência, capacitância e indutância, que podem ser combinados em série ou em paralelo para se adequarem aos dados experimentais (BONANNI; DEL VALLE, 2010; BROSEL-OLIU *et al.*, 2019). A Figura 1 é um esquema ilustrativo do circuito equivalente utilizado nesse trabalho. “ $C_p$ ” é a capacitância do filme, “ $C_g$ ” é a capacitância geométrica do eletrodo, “ $C_d$ ” é a capacitância de dupla camada. “ $R_p$ ”, “ $R_t$ ” e “ $R_d$ ” são as condutâncias.

**Figura 1:** Esquema ilustrativo do Circuito equivalente utilizado nesse trabalho.



Fonte: Adaptado de Taylor (1987) e Riul *et al* (2002).

Ao analisar a resposta de impedância de um sistema, é possível obter informações detalhadas sobre as propriedades elétricas e eletroquímicas. Essas informações podem incluir a resistência, capacitância, indutância, polarização de interface, taxa de reação e difusão iônica, entre outras (LVOVICH, 2015). Nesse trabalho, foram usadas as informações de capacitância e resistência.

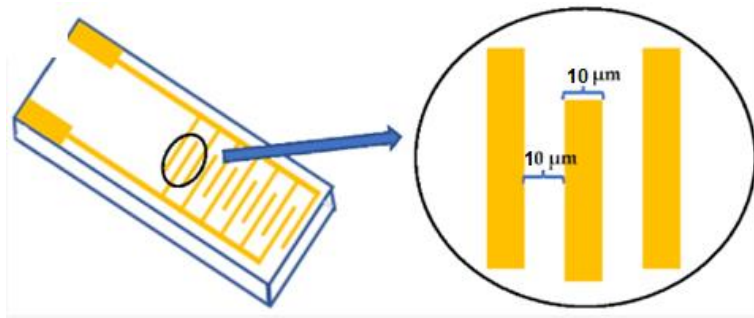
## 2.4 Sensores químicos

Sensores químicos são dispositivos que convertem uma informação química, desde a concentração de uma determinada substância até a composição total de uma amostra, em um sinal analiticamente mensurável (óptico elétrico, mecânico, magnético, gravimétrico etc.) (HULANICKI; GLAB; INGMAN, 1991). Distinguem-se dos instrumentos de largo porte nos quesitos de tamanho, sensibilidade, seletividade, portabilidade e custo (LOWINSOHN; BERTOTTI, 2006).

Como descrito anteriormente, a língua eletrônica é um sistema multissensorial, contendo um arranjo de sensores químicos individuais, geralmente entre 5 e 10 sensores. O sensor do tipo impedimétrico, utilizado nesse trabalho, é construído sobre um substrato de vidro contendo um par de eletrodos interdigitados (IDEs) de ouro e modificado com um filme ultrafino de nanomateriais diversos, conforme ilustrado na Figura 2. O IDE transdutor, uma vez converte a informação química no meio em um sinal elétrico. Entretanto, um arranjo de IDEs sem modificação não exibe sensibilidade cruzada suficiente para discriminar amostras e, por isso, é necessário modificá-los com filmes de diferentes nanomateriais. A sensibilidade se dá justamente por essa camada sensotiva, que aumenta a interação interfacial sensor-solução. A geometria do IDE consiste em trilhas paralelas micrométricas, chamadas de dígitos, que são conectadas a uma trilha principal e um contato. Eles são usados para aumentar a área efetiva e, conseqüentemente, aumentar o sinal elétrico entre os eletrodos em áreas pequenas (MAMISHEV *et al.*, 2004).



**Figura 2:** Ilustração esquemática de um microeletrodo interdigitado.



Fonte: Zagalo; Ribeiro; Raposo (2022).

## 2.5 Análise de Componentes Principais (PCA) e de Agrupamento Hierárquico de Clusters (HCA)

A Análise de Componentes Principais ou PCA (*Principal Components Analysis*) é um método matemático estatístico multivariado que consiste em transformar um conjunto de variáveis originais em um outro conjunto de variáveis de mesma dimensão denominadas de componentes principais. É um método de redução utilizado quando há uma grande quantidade de atributos correlacionados.

O algoritmo de PCA decompõe os dados da matriz principal  $\mathbf{X}$  ( $I \times J$ ) em duas matrizes menores, chamadas de score  $\mathbf{T}$  ( $I \times R$ ) e pesos  $\mathbf{P}$  ( $J \times R$ ), e uma matriz residual  $\mathbf{E}$  ( $I \times J$ ), conforme equação 2, em que  $R$  é o número de componentes principais (PCs) escolhidos para o modelo de dados.

$$\mathbf{X} = \mathbf{t}_1\mathbf{p}_1^T + \mathbf{t}_2\mathbf{p}_2^T + \dots + \mathbf{t}_R\mathbf{p}_R^T + \mathbf{E} = \mathbf{TP}^T + \mathbf{E} \quad (2)$$

O par de vetores de scores e pesos,  $\mathbf{t}_R\mathbf{p}_R$ , define o componente principal (PC), que é a direção contendo as principais informações dos dados e retendo a maior parte da variância dos dados. A similaridade ou a diferença entre as amostras e variáveis pode ser observada através de gráficos nos quais scores e pesos, respectivamente, são plotados um contra o outro e cujas posições são delimitadas pelos eixos dos PCs. Em resumo, scores e pesos localizados nas vizinhanças são considerados similares ou relacionados, enquanto aqueles localizados em posições distantes são considerados não relacionados (JOLLIFE, 2002).

A análise de agrupamento hierárquico de *clusters* (HCA – *Hierarchical Cluster Analysis*) também desempenha um papel importante no reconhecimento de padrões. A HCA pode ser realizada nos dados pré-processados ou usando os scores do PCA. O resultado é um

gráfico bidimensional (dendrograma), no qual as amostras ou variáveis são divididas em grupos (também chamados *clusters*) de acordo com sua similaridade.

No presente trabalho, a HCA foi aplicada para confirmar a presença de grupos entre as amostras e para verificar se os mesmos grupos estariam presentes nos padrões observados pelos scores do PCA. Toda a HCA foi feita utilizando a distância euclidiana de similaridade e o critério de ligação Ward (OTTO, 1999). Este critério minimiza a variância dentro dos grupos, gerando grupos mais homogêneos, sendo de tamanhos mais aproximados e menos sensível a *outliers*. Isso é feito calculando-se a soma dos quadrados dos desvios das médias dos grupos (WARD Jr., 1963).

Dentro do Python há pacotes que reconhecem a matriz de dados e aplicam o método matemático retornando a auto escalação dos valores e extraíndo as matrizes de covariância, fornecendo os autovalores e autovetores.

## 2.6 Coeficiente de Correlação de Pearson

Outra ferramenta utilizada neste trabalho em conjunto com PCA e HCA é o coeficiente de correlação, mais especificamente o elaborado por Karl Pearson e Francis Galton em meados de 1880 (FILHO; JÚNIOR, 2009). O coeficiente de correlação de Pearson (CCP), definido como “*r*”, é uma medida de associação linear entre variáveis, que mede a variância compartilhada entre duas variáveis. O valor de “*r*” pode ser definido, matematicamente, pela equação 3.

$$r = \frac{1}{n-1} \sum \left( \frac{x_i - \bar{X}}{sx} \right) \left( \frac{y_i - \bar{Y}}{sy} \right) \quad (3)$$

Na equação 3, *n* é tamanho da amostra; *x<sub>i</sub>* e *y<sub>i</sub>* são pontos individuais da amostra;  $\bar{X}$  e  $\bar{Y}$  são as médias da amostra e *s* como o desvio padrão. X e Y são considerados um par de variáveis aleatórias da amostra.

O valor de *r* varia de -1 a 1, onde o sinal indica a direção positiva ou negativa do relacionamento e o valor, o quanto são relacionáveis as variáveis. Quanto mais perto de 1, independentemente do sinal, maior é o grau de dependência estatística linear entre as variáveis. e, conseqüentemente, quanto menor o valor, menor é a relação. Importante ressaltar que o

coeficiente não diferencia variáveis dependentes ou independentes, apenas o valor da sua correlação (FIGUEIREDO FILHO; SILVA JÚNIOR, 2009).

Nesse trabalho, o CCP foi aplicado para guiar e ajudar na interpretação das variáveis e como elas se relacionam. O coeficiente foi utilizado nos gráficos de mapa de calor com o intuito de ajudar na visualização do relacionamento entre as variáveis. Foi utilizado também para verificar como essas relações apareceriam no HCA e no PCA.

## 2.7 Aprendizado de Máquina

O aprendizado de máquina compreende a habilidade do computador, por meio de algoritmos, em aprender pelos erros e fazer previsões sobre dados (HOSCH, 2023). O aprendizado automático se baseia na construção de algoritmos que constroem um modelo a partir de um “*input*” amostral com o objetivo de fazer previsões ou decisões guiados pelos dados ao invés de simplesmente fazê-lo seguindo as instruções programadas (HORNİK *et al.*, 2009; ALONSO *et al.*, 2010; OU-YANG *et al.*, 2011; CIOSEK *et al.*, 2006; STOJIĆ *et al.*, 2010; AITKENHEAD *et al.*, 2014).

O aprendizado de máquina é classificado em 3 tipos de acordo com a natureza do “sinal” de aprendizado disponível para um sistema de aprendizado:

- Aprendizado supervisionado: são oferecidos ao computador exemplos de entrada e saída desejadas fornecidos por um supervisor.
- Aprendizado não supervisionado nenhum tipo de informação é dado ao algoritmo de aprendizado, que é deixado sozinho para encontrar os padrões e a estruturas nas entradas fornecidas. Este é mais utilizado quando se deseja encontrar padrões.
- Aprendizado por reforço: o computador interage com o ambiente e é fornecido ao programa um feedback com premiações e punições na medida em que ele interage com o ambiente.

O aprendizado de máquina está ligado com as redes neurais artificiais que são inspirados pelo sistema nervoso central como o cérebro. (HORNİK *et al.*, 2009; ALONSO *et al.*, 2010; OU-YANG *et al.*, 2011; CIOSEK *et al.*, 2006; STOJIĆ *et al.*, 2010; AITKENHEAD *et al.*, 2014).

Nesse trabalho, o aprendizado foi feito de forma não supervisionada, quando nenhuma informação foi fornecida e esperava-se observar o reconhecimento de padrões com o algoritmo, e aprendizado supervisionado, quando foram fornecidas as entradas necessárias e as saídas

desejadas. Para treinamento, os dados são divididos de forma que uma parte seja utilizada para o aprendizado e a outra parte para verificação. O algoritmo, então, é treinado  $k$  vezes e o aprendizado é avaliado. Quanto mais dados e entradas forem fornecidas, mais eficiente se torna esse aprendizado.

O aprendizado de máquina foi implementado na língua eletrônica com o intuito de desenvolver um sistema multissensorial completo com capacidade de leitura das amostras e o reconhecimento de padrões. O treinamento de máquina fornece dados suficientes para o sistema para que ele não precise de calibração e se autorregule. Isso também visa a verificação e a validação dos dados, pois compara os resultados reais com aqueles alcançados com a máquina e, assim, atesta a confiabilidade do sistema.

### 2.7.1 Regressão Logística

A regressão logística é um modelo estatístico muito utilizado para classificação e análise preditiva. Ela estima a probabilidade de ocorrência de um evento com base em um determinado conjunto de dados de variáveis independentes. Geralmente, a ocorrência do evento de interesse assume uma codificação binária de “1”, como sendo a chance de o evento ocorrer e “0”, de não ocorrência, podendo ser também “sim” ou “não”, “verdadeiro” ou “falso”, etc. (IBM, 2024). Ela é matematicamente mais simples que outros métodos matemáticos de classificação e pode ser utilizado em diversos cenários como marketing, saúde, financeiro e muitos outros.

A regressão logística modela a relação entre a variável dependente e as variáveis independentes através de uma função logística, que transforma a saída da regressão linear em uma escala logarítmica, produzindo então valores entre 0 e 1, representando as probabilidades. O modelo da regressão logística segue, então, a equação 4.

$$P(D = 1 | X_1, X_2, X_3, \dots, X_k) = P(\mathbf{X})$$
$$P(\mathbf{X}) = \frac{1}{1 + e^{-(\alpha + \sum \beta_i X_i)}} \quad (4)$$

Na equação 4,  $P(\mathbf{X})$  é a probabilidade condicional de que a variável  $D$  seja igual a 1 dado um conjunto de valores variáveis independentes  $X$ . “ $e$ ” é a base do logaritmo natural, “ $\beta_i$ ” são os coeficientes a serem estimados e  $X_1, X_2, X_3, \dots, X_k$ , são os valores das variáveis independentes (KLEINBAUM; KLEIN, 2010).

### **3 OBJETIVOS**

Este trabalho tem por objetivo geral avaliar o desempenho de uma língua eletrônica impedimétrica com algoritmo de aprendizado de máquina na análise de águas residuais de diferentes estações de tratamento do Distrito Federal. Os objetivos específicos são descritos abaixo:

- Construir e configurar os sensores da língua eletrônica a partir de nanomateriais e utilizando a técnica de automontagem;
- Determinar o perfil de resposta dos sensores para amostras de água residual;
- Processar os dados obtidos (sinais dos sensores) por meio de análise de componentes principais (PCA);
- Implementação algoritmo de aprendizado de máquina supervisionado e não supervisionado;
- Avaliar o desempenho da língua eletrônica na discriminação das Estações de Tratamento utilizando a análise de componentes principais PCA e o dendrograma
- Avaliar o desempenho da língua eletrônica a partir de um algoritmo classificatório e verificar o índice de acerto de alguns parâmetros físico-químicos.

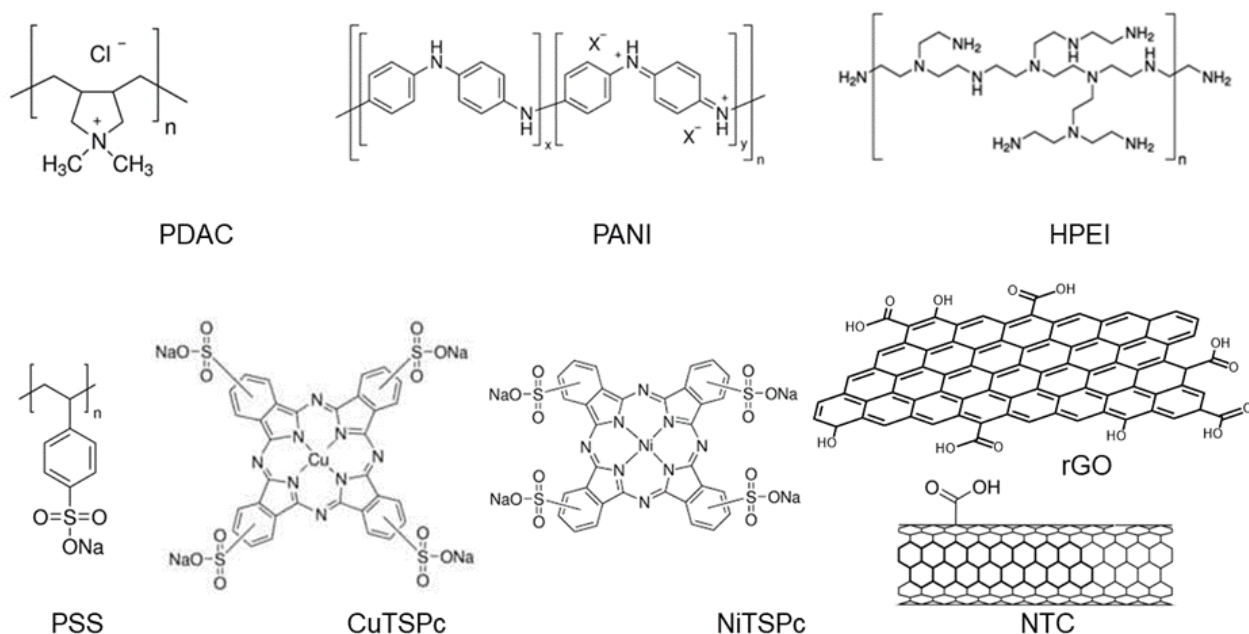
## 4 MATERIAIS E MÉTODOS

### 4.1 Materiais

Ftalocianinas de níquel (II) e cobre(II) tetrasulfonadas (Ftc-Ts-Ni e Ftc-Ts-Cu), poli(etilenoimina) hiperramificada (HPEI, 50%,  $M_w$  750.000 g/mol), hidrocloreto de poli (dialildimetilâmonio) (PDAC, 20%,  $M_w$  450.000 g/mol) e poli (estireno sulfonato de sódio) (PSS,  $M_w$  70.000 g/mol) e ácido 3-mercaptopropanosulfônico (3- MPS) foram adquiridos da Sigma Aldrich e utilizados como recebidos.

Óxido de grafeno e óxido de grafeno reduzido (GO e rGO), nanotubos de carbono de paredes múltiplas oxidado (NTC) e nanopartículas de ouro (Au-NP), magnetita (MAG) e óxido de cério ( $CeO_2$ ) foram fornecidos pelo Laboratório de Pesquisa em Polímeros e Nanomateriais do IQ-UnB (LABPOLN). A fórmula estrutural de alguns dos materiais empregados na modificação dos eletrodos é apresentada na Figura 3.

**Figura 3:** Fórmula estrutural dos materiais empregados na fabricação dos sensores.



Fonte: A autora (2024).

Os IDEs de ouro com 50 pares de dígitos com espaçamento de 10  $\mu\text{m}$ , largura 10  $\mu\text{m}$  e altura de 0,2  $\mu\text{m}$ , impressos sobre vidro boros-silicato (1 mm x 13 mm x 25 mm) foram fornecidos pelo Laboratório de Microeletrônica da Escola Politécnica da USP. A água utilizada para preparação de soluções, enxágue e outros procedimentos foi do tipo ultrapura (resistividade: 18 Mohm.cm), produzida por um purificador de água da Milli-Pore modelo Milli-Q. O ar comprimido utilizado para a secagem das lâminas, foi produzido por um moto compressor de ar da marca BR Motors<sup>®</sup>, modelo BRC 8,8 24L.

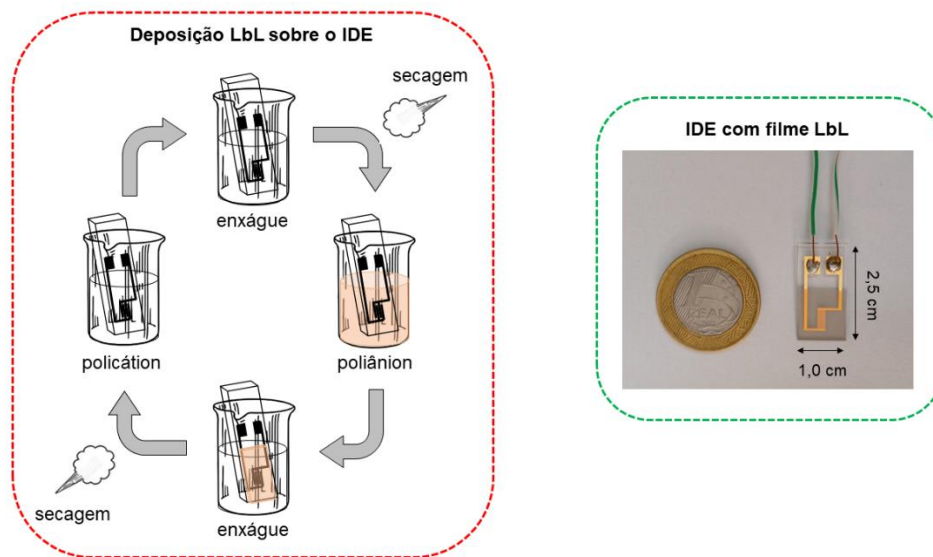
## 4.2 Metodologia

### 4.2.1 Fabricação dos sensores impedimétricos

Os sensores impedimétricos consistiram do IDE descrito na seção 4.1 e modificados com filmes ultrafinos de nanomateriais diversos depositados pela técnica de automontagem LbL (layer-by-layer) no modo eletrostático. O primeiro passo da modificação foi a funcionalização da superfície dos dígitos de ouro com grupos sulfônicos ( $-\text{SO}_3^-$ ) a partir da imersão do IDE em uma solução de 3-MPS 1g/L a temperatura ambiente e por 3h. Com esse procedimento, a superfície do ouro se torna negativamente carregada, aumentando o número de sítios para adsorção dos nanomateriais subsequentes.

Após funcionalização com 3-MPS, os IDEs foram modificados por automontagem de polieletrólitos e nanomateriais conforme ilustrado esquematicamente na Figura 4. O IDE é inicialmente imerso numa solução de polieletrólito ou nanomaterial catiônico, onde permanece submerso por 3 minutos. Em seguida, é transferido para uma solução de enxágue com agitação magnética e nela mantido por 20 segundos. Por fim, é secado com um jato de ar comprimido. O processo é, então, repetido para o polieletrólito/nanomaterial aniônico. Ao fim do ciclo é obtido um filme com uma bicamada (material catiônico/material aniônico) e o ciclo pode ser repetido até o número planejado de bicamadas. Todos os sensores foram produzidos com 10 bicamadas cada. Foram produzidos 8 sensores (S), cujas arquiteturas são descritas na Tabela 1. Todos foram iniciados com uma camada de PDAC para melhorar a adesão dos materiais. As combinações de polieletrólitos e nanomateriais escolhidas foram desenvolvidas e testadas em trabalhos prévios (PATERNO *et al.*, 2009; CORREA *et al.*, 2014; MEDEIROS *et al.*, 2012a; MEDEIROS *et al.*, 2012b).

**Figura 4:** Ilustração esquemática da deposição dos filmes pela técnica LbL.



Fonte: A autora (2024).

**Tabela 1:** Descrição dos sensores impedimétricos

SENSORES	BICAMADA	CONDIÇÕES DE DEPOSIÇÃO
S1	Ftc-Ts-Ni / PANI	1 g/L; enxágue pH=2,7
S2	HPEI / Ftc-Ts-Cu	1 g/L; enxágue pH=4,0
S3	PSS / PANI	1 g/L; enxágue pH=2,7
S4	PDAC / RGO	1 g/L/0,5 g/L; enxágue pH= 7,0 redução com hidrazina 1:1 por 10 minutos
S5	PDAC / CeO <sub>2</sub> -NP	1 g/L/1,5 g/L; enxágue pH= 7,0
S6	PDAC / NTC	1 g/L/0,3 g/L; enxágue pH= 5,5
S7	PSS / Mag-NP	1 g/L/1,2 g/L; enxágue pH=4,0
S8	PDAC / Au-NP	1 g/L/0,2 g/L; enxágue pH= 7,0

Fonte: A autora (2024).

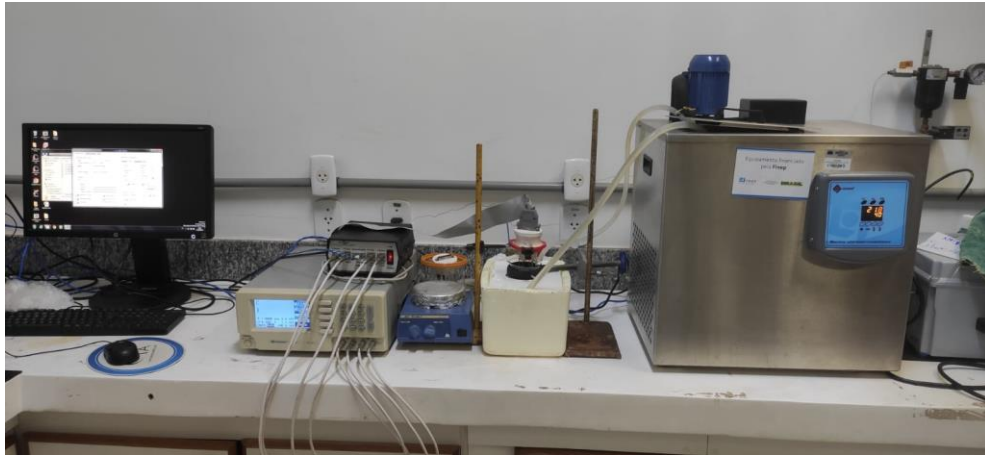
#### 4.2.2 Sistema de medidas da língua eletrônica e procedimento de medição

As Figura 5 e 6 ilustram o sistema de medidas da língua eletrônica usado nesse trabalho. O sistema é composto pelas seguintes partes, conforme indicado: computador com programa para controle e aquisição de dados (software da língua eletrônica em Labview), medidor LCR



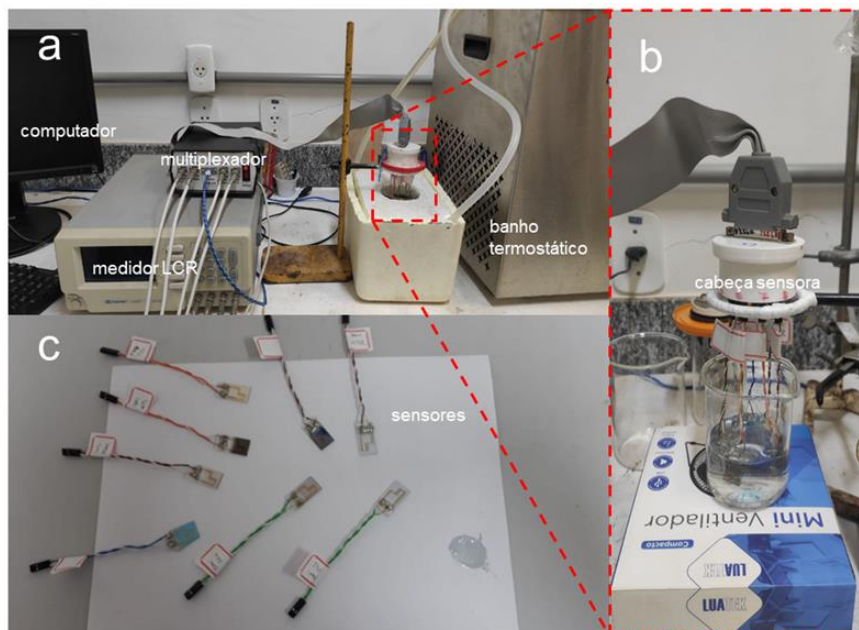
Instek modelo 816, multiplexador de 10 canais, banho termostático Quimis modelo Q214, cabeça sensora contendo 8 sensores conectada ao multiplexador com um conector db25 e cabo flat. Para tratamento dos dados, foi utilizado o Excel® e o Google Colab®. Foi também usada a linguagem Python® e os pacotes matemáticos oferecidos por essa linguagem.

**Figura 5:** Visão geral do sistema de medição da língua eletrônica.



Fonte: A autora (2024)

**Figura 6:** Sistema de medição da língua eletrônica impedimétrica: (a) visão geral e partes indicadas, (b) visão ampliada da cabeça sensora, (c) sensores.



Fonte: A autora (2024)

Todas as análises foram realizadas com controle de temperatura das amostras em  $25,0 \pm 0,5$  °C. Para cada amostra analisada, foram registrados 10 valores de capacitância e 10 valores de resistência de cada sensor, resultando em um total de 80 valores para capacitância e 80 para resistência, em 1 kHz, amplitude de 250 mV. Inicialmente, o funcionamento da língua eletrônica foi avaliado com leituras alternadas em amostras de solução salina (NaCl) de 1  $\mu$ M e água ultrapura. Esse ensaio foi repetido durante duas semanas, medindo uma vez por dia. Posteriormente, foram analisadas as amostras de água da CAESB conforme descrição a seguir.

### **4.2.3 Coleta e preparo das amostras CAESB**

As amostras de água residual foram coletadas das estações de tratamento da CAESB. As estações são divididas em regiões norte e sul. As amostras coletadas são do efluente final, com filtração na membrana 0,45 $\mu$ m, após todo o tratamento. De duas estações, Norte e Sul, também foram coletadas amostras do efluente biológico com filtração de membrana. O efluente biológico é um ponto intermediário, antes do ponto final. As amostras foram coletadas das seguintes estações:

#### **Sul**

- Melchior – efluente final
- Sul – efluente biológico e final
- Águas Lindas – efluente final
- Brazlândia – efluente final
- Riacho fundo – efluente final

#### **Norte**

- Norte – efluente biológico e final
- Alagado – efluente final

A estação Riacho Fundo conta ainda com 3 reatores de onde se coletam as amostras. Dessa forma, dessa estação, saem 3 amostras de cada dia.

As amostras que chegam nas estações de tratamento são do dia anterior à análise e são analisadas nos laboratórios das estações. Após as análises, são produzidos relatórios com os resultados obtidos e analisados de acordo com os padrões estabelecidos pelas legislações

vigentes. A CAESB então disponibiliza os relatórios dos dias analisados para que as informações possam ser adicionadas àquelas registradas com a língua eletrônica.

As amostras eram coletadas nas estações de tratamento com frascos especiais de plástico de 125 mL que eram lavados com detergente sem sulfato, para não haver interferência, além de serem protegidos da luz com papel alumínio. Os frascos com as águas eram então colocados em uma caixa de isopor com gelo gel rígido para manter a temperatura e o condicionamento. A caixa com as amostras era levada para o LABPOLN onde foram analisadas com a língua eletrônica.

O processo de medição começava com uma alíquota de 75 mL de água ultrapura, seguida de 75 mL de água ultrapura usada como controle. A amostra de controle era colocada em um mesmo tipo de frasco que as amostras e isso era feito para monitorar a eventual contaminação das amostras durante o transporte. A rotina de medição consistia em registrar 10 valores de capacitância e 10 valores de resistência de cada amostra com cada sensor, sob tensão alternada de 250mV de amplitude, frequência de 1 kHz, em configuração de circuito paralelo para um volume de 75mL. Os arquivos gerados eram no formato txt.

Ao final da análise de cada amostra, os sensores eram enxaguados com água ultrapura sob agitação magnética, por 3 minutos.

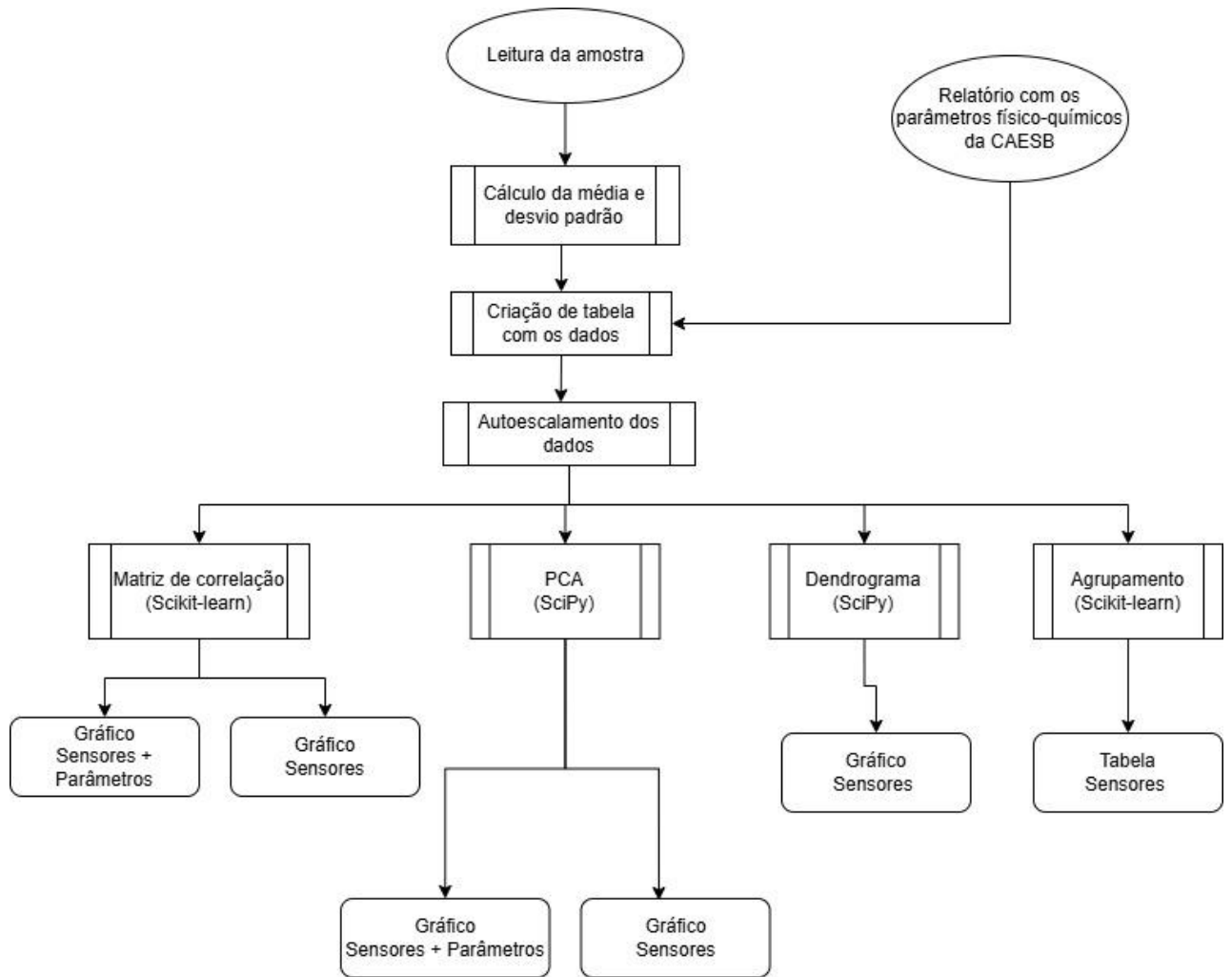
Após as leituras de todas as amostras do dia, era realizada uma leitura em 75 mL de água ultrapura para verificar a recuperação (reset) da resposta dos sensores e avaliar a eficiência dos enxágues (Anexo B). No total, foram analisadas 212 amostras coletadas nos meses de novembro e dezembro de 2022, janeiro e de setembro a dezembro de 2023.

#### **4.2.4 Programação, tratamento dos dados e aprendizado de máquina**

Em posse dos dados registrados com a língua eletrônica e dos relatórios da CAESB, os dados foram inicialmente organizados, depois calculadas as médias das leituras de cada amostra com o desvio padrão. Com o Google Colab, na linguagem Python, foram feitas as análises de PCA, HCA e a programação de aprendizado de máquina utilizando PCA e gráfico de erros e acertos da máquina.

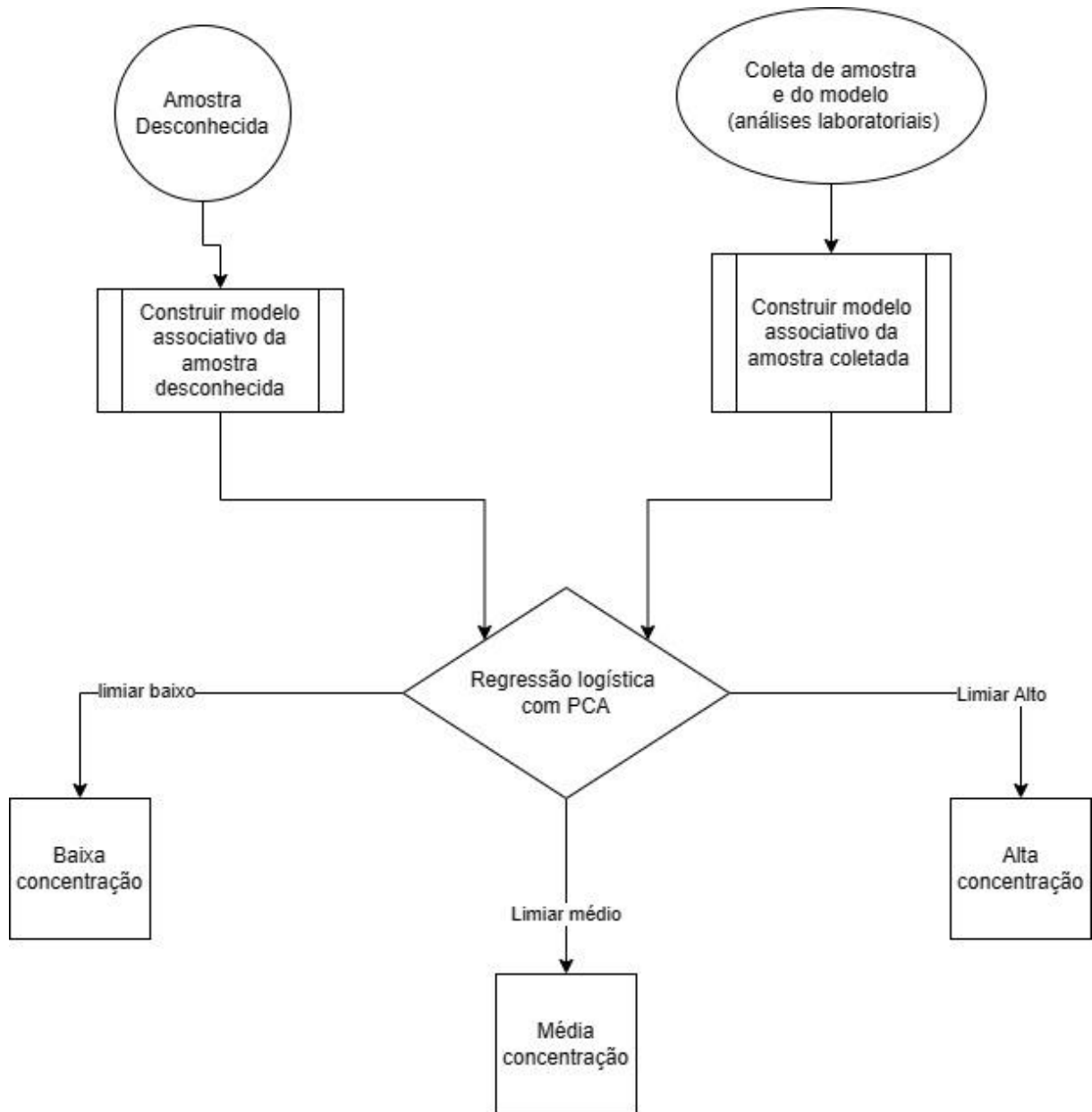
As Figura 7 e 8 mostram os fluxogramas do tratamento de dados para as análises feitas nesse trabalho e do estudo dos indicadores, respectivamente. Todos feitos para capacitância, resistência e capacitância e resistência.

**Figura 7:** Fluxograma do tratamento dos dados com o uso da linguagem Python.



Fonte: A autora (2024).

**Figura 8:** Fluxograma do estudo dos indicadores.



Fonte: A autora (2024).

## **5 RESULTADOS E DISCUSSÃO**

### **5.1 Tratamento dos dados**

Os dados obtidos foram organizados numa tabela em que as linhas representam as amostras, identificadas pelo local e dia de tratamento, enquanto as colunas representam os sensores. Cada amostra era lida 10 vezes e a média e o desvio padrão foram feitos para cada amostra. O desvio padrão observado foi baixo, indicando pouca variação nas leituras de cada amostra.

Além dos valores médios/desvios de capacitância (Anexo A) e de resistência, a tabela foi incrementada com atributos físico-químicos documentados nos relatórios fornecidos pela CAESB. Nesses relatórios, são registrados diversos parâmetros, medidos segundo às normas estabelecidas. Entretanto, alguns parâmetros não eram lidos em todas as amostras, uns eram registrados em algumas estações e em outras não. Assim, os parâmetros com maior quantidade de resultados foram escolhidos: DQO, SST, Nitrato, Nitrogênio de Amônia, NT e PHOSP. A ausência de algum dado do relatório pela falta de análise, foi preenchida com a média de todas as leituras.

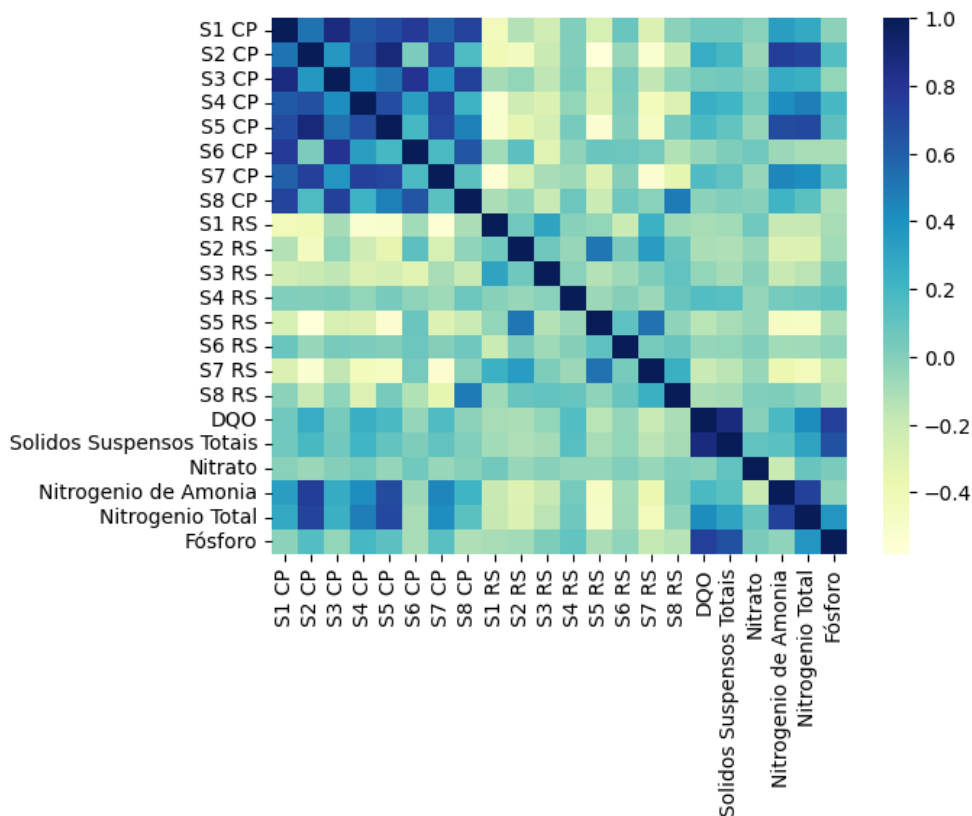
Em seguida, os dados foram autoescalados para haver uniformidade. Os dados são centrados na média.

### **5.2 Mapa de Calor**

Cada sensor gera um resultado diferente de capacitância e resistência pois, pelo princípio de funcionamento da língua eletrônica, cada sensor responde com maior ou menor sensibilidade à composição global da amostra, e não a uma substância em específico, resultando em algo semelhante a uma “impressão digital” para aquela amostra (Anexo D).

Num primeiro momento, foram calculados os coeficientes de correlação de Pearson e plotados em um mapa de correlação, para verificar a relação dos sensores, com os respectivos valores de capacitância (CP) e resistência (RS), com os parâmetros físico-químicos da água. Esse mapa, conhecido também como mapa de calor, é apresentado na Figura 9. O intuito é perceber como é a correlação entre cada sensor com os valores de resistência e capacitância, bem como se correlacionam os sensores com os parâmetros físico-químicos e ver se a tendência vai ser observada em outras análises.

**Figura 9:** Mapa de correlação entre sensores e parâmetros



Fonte: A autora (2024).

Os valores do mapa de calor variam de +1 até -1, sendo que valores positivos indicam uma associação positiva (quando uma variável aumenta, a outra também aumenta), enquanto valores negativos indicam uma associação negativa (quando uma variável aumenta, a outra diminui). Assim, valores próximos a zero indicam correlações fracas enquanto valores mais próximos de +1 (mais escuro) ou -1 (mais claro), indicam correlações fortes.

No mapa apresentado na Figura 9, os valores de capacitância dos sensores têm uma boa correlação entre si, tendo correlação fraca entre os sensores 2 e 6 (S2 CP e S6 CP). Já os valores de capacitância dos sensores 2 e 5 (S2 CP e S5 CP) têm uma boa correlação com o nitrogênio de amônia e o NT. Essa maior correlação se manteve desde a última análise dos dados, feita na primeira parte do trabalho com menos amostras analisadas (79). O sensor 2 tem em sua composição a Ftalocianina de Cobre, que tem afinidade com compostos nitrogenados (VIDAL-MADJAR; GUIOCHON, 1971), uma possível explicação para a correlação com esses dois parâmetros. Isso faz que mais componentes sejam adsorvidos na superfície do sensor, aumentando o sinal de resposta. Porém, deveria ter sido observado também a mesma correlação no sensor 1, feito com Ftalocianina de Níquel, e isso não é observado, mesmo essas duas

ftalocianinas tendo boa afinidade com compostos nitrogenados. Uma razão para isso, é que mesmo tendo estruturas e comportamentos parecidos, o filme com Cobre tem uma superfície mais rugosa se comparada com o filme de níquel, cuja superfície é mais uniforme (RIDHI *et al.*, 2018). A rugosidade aumenta a porosidade que, por sua vez, aumenta a quantidade de espaços para portadores de carga, aumentando assim a condutividade (ALCANTARA *et al.*, 2013). Em relação ao sensor 5, tem nanopartículas de Cério em sua composição. Essas nanopartículas também têm alta afinidade com compostos nitrogenados (ROZHIN *et al.*, 2021). Assim, há uma alta adsorção desses compostos na superfície do sensor que por consequência aumenta a resposta do sinal. Porém, ao observar o nitrato, não há a correlação forte observada nos sensores 2 e 5. O motivo deve ser devido a forma iônica e/ou também pelo método diferenciado de análise, uma vez que o nitrogênio de amônia e o nitrogênio total são analisados por espectrofotometria no UV-VIS enquanto o nitrato é determinado por cromatografia iônica.

Quanto aos valores de resistência, os sensores mostraram baixa correlação entre si. Alguns poucos apresentam uma correlação baixa entre si, como os sensores 2 e 5 (S2 RS e S5 RS) e os sensores 2 e 7 (S2 RS e S7 RS), essa correlação entre esses pares de sensores também foi observada no caso da capacitância. A resistência apresenta uma correlação negativa, próxima de -0,4, em relação ao nitrogênio de amônia e ao nitrogênio total no caso dos sensores 5 e 7 (S5 RS e RS7), apresentando valores mais claros.

Em relação aos valores de resistência e capacitância, observa-se que para a maioria dos sensores a correlação é negativa e baixa, ou seja, os valores não se podem ser relacionados diretamente. Alguns apresentam uma correlação negativa maior, próxima de -0,6, como no caso de S5 RS e S2 CP, S7 RS e S2 CP, S1 RS e S4 CP, S1 RS e S5 CP, S1 RS e S7 CP, S5 RS e S5 CP.

No caso dos parâmetros, percebe-se boa correlação positiva entre DQO e SST, DQO e fósforo total e nitrogênio de amônia com nitrogênio total. Há uma baixa correlação entre nitrato e nitrogênio de amônia e, praticamente, nenhuma correlação com nitrogênio total.

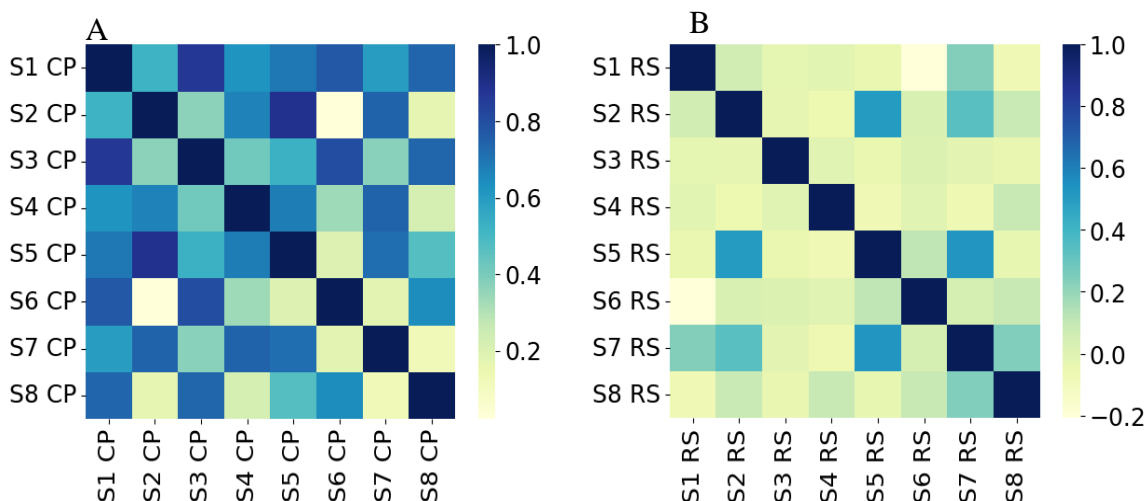
Também foram construídos mapas de calor somente dos dados da língua eletrônica (capacitância e resistência), mostrados na Figura 10. O objetivo é avaliar o quanto os sensores estão correlacionados entre si e estudar a possibilidade da retirada e/ou substituição de algum deles, caso haja muita redundância nas respostas. Nesse caso, a escolha se dá pelo sensor de menor custo de fabricação e maior estabilidade. Quanto maior a correlação entre os sensores, mais próximas são suas respostas para uma mesma amostra.



Como pode ser visto na Figura 10a, os sensores S2 CP e S6 CP continuam com baixíssima correlação, próxima de zero, o que conseqüentemente leva a uma resposta diferentes entre eles em relação às amostras. Assim sendo, esses dois sensores devem ser mantidos na língua eletrônica. Já S3 CP e o S1 CP têm uma correlação de um pouco mais de 0,8. Contudo, ainda não é uma correlação alta o suficiente para gerar leituras redundantes, além de que a resposta para a resistência é diferente como também para os parâmetros físico-químicos (Figura 9). Portanto, esse sensor foi mantido durante todas as análises.

Quanto aos valores de resistência (Figura 10b), as correlações se mantiveram as mesmas das que as observadas anteriormente. Porém, pode se perceber uma correlação negativa mais clara, próxima de -0,2, para S1 RS e S6 RS.

**Figura 10:** Mapa de correlação entre sensores. (a) Capacitância. (b) Resistência.



Fonte: A autora (2024).

### 5.3 Análise dos dados por PCA

A variância de cada componente explica quanto cada componente contém de informação do conjunto de dados. A variância acumulada mostra quantos componentes principais são importantes para explicar o quanto vai se acumulando de informações a cada componente principal e, assim, compactar o conjunto de dados ao máximo, mas sem que haja perda de informação. As Tabela 2, 3 e 4 mostram a variância explicada dos dois primeiros componentes principais e do componente principal cuja variância acumulada consegue explicar mais de 95% de informação do sistema. Essas tabelas também consideram tanto os sensores como os parâmetros físico-químicos. Sendo assim, as informações que constam na Tabela 2

consideram os 8 sensores mais 6 parâmetros físico-químicos, totalizando 14. Da mesma forma acontece para a Tabela 3. Para a Tabela 4, são consideradas as informações de capacitância e resistência juntos, totalizando 22. Para maiores detalhes sobre a variância de cada componente principal vide Anexo E.

**Tabela 2:** Variância explicada para capacitância.

<b>Componente Principal</b>	<b>Variância (%)</b>	<b>Variância Acumulada (%)</b>
PC1	39,98	39,98
PC2	20,96	60,93
PC8	1,91	96,19

Fonte: A autora (2024).

**Tabela 3:** Variância explicada para resistência.

<b>Componente Principal</b>	<b>Variância (%)</b>	<b>Variância Acumulada (%)</b>
PC1	24,8	24,80
PC2	14,63	39,43
PC11	2,93	96,48

Fonte: A autora (2024).

**Tabela 4:** Variância explicada para capacitância e resistência.

<b>Componente Principal</b>	<b>Variância (%)</b>	<b>Variância Acumulada (%)</b>
PC1	29,97	29,97
PC2	14,36	44,33
PC13	1,41	95,49

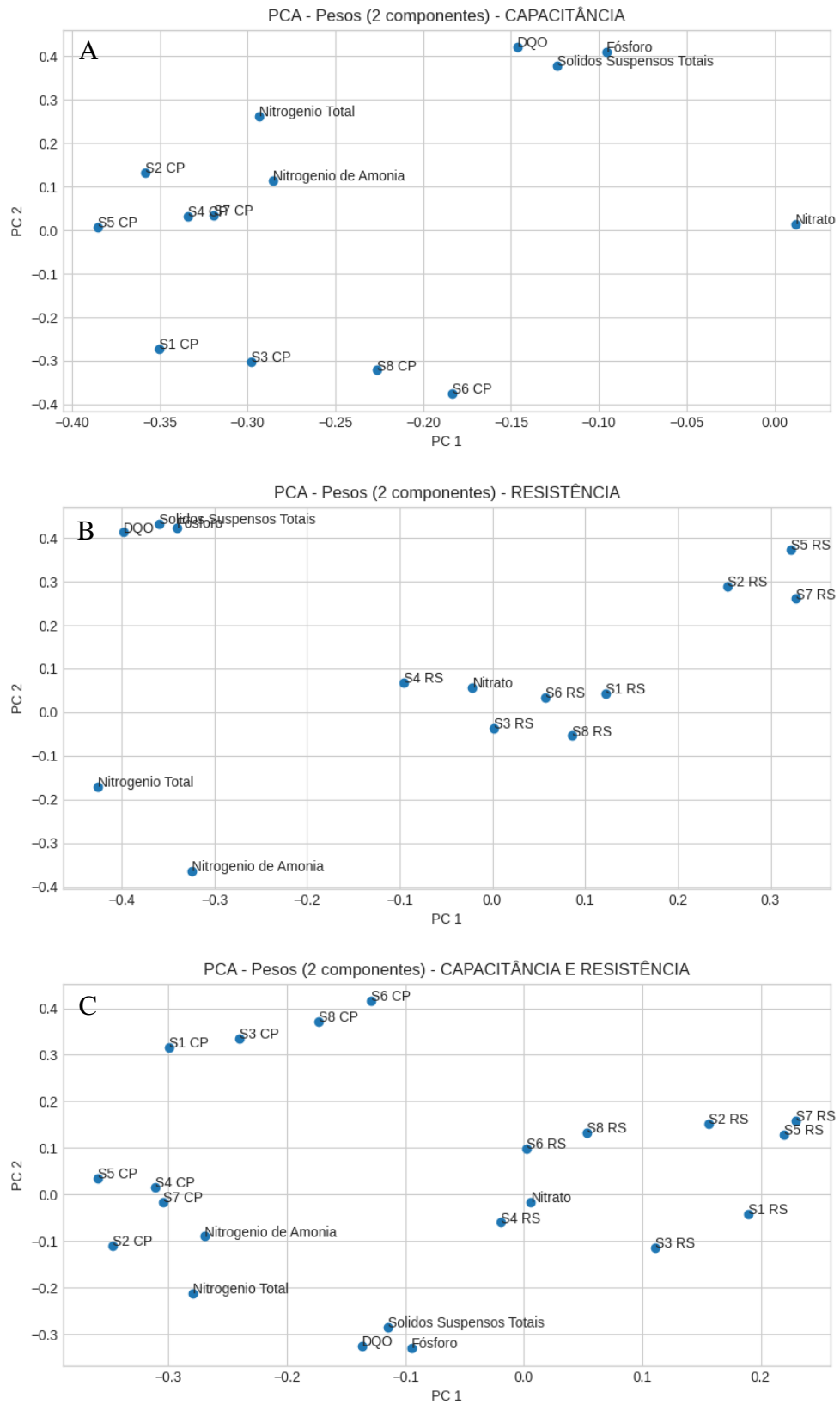
Fonte: A autora (2024).

Os dois primeiros componentes principais explicam a maior parte das informações do sistema, mas eles não são o suficiente, uma vez que eles somados explicam bem menos que 95% das informações. A capacitância (Tabela 2) consegue explicar mais com menos componentes principais (PCs), quase 61% com dois PCs e 8 para explicar mais de 95%. Enquanto a resistência (Tabela 3), os dois primeiros componentes principais explicam bem

menos de 50% e ela necessita de 11 PCs para explicar mais de 95%. Com a resistência e a capacitância (Tabela 4), há um balanceamento.

Um gráfico de dois (Anexo F) ou três componentes principais pode ser montado para tentar visualizar eventuais grupos. Porém, como o foco é o desempenho da língua eletrônica, esses gráficos serão melhor explorados mais à frente. Em seguida, é mostrado os gráficos dos pesos para dois componentes principais (Figura 11).

**Figura 11:** Gráfico dos pesos de 2 componentes principais, sensores e parâmetros. (a) Capacitância. (b) Resistência. (c) Capacitância e Resistência.



Fonte: A autora (2024).

No gráfico de pesos da capacitância (Figura 11a), percebe-se que os sensores estão mais à esquerda e abaixo enquanto os parâmetros estão afastados mais à direita e acima, com exceção do nitrogênio total e o nitrogênio de amônia que se encontram mais próximos dos sensores. O nitrato está muito afastado e isolado do lado direito. No caso da resistência, (Figura 11b), acontece a mesma separação entre sensores e parâmetros, mas nesse caso, o nitrato está mais próximo dos sensores, os sensores à direita e os parâmetros à esquerda, com o nitrato muito próximo de um grupo de sensores. No caso da Figura 11c, há uma separação entre três grupos: capacitância à esquerda, resistência à direita e parâmetros na parte de baixo do gráfico. O nitrato foi o único parâmetro que ficou do lado mais direito do gráfico, junto com a resistência. Algo esperado, uma vez que as leituras de capacitância e resistência têm valores distintos que também se distinguem das leituras dos parâmetros, com exceção do nitrato que parece exibir uma afinidade com os sensores no caso da resistência. Ele provavelmente, se liga à superfície do filme, mas ao invés de aumentar sua condutividade, provavelmente aumenta a resistência por impossibilitar o transporte de carga (ALCANTARA *et al.*, 2013).

#### **5.4 Análise dos sensores - PCA**

As análises feitas até esse momento consideravam as respostas dos sensores junto com a resposta dos parâmetros. Nesse momento, os sensores serão avaliados de forma separada, uma vez que o objetivo é analisar o desempenho da língua eletrônica e verificar se ela consegue identificar padrões.

As Tabela 5, 6 e 7 mostram a variância explicada dos dois primeiros componentes principais e do componente principal cuja variância acumulada consegue explicar mais de 95% de informação do sistema. Dessa vez, apenas as respostas dos sensores são consideradas para obter a informação do sistema. Assim, para as Tabela 5 e 6 há um total de 8 componentes principais, enquanto a Tabela 7 apresenta 16 componentes principais.

**Tabela 5:** Variância explicada para capacitância - Sensores.

<b>Componente Principal</b>	<b>Variância (%)</b>	<b>Variância Acumulada (%)</b>
PC1	59,80	59,80
PC2	24,00	83,80
PC5	2,80	97,60

Fonte: A autora (2024).

**Tabela 6:** Variância explicada para resistência - Sensores.

<b>Componente Principal</b>	<b>Variância (%)</b>	<b>Variância Acumulada (%)</b>
PC1	24,90	24,90
PC2	15,90	40,80
PC7	7,50	96,20

Fonte: A autora (2024).

**Tabela 7:** Variância explicada para capacitância e resistência - Sensores.

<b>Componente Principal</b>	<b>Variância (%)</b>	<b>Variância Acumulada (%)</b>
PC1	35,20	35,20
PC2	16,30	51,50
PC10	1,90	96,10

Fonte: A autora (2024).

Os valores das duas primeiras PCs para a capacitância (Tabela 5) conseguem explicar mais de 80% das informações, melhorando a compactação das informações nos primeiros componentes. Esse valor é ainda maior que aquele obtido quando comparamos apenas os parâmetros físico-químicos em que PC1 e PC2 juntos explicam 72,8% (Anexo G). A variância também é mais de 95% explicada com 5 componentes principais.

Em contrapartida, o PCA para a resistência (Tabela 6) melhora muito pouco e só é possível reduzir o sistema em apenas uma variável, ficando com 7 componentes principais. Por fim, com a capacitância e a resistência juntas (Tabela 7), há novamente o balanceamento, não havendo uma melhora significativa nos dados juntos.

As Figura 12, 13 e 14 mostram os gráficos dos dois PCs com maiores informações do sistema, PC1 e PC2. Mesmo sabendo que 2 componentes principais explicam menos de 95%,

é possível observar uma tendência de formação de grupos de amostras que pertencem a mesma ETE. As elipses pontilhadas em cinza claro mostram alguns desses grupos.

No gráfico de capacitância, Figura 12, os escores das amostras das ETEs Norte e Sul ficam aglomerados mais próximos uns dos outros. As ETEs Norte e Sul são consideradas estações “irmãs” pois as águas residuais são submetidas a um mesmo tipo de processo. Assim, a língua eletrônica conseguiu identificar essa tendência sem nenhum tipo de informação prévia. Os escores das amostras de Riacho Fundo também ficaram próximos, formando outro grupo. Percebe-se também um afastamento das amostras de Brazlândia em relação às amostras das estações Norte e Sul. De acordo com a CAESB, as amostras produzidas pela ETE Brazlândia têm um comportamento distinto dessas duas, o que foi percebido pela língua eletrônica. Esses grupos estão mostrados pelas elipses tracejadas. Pode-se notar que as estações Norte e Sul tendem a ficar mais à esquerda do gráfico, enquanto as outras estações tendem a ficar mais à direita. Melchior não apresenta uma tendência clara.

O uso de PCA com a língua eletrônica como método para encontrar grupos que pertencem a uma mesma classe já é bem consolidado na literatura (QUEIROZ *et al.*, 2016). Para amostras com apenas um analito, os grupos são bem definidos e conforme a complexidade da amostra vai aumentando, devido a presença de uma diversidade de analitos, os grupos no PCA começam a ficar pouco definidos, porém, ainda visíveis (QUEIROZ *et al.*, 2016; RIUL, A *et al.*, 2004).

O gráfico de resistência, Figura 13, de dois componentes não identificou nenhum grupo visível, houve um aglomerado dos escores de todas as amostras, sugerindo que a resistência não é um parâmetro muito adequado para analisar esse tipo de amostra.

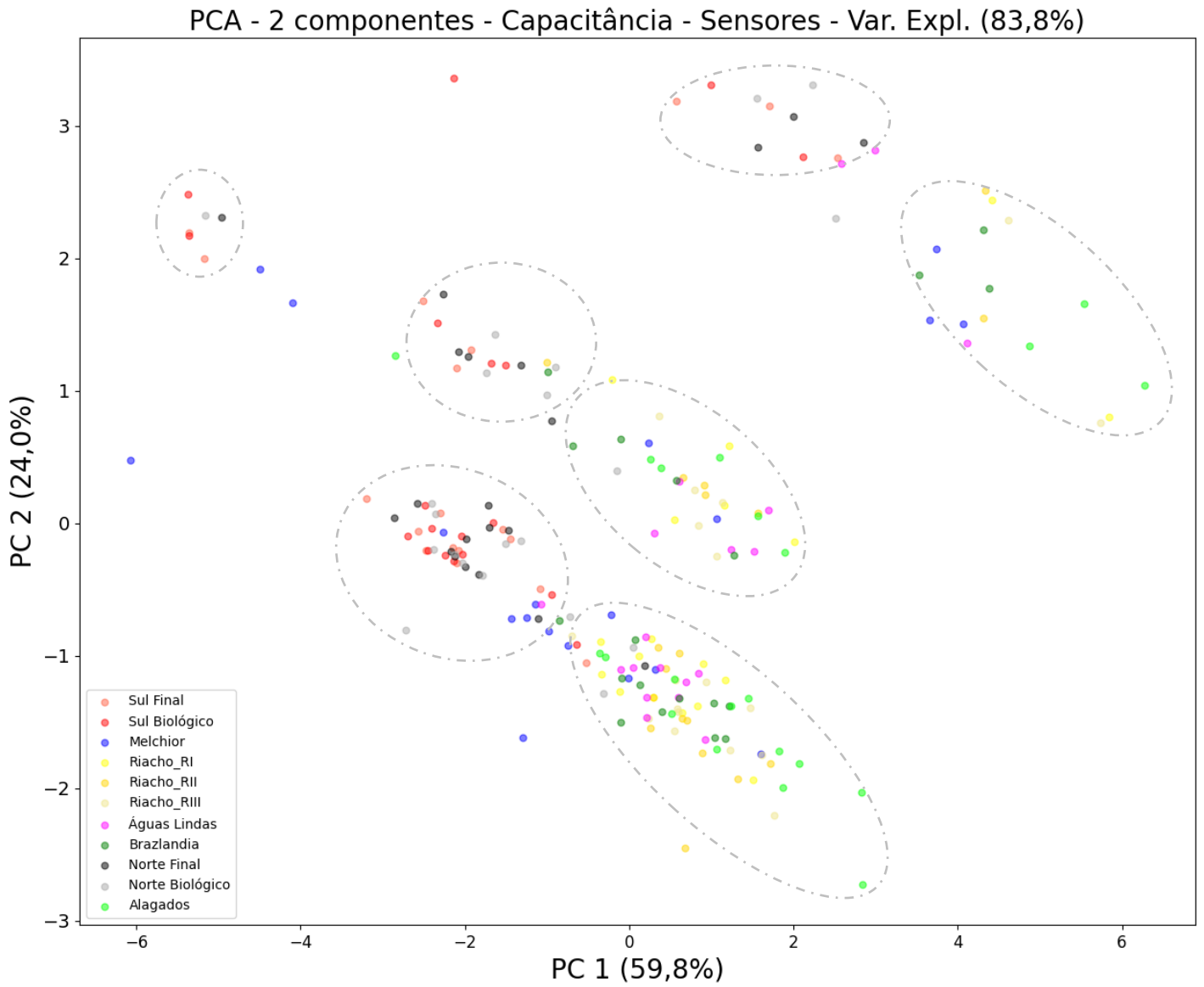
Na Figura 14, no gráfico PCA de capacitância e resistência, a mistura e o espalhamento entre as amostras das ETEs foi um pouco maior se comparada com o gráfico apenas de capacitância. Mesmo assim, é possível observar os grupos de amostras de Riacho Fundo juntos e próximos de Brazlândia. As amostras de Norte e Sul também estão próximas, mas afastadas de outros grupos. Esses conjuntos estão representados pelas elipses tracejadas em cinza claro. Diferente da Figura 12, Norte e Sul tendem a ficar mais à direita do gráfico, enquanto as outras estações tendem a ficar mais à esquerda.

A partir dessas observações, é possível notar que a capacitância é um melhor descritor comparativamente a resistência e é melhor sozinha, uma vez que junto com a resistência não há melhora significativa. O uso da capacitância para a língua eletrônica impedimétrica de fato é melhor como já mostrado em diversos trabalhos feitos sobre isso (Riul Jr. *et al.*, 2010;

Riul Jr. *et al.*, 2013). As leituras feitas em 1kHz potencializam as propriedades dos filmes contendo nanopartículas e aumentam a efetividade da capacitância, devido ao efeito da dupla camada formada quando o eletrodo com o filme é colocado em uma solução eletrolítica, que domina o sinal em frequências de  $10^2$  a  $10^4$  Hz. (TAYLOR, MACDONALD, 1987, RIUL JÚNIOR. *et al.*, 2003). Além disso, capacitância aumenta significativamente em filmes com nanopartículas em sistemas eletrolíticos devido ao aumento de carga na superfície, enquanto a resistência geralmente diminui e o grau de decréscimo depende da rugosidade do nanomaterial. Quanto mais rugoso, maior a porosidade, o que leva a íons dentro desses poros e aumenta o número de portadores de carga que contribuem para a condutividade do filme (ALCANTARA *et al.*, 2013).

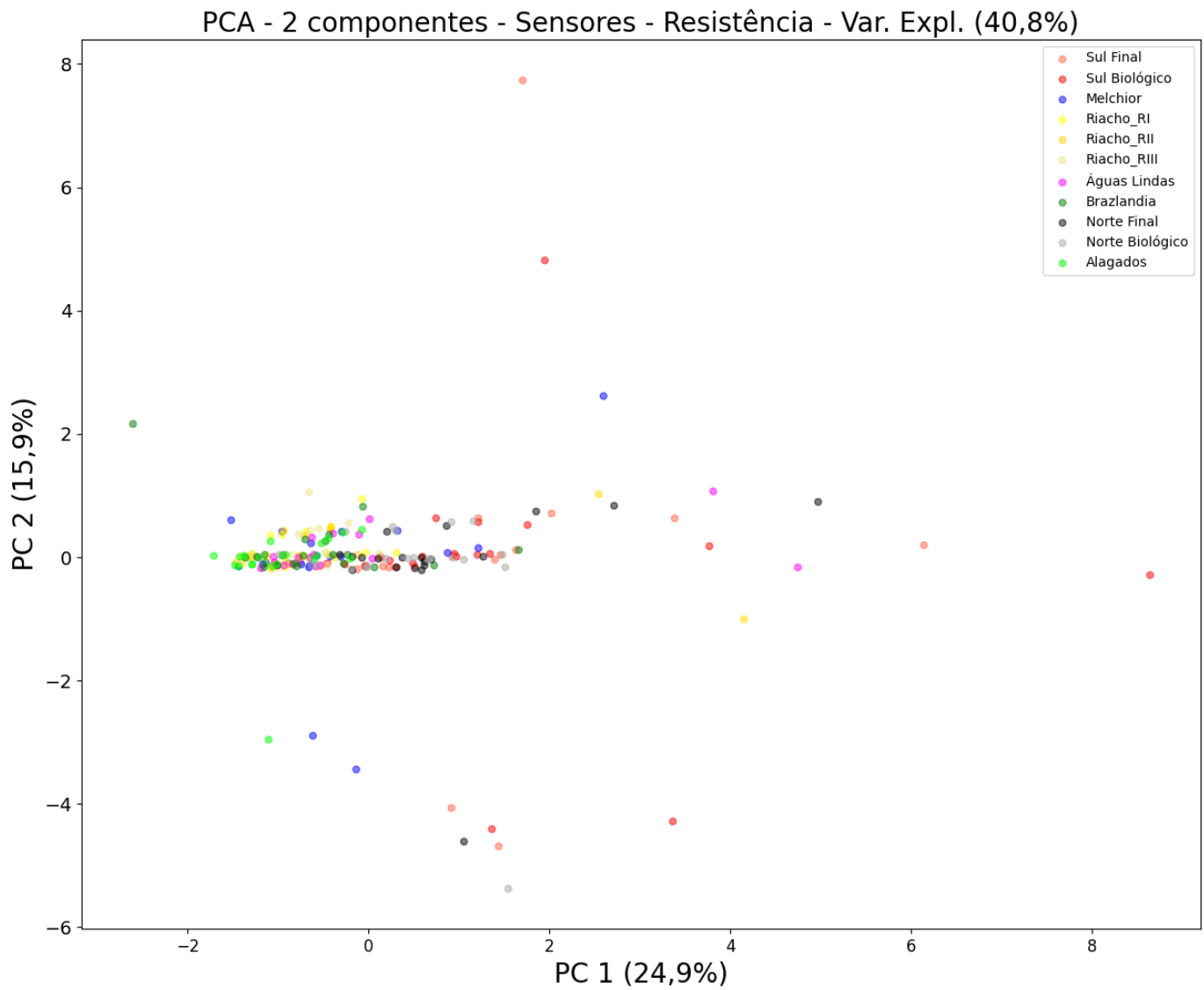


**Figura 12:** Gráfico 2D de dois componentes principais dos sensores de todas as amostras para a Capacitância.



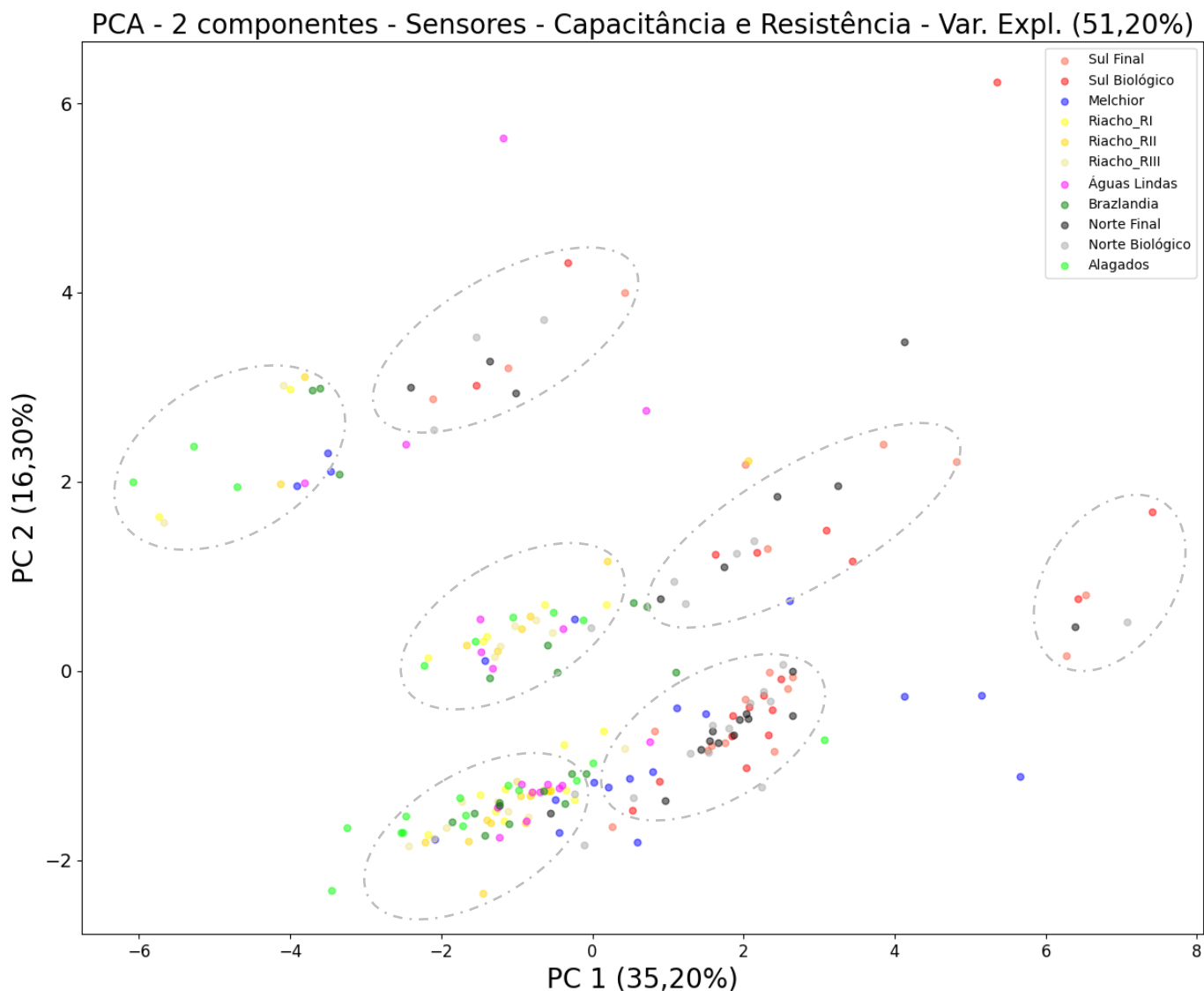
Fonte: A autora (2024).

**Figura 13:** Gráfico 2D de dois componentes principais dos sensores de todas as amostras para a Resistência.



Fonte: A autora (2024).

**Figura 14:** Gráfico 2D de dois componentes principais dos sensores de todas as amostras para a Capacitância e Resistência.



### 5.4.1 Dendrograma e agrupamento

O dendrograma verifica a similaridade entre as amostras por meio da medida de dissimilaridade, também chamada de distância, entre cada ponto. A correlação entre pontos (amostras) é maior quanto menor essa distância entre eles. Assim, é possível fazer a análise de agrupamento hierárquico (HCA). A distância utilizada é a Euclidiana.

As Figura 15, 16 e 17 mostram 3 dendrogramas: de capacitância, de resistência e capacitância e resistência. As ETEs estão codificadas em: Alagado (ALA), Águas Lindas

(AGL), Brazlândia (BRAZ), Melchior (MEL) e Riacho Fundo (RFR) podendo ser RI, RII ou RIII. Para as ETEs Sul (SUL) e Norte (NOR), foi ainda acrescentado BIO para biológico e FIN para final. O número ao final significa o número da amostra relativo àquela estação.

Pelos dendrogramas, há uma mistura maior das amostras e não há uma tendência clara em amostras de mesma ETE ficarem mais próximas. Na Figura 15, há uma junção de algumas amostras de Águas Lindas e, também, as estações Norte e Sul também tendem a ficar um pouco mais próximas na parte laranja.

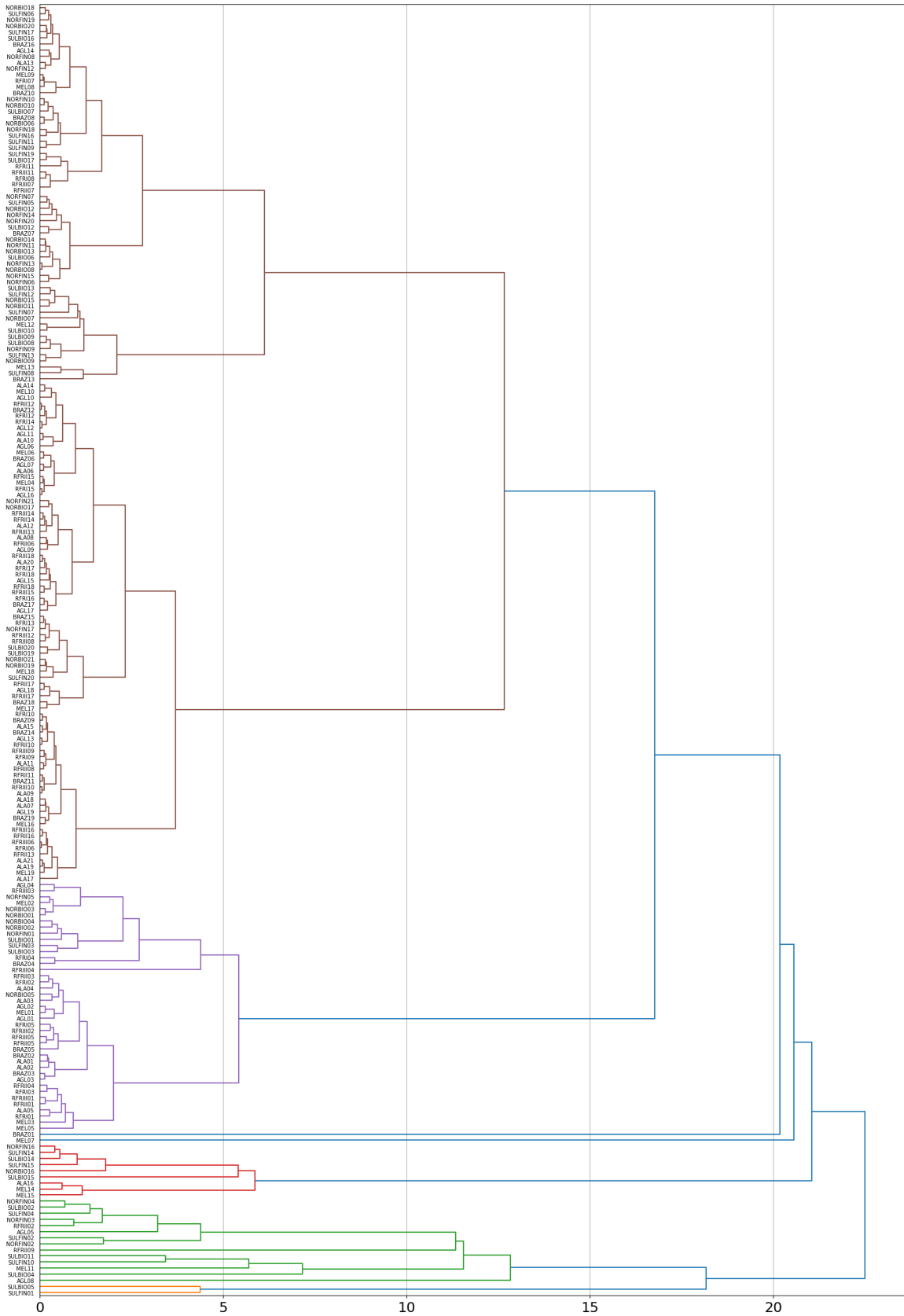
Grupos mal definidos também são formados no dendrograma de resistência, Figura 16. Há algumas amostras de Norte Biológico que se juntam mais próximas, isso acontece também com Sul Final e algumas amostras de Riacho Fundo. Há também umas discrepâncias com relação às amostras finais. Pode-se ver que há mais separação de grupos do que em relação ao dendrograma de capacitância. No dendrograma de resistência (Figura 16), há a formação de 6 grupos contra 3 grupos formados com os dados de capacitância (Figura 15). Contudo, para a capacitância, os grupos são mais definidos. Na parte vermelha, há a maior concentração de amostras pertencentes a Riacho Fundo, Brazlândia e Águas Lindas e, como já dito, Norte e Sul na parte laranja.

O mesmo acontece para o dendrograma de capacitância e resistência (Figura 17). Nesse caso, a separação ainda foi um pouco pior em relação aos anteriores, há uma mistura maior entre as amostras. Pode-se ver 4 grupos maiores, maior separação em relação a capacitância, contudo grupos mal definidos. Pode-se ver poucas amostras e Águas Lindas juntas, bem como amostras de Sul próximas entre si e Norte também juntas.

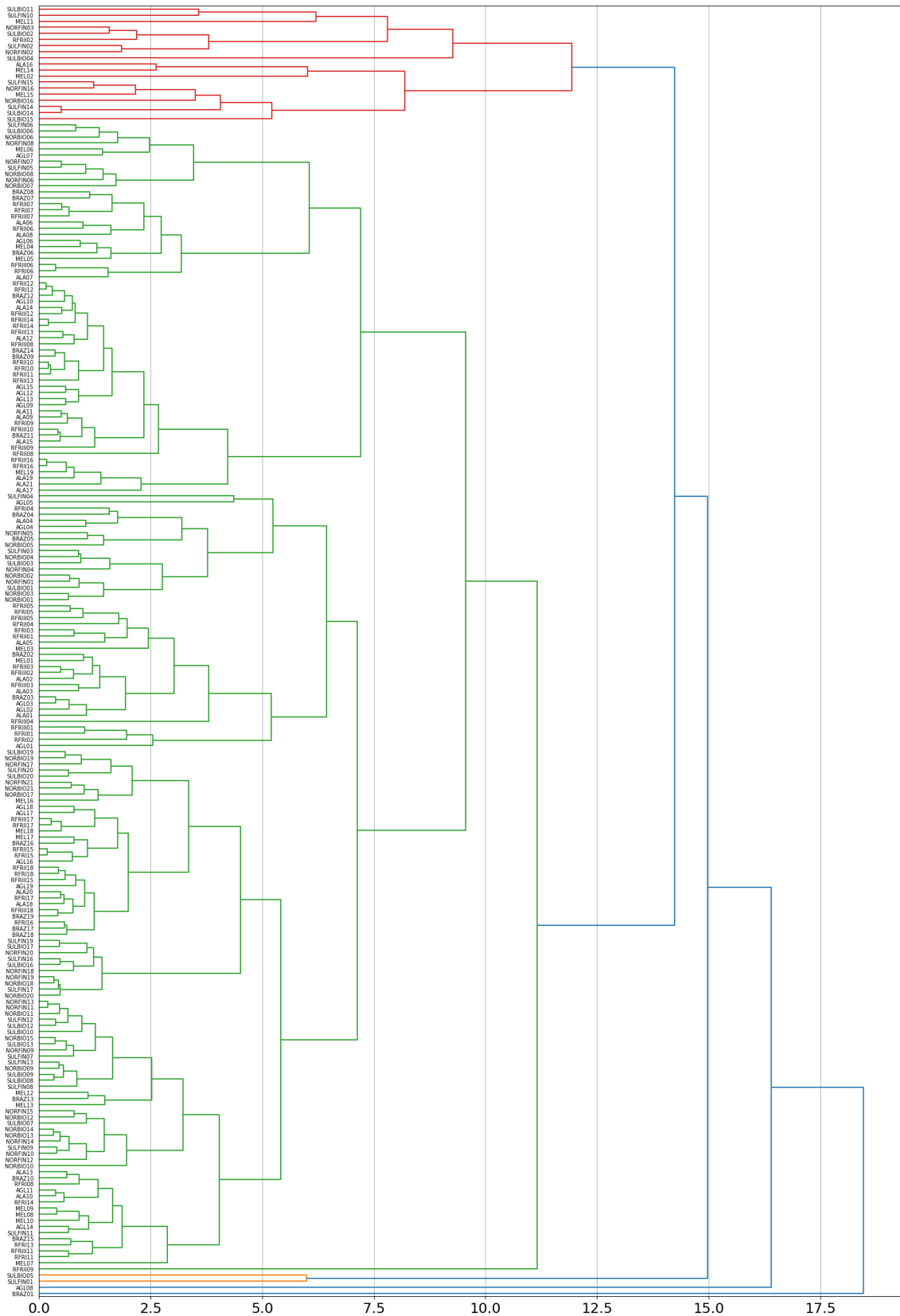
O dendrograma não foi muito eficaz para mostrar separação visual entre grupos de amostras pertencentes a uma mesma estação. Um motivo também pode se justificar pela grande quantidade de amostras, o que torna a visualização pelo dendrograma mais difícil.



**Figura 16:** Dendrograma dos sensores com valores autoescalados para resistência.



**Figura 17:** Dendrograma dos sensores com valores autoescalados para a (c) Capacitância e Resistência.



Fonte: A autora (2024).

O agrupamento mescla recursivamente pares de grupos de dados da amostra e usa a distância Euclidiana e o método de ligação Ward. O agrupamento é feito de baixo para cima, cada observação começa em seu próprio grupo e estes são mesclados sucessivamente. Assim, nesse método é permitido escolher em quantos grupos se deseja separar as amostras. Foi escolhido formar 7 grupos com base no número de estações de tratamento, de modo a avaliar se a língua consegue discriminar as amostras de água residual pelas regiões das ETEs.

**Quadro 1:** Formação de 7 grupos para capacitância. Dados autoescalados.

<b>Grupo 1</b>	'BRAZ01', 'BRAZ02', 'BRAZ03', 'BRAZ04', 'BRAZ05', 'BRAZ09', 'AGL09', 'BRAZ10', 'AGL10', 'BRAZ11', 'AGL11', 'BRAZ12', 'AGL12', 'BRAZ13', 'AGL13', 'BRAZ14', 'AGL14', 'BRAZ15', 'AGL15', 'BRAZ16', 'AGL16', 'BRAZ17', 'AGL17', 'BRAZ18', 'AGL18', 'BRAZ19', 'AGL19'
<b>Grupo 2</b>	'AGL01', 'RFRI01', 'RFRII01', 'RFRIII01', 'ALA01', 'ALA02', 'AGL02', 'RFRI02', 'RFRII02', 'ALA03', 'RFRI03', 'RFRII03', 'RFRIII03', 'AGL04', 'RFRI04', 'RFRII04', 'RFRIII04', 'ALA04', 'AGL05', 'RFRI05', 'RFRII05', 'RFRIII05', 'ALA05', 'RFRI08', 'RFRII08', 'RFRIII08', 'ALA09', 'RFRI09', 'RFRII09', 'RFRIII09', 'ALA10', 'RFRI10', 'RFRII10', 'RFRIII10', 'ALA11', 'RFRI11', 'RFRII11', 'RFRIII11', 'ALA12', 'RFRI12', 'RFRII12', 'RFRIII12', 'ALA13', 'RFRI13', 'RFRII13', 'RFRIII13', 'ALA14', 'RFRI14', 'RFRII14', 'RFRIII14', 'ALA15', 'ALA17', 'RFRI15', 'RFRII15', 'RFRIII15', 'RFRI16', 'RFRII16', 'RFRIII16', 'ALA18', 'RFRI17', 'RFRII17', 'RFRIII17', 'ALA19', 'ALA20', 'RFRI18', 'RFRII18', 'RFRIII18', 'ALA21'
<b>Grupo 3</b>	'SULFIN01', 'SULBIO01', 'NORBIO01', 'NORFIN01', 'NORBIO02', 'NORFIN02', 'SULBIO02', 'SULFIN02', 'RFRII02', 'NORBIO03', 'NORFIN03', 'NORBIO04', 'NORFIN04', 'SULBIO03', 'SULFIN03', 'NORBIO05', 'NORFIN05', 'SULBIO04', 'SULBIO07', 'SULFIN07', 'NORBIO09', 'NORFIN09', 'SULBIO08', 'SULFIN08', 'NORBIO10', 'NORFIN10', 'SULBIO09', 'SULFIN09', 'NORBIO11', 'NORFIN11', 'SULBIO10', 'SULFIN10', 'NORBIO12', 'NORFIN12', 'SULBIO11', 'SULFIN11', 'NORBIO13', 'NORFIN13', 'SULBIO12', 'SULFIN12', 'NORBIO14', 'NORFIN14', 'SULBIO13', 'SULFIN13', 'NORBIO15', 'NORFIN15', 'ALA16', 'NORBIO17', 'NORFIN17', 'SULBIO16', 'SULFIN16', 'NORBIO18', 'NORFIN18', 'SULBIO17', 'SULFIN17', 'NORBIO19', 'NORFIN19', 'SULBIO19', 'SULFIN19', 'NORBIO20', 'NORFIN20', 'SULBIO20', 'SULFIN20', 'NORBIO21', 'NORFIN21'
<b>Grupo 4</b>	'MEL04', 'AGL06', 'BRAZ06', 'RFRI06', 'RFRII06', 'RFRIII06', 'ALA06', 'MEL05', 'BRAZ07', 'ALA07', 'MEL06', 'BRAZ08', 'RFRI07', 'RFRII07', 'RFRIII07', 'ALA08'
<b>Grupo 5</b>	'MEL01', 'AGL03', 'MEL03', 'MEL07', 'MEL08', 'MEL09', 'MEL10', 'MEL11', 'MEL12', 'MEL13', 'MEL16', 'MEL17', 'MEL18', 'MEL19'
<b>Grupo 6</b>	'MEL02', 'MEL14', 'SULBIO14', 'SULFIN14', 'NORBIO16', 'NORFIN16', 'MEL15', 'SULBIO15', 'SULFIN15'
<b>Grupo 7</b>	'SULFIN04', 'NORBIO06', 'NORFIN06', 'SULBIO05', 'SULFIN05', 'AGL07', 'NORBIO07', 'NORFIN07', 'SULBIO06', 'SULFIN06', 'AGL08', 'NORBIO08', 'NORFIN08'

Fonte: A autora (2024).



**Quadro 2:** Formação de 7 grupos para resistência. Dados autoescalados.

<b>Grupo 1</b>	'MEL01', 'AGL01', 'RFRI01', 'RFRII01', 'RFRIII01', 'ALA01', 'ALA02', 'MEL02', 'AGL02', 'RFRI02', 'RFRII02', 'RFRIII02', 'ALA03', 'AGL03', 'RFRI03', 'RFRII03', 'RFRIII03', 'AGL04', 'RFRI04', 'RFRII04', 'RFRIII04', 'ALA04', 'MEL03', 'RFRI05', 'RFRII05', 'RFRIII05', 'ALA05', 'MEL04', 'AGL06', 'RFRI06', 'RFRII06', 'RFRIII06', 'ALA06', 'MEL05', 'AGL07', 'ALA07', 'MEL06', 'RFRI07', 'RFRII07', 'RFRIII07', 'ALA08', 'RFRI08', 'RFRII08', 'RFRIII08', 'ALA09', 'MEL08', 'RFRI09', 'RFRII09', 'RFRIII09', 'ALA10', 'MEL09', 'RFRI10', 'RFRII10', 'RFRIII10', 'ALA11', 'MEL10', 'RFRI11', 'RFRII11', 'RFRIII11', 'ALA12', 'RFRI12', 'RFRII12', 'RFRIII12', 'ALA13', 'RFRI13', 'RFRII13', 'RFRIII13', 'ALA14', 'RFRI14', 'RFRII14', 'RFRIII14', 'ALA15', 'ALA17', 'MEL16', 'RFRI15', 'RFRII15', 'RFRIII15', 'RFRI16', 'RFRII16', 'RFRIII16', 'ALA18', 'MEL17', 'RFRI17', 'RFRII17', 'RFRIII17', 'ALA19', 'MEL18', 'ALA20', 'MEL19', 'RFRI18', 'RFRII18', 'RFRIII18', 'ALA21'
<b>Grupo 2</b>	'SULBIO01', 'NORBIO01', 'NORFIN01', 'NORBIO02', 'NORFIN02', 'SULBIO02', 'SULFIN02', 'RFRII02', 'NORBIO03', 'NORFIN03', 'NORBIO04', 'NORFIN04', 'SULBIO03', 'SULFIN03', 'AGL05', 'NORBIO05', 'NORFIN05', 'SULBIO04', 'SULFIN04', 'NORBIO06', 'NORFIN06', 'SULFIN05', 'NORBIO07', 'NORFIN07', 'SULBIO06', 'SULFIN06', 'AGL08', 'NORBIO08', 'NORFIN08', 'SULBIO07', 'SULFIN07', 'NORBIO09', 'NORFIN09', 'SULBIO08', 'SULFIN08', 'RFRII09', 'NORBIO10', 'NORFIN10', 'SULBIO09', 'SULFIN09', 'NORBIO11', 'NORFIN11', 'SULBIO10', 'SULFIN10', 'NORBIO12', 'NORFIN12', 'MEL11', 'SULBIO11', 'SULFIN11', 'NORBIO13', 'NORFIN13', 'MEL12', 'SULBIO12', 'SULFIN12', 'NORBIO14', 'NORFIN14', 'MEL13', 'SULBIO13', 'SULFIN13', 'NORBIO15', 'NORFIN15', 'NORFIN17', 'SULBIO16', 'SULFIN16', 'NORBIO18', 'NORFIN18', 'SULBIO17', 'SULFIN17', 'NORBIO19', 'NORFIN19', 'SULBIO19', 'SULFIN19', 'NORBIO20', 'NORFIN20', 'SULBIO20', 'SULFIN20', 'NORFIN21'
<b>Grupo 3</b>	'BRAZ02', 'BRAZ03', 'BRAZ04', 'BRAZ05', 'BRAZ06', 'BRAZ07', 'BRAZ08', 'BRAZ09', 'AGL09', 'BRAZ10', 'AGL10', 'BRAZ11', 'AGL11', 'BRAZ12', 'AGL12', 'BRAZ13', 'AGL13', 'BRAZ14', 'AGL14', 'BRAZ15', 'AGL15', 'NORBIO17', 'BRAZ16', 'AGL16', 'BRAZ17', 'AGL17', 'BRAZ18', 'AGL18', 'BRAZ19', 'AGL19', 'NORBIO21'
<b>Grupo 4</b>	'MEL07'
<b>Grupo 5</b>	'MEL14', 'SULBIO14', 'SULFIN14', 'ALA16', 'NORBIO16', 'NORFIN16', 'MEL15', 'SULBIO15', 'SULFIN15'
<b>Grupo 6</b>	'BRAZ01'
<b>Grupo 7</b>	'SULFIN01', 'SULBIO05'

Fonte: A autora (2024).

**Quadro 3:** Formação de 7 grupos para capacitância e resistência. Dados autoescalados.

<b>Grupo 1</b>	'SULBIO01', 'NORBIO01', 'NORFIN01', 'NORBIO02', 'NORFIN02', 'SULBIO02', 'SULFIN02', 'RFRII02', 'NORBIO03', 'NORFIN03', 'NORBIO04', 'NORFIN04', 'SULBIO03', 'SULFIN03', 'AGL05', 'NORBIO05', 'NORFIN05', 'SULBIO04', 'SULFIN04', 'SULBIO07', 'SULFIN07', 'NORBIO09', 'NORFIN09', 'SULBIO08', 'SULFIN08', 'NORBIO10', 'NORFIN10', 'SULBIO09', 'SULFIN09', 'NORBIO11', 'NORFIN11', 'SULBIO10', 'SULFIN10', 'NORBIO12', 'NORFIN12', 'MEL11', 'SULBIO11', 'SULFIN11', 'NORBIO13', 'NORFIN13', 'MEL12', 'SULBIO12', 'SULFIN12', 'NORBIO14', 'NORFIN14', 'MEL13', 'SULBIO13', 'SULFIN13', 'NORBIO15', 'NORFIN15', 'NORFIN17', 'SULBIO16', 'SULFIN16', 'NORBIO18', 'NORFIN18', 'SULBIO17', 'SULFIN17', 'NORBIO19', 'NORFIN19', 'SULBIO19', 'SULFIN19', 'NORBIO20', 'NORFIN20', 'SULBIO20', 'SULFIN20', 'NORFIN21'
<b>Grupo 2</b>	'MEL01', 'AGL01', 'RFRI01', 'RFRII01', 'RFRIII01', 'ALA01', 'ALA02', 'AGL02', 'RFRI02', 'RFRII02', 'ALA03', 'AGL03', 'RFRI03', 'RFRII03', 'RFRIII03', 'AGL04', 'RFRI04', 'RFRII04', 'RFRIII04', 'ALA04', 'MEL03', 'RFRI05', 'RFRII05', 'RFRIII05', 'ALA05', 'MEL07', 'RFRI08', 'RFRII08', 'RFRIII08', 'ALA09', 'MEL08', 'RFRI09', 'RFRII09', 'RFRIII09', 'ALA10', 'MEL09', 'RFRI10', 'RFRII10', 'RFRIII10', 'ALA11', 'MEL10', 'RFRI11', 'RFRII11', 'RFRIII11', 'ALA12', 'RFRI12', 'RFRII12', 'RFRIII12', 'ALA13', 'RFRI13', 'RFRII13', 'RFRIII13', 'ALA14', 'RFRI14', 'RFRII14', 'RFRIII14', 'ALA15', 'ALA17', 'MEL16', 'RFRI15', 'RFRII15', 'RFRIII15', 'RFRI16', 'RFRII16', 'RFRIII16', 'ALA18', 'MEL17', 'RFRI17', 'RFRII17', 'RFRIII17', 'ALA19', 'MEL18', 'ALA20', 'MEL19', 'RFRI18', 'RFRII18', 'RFRIII18', 'ALA21'
<b>Grupo 3</b>	'MEL04', 'AGL06', 'BRAZ06', 'RFRI06', 'RFRII06', 'RFRIII06', 'ALA06', 'NORBIO06', 'NORFIN06', 'MEL05', 'SULFIN05', 'AGL07', 'BRAZ07', 'ALA07', 'NORBIO07', 'NORFIN07', 'MEL06', 'SULBIO06', 'SULFIN06', 'AGL08', 'BRAZ08', 'RFRI07', 'RFRII07', 'RFRIII07', 'ALA08', 'NORBIO08', 'NORFIN08'
<b>Grupo 4</b>	'BRAZ02', 'BRAZ03', 'BRAZ04', 'BRAZ05', 'BRAZ09', 'AGL09', 'BRAZ10', 'AGL10', 'BRAZ11', 'AGL11', 'BRAZ12', 'AGL12', 'BRAZ13', 'AGL13', 'BRAZ14', 'AGL14', 'BRAZ15', 'AGL15', 'NORBIO17', 'BRAZ16', 'AGL16', 'BRAZ17', 'AGL17', 'BRAZ18', 'AGL18', 'BRAZ19', 'AGL19', 'NORBIO21'
<b>Grupo 5</b>	'MEL02', 'MEL14', 'SULBIO14', 'SULFIN14', 'ALA16', 'NORBIO16', 'NORFIN16', 'MEL15', 'SULBIO15', 'SULFIN15'
<b>Grupo 6</b>	'BRAZ01'
<b>Grupo 7</b>	'SULFIN01', 'SULBIO05'

Fonte: A autora (2024).

Com o agrupamento mostrado nos Quadro 1, 2 e 3, pode-se ver que ficou mais fácil a visualização da divisão das amostras nos grupos. Era esperado que as amostras de mesma estação ficassem nos mesmos grupos. Apesar de não ter sido assim, muitos grupos contêm muitas amostras da mesma estação.

No grupamento de capacitância (Quadro 1), o grupo 1 ficou com grande maioria de amostras de Brazlândia e Águas lindas. Regionalmente, as duas estações são próximas entre si,

o que pode explicar ficarem no mesmo grupo. Águas Lindas próxima de Brazlândia também foi observado no gráfico de PCA. No grupo 2, há bastante amostras de Riacho Fundo, além de algumas amostras de Alagado. Essas duas estações têm também uma certa proximidade regional. Os grupos 3 e 7 contam com maioria de amostras que fazem parte das estações Norte e Sul, as duas estações “irmãs” e o grupo 5 com mais amostras de Melchior.

No Quadro 2, os grupos formados não foram tão bons, houve mais misturas em relação a capacitância e mais grupos descompensados. Porém, o grupo 1 conta com mais amostras de Riacho Fundo e Alagado, o grupo 2 com mais amostras de Sul e Norte e o grupo 3 com amostras de Brazlândia e Águas Lindas.

Com os dados de capacitância e resistência (Quadro 3), o agrupamento melhorou em relação à resistência sozinhas, mas ainda foi inferior à capacitância. O grupo 1 tem maioria de amostras das estações Norte e Sul. Já o grupo 2 tem algumas amostras de Melchior e Águas Lindas no meio da grande maioria de amostras de Alagados e Riacho Fundo. No grupo 4, a maioria das amostras são de Águas Lindas e Brazlândia e o grupo 7, apesar de 2 amostras somente, contém somente estação Sul.

À vista disso, o grupamento foi satisfatório no geral para discriminar as amostras nas estações, além de visualmente melhor. O dendrograma explica todas as informações contidas em um gráfico 2D, porém, com o aumento da quantidade de amostras, a análise se torna mais complexa.

Assim sendo, a língua eletrônica conseguiu identificar o padrão de estações de tratamento sem informações prévias, apenas com os sensores e os resultados de capacitância. Ainda que os resultados não tenham atingido índice de acerto de 100%, vale ressaltar que as amostras de água são sistemas muito complexos e que dependem de inúmeros fatores e, assim, os resultados foram bastante satisfatórios. Isso abre possibilidades para identificação de amostras desconhecidas de água de tratamento ou outras utilizações.

## **5.5 Estudo dos indicadores e aprendizado de máquina**

O próximo estudo realizou o aprendizado de máquina para que, a língua eletrônica pudesse responder qualitativamente se uma amostra era “conforme” ou “não conforme” dentro de um padrão escolhido. Por isso, foram escolhidos 4 parâmetros físico-químicos com maior disponibilidade de valores nos relatórios da CAESB, que eram de maior interesse e eram comuns a todas as estações: DQO, SST, NT e PHOSP.

Uma tabela foi elaborada com os valores dos parâmetros físico-químicos fornecidos pela CAESB, a data da análise e a estação de tratamento. Para o aprendizado de máquina, as faixas de aceitação de cada parâmetro também foram obtidas a partir dos protocolos da CAESB de análise de água. Com isso, foi determinado um *score* para ajudar no aprendizado da máquina, conforme descrito no Quadro 4.

**Quadro 4:** Indicadores e faixa de ótimo

INDICADOR	SCORE	FAIXA (mg/L)	PARÂMETRO
Baixo	0	menor que 30	DQO
Médio	1	entre 30 e 70	DQO
Alto	2	maior que 70	DQO
Baixo	0	menor que 15	SST
Médio	1	entre 15 e 40	SST
Alto	2	maior que 40	SST
Baixo	0	menor que 10	NT
Médio	1	entre 10 e 25	NT
Alto	2	maior que 25	NT
Baixo	0	menor que 0,3	PHOSP
Médio	1	entre 0,3 e 1,0	PHOSP
Alto	2	maior que 1,0	PHOSP

Fonte: A autora (2024).

Quatro grandes tabelas foram organizadas, contendo os valores as saídas desejadas, os dados de capacitância e de resistência. A partir dessas tabelas, os dados dos sensores foram autoescalados para que não houvesse problema de unidades e medidas muito diferentes. Em seguida, os dados foram divididos em partes para o aprendizado de máquina. Dos 212 resultados, 85% foram usados para treino e 15% foram separados para teste e 800 iterações para o aprendizado de máquina.

O método de PCA foi utilizado para aumentar a eficiência, uma vez que ele compacta as informações. No estudo de PCA dos sensores, foi observado que, para a capacitância, 5 componentes principais explicam 97,10% das informações. Para a resistência, 7 componentes principais explicam 96,20% e para a capacitância e resistência, 10 componentes explicam 96,10%. O método de regressão, um método de classificação supervisionado, foi escolhido para treinar a máquina. Assim, para a implementação da regressão logística com PCA, essas informações dos componentes principais foram usadas.

Primeiramente, com a tabela dos *scores* foi determinado a relação da quantidade cada *score* com as amostras, mostradas na Tabela 8.

**Tabela 8:** Relação de quantidade de cada *score*.

	<b>0 (baixo)</b>	<b>1 (médio)</b>	<b>2 (alto)</b>
<b>DQO</b>	53	111	48
<b>SST</b>	123	53	36
<b>NT</b>	73	5	134
<b>PHOSP</b>	75	60	77

Fonte: A autora (2024).

Pode-se perceber que os dados das quantidades não são bem homogêneos, pois há uma descompensação de quantidade. A maior descompensação está em relação ao NT, no qual há uma grande quantidade de amostras com NT alto e pouquíssimas com valores médios. Essa descompensação pode gerar um problema, o treinamento da máquina fica prejudicado, uma vez que os dados escolhidos para treinamento foram escolhidos de maneira aleatória. Por sua vez, o valor do PHOSP é bem homogêneo, sendo bom para o treinamento da máquina.

Os índices de acerto da língua eletrônica no reconhecimento de cada um dos parâmetros físico-químicos treinados são apresentados no Quadro 5.

**Quadro 5:** Porcentagem de acerto da língua eletrônica para os indicadores.

<b>Indicador</b>	<b>Porcentagem de Acerto (%)</b>		
	<b>Capacitância</b>	<b>Resistência</b>	<b>Capacitância e Resistência</b>
DQO	46,87	50,00	59,37
SST	50,00	62,50	56,25
NT	62,50	75,00	71,87
PHOSP	53,12	31,25	46,87

Fonte: A autora (2024).

Apesar da porcentagem baixa de acerto para os quatro parâmetros, com nenhum chegando a um índice de acerto maior que 90%, o resultado se mostrou bastante interessante. Os valores variaram em relação à análise feita na primeira parte do trabalho com 79 amostras, mostradas no Quadro 6.

**Quadro 6:** Porcentagem de acerto língua eletrônica para os indicadores feita na primeira parte do trabalho.

<b>Indicador</b>	<b>Capacitância - Porcentagem de Acerto (%)</b>
DQO	33,34 %
SST	66,67%
NT	91,67%

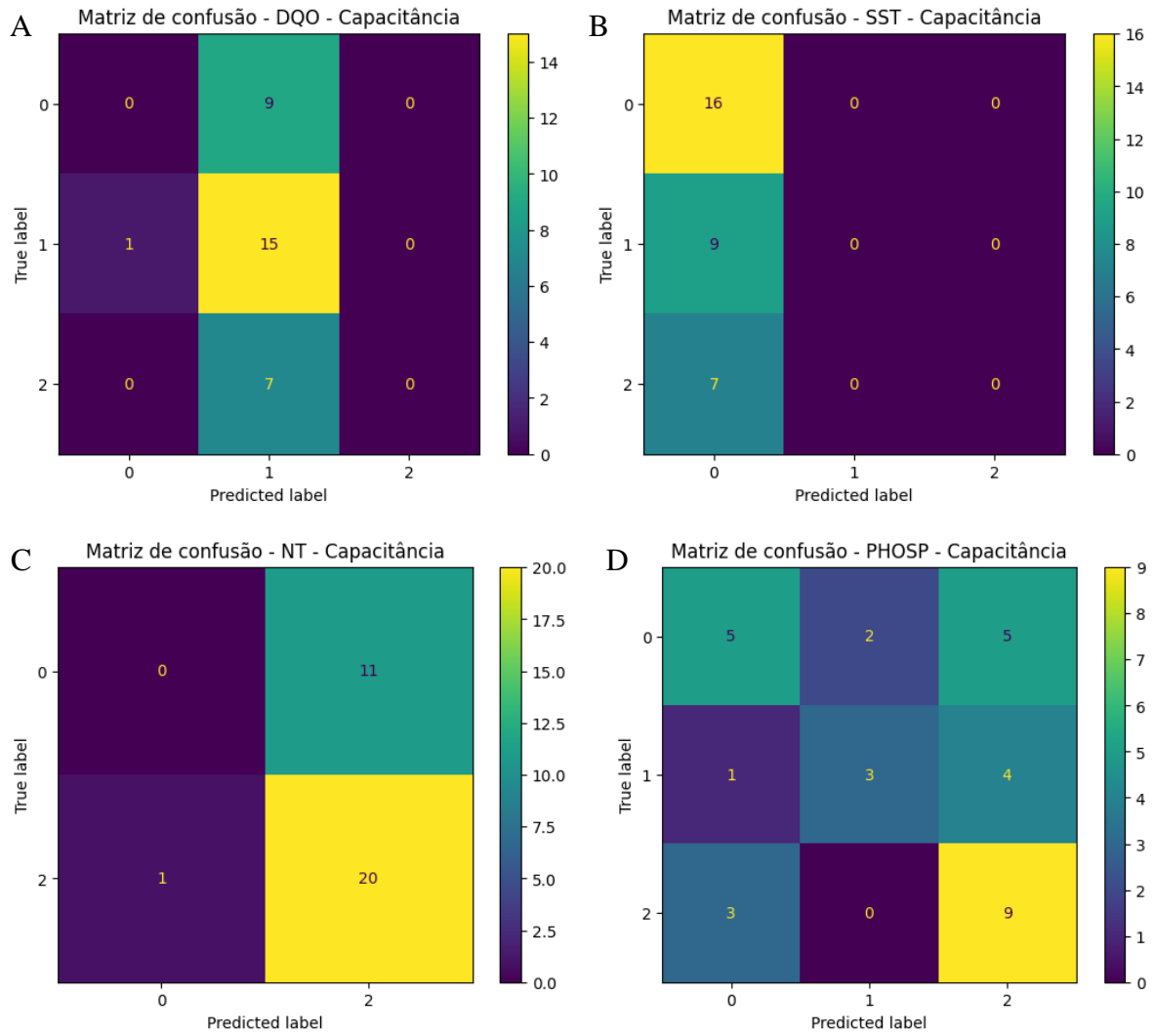
Fonte: A autora (2024).

Para DQO, houve um aumento bem significativo de acerto. O aumento do acerto do SST foi pequeno, enquanto que o acerto do NT diminuiu significativamente em relação a capacitância. Contudo, o valor de NT tem que ser avaliado com cuidado, uma vez que a discrepância de quantidade de *scores* pode fazer essa porcentagem de acerto variar muito.

O valor de acerto do PHOSP deve ser considerado mais consistente devido à homogeneidade dos dados. Vale ressaltar que esse tipo de sistema é completamente novo, não havendo dado sobre ele. Essas informações já permitiram a criação de um banco de dados que já o primeiro passo para o aprimoramento. 212 amostras é um número ainda considerado baixo para o treinamento da máquina, geralmente são necessárias na casa dos milhares de dados. Contudo, o desempenho do treinamento foi muito positivo, indicando que a língua eletrônica pode ser treinada para reconhecer padrões e se tornar um equipamento de uso prático.

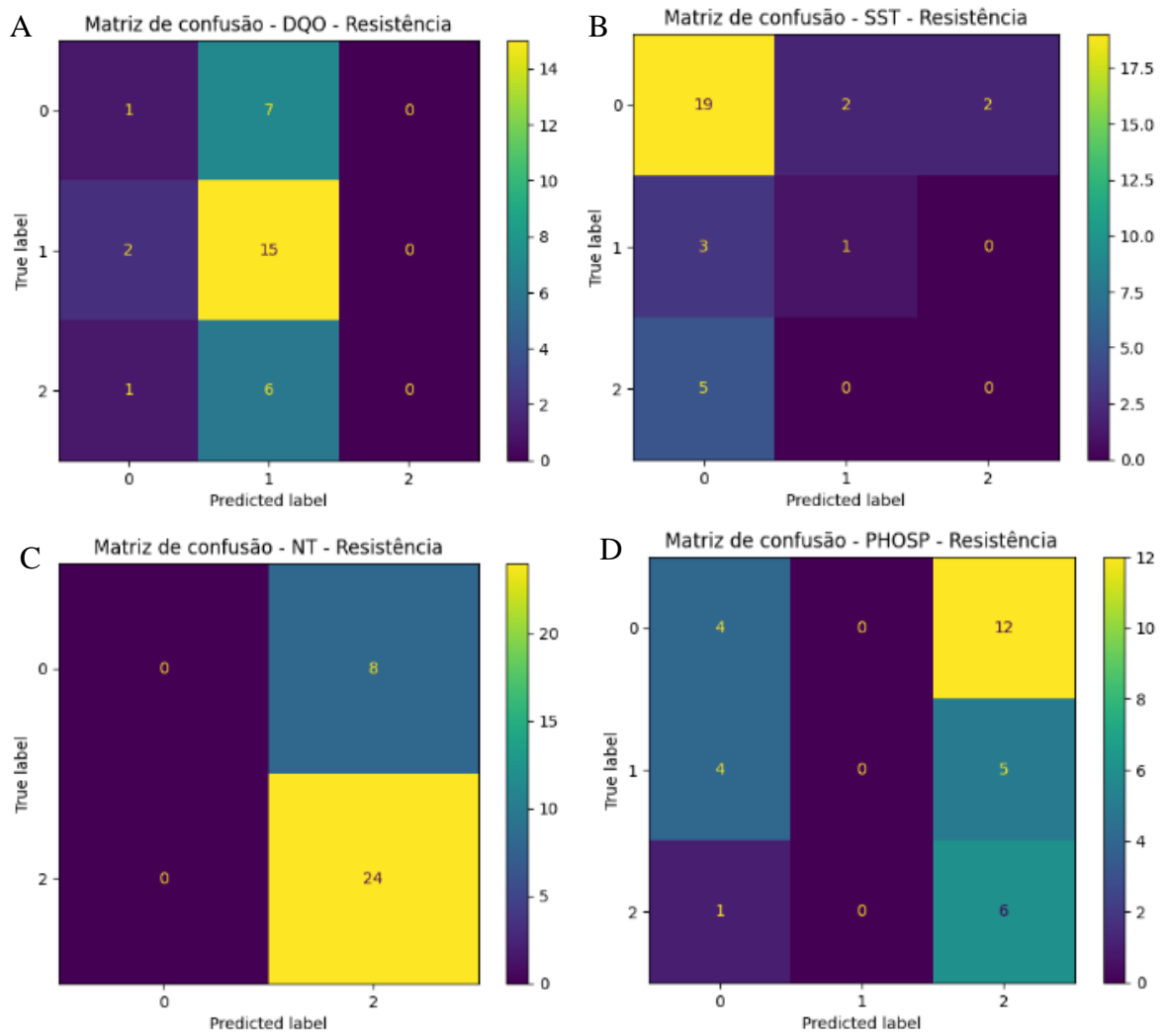
Nas Figura 18, 19 e 20 são apresentadas as matrizes de confusão para os quatro parâmetros físico-químicos treinados na língua eletrônica. A matriz de confusão mostra a resposta do sistema com o aprendizado de máquina implementado. O ideal seria a diagonal da matriz ser toda clara e as outras cores escuras. Pela matriz é possível constatar o número baixo de acertos (cores com tons mais claros fora da diagonal).

**Figura 18:** Matrizes de confusão para a capacitância. Acertos e erros do algoritmo de aprendizado de máquina. (a) DQO, (b) SST, (c) NT e (d) PHOSP.



Fonte: A autora (2024).

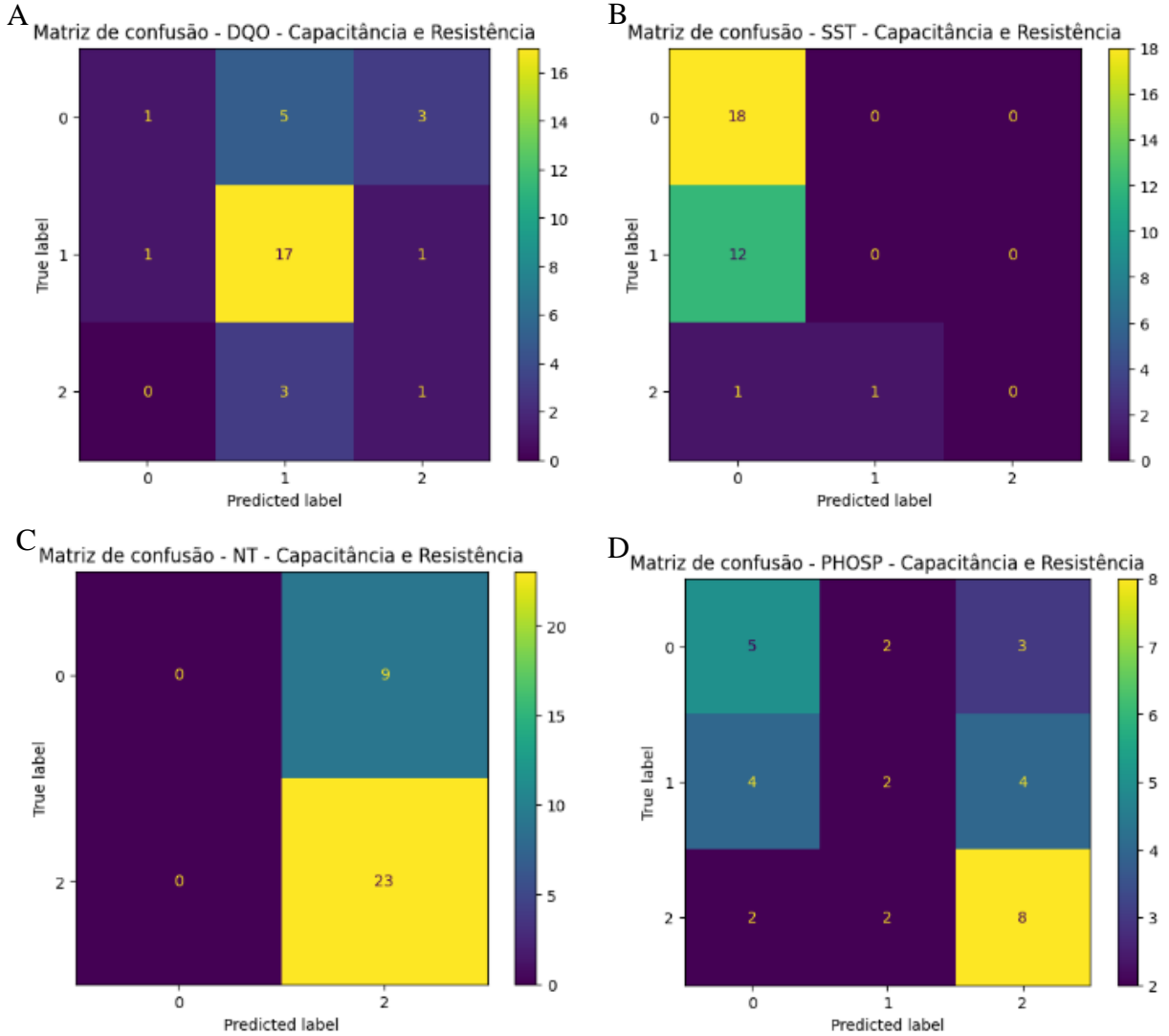
**Figura 19:** Matrizes de confusão para a resistência. Acertos e erros do algoritmo de aprendizado de máquina. (a) DQO, (b) SST, (c) NT e (d) PHOSP.



Fonte: A autora (2024).



**Figura 20:** Matrizes de confusão para a resistência. Acertos e erros do algoritmo de aprendizado de máquina para os quatro indicadores analisados: (a) DQO, (b) SST, (c) NT e (d) PHOSP.



Fonte: A autora (2024).

Um gráfico PCA de dois componentes mostra visualmente como se distribuem as amostras em relação aos *scores* estabelecidos. Por esses gráficos, mostrados nas Figura 21, 22 e 23, pode-se perceber que os grupos de amostras de baixa, média e alta concentração estão bem misturados. Isso indica porque o treinamento da máquina teve uma porcentagem de acerto baixa para os parâmetros. Os PCAs para nitrogênio mostraram uma boa separação entre as amostras, porém, pode-se notar também a baixa presença de amostras de média concentração, que influencia diretamente no treinamento.

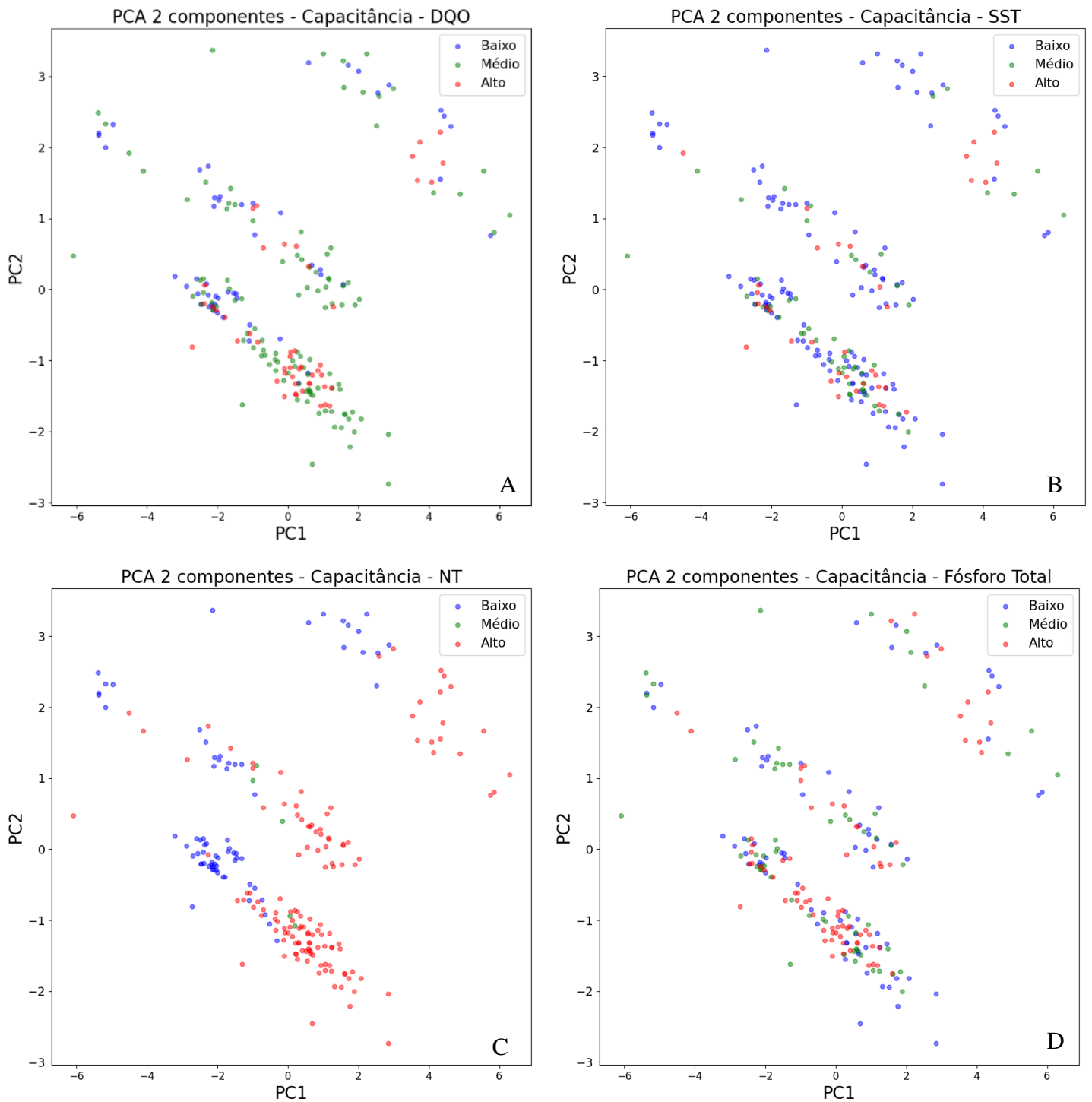
Os gráficos de capacitância para os indicadores mostram uma menor distância entre as amostras e a formação de 3 grandes grupos. Na Figura 23, os gráficos capacitância e resistência mostram um comportamento parecido com o da Figura 21, porém, há um espalhamento maior das amostras. Já para os gráficos de somente resistência, Figura 22, as amostras se uniram em um grande grupo, sem distinção visível entre as amostras. Contudo, é interessante notar que para o gráfico de nitrogênio, é possível distinguir entre baixo e alto.

A partir dessas observações, é importante perceber que a capacitância continua se mostrando o parâmetro mais eficaz para a língua eletrônica e a resistência tendo baixa eficácia, corroborando com os estudos já vistos na literatura sobre o uso majoritário da capacitância para língua eletrônica impedimétrica (TAYLOR, MACDONALD, 1987, RIUL JÚNIOR. *et al.*, 2003, ALCANTARA *et al.*, 2013).

Cada amostra de água residual é um sistema muito complexo com uma diversidade de moléculas e interações entre elas. Isso pode ser notado na Figura 24, que mostra a relação da capacitância com os valores de DQO. Para nenhum dos sensores observa-se qualquer relação direta entre essas variáveis.

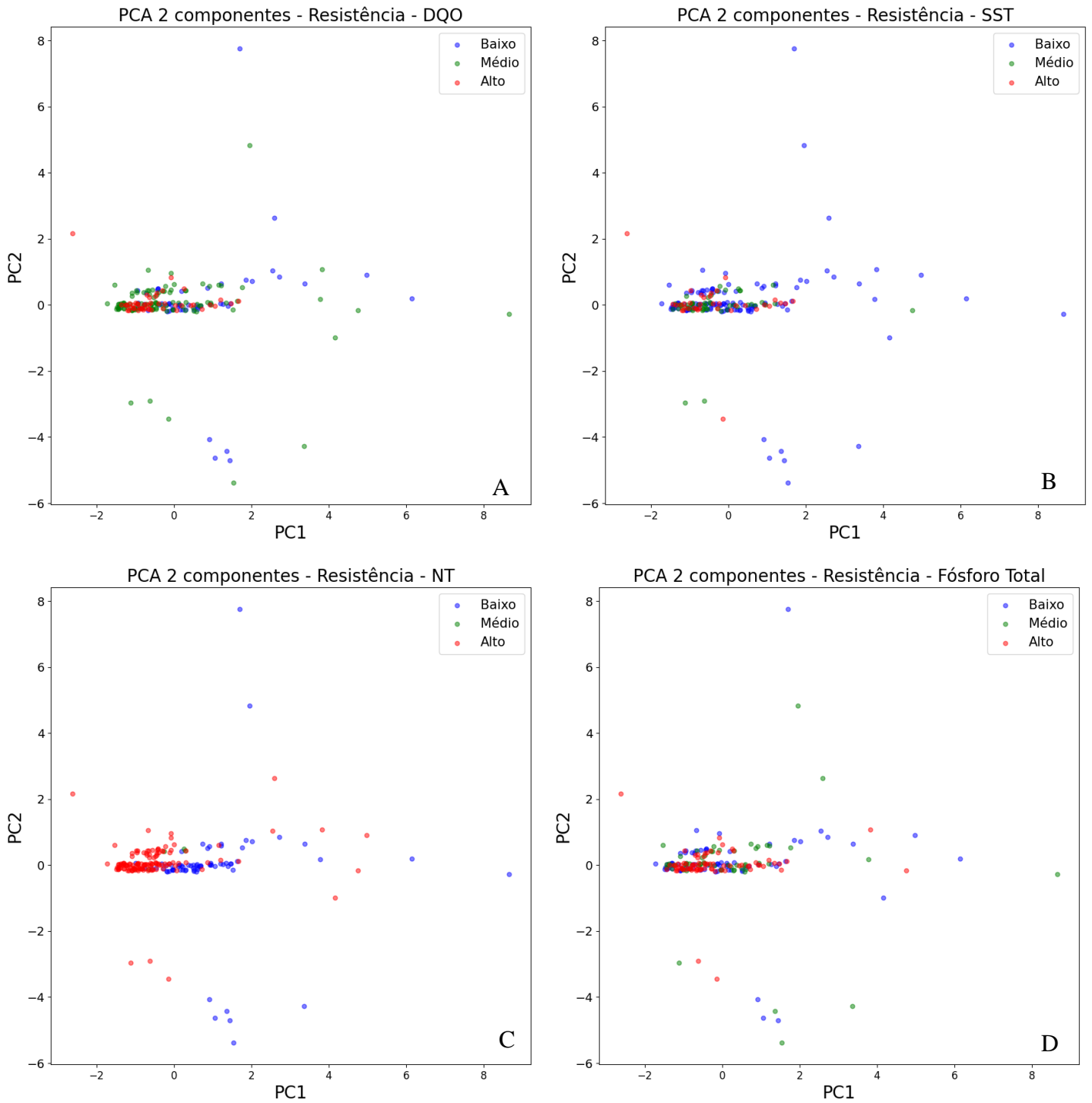
No entanto, a língua eletrônica permitiu obter resultados satisfatórios, mesmo usando sensores inespecíficos. Num futuro, os estudos podem abranger outros parâmetros ou pode-se diminuir a quantidade de parâmetros avaliados.

**Figura 21:** PCA 2 componentes para a capacitância dos quatro indicadores analisados: (a) DQO, (b) SST, (c) NT e (d) PHOSP.



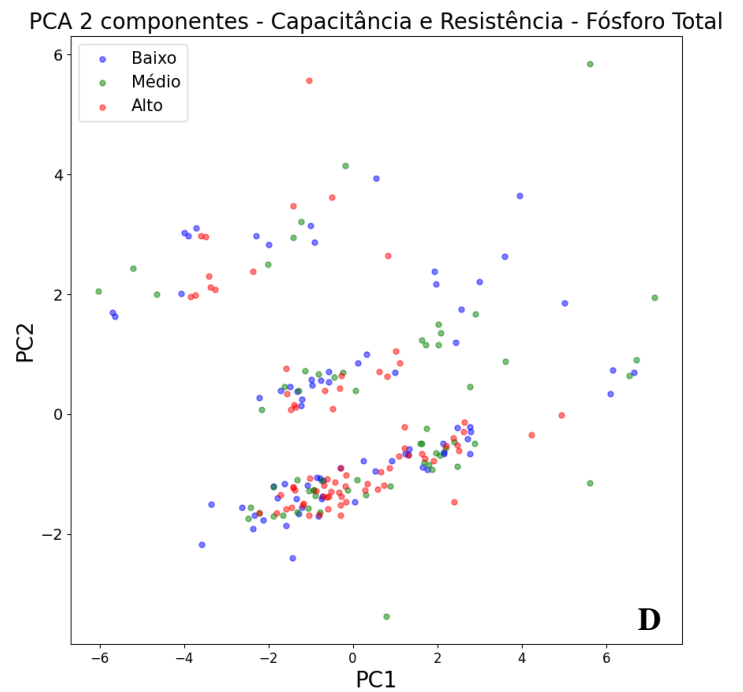
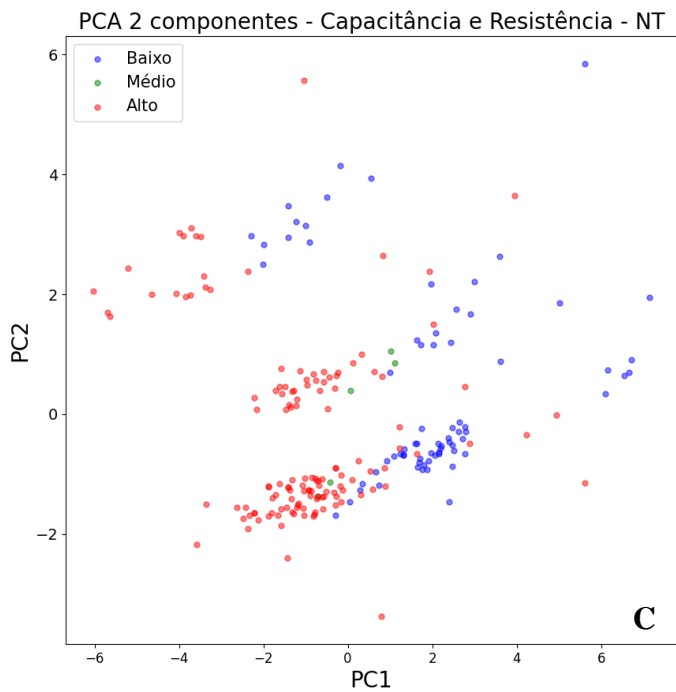
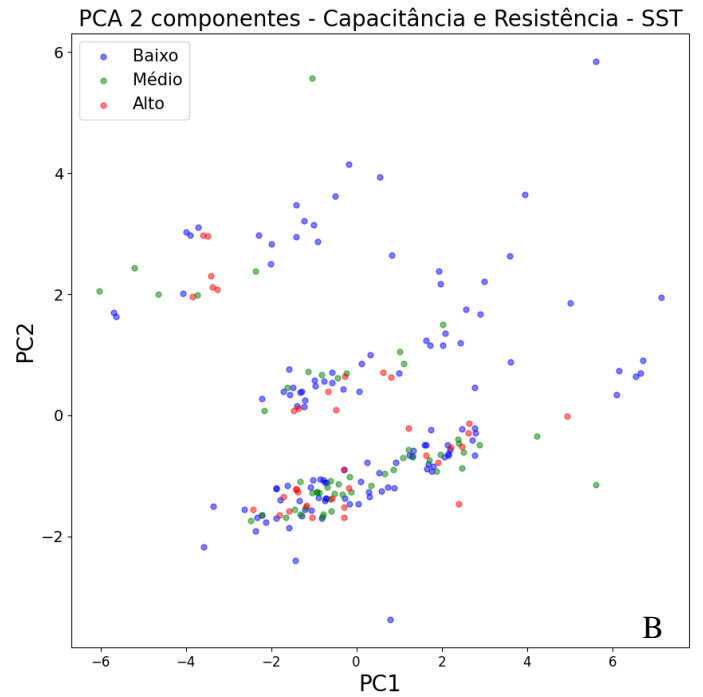
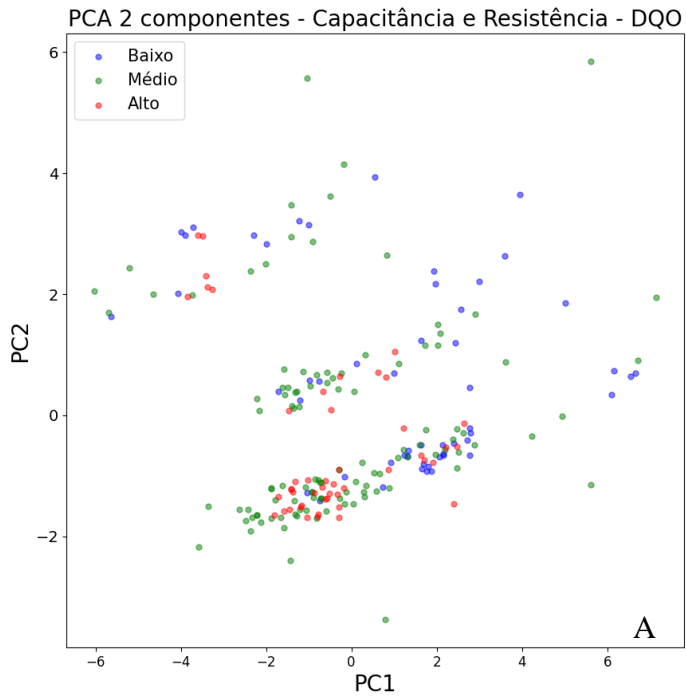
Fonte: A autora (2024).

**Figura 22:** PCA 2 componentes para a resistência dos quatro indicadores analisados: (a) DQO, (b) SST, (c) NT e (d) PHOSP.



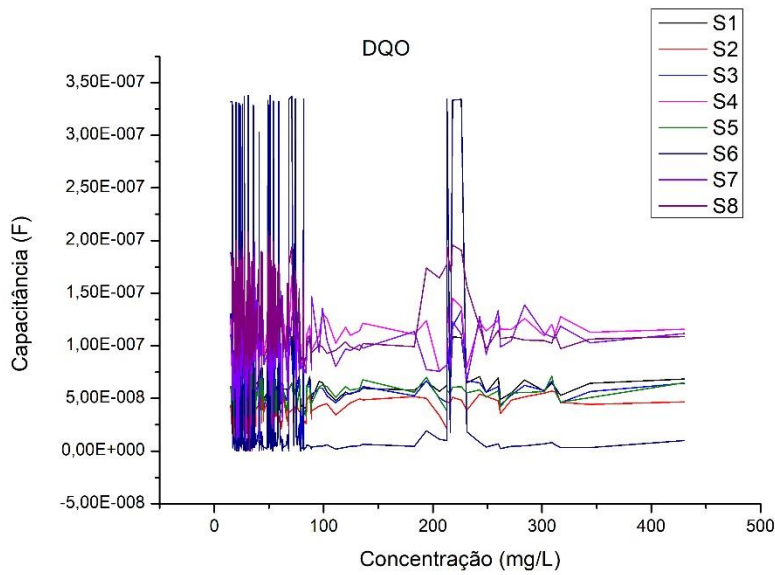
Fonte: A autora (2024).

**Figura 23:** PCA 2 componentes para a capacitância e resistência dos quatro indicadores analisados: (a) DQO, (b) SST, (c) NT e (d) PHOSP.



Fonte: A autora (2024).

**Figura 24:** Capacitância pela concentração de DQO.



Fonte: A autora (2024).

## 5.6 Estudo das Estações Norte e Sul

Como visto anteriormente, pelo estudo com algoritmo de aprendizado sem supervisão, a língua eletrônica foi capaz de encontrar alguns padrões e determinar alguns grupos de amostras que pertencem a mesma estação de tratamento. O objetivo agora é implementar o algoritmo supervisionado para as estações.

Foram escolhidas as estações Norte e Sul pela similaridade no processo de tratamento e o intuito é verificar se a língua distingue processos ou ETEs. A regressão logística com PCA foi implementada somente para as estações Norte e Sul final, recebendo cada uma dela o *score* 0 e 1 para Sul e Norte respectivamente conforme Quadro 7.

**Quadro 7:** Score estações e quantidades.

ESTAÇÃO	SCORE	QUANTIDADE
Sul	0	19
Norte	1	21

Fonte: A autora (2024).

Três tabelas contendo as saídas desejadas e as leituras das capacitâncias e resistência de cada sensor para a amostra do dia foram elaboradas para a implementação do PCA com a

regressão logística. Também foram usadas as informações anteriores sobre os PCs para cada tabela, além do autoescalamento dos dados. O Quadro 7 também mostra como as quantidades são mais homogêneas.

Com a redução das estações para apenas duas, a quantidade de amostras foi reduzida para 40. Para o treinamento da máquina, 90% dos dados foram usados para treinos e 10% para teste, além de o máximo de 800 iterações para o treino.

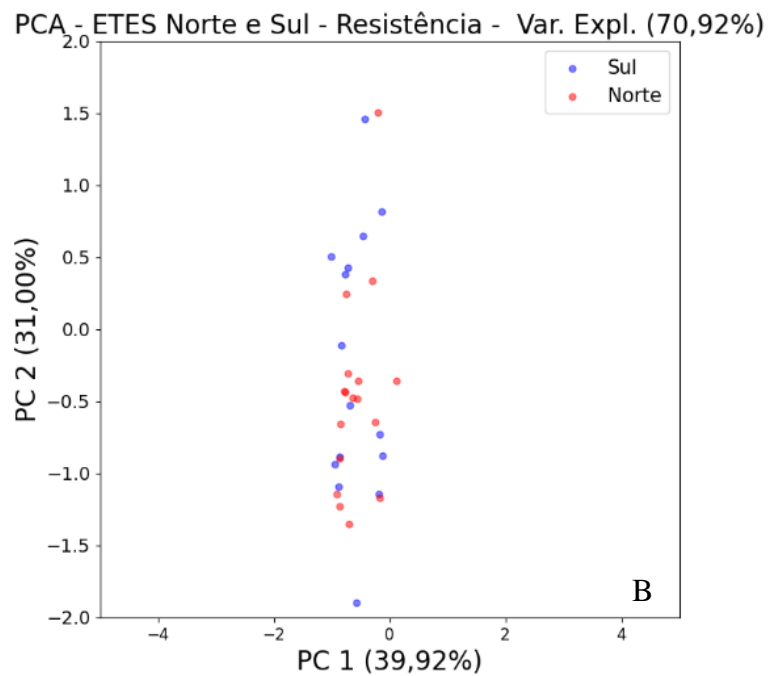
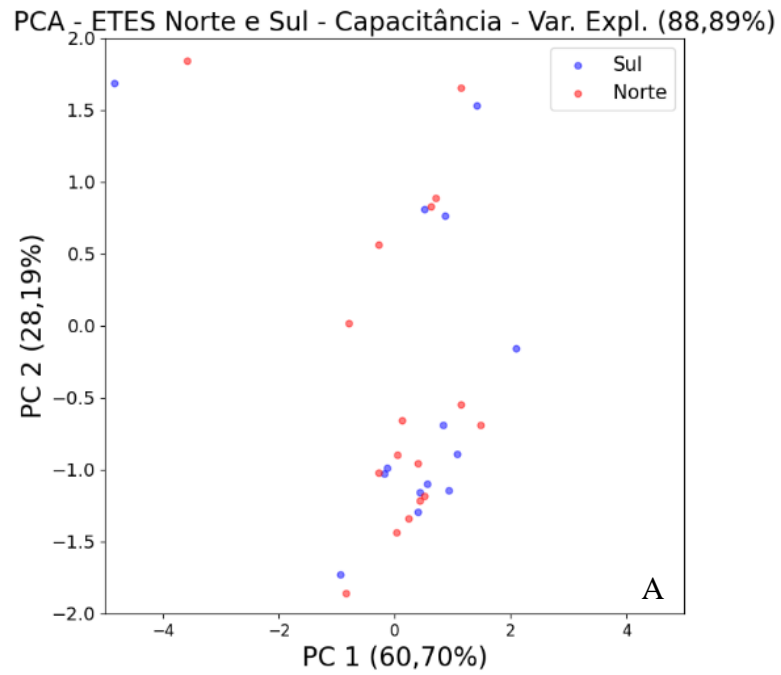
**Quadro 8:** Porcentagem de acerto da língua eletrônica para as estações Sul e Norte.

Estação	Porcentagem de Acerto (%)		
	Capacitância	Resistência	Capacitância e Resistência
SUL	50,00	50,00	50,00
NORTE	50,00	50,00	50,00

Fonte: A autora (2024).

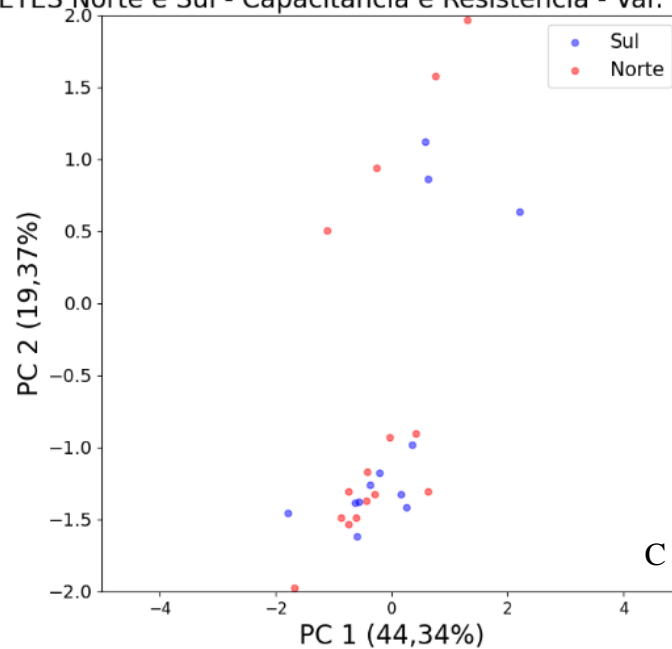
Todos os resultados atingiram 50% de acerto, não sendo efetivo para nenhum tipo de análise. Esse resultado mostra que a língua eletrônica detecta, preferencialmente, em padrão determinado pelo tipo de processo de tratamento de água que é feito por cada ETE. Os gráficos de PCA para dois componentes principais também foram feitos e mostrados na Figura 25.

**Figura 25:** PCA dois componentes para as estações Norte e Sul. (a) Capacitância. (b) Resistência. (c) Capacitância e Resistência.





PCA - ETES Norte e Sul - Capacitância e Resistência - Var. Expl. (63,71%)



Fonte: A autora (2024).

A variância explicada da capacitância (Figura 25a) para dois componentes explica quase 90% das informações, contra 70,92% da resistência (Figura 25b) e 63,71% da capacitância e resistência (Figura 25c). E esse foi o caso em que, pela primeira vez, a resistência sozinha é melhor do que junto com a capacitância. Entretanto, ressaltando mais uma vez que a capacitância explica mais do sistema com menos componentes principais.

Nos três gráficos, há a mistura dos grupos Norte e Sul, sem separação clara. Pode-se notar que essa mistura também explica o porquê o aprendizado de máquina não foi eficiente e confirmando que a língua eletrônica detecta o processo de tratamento realizado na estação. Ela é assertiva em observar que o processo de tratamento feito nessas duas estações é eficiente, uma vez que a língua eletrônica não consegue distinguir de qual estação pertence o produto final de tratamento. Isso também foi observado durante todo o trabalho, onde as estações Norte e Sul costumam ficar próximas. O mesmo acontece para as estações de Riacho Fundo. Dessa forma, a língua eletrônica mostrou um bom desempenho para a identificação do processo mostrando seu positivo potencial para classificação e identificação de amostras, aliado com sua versatilidade, praticidade, baixo custo e rápida resposta.

## 6 CONCLUSÕES

Em linhas gerais, a língua eletrônica desenvolvida nesse trabalho provou ser um sistema multissensorial promissor para auxiliar o controle de qualidade de tratamento de água residual. A resposta dos sensores foi estável ao longo do tempo, assegurando confiabilidade dos resultados obtidos. Os sensores também se mostraram duráveis pois, as amostras não passaram por nenhum tipo de procedimento prévio e as análises foram todas realizadas a temperatura ambiente ( $25,0 \pm 0,5$  °C). A quantidade de amostras analisadas tornou possível o estudo e uma compreensão do processo, mas ainda insuficiente para um aprendizado de máquina mais efetivo.

Em particular, observou-se que a língua eletrônica foi capaz de perceber que algumas estações têm comportamentos muito distintos, como Brazlândia e Norte, Brazlândia e Sul. Enquanto outras tem comportamento muito parecidos, como Norte e Sul. Com a análise dos dados e implementação de um modelo classificatório não supervisionado, a língua eletrônica desenvolvida consegue distinguir as estações de tratamento pelo tipo de processo pelo qual passa a água residual e não pela localização da estação.

Observou-se ainda que os dados de capacitância são melhores para a discriminação e classificação das amostras pois, após análise estatística, foram aqueles que apresentaram maiores valores de variância explicada, exigindo um número menor de componentes principais. O efeito de dupla camada para frequência de 1kHz e aumento de carga superficial em filmes com nanopartículas aumenta consideravelmente a capacitância, levando-a a ser um parâmetro melhor que a resistência, algo que foi percebido durante todo o trabalho.

A implementação do aprendizado de máquina supervisionado mostrou resultados interessantes, permitindo a classificação das amostras por meio de alguns parâmetros físico-químicos. Apesar do número relativamente pequeno de amostras, ressaltando ainda a complexidade de amostras desse tipo, o índice de acerto foi maior que 50% para a maioria dos parâmetros analisados e, em alguns casos, como para o nitrogênio total, o índice de acerto da língua eletrônica foi superior a 70%.

Em resumo, conclui-se que a língua eletrônica impedimétricas com sensores de nanomateriais e algoritmo de aprendizado de máquina tem um futuro promissor como ferramenta para auxiliar o controle de qualidade de águas residuais, permitindo a análises com diferentes perspectivas. Ainda oferece a possibilidade de monitoramento contínuo do processo de tratamento, com uma resposta mais rápida e de custo menor que aquelas obtidas por métodos

convencionais de análise. A inespecificidade dos sensores a torna um equipamento versátil capaz de analisar qualquer parâmetro desejado. O conjunto de dados criados para esse sistema multissensorial é o primeiro desse tipo e equipamentos futuros poderão se beneficiar desse conhecimento. Entretanto, fica claro que é preciso um aumento substancial no número de amostras avaliadas para que o algoritmo seja treinado com um universo mais abrangente de características físico-químicas.

## 7 PERSPECTIVAS

- I) Análise de mais amostras;
- II) Implementação de outros tipos de aprendizado de máquina, como redes neurais artificiais (RNAs);
- III) Explorar mais profundamente um indicador com uma variedade maior de concentrações;
- IV) Fazer análise quantitativa da concentração do indicador;
- V) Explorar novas composições de sensores com diferentes nanomateriais almejando um analito específico;
- VI) Desenvolver outra língua eletrônica e observar se os resultados se mantêm consistentes.

## REFERÊNCIAS

AGÊNCIA NACIONAL DE ÁGUAS E SANEAMENTO BÁSICO (ANA). Saneamento. **Gov.br. Ministério do Meio Ambiente e Mudança do Clima**. [2023a]. Disponível em: <https://www.ana.gov.br/saneamento/>. Acesso em: 28 jun. 2023.

AGÊNCIA NACIONAL DE ÁGUAS E SANEAMENTO BÁSICO (ANA). Usos da água. **Gov.br. Ministério do Meio Ambiente e Mudança do Clima**. [2023b]. Disponível em: <https://www.gov.br/ana/pt-br/assuntos/gestao-das-aguas/usos-da-agua>. Acesso em: 18 maio 2023.

AITKENHEAD, M. J. J. *et al.* Neural network integration of field observations for soil endocrine disruptor characterisation, **Sci. Total Environ.** [S. l.], n. 468–469, p. 240–248, 2014. Disponível em: <https://doi.org/10.1016/j.scitotenv.2013.08.007>. Acesso em: 28 jun. 2023.

ALCANTARA, G; PATERNO, L. G.; FONSECA, F. J.; SILVA, M. A. P.; MORAIS, P. C.; SOLER, M.A.G.. Layer-by-Layer Assembled Cobalt Ferrite Nanoparticles for Chemical Sensing, **Journal of Nanofluids**. Vol. 2, pp. 175–183. ISSN: 2169432X, 141. 2013

ALONSO, G.A. *et al.* Artificial neural network implementation in single low-cost chip for the detection of insecticides by modeling of screen-printed enzymatic sensors response. **Comput. Electron. Agric.** [S. l.], v. 74, p. 223–229, 2010. Disponível em: <https://doi.org/10.1016/j.compag.2010.08.003>. Acesso em: 28 jun. 2023.

ASTREE, taste sensors electronic tongue. **Alpha MOS**, Paris, 2023. Disponível em: <https://www.alpha-mos.com/taste-analysis-astree-electronic-tongue>. Acesso em: 11 jul. 2023.

BONANNI, A., DEL VALLE, M. Use of nanomaterials for impedimetric DNA sensors: a review. **Analytica Chimica Acta**, 678(1), 7–17. 2010. DOI:10.1016/j.aca.2010.08.022

BRASIL. Conselho Nacional do Meio Ambiente. Resolução Conama n° 357, de 17 de março de 2005. **Diário Oficial da União**, Brasília, 18 mar. 2005.

BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. **Vigilância e controle da qualidade da água para consumo humano/ Ministério da Saúde, Secretaria de Vigilância em Saúde**. Brasília: Ministério da Saúde, 2006. 212 p. (Série B. Textos Básicos de Saúde). Disponível em: [https://bvsmms.saude.gov.br/bvs/publicacoes/vigilancia\\_controle\\_qualidade\\_agua.pdf](https://bvsmms.saude.gov.br/bvs/publicacoes/vigilancia_controle_qualidade_agua.pdf). Acesso em: 15 jan. 2024.

BRASIL. Ministério da Saúde. **Portaria n° 2914, de 12 de dezembro de 2011**. Brasília: Ministério da Saúde, 2011. Disponível em: [https://bvsmms.saude.gov.br/bvs/saudelegis/gm/2011/prt2914\\_12\\_12\\_2011.html](https://bvsmms.saude.gov.br/bvs/saudelegis/gm/2011/prt2914_12_12_2011.html). Acesso em: 29 jun. 2023.

BROSEL-OLIU, S., ABRAMOVA, N., URIA, N., BRATOV, A. Impedimetric transducers based on interdigitated electrode arrays for bacterial detection – A review. **Analytica Chimica Acta**, 1088, 1–19. 2019. Disponível em: <https://doi.org/10.1016/j.aca.2019.09.026>.

CAESB. Análises e Resultados - Sistema de Esgotamento. **CAESB - Companhia de Saneamento Ambiental do Distrito Federal**. [2023]. Disponível em: <https://www.caesb.df.gov.br/esgoto/sistemas-de-esgotamento.html>. Acesso em: 18 maio 2023.

CHRISTINELLI, W. A. *et al.* Two-dimensional MoS<sub>2</sub>-based impedimetric electronic tongue for the discrimination of endocrine disrupting chemicals using machine learning. **Sensors and Actuators: B. Chemical**, [S. I.], v. 336, jun. 2021. Disponível em: <https://doi.org/10.1016/j.snb.2021.129696>. Acesso em: 29 jun. 2023.

CIOSEK, P. *et al.* Milk classification by means of an electronic tongue and Support Vector Machine neural network. **Meas. Sci. Technol.** Bristol, v. 17, 2006. Disponível em: <https://doi.org/10.1088/0957-0233/17/6/014>. Acesso em: 29 jun. 2023.

CODEPLAN. **Um Panorama das Águas no Distrito Federal**. Brasília: Governo do Distrito Federal, 2020.

CORREA, D. S., MEDEIROS, E. S., OLIVEIRA, J. E., PATERNO, L. G., MATTOSO, L. H. C. Nanostructured Conjugated Polymers in Chemical Sensors: Synthesis, Properties and Applications. **Journal of Nanoscience and Nanotechnology**. Vol. 14, 1–19, 2014

COSTA, I.A.; GROSS, M.A.; ALVES, E.D.O.; FONSECA, F.J.; PATERNO, L.G. An Impedimetric E-Tongue Based on CeO<sub>2</sub> -Graphene Oxide Chemical Sensors for Detection of Glyphosate and Its Potential Interferents. **J. Electroanal. Chem.** [S. I.], 922, 116719, 2022.

FIGUEIREDO FILHO, D. B.; SILVA JÚNIOR, J. A. Desvendando os Mistérios do Coeficiente de Correlação de Pearson (r). **Revista Política Hoje**, Recife, v. 18, n. 1, 2009.

GÉRON, A. **Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems**. 1. ed. Paris: O'Reilly, 2017.

GROSS, M. A. *et al.* Reduced graphene oxide multilayers for gas and liquid phases chemical sensing. **RSC Advances**, Cambridge, v. 4, n. 34, p. 17917-17924, 2014. Disponível em: <http://hdl.handle.net/11449/111796>. Acesso em: 29 jun. 2023.

HORNIK, K. *et al.* Open-source machine learning: R meets Weka. **Comput. Stat.** [S. I.], v. 24, p. 225–232, 2009. Disponível em: <https://doi.org/10.1007/s00180-008-0119-7>.

HOSCH, William L. "Machine Learning". **Encyclopedia Britannica**, 21 jun. 2023. Disponível em: <https://www.britannica.com/technology/machine-learning>. Acesso em: 29 jun. 2023.

HULANICKI, A.; GLAB, S.; INGMAN, F. Chemical sensors definitions and classification. **Pure & Appl. Chem.**, Londres, v. 63, n. 9, p. 1247-1250, 1991.

IBM. O que é regressão logística? *In*: IBM. [S. l.: s. n., 2024]. Disponível em: <https://www.ibm.com/br-pt/topics/logistic-regression>. Acesso em: 29 jun. 2023.

JOLLIFFE, I. T. **Principal Component Analysis**. 2. ed. Nova Iorque: Springer, 2002.

KLEINBAUM, D. G.; KLEIN, M. **Logistic regression: a self-learning text**. 3. ed. Nova Iorque: Springer, 2010. Disponível em: 10.1007/978-1-4419-1742-3. Acesso em: 29 jun. 2023.

LOWINSOHN, D.; BERTOTTI, M. Sensores eletroquímicos: considerações sobre mecanismos de funcionamento e aplicações no monitoramento de espécies químicas em ambientes microscópicos. **Quím. Nova**, São Paulo, v. 29, n. 6, dez. 2006. Disponível em: <https://doi.org/10.1590/S0100-40422006000600029>. Acesso em: 29 jun. 2023.

LVOVICH, V. F. **Impedance Spectroscopy: Applications to Electrochemical and Dielectric Phenomena**. 1. ed. [S. l.]: Wiley, 2015.

MACHADO, J. C. *et al.* Efficient Praziquantel Encapsulation into Polymer Microcapsules and Taste Masking Evaluation Using an Electronic Tongue. **Bulletin of the Chemical Society of Japan**, Tóquio, v. 91, n. 6, p. 865–874, 2018.

MAMISHEV, A. V. *et al.* Interdigital sensors and transducers. **Proceedings of the IEEE**, [S. l.], v. 92, n. 5, p. 808–845, 2004.

MEDEIROS, E. S., OLIVEIRA, J. E., CONSOLIN-FILHO, N., PATERNO, L. G., MATTOSO, L. H. C. Uso de Polímeros Condutores em Sensores. Parte 1: Introdução aos Polímeros Condutores. **Revista Eletrônica de Materiais e Processos**, v.7.2, 62 – 77. ISSN 1809-8797, 2012a.

MEDEIROS, E. S., OLIVEIRA, J. E., CONSOLIN-FILHO, N., PATERNO, L. G., MATTOSO, L. H. C. Uso de Polímeros Condutores em Sensores. Parte 2: Aplicações em Sensores. **Revista Eletrônica de Materiais e Processos**, v.7.3, 144–158. ISSN 1809-8797, 2012b

MISRA, S. K. *et al.* Comparative study using spheres, rods and spindle-shaped nanoplatelets on dispersion stability, dissolution and toxicity of CuO nanomaterials. **Nanotoxicology**, [S. l.], v. 8, n. 4, p. 422-432, 2013. Disponível em: <https://doi.org/10.3109/17435390.2013.796017>. Acesso em: 29 jun. 2023.

MONITORAMENTO DA QUALIDADE DA ÁGUA DO RIO PIRAIÁ À MONTANTE DO TÚNEL DE TÓCOS, NA REGIÃO HIDROGRÁFICA II – GUANDU, EM CONFORMIDADE COM O PLANO ESTRATÉGICO DE RECURSOS HÍDRICOS DAS BACIAS HIDROGRÁFICAS DOS RIOS GUANDU, DA GUARDA EM GUANDU MIRIM. **Relatório Consolidado – Janeiro a Dezembro de 2016**. Relatório SP 627/2017 Março de 2017. Disponível em: [http://sigaceivap.org.br/publicacoesArquivos/guandu/arq\\_pubMidia\\_Processo\\_070-2014\\_P4.pdf](http://sigaceivap.org.br/publicacoesArquivos/guandu/arq_pubMidia_Processo_070-2014_P4.pdf). Acesso em: 15 jan. 2024.

NAÇÕES UNIDAS. Água. **Nações Unidas – Centro Regional de informação para a Europa Ocidental**. [2023]. Disponível em: <https://unric.org/pt/agua/>. Acesso em: 18 maio 2023.

NILSSON, J. W.; RIEDEL, S. A. **Circuitos Elétricos**. 10. ed. São Paulo: Pearson, 2015.

OLIVEIRA JÚNIOR, O. N.; OLIVEIRA, M. C. F. Sensing and Biosensing in the World of Autonomous Machines and Intelligent Systems. **Frontiers in Sensors**, Dublin, v. 2, p. 1-7, set. 2021. Disponível em: <https://doi.org/10.3389/fsens.2021.752754>. Acesso em: 29 jun. 2023.

OTTO, M. **Chemometrics: statistics and computer application in analytical chemistry**. [S. l.]: Weinheim by Wiley-VCH, 1999. v. 16. 314 p.

OU-YANG, Q. *et al.* Study on classification of soy sauce by electronic tongue technique combined with artificial neural network. **Journal Food Sci.** [S. l.], v. 76, p. S523–S527, 2011. Disponível em: <https://doi.org/10.1111/j.1750-3841.2011.02382.x>. Acesso em: 29 jun. 2023.

PATERNIO, L. G. *et al.* Fabrication and characterization of nanostructured conducting polymer films containing magnetic nanoparticles. **Thin Solid Films**, [S. l.], v. 517, p. 1753-1758, 2009.

PATERNIO, L. G.; SOLER, M. A. G. Layer-by-Layer Enabled Nanomaterials for Chemical Sensing and Energy Conversion. **JOM**, [S. l.], v. 65, p. 709-719, 2013. Disponível em: <https://doi.org/10.1007/s11837-013-0608-1>. Acesso em: 29 jun. 2023.

PAULOVICH, F. V. *et al.* Information visualization techniques for sensing and biosensing. **Analyst**, [S. l.], v. 136, n. 7, p. 1344-1350, 2011. Disponível em: <https://doi.org/10.1039/C0AN00822B>. Acesso em: 29 jun. 2023.

PEREGRINO, P. P. *et al.* Starch-Mediated Immobilization, Photochemical Reduction, and Gas Sensitivity of Graphene Oxide Films. **ACS Applied Materials and Interfaces**, [S. l.], v. 5, n. 10, p. 5001-5012, mar. 2020. Disponível em: <https://doi.org/10.1021/acsomega.9b03892>. Acesso em: 29 jun. 2023.

QUEIROZ, D. P., FLORENTINO, A. O., BRUNO, J. C., SILVA, J. H. D., Riul ,A., GIACOMETTI, J. A. The use of an e-tongue for discriminating ethanol/water mixtures and



determination of their water content. **Sensors and Actuators B: Chemical**. V. 230, p. 566-570. ISSN 0925-4005, 2016.

RIBEIRO, C. L. *et al.* Electrochemical oxidation of salicylic acid at ITO substrates modified with layer-by-layer films of carbon nanotubes and iron oxide nanoparticles. **Journal of Electroanalytical Chemistry**, [S. I.], v. 805, p. 53-59, 2017. Disponível em: <https://doi.org/10.1016/j.jelechem.2017.09.063>. Acesso em: 29 jun. 2023.

RIDHI, R., SINGH, S., SAINI, G. S. S., TRIPATHI, S. K. Comparison of interaction mechanisms of copper phthalocyanine and nickel phthalocyanine thin films with chemical vapours. **Journal of Physics and Chemistry of Solids**. V. 115, p. 119-126. 2018. <https://doi.org/10.1016/j.jpics.2017.10.046>

RIUL, A., *et al.* Artificial taste sensor: Efficient combination of sensors made from Langmuir-Blodgett films of conducting polymers and a ruthenium complex and self-assembled films of an azobenzene-containing polymer. **Langmuir**, v. 18, n. 1, p. 239–245, 2002.

RIUL, A., SOUSA, H. C., MALMEGRIM, R. R., SANTOS, D. S., CARVALHO, A. C. P. L. F., FONSECA, F. J., OLIVEIRA, O. N., MATTOSO, L. H. C. Wine classification by taste sensors made from ultra-thin films and using neural networks. **Sensors and Actuators B: Chemical**. V. 98, Issue 1, Pages 77-82. ISSN 0925-4005, 2004.

RIUL JÚNIOR., A.; SOTO, A.G.; MELLO, S.; BONE, S.; TAYLOR, D.; MATTOSO, L. An electronic tongue using polypyrrole and polyaniline. **Synth. Met.** 132, 109–116. 2003.

RIUL JÚNIOR, A. *et al.* The Use of Nanostructured Films in Sensing Applications. In: LVOVA, Larisa *et al.* (org.). **Multisensor Systems for Chemical Analysis**. Stanford: Jenny Stanford Publishing, v. 1, p. 303-331. 2013.

RIUL JÚNIOR, A. *et al.* Recent advances in electronic tongues. **Analyst**, [S. I.], v. 135, n. 10, p. 2481-2495, 2010. Disponível em: <https://doi.org/10.1039/C0AN00292E>. Acesso em: 29 jun. 2023.

ROZHIN, P.; MELCHIONNA, M.; FORNASIERO, P.; MARCHESAN, S. Nanostructured Ceria: Biomolecular Templates and (Bio)applications. **Nanomaterials**, 11, 2259. 2021. Disponível em: <https://doi.org/10.3390/nano11092259>

SHIMIZU, F. M.; BRAUNGER, M. L.; RIUL, A. Heavy Metal/Toxins detection using electronic tongues. **Chemosensors**, Basel, v. 7, n 3, p. 36, 2019. Disponível em: <https://doi.org/10.3390/chemosensors7030036>. Acesso em: 29 jun. 2023.

SHIMIZU, F. M., BRAUNGER, M. L., RIUL., A. Electronic Tongues Fundamentals and recent advances. **IOP Publishing**. [S. I.], nov. 2021. Disponível em: <https://dx.doi.org/10.1088/978-0-7503-3687-1>. Acesso em: 12 jul. 2023.

SHIMIZU, F. M. *et al.* Functionalization-free microfluidic electronic tongue based on a single response. **ACS Sensors**, [S. I.], v. 2, n. 7, p. 1027–1034, 2017.

STOJIC, N. *et al.* Prediction of toxicity and data exploratory analysis of estrogen-active endocrine disruptors using counter-propagation artificial neural networks. **J. Mol. Graph.**, [S. I.], v. 29, n. 3, p. 450–460, nov. 2010. Disponível em: <https://doi.org/10.1016/j.jmgm.2010.09.001>.

TAYLOR, D.M., MACDONALD, A.G. AC admittance of the metal/insulator/electrolyte interface. **J. Phys. D.: Appl. Phys.**, 20, pp. 1277–1283. 1987.

TOKO, K. Taste sensor with global selectivity. **Materials Science and Engineering: C**, [S. I.], v. 4, p. 69–82. 1996. Disponível em: [https://doi.org/10.1016/0928-4931\(96\)00134-8](https://doi.org/10.1016/0928-4931(96)00134-8). Acesso em: 11 jul. 2023.

VIDAL-MADJAR, C.; GUIOCHON, G. The Analysis of Complex Organic Compounds by Gas Solid Chromatography on Various Phthalocyanines. **Journal of Chromatographic Science**, V. 9, Issue 11, P. 664–672. 1971. Disponível em: <https://doi.org/10.1093/chromsci/9.11.664>

VLASOV, Y. *et al.* Nonspecific sensor arrays (“electronic tongue”) for chemical analysis of liquids. **Pure Appl. Chem.**, Berlim, v. 77, n. 11, p. 1965–1983, 2005. Disponível em: [10.1351/pac200577111965](https://doi.org/10.1351/pac200577111965). Acesso em: 29 jun. 2023.

VLASOV, Y, LEGIN, A. Non-selective chemical sensors in analytical chemistry: from “electronic nose” to “electronic tongue”. **Fresenius: Journal of Analytical Chemistry**, v. 361(3), p. 255–260, 1998.

VOLPATI, D. *et al.* Toward the Optimization of an e-Tongue System Using Information Visualization: A Case Study with Perylene Tetracarboxylic Derivative Films in the Sensing Units. **Langmuir**, Toronto, v. 28, n. 1, p. 1029–1040. 2012. Disponível em: [10.1021/la203641a](https://doi.org/10.1021/la203641a). Acesso em: 29 jun. 2023.

WARD Jr., J. H. Hierarchical Grouping to Optimize an Objective Function. **Journal of the American Statistical Association**, [S. I.], v. 58, n. 301, p. 236–244, 1963. Disponível em: [10.1080/01621459.1963.10500845](https://doi.org/10.1080/01621459.1963.10500845). Acesso em: 29 jun. 2023.

WINQUIST, F.; KRANTZ-RÜLCKER, C.; WIDE, P.; LUNDSTRÖM, I. Monitoring of freshness of milk by an electronic tongue on the basis of voltammetry. **Measurement Science and Technology**, Bristol, v. 9, n. 12, p. 1937–1946, 1998.

WINQUIST, F.; WIDE, P.; LUNDSTRÖM, I. An electronic tongue based on voltammetry. **Analytica Chimica Acta**, [S. I.], v. 357, n. 1–2, p. 21–31, 1997.

ZAGALO, Paulo M.; RIBEIRO, Paulo A.; RAPOSO, Maria. Effect of Applied Electrical Stimuli to Interdigitated Electrode Sensors While Detecting 17 $\alpha$ -Ethinylestradiol in Water Samples. **MDPI**, [S. l.], v. 10, n. 3, p. 114, 2022. Disponível em: <https://www.mdpi.com/2227-9040/10/3/114>. Acesso em: 29 jun. 2023.

ZHANG, F. *et al.* Cerium oxide nanoparticles: Size-selective formation and structure analysis. **Applied Physics Letters**, Nova Iorque, v. 80, n. 1, p. 127-129, 2002. Disponível em: <https://doi.org/10.1063/1.1430502>. Acesso em: 29 jun. 2023.

# ANEXOS

## ANEXO A - Tabela dos dados coletados.

Valores de capacitância para as médias. Valores em vermelho são as médias para substituir dados faltantes. A tabela só mostra uma parte dos dados uma vez que foram 212 amostras analisadas.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC
Amostra	ID	COD.	ESPEC.	DATA	ESPEC.	ESPEC.	ESPEC.	ESPEC.	ESPEC.	ESPEC.	ESPEC.	ESPEC.	ESPEC.	ESPEC.	ESPEC.	ESPEC.	ESPEC.	ESPEC.	ESPEC.	ESPEC.	ESPEC.	ESPEC.	ESPEC.	Soluções	Substâncias	Níveis de Amônia	Nitrogênio Total	Fósforo Total
1	Meichlor	SULFINO	SUL	07.11.22	5.05E-08	3.75E-08	6.95E-08	5.92E-08	8.85E-08	3.65E-08	4.79E-08	7.26E-08	1.34E-07	1.47E-11	1.23E-10	1.39E-10	1.68E-11	3.77E-11	8.70E-11	1.76E-10	1.14E-10	1.14E-10	166	17.9	2.38	0.8	57	0.89
2	Sul Final	SULFINO	SUL	07.11.22	5.05E-08	3.75E-08	6.95E-08	5.92E-08	8.85E-08	3.65E-08	4.79E-08	7.26E-08	1.34E-07	1.47E-11	1.23E-10	1.39E-10	1.68E-11	3.77E-11	8.70E-11	1.76E-10	1.14E-10	1.14E-10	177	20.6	2.95	1.0	62	1.04
3	Sul Biológico	SULFINO	SUL	07.11.22	5.05E-08	3.75E-08	6.95E-08	5.92E-08	8.85E-08	3.65E-08	4.79E-08	7.26E-08	1.34E-07	1.47E-11	1.23E-10	1.39E-10	1.68E-11	3.77E-11	8.70E-11	1.76E-10	1.14E-10	1.14E-10	188	23.3	3.22	1.1	67	1.17
4	4 AguasLinda	AGL001	SUL	08.11.22	5.48E-08	2.35E-08	9.75E-08	8.22E-08	6.52E-08	1.83E-08	6.62E-08	1.62E-07	1.88E-07	6.01E-11	2.65E-10	1.90E-09	4.07E-11	1.40E-10	5.74E-09	4.35E-10	2.59E-10	2.59E-10	68.04439	8.5	3.47828	22.9864969	35.61931707	2.015143069
5	5 Brezandria	BRZ001	SUL	08.11.22	8.1E-08	6.69E-08	8.9E-08	8.7E-11	7.8E-08	1.63E-08	1.2E-07	1.35E-07	1.86E-07	2.62E-10	3.27E-10	2.68E-10	2.39E-10	3.38E-10	1.20E-10	3.28E-10	2.92E-10	2.92E-10	32	12	0.12	0.22	54	0.24
6	6 Brezandria	BRZ001	SUL	08.11.22	8.1E-08	6.69E-08	8.9E-08	8.7E-11	7.8E-08	1.63E-08	1.2E-07	1.35E-07	1.86E-07	2.62E-10	3.27E-10	2.68E-10	2.39E-10	3.38E-10	1.20E-10	3.28E-10	2.92E-10	2.92E-10	168	1.999	0.12	0.22	54	0.24
7	Riacho_RII	RFR001	SUL	08.11.22	8E-08	4.98E-08	8.2E-08	9.8E-08	7.1E-08	2.35E-08	2.35E-08	8.8E-08	1.6E-07	7.14E-11	3.90E-10	2.89E-10	4.79E-10	2.17E-10	1.81E-10	3.54E-10	2.08E-10	2.08E-10	30	13.3	2.8	0.66	5.2	0.44
8	Riacho_RII	RFR001	SUL	08.11.22	8E-08	4.98E-08	8.2E-08	9.8E-08	7.1E-08	2.35E-08	2.35E-08	8.8E-08	1.6E-07	7.14E-11	3.90E-10	2.89E-10	4.79E-10	2.17E-10	1.81E-10	3.54E-10	2.08E-10	2.08E-10	17.7	16.5	3.75	1.19	49.3	0.58
9	Alagados	ALAO01	SUL	08.11.22	7.8E-08	3.31E-08	7.3E-08	1E-07	7.7E-08	5.2E-08	6.5E-08	1.8E-07	9.36E-11	9.36E-11	4.85E-10	3.90E-10	3.45E-10	3.15E-10	2.07E-10	1.71E-10	5.30E-10	5.30E-10	32	15.6	3.75	1.19	49.3	0.58
10	Alagados	ALAO01	SUL	08.11.22	7.8E-08	3.31E-08	7.3E-08	1E-07	7.7E-08	5.2E-08	6.5E-08	1.8E-07	9.36E-11	9.36E-11	4.85E-10	3.90E-10	3.45E-10	3.15E-10	2.07E-10	1.71E-10	5.30E-10	5.30E-10	32	15.6	3.75	1.19	49.3	0.58
11	10 Norte Biológico	NORBI001	NORTE	09.11.22	6.6E-08	2.75E-08	6.9E-08	7.1E-08	4.4E-08	1.83E-08	6.6E-08	1.6E-07	7.5	4.22E-11	7.11E-11	2.96E-10	1.77E-10	1.60E-10	5.96E-11	5.96E-11	5.96E-11	5.96E-11	75	28.7	0.14	4.4	10.4	5.92
12	11 Norte Biológico	NORBI001	NORTE	09.11.22	6.6E-08	2.75E-08	6.9E-08	7.1E-08	4.4E-08	1.83E-08	6.6E-08	1.6E-07	7.5	4.22E-11	7.11E-11	2.96E-10	1.77E-10	1.60E-10	5.96E-11	5.96E-11	5.96E-11	5.96E-11	191	7	0.25	4.92	7.7	6.65
13	12 Alagados	ALAO2	NORTE	16.11.22	7.8E-08	4.33E-08	7.9E-08	9E-08	6.7E-08	1.73E-08	7.9E-08	1.7E-07	1.34E-10	1.34E-10	4.17E-10	6.78E-10	2.04E-10	1.56E-10	4.86E-09	3.41E-10	3.46E-11	3.46E-11	55	20	1.46	4.6	35.61931707	0.55
14	13 Norte Biológico	NORBI002	NORTE	16.11.22	5.1E-08	3.07E-08	5.6E-08	6.5E-08	3.7E-08	1.39E-08	6.5E-08	1.5E-07	1.77E-10	1.77E-10	1.90E-10	4.22E-10	8.81E-10	3.95E-10	1.51E-10	6.31E-10	1.12E-10	1.12E-10	61	18.7	0.23	2.64	35.61931707	0.72
15	14 Norte Final	MEL002	SUL	21.11.22	1.9E-08	1.35E-08	1.4E-08	2.4E-08	3E-08	1.1E-08	3.2E-08	3.1E-11	6.37E-09	6.37E-09	4.44E-11	2.94E-11	1.01E-10	1.55E-11	8.17E-13	5.94E-13	6.72E-13	6.72E-13	32.1	6	0.33	2.85	35.61931707	0.26
16	15 Meichlor	MEL002	SUL	21.11.22	4.6E-08	1.07E-08	4.7E-08	6.1E-08	3.4E-08	1.08E-08	5.1E-08	1.5E-07	9.36E-11	9.36E-11	8.52E-11	1.32E-10	1.00E-10	1.10E-10	9.17E-11	1.94E-10	2.07E-10	2.07E-10	30	13.3	2.8	0.66	5.2	0.44
17	16 Sul Biológico	SULBI002	SUL	21.11.22	4.6E-08	1.07E-08	4.7E-08	6.1E-08	3.4E-08	1.08E-08	5.1E-08	1.5E-07	9.36E-11	9.36E-11	8.52E-11	1.32E-10	1.00E-10	1.10E-10	9.17E-11	1.94E-10	2.07E-10	2.07E-10	17.7	16.5	3.75	1.19	49.3	0.58
18	17 Norte Biológico	NORBI002	SUL	22.11.22	7.2E-08	1.59E-08	7.1E-08	1.2E-07	6.9E-08	1.74E-08	7.4E-08	1.8E-07	1.23E-10	1.23E-10	8.92E-10	1.98E-10	8.25E-10	2.41E-10	2.37E-10	4.63E-10	5.21E-11	5.21E-11	68.04439	12	4.47828	22.9864969	35.61931707	2.015143069
19	18 Brezandria	BRZ002	SUL	22.11.22	6.6E-08	4.62E-08	6.8E-08	9.8E-08	5.5E-08	1.73E-08	7.7E-08	1.7E-07	2.59E-11	2.59E-11	1.84E-10	1.02E-10	3.09E-10	9.45E-11	1.08E-10	8.21E-11	5.29E-11	5.29E-11	216	7.5	0.49	29.3	55	7.04
20	19 Brezandria	BRZ002	SUL	22.11.22	6.6E-08	4.62E-08	6.8E-08	9.8E-08	5.5E-08	1.73E-08	7.7E-08	1.7E-07	2.59E-11	2.59E-11	1.84E-10	1.02E-10	3.09E-10	9.45E-11	1.08E-10	8.21E-11	5.29E-11	5.29E-11	216	7.5	0.49	29.3	55	7.04
21	20 Riacho_RII	RFR002	SUL	22.11.22	5.8E-08	2.21E-08	6E-08	8.5E-08	6.4E-08	1.73E-08	8E-08	1.7E-07	3.05E-11	3.05E-11	6.08E-11	9.44E-11	3.80E-09	3.79E-11	5.27E-11	1.01E-10	5.99E-11	5.99E-11	24.8	4.4	0.499	40.4	45	0.099
22	21 Riacho_RII	RFR002	SUL	22.11.22	5.8E-08	2.21E-08	6E-08	8.5E-08	6.4E-08	1.73E-08	8E-08	1.7E-07	3.05E-11	3.05E-11	6.08E-11	9.44E-11	3.80E-09	3.79E-11	5.27E-11	1.01E-10	5.99E-11	5.99E-11	24.8	4.4	0.499	40.4	45	0.099
23	22 Riacho_RII	RFR002	SUL	22.11.22	5.8E-08	2.21E-08	6E-08	8.5E-08	6.4E-08	1.73E-08	8E-08	1.7E-07	3.05E-11	3.05E-11	6.08E-11	9.44E-11	3.80E-09	3.79E-11	5.27E-11	1.01E-10	5.99E-11	5.99E-11	24.8	4.4	0.499	40.4	45	0.099
24	23 Alagados	ALAO3	NORTE	23.11.22	5.9E-08	1.04E-08	6.8E-08	9.1E-08	5.7E-08	1.76E-08	7.8E-08	1.6E-07	2.89E-11	2.89E-11	1.10E-10	2.10E-10	1.40E-10	4.01E-11	5.61E-11	1.20E-10	3.99E-10	3.99E-10	40.8	8.8	0.499	38.6	44	0.099
25	24 Norte Biológico	NORBI003	NORTE	23.11.22	5.9E-08	1.04E-08	6.8E-08	9.1E-08	5.7E-08	1.76E-08	7.8E-08	1.6E-07	2.89E-11	2.89E-11	1.10E-10	2.10E-10	1.40E-10	4.01E-11	5.61E-11	1.20E-10	3.99E-10	3.99E-10	56	24.5	2.3	3.3	48	0.67
26	25 Norte Biológico	NORBI003	NORTE	23.11.22	5.9E-08	1.04E-08	6.8E-08	9.1E-08	5.7E-08	1.76E-08	7.8E-08	1.6E-07	2.89E-11	2.89E-11	1.10E-10	2.10E-10	1.40E-10	4.01E-11	5.61E-11	1.20E-10	3.99E-10	3.99E-10	67	20	1	7.78	11.7	1.8
27	26 Brezandria	BRZ003	SUL	29.11.22	8.7E-08	5.03E-08	9.7E-08	1.2E-07	6.7E-08	1.9E-08	8.1E-08	1.8E-07	1.66E-10	1.66E-10	5.10E-10	1.31E-10	4.46E-09	3.64E-10	1.18E-10	3.70E-10	7.75E-11	7.75E-11	32.4	13.5	5.6	1.0	76	2.015143069
28	27 Brezandria	BRZ003	SUL	29.11.22	8.7E-08	5.03E-08	9.7E-08	1.2E-07	6.7E-08	1.9E-08	8.1E-08	1.8E-07	1.66E-10	1.66E-10	5.10E-10	1.31E-10	4.46E-09	3.64E-10	1.18E-10	3.70E-10	7.75E-11	7.75E-11	194	94.5	0.49	33.4	49	7.64
29	28 Riacho_RII	RFR003	SUL	29.11.22	7.8E-08	3.29E-08	8.3E-08	1.1E-07	7.3E-08	2.14E-08	1E-07	1.8E-07	2.33E-11	2.33E-11	1.07E-10	4.58E-11	1.59E-10	8.92E-11	8.42E-11	1.20E-10	1.14E-10	1.14E-10	34.4	3.7	0.499	51.7	53	0.099
30	29 Riacho_RII	RFR003	SUL	29.11.22	7.8E-08	3.29E-08	8.3E-08	1.1E-07	7.3E-08	2.14E-08	1E-07	1.8E-07	2.33E-11	2.33E-11	1.07E-10	4.58E-11	1.59E-10	8.92E-11	8.42E-11	1.20E-10	1.14E-10	1.14E-10	26	5.9	0.499	47.9	51	0.15
31	30 Riacho_RII	RFR003	SUL	29.11.22	7.8E-08	3.29E-08	8.3E-08	1.1E-07	7.3E-08	2.14E-08	1E-07	1.8E-07	2.33E-11	2.33E-11	1.07E-10	4.58E-11	1.59E-10	8.92E-11	8.42E-11	1.20E-10	1.14E-10	1.14E-10	50	18	0.499	53.7	61	0.39
32	31 AguasLinda	AGL04	SUL	06.12.22	6.4E-08	4.4E-08	7.1E-08	9.4E-08	6.3E-08	1.58E-08	8.3E-08	1.8E-07	2.49E-11	2.49E-11	1.22E-10	3.90E-11	1.99E-11	1.04E-10	7.45E-11	1.22E-10	1.69E-10	1.69E-10	209	102	0.49	31.6	48	6.55
33	32 Brezandria	BRZ04	SUL	06.12.22	5.7E-08	3.98E-08	5E-08	7.9E-08	4.9E-08	1.12E-08	7.6E-08	1.6E-07	1.37E-10	1.37E-10	3.95E-10	3.79E-10	3.06E-10	1.45E-10	2.90E-10	2.01E-10	2.01E-10	31	12.6	0.99	31.8	48	6.13	
34	33 Riacho_RII	RFR04	SUL	06.12.22	5.4E-08	3.07E-08	5.1E-08	9.8E-08	6.7E-08	3.47E-08	6.4E-08	1.3E-07	2.83E-10	2.83E-10	4.85E-11	4.29E-11	3.88E-11	1.25E-10	3.19E-11	2.67E-10	2.67E-10	21.4	0.299	0.12	0.22	54	0.24	
35	34 Brezandria	BRZ04	SUL	06.12.22	5.4E-08	3.07E-08	5.1E-08	9.8E-08	6.7E-08	3.47E-08	6.4E-08	1.3E-07	2.83E-10	2.83E-10	4.85E-11	4.29E-11	3.88E-11	1.25E-10	3.19E-11	2.67E-10	2.67E-10	40.4	12.4	0.499	53.8	61	0.28	
36	35 Riacho_RII	RFR04	SUL	06.12.22	6.6E-08	4.37E-08	6.3E-08	1E-07	6.9E-08	1.71E-08	8.2E-08	1.8E-07	4.38E-10	4.38E-10	1.97E-10	1.18E-10	4.0											

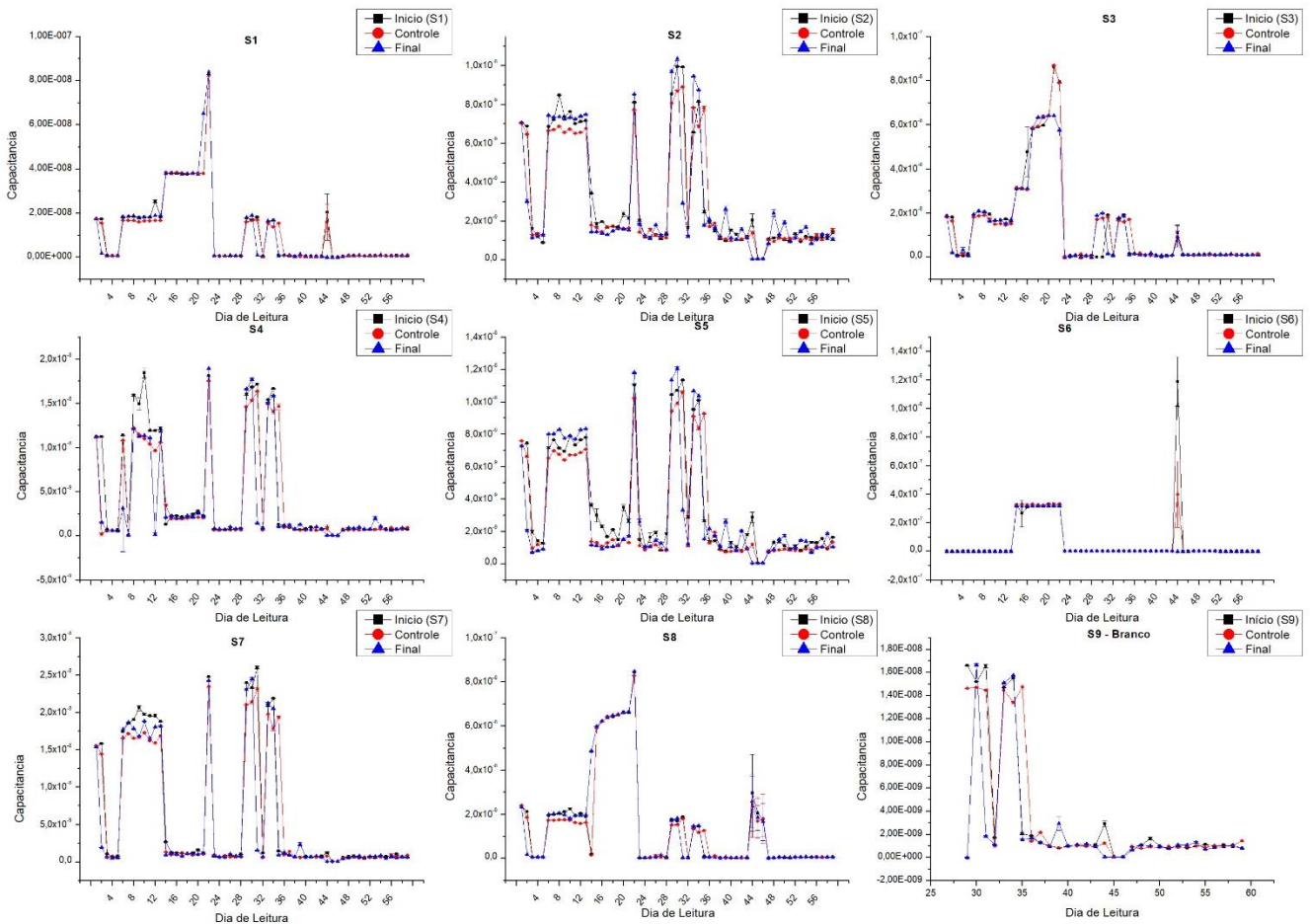
## ANEXO B - Valores de resistência e capacitância de cada sensor.

	Capacitância (F)	Resistência ( $\Omega$ )
S1	6,33E-11	463,2233
S2	5,45E-11	0
S3	5,41E-11	0
S4	7,73E-11	9415,433
S5	5,36E-11	0
S6	3,94E-10	42,08733
S7	5,43E-11	0
S8	6,61E-11	259,59

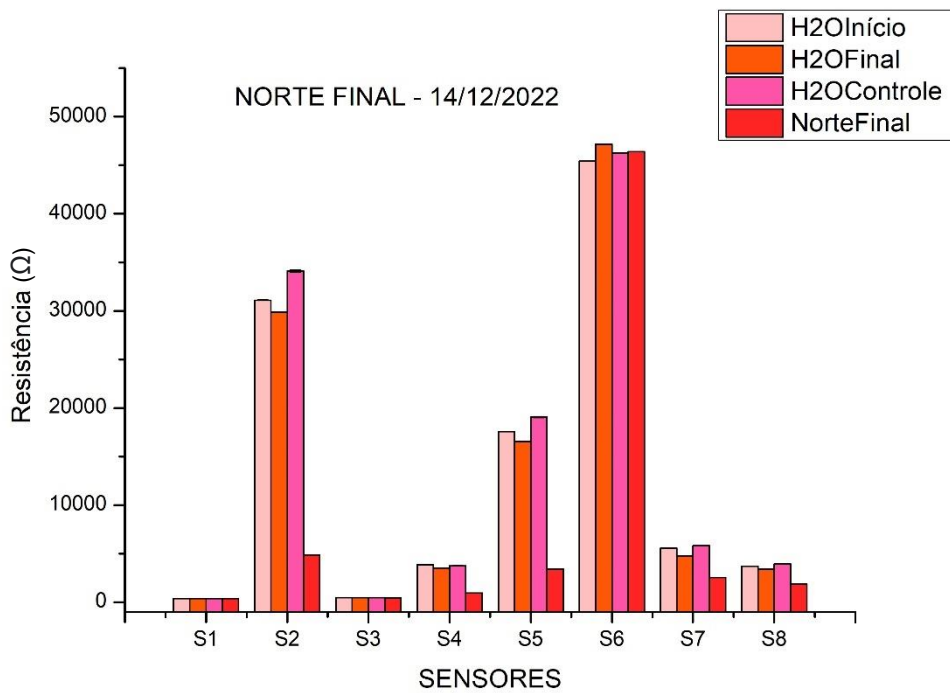
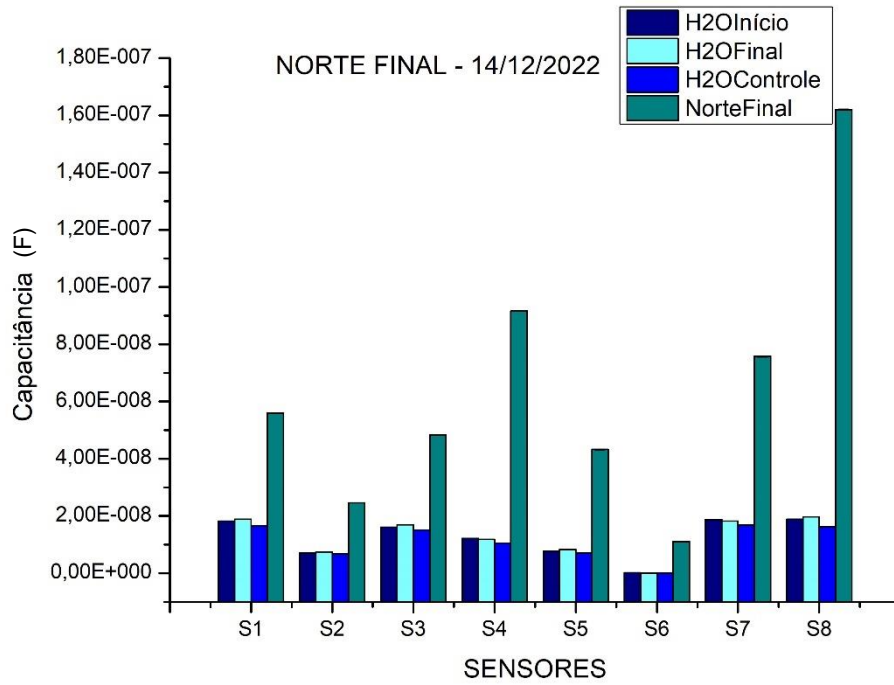
Obtido através da média de 30 leituras para cada, tanto capacitância quanto resistência.

## ANEXO C- Leituras de capacitância da água para verificação do reset.

Leituras de água inicial, de controle e final nos dias de leitura das amostras.



**ANEXO D - Gráfico de barras de capacitância e resistência para amostra da ETE Norte Final do dia 14/12/2022**



**ANEXO E - Quadros de variância (Capacitância, Resistência, Capacitância e Resistência).**

**Variância sensores e parâmetros CAPACITÂNCIA:**

	<b>Variância explicada (%)</b>	<b>Variância Acumulada (%)</b>
PC1	39,98	39,98
PC2	20,96	60,93
PC3	14,02	74,95
PC4	7,48	82,44
PC5	6,44	88,87
PC6	3,01	91,89
PC7	2,39	94,27
<b>PC8</b>	<b>1,91</b>	<b>96,19</b>
PC9	1,12	97,31
PC10	0,81	98,11
PC11	0,71	98,82
PC12	0,60	99,42
PC13	0,31	99,74
PC14	0,26	100

**Variância sensores e parâmetros RESISTÊNCIA:**

	<b>Variância explicada (%)</b>	<b>Variância Acumulada (%)</b>
PC1	24,80	24,80
PC2	14,63	39,43
PC3	9,54	48,97
PC4	8,90	57,87
PC5	7,94	65,80
PC6	6,82	72,63
PC7	6,34	78,97
PC8	5,49	84,46
PC9	4,97	89,43
PC10	4,12	93,55
<b>PC11</b>	<b>2,93</b>	<b>96,48</b>
PC12	2,08	98,56
PC13	0,81	99,36
PC14	0,63	100

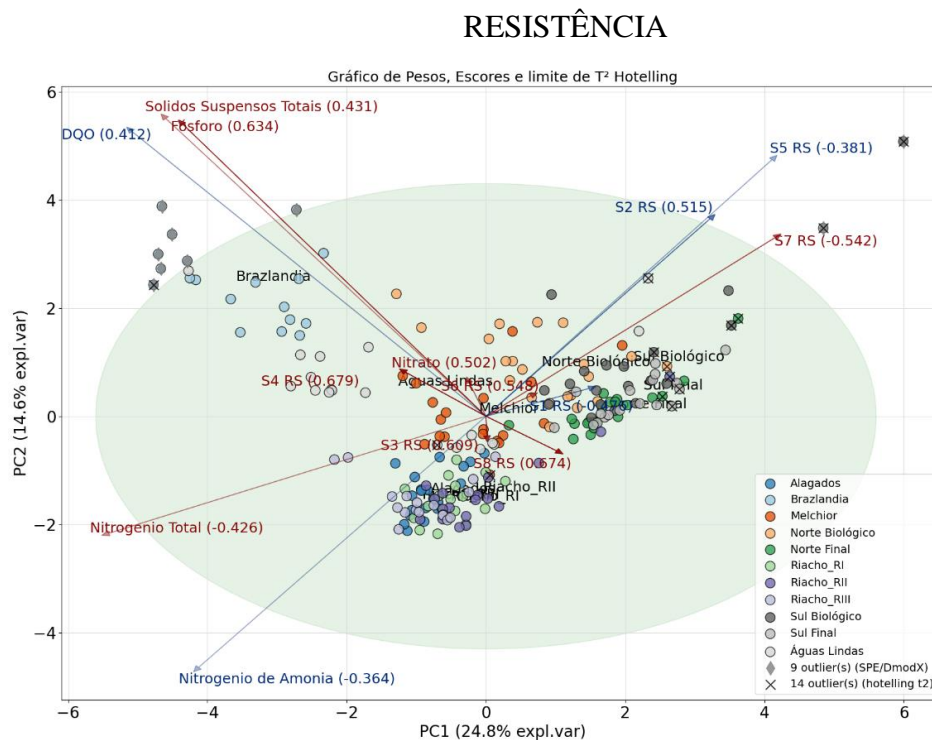
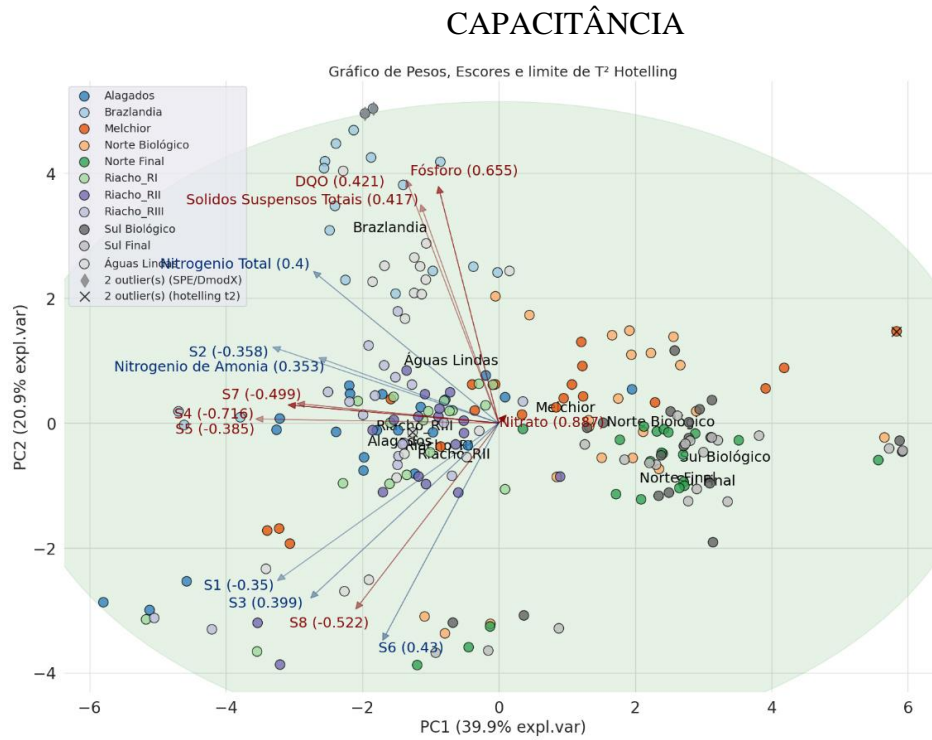
**Variância sensores e parâmetros CAPACITÂNCIA E RESISTÊNCIA:**

	<b>Variância explicada (%)</b>	<b>Variância Acumulada (%)</b>
PC1	29,97	29,97
PC2	14,36	44,33
PC3	9,89	54,22
PC4	7,95	62,17
PC5	6,37	68,54
PC6	5,09	73,63
PC7	4,54	78,17
PC8	4,41	82,57
PC9	3,87	86,45
PC10	3,32	89,76
PC11	2,42	92,18
PC12	1,89	94,08
<b>PC13</b>	<b>1,41</b>	<b>95,49</b>
PC14	1,18	96,67
PC15	0,95	97,61
PC16	0,58	98,19
PC17	0,47	98,66
PC18	0,43	99,09
PC19	0,35	99,44
PC20	0,32	99,75
PC21	0,14	99,89
PC22	0,10	100

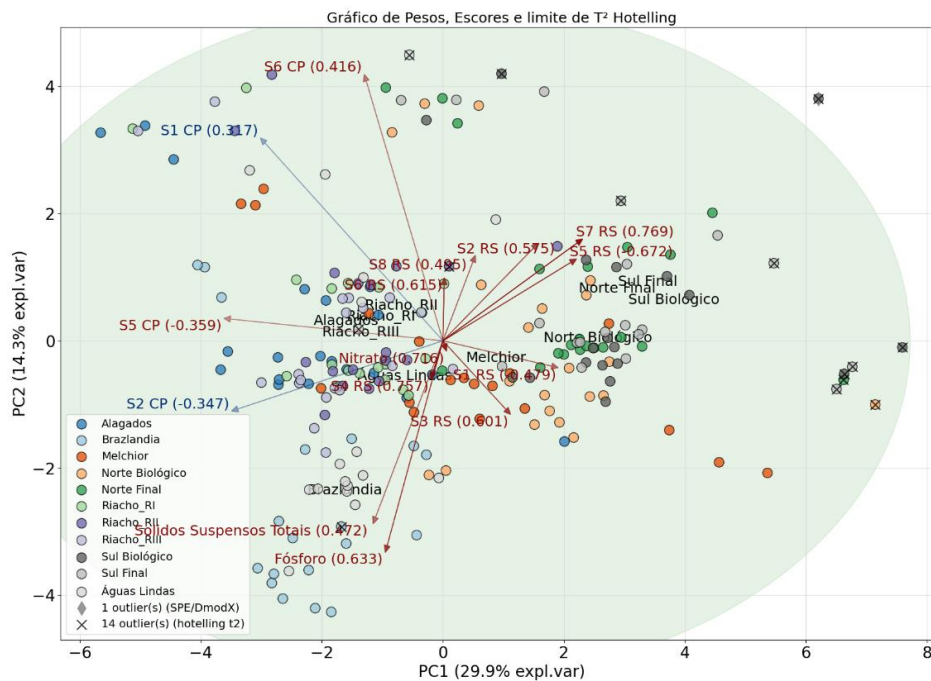


## ANEXO F - Gráficos 2D dos componentes principais dos sensores e parâmetros.

O gráfico 2D mostra a região que contempla 95% de confiança, elipse T<sup>2</sup> Hotelling, além dos vetores que mostram os pesos e as ETES.



# CAPACITÂNCIA E RESISTÊNCIA



**ANEXO G - Quadros de variância para os sensores**

**Variância sensores - CAPACITÂNCIA:**

	<b>Variância explicada (%)</b>	<b>Variância Acumulada (%)</b>
PC1	59,80	59,80
PC2	24,00	83,80
PC3	7,50	91,30
PC4	3,40	94,80
<b>PC5</b>	<b>2,80</b>	<b>97,60</b>
PC6	1,30	98,80
PC7	0,60	99,50
PC8	0,50	100

**Variância sensores - RESISTÊNCIA:**

	<b>Variância explicada (%)</b>	<b>Variância Acumulada (%)</b>
PC1	24,90	24,90
PC2	15,90	40,80
PC3	13,90	54,70
PC4	12,40	67,10
PC5	11,50	78,60
PC6	10,00	88,70
<b>PC7</b>	<b>7,50</b>	<b>96,20</b>
PC8	3,80	100

**Variância sensores – CAPACITÂNCIA E RESISTÊNCIA:**

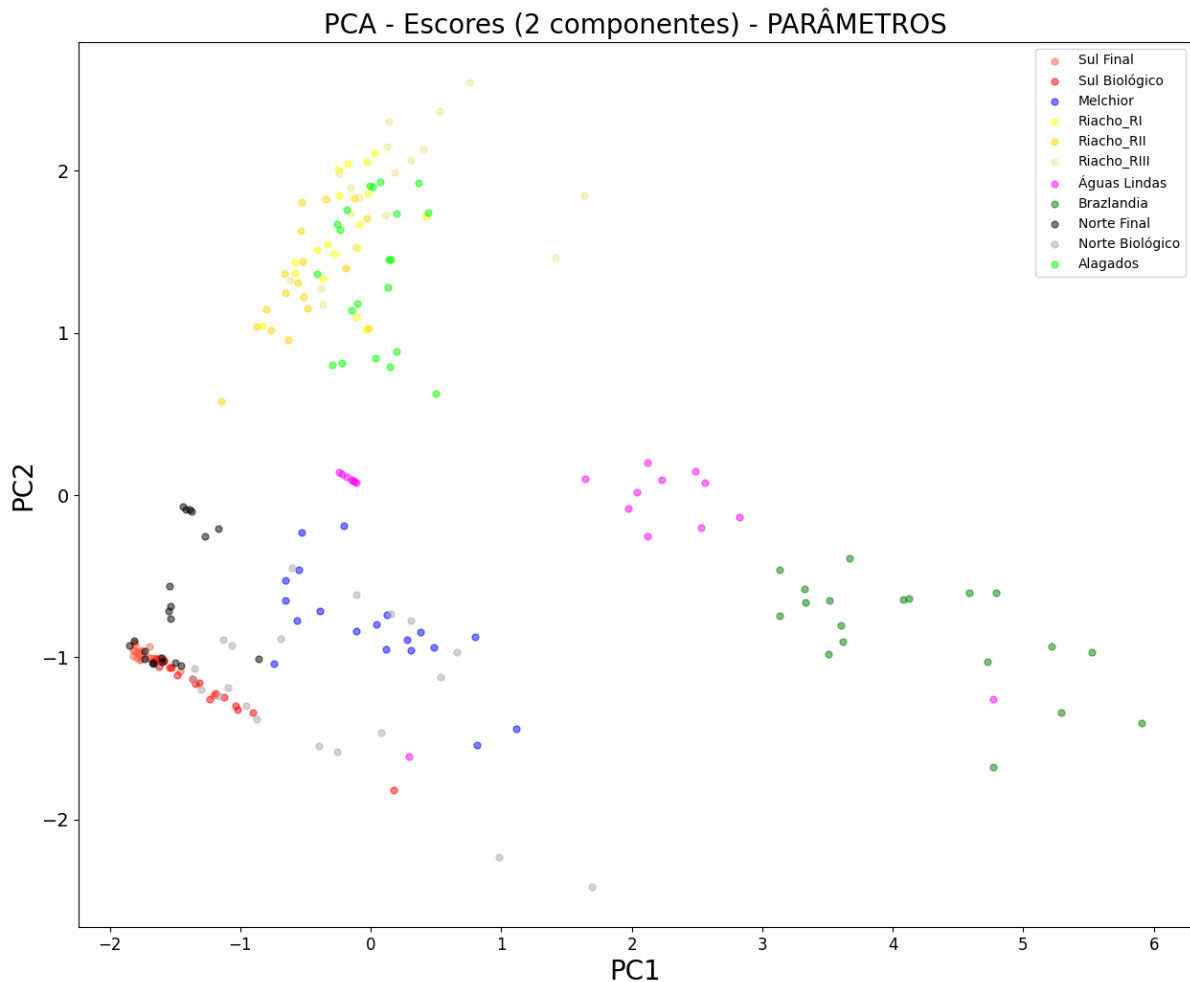
	<b>Variância explicada (%)</b>	<b>Variância Acumulada (%)</b>
PC1	35,20	35,20
PC2	16,30	51,50
PC3	10,40	61,90
PC4	8,50	70,40
PC5	6,20	76,70
PC6	5,70	82,40
PC7	5,50	87,90
PC8	3,40	91,30
PC9	3,00	94,30
<b>PC10</b>	<b>1,90</b>	<b>96,10</b>
PC11	1,50	97,60
PC12	1,00	98,60
PC13	0,60	99,20
PC14	0,50	99,60
PC15	0,20	99,80
PC16	0,20	100

## ANEXO H - Análise dos parâmetros físicos.

A tabela a seguir mostra a variância explicada apenas para os parâmetros físico-químicos, que totalizam 6.

### Variância Parâmetros:

	Variância explicada (%)	Variância Acumulada (%)
PC1	47,80	47,80
PC2	25,00	72,80
PC3	16,80	89,60
PC4	7,00	96,60
PC5	1,90	98,5
PC6	1,50	100




O gráfico de 2 componentes mostra grupos bem divididos das amostras em regiões de mesma ETE. Isso é possível graças a diversidade de análises realizadas no laboratório da estação.

**DECLARAÇÃO DE ORIGINALIDADE DE DISSERTAÇÃO DE MESTRADO OU TESE DE DOUTORADO**

Declaro que a presente dissertação/tese é original, elaborada especialmente para este fim, não tendo sido apresentada para obtenção de qualquer título e que identifico e cito devidamente todas as autoras e todos os autores que contribuíram para o trabalho, bem como as contribuições oriundas de outras publicações de minha autoria.

Declaro estar ciente de que a cópia ou o plágio podem gerar responsabilidade civil, criminal e disciplinar, consistindo em grave violação à ética acadêmica.

Brasília DF, 03 de julho de 2024  
Brasília, (dia) de (mês) de (ano).

Documento assinado digitalmente  
 CAMILA CALANDRINY ROCHA DA COSTA  
Data: 03/07/2024 11:07:40-0300  
Verifique em <https://validar.iti.gov.br>

Assinatura do/a discente: \_\_\_\_\_

Programa: Pós-graduação em Química

Nome completo: Camila Calandriny Rocha da Costa

Título do Trabalho: Desempenho de uma Língua Eletrônica Impedimétrica com Algoritmo de Aprendizado de Máquina na Análise de Águas Residuais

Nível:  Mestrado     Doutorado

Orientador/a: Professor Dr. Leonardo Giordano Paterno