

**UNIVERSIDADE FEDERAL DE ALAGOAS
UNIDADE ACADÊMICA CENTRO DE TECNOLOGIA
CURSO DE ENGENHARIA QUÍMICA**

RODRIGO LOPES DA CRUZ MAGALHÃES

**ANÁLISE DO POTENCIAL DE OXIRREDUÇÃO DO HIPOCLORITO DE
SÓDIO: DESENVOLVIMENTO E APLICAÇÃO DE UM MÉTODO DE
INFERÊNCIA**

Maceió
2025

RODRIGO LOPES DA CRUZ MAGALHÃES

**ANÁLISE DO POTENCIAL DE OXIRREDUÇÃO DO HIPOCLORITO DE
SÓDIO: DESENVOLVIMENTO E APLICAÇÃO DE UM METÓDO DE
INFERÊNCIA**

Trabalho de Conclusão de Curso apresentado
ao curso de Engenharia Química da
Universidade Federal de Alagoas como
Requisito para obtenção do título de Bacharel
em Engenharia Química.

Orientador: Prof. Dr. Wagner Roberto de
Oliveira Pimentel

Maceió
2025

Catálogo na Fonte
Universidade Federal de Alagoas
Biblioteca Central
Divisão de Tratamento Técnico

Bibliotecário: Jone Sidney A. de Oliveira – CRB-4 – 1485

M188a Magalhães, Rodrigo Lopes da Cruz.

Análise do potencial de oxirredução do hipoclorito de sódio: desenvolvimento e aplicação de um método de inferência / Rodrigo Lopes da Cruz Magalhães. – 2025.

53 f. : il.

Orientadora: Wagner Roberto de Oliveira Pimentel

Monografia (Trabalho de Conclusão de Curso em Engenharia Química: Bacharelado) – Universidade Federal de Alagoas. Instituto de Química e Biotecnologia. Maceió, 2025.

Bibliografia: f. 50-51

Anexo: f. 52- 53.

1. Hipoclorito de Sódio. 2. Analisadores ORP. 3. Potencial de Oxirredução. 4. Regressão Linear - Não Linear. I. Título.

CDU:542

*Dedico este trabalho a minha Mãe, R.
Maria, que sob muito sol e esforço, me fez
chegar até aqui, na sombra.*

AGRADECIMENTOS

Agradeço a Universidade Federal de Alagoas, professores e funcionários, que se empenharam integralmente para oferecer o curso superior em alto nível científico, marcado, sobretudo, pela diversidade intelectual e pela qualificação do seu corpo docente.

Aos professores Cristiane Sodré, Eduardo Lucena, Livia Oliveira, Rodolfo Brandão e Maritza Montoya pela orientação, apoio e confiança durante os momentos cruciais da graduação; suporte imprescindível para meu êxito.

A Braskem e Origem, por ter propiciado experiência profissional de excelência, além de presentear minha vida com pessoas incríveis, mais que profissionais, verdadeiros amigos de trabalho: Adryelle, Alvaro, Gabriel, Gabrielle, Hada, Sabryna e Sarah que tanto me ajudaram durante o processo de aprendizagem em ambas as empresas.

Ao laboratório de controle Ambiental por ter me proporcionado minha primeira Experiência com iniciação científica.

Aos meus amigos e quase irmãos: Luiz Antonio e Cefras José, que torcem por mim e vibram cada conquista dos meus sonhos. Agradeço pela amizade, conselhos, estímulo e experiências compartilhadas.

Ao meu orientador Prof. Dr. Wagner Roberto, singular e especial gratidão; desde o aceite de orientação, o empenho dedicado para nortear este trabalho, e as incontáveis orientações que sedimentam meu êxito.

Aos meus pais Risolândia Maria Lopes de Magalhães e José Arnaldo da Cruz Magalhães, pelo amor, incentivo e apoio incondicional. Sei que não mediram esforços para que este sonho se realizasse, sem a compreensão, ajuda e confiança de vocês, nada disso seria possível. Vocês são a força que me permite superar todo obstáculo, pois foi ao enxergar o exemplo de vocês que aprendi a perseverar e ser paciente.

Aos meus Irmãos, Rivonnaldo e Ronicleudson Magalhães, pela irmandade e por acreditarem no meu sonho.

Por fim, a todos que direta ou indiretamente fizeram parte da minha formação, o meu muito obrigado.

RESUMO

As Unidades Cloro-Soda produzem hipoclorito de sódio (NaClO) por meio de um processo em batelada, utilizando análises titulométricas periódicas para controle de qualidade, com foco nas concentrações de hipoclorito de sódio e soda cáustica (NaOH). Tais análises demandam tempo significativo, esforço operacional e grande consumo de reagentes e materiais, sendo intensificadas próximo ao fim da batelada para garantir a estabilidade do produto. Como alternativa mais eficiente, este trabalho explora métodos potenciométricos baseados em medições do Potencial de Oxidação e Redução (ORP), que permitem o controle indireto das concentrações de NaClO. O uso de medições inferenciais via ORP pode proporcionar maior eficiência operacional, redução de custos e segurança, minimizando a exposição dos operadores aos riscos do processo. Os objetivos principais deste estudo foram estabelecer um plano de amostragem baseado em instruções técnicas padronizadas e construir modelos de regressão, lineares ou não, utilizando redes neurais artificiais. Foram utilizados dados experimentais de um processo real, aplicando técnicas estatísticas para descrever com precisão a relação entre a concentração de cloro e o Potencial de Oxirredução. Os resultados indicaram uma relação exponencial inversa entre o ORP e a concentração de cloro ativo, com um coeficiente de determinação ($R^2 = 0,985$) no modelo ajustado via Excel. Adicionalmente, modelos de redes neurais artificiais (RNA), com configurações de 20 e 36 neurônios, apresentaram alto desempenho, com valores de acurácia acima de 96% e erros quadráticos médios (MSE) reduzidos em comparação ao modelo exponencial. O teste estatístico confirmou a significância da relação entre as variáveis, com valor- $p < 0,05$, validando a robustez dos modelos desenvolvidos.

Palavras chaves: Hipoclorito de Sódio, Analisadores ORP, Potencial de Oxirredução, Regressão Linear e Não Linear.

ABSTRACT

The Chlorine-Soda Units produce sodium hypochlorite (NaClO) through a batch process, employing periodic titrimetric analyses for quality control, with a focus on sodium hypochlorite and caustic soda (NaOH) concentrations. These analyses demand significant time, operational effort, and extensive consumption of reagents and materials, becoming more frequent toward the end of the batch to ensure product stability. As a more efficient alternative, this study explores potentiometric methods based on Oxidation-Reduction Potential (ORP) measurements, allowing for the indirect control of NaClO and Cl_2 concentrations. Inferential ORP measurements can provide greater operational efficiency, cost reduction, and safety, minimizing operators' exposure to process-related risks. The primary objective of this study is to establish a sampling plan based on standardized technical instructions and construct regression models, both linear and non-linear, using Artificial Neural Networks (ANN). Experimental process data were utilized, applying statistical techniques to accurately describe the relationship between chlorine concentration and Oxidation-Reduction Potential (ORP). The results indicated an inverse exponential relationship between ORP and active chlorine concentration, with a coefficient of determination ($R^2 = 0.985$) in the model adjusted via Excel. Additionally, Artificial Neural Network models with configurations of 20 and 36 neurons demonstrated high performance, with accuracy values above 96% and reduced mean absolute errors (MSE) compared to the Excel model. Statistical testing confirmed the significance of the relationship between the variables, with $p\text{-value} < 0.05$, validating the robustness of the developed models.

Keywords: Sodium Hypochlorite, ORP Analyzers, Oxidation-Reduction Potential, Linear and Non-Linear Regression.

LISTA DE ILUSTRAÇÕES

Figura 1 - Esquema simplificado do processo de uma indústria cloro e soda através da Eletrolise.	16
Figura 2 - Fluxograma simplificado do processo de produção de Hipoclorito de Sódio.	18
Figura 3 - Distribuição T.....	24
Figura 4 - Exemplo de Rede Neural Artificial	26
Figura 5 - Representação de um Neurônio artificial	27
Figura 6 - Modelo de Regressão (Excel).....	39
Figura 7 – Curva de regressão (10 neurônios)	41
Figura 8 – Loss (10 neurônios)	41
Figura 9 - Curva de regressão (20 neurônios)	42
Figura 10 - Loss (20 neurônios)	42
Figura 11 - Curva de regressão (36 neurônios)	43
Figura 12 - Loss (36 neurônios)	43
Figura 13 - Curva de regressão (64 neurônios).....	44
Figura 14 - Loss (64 neurônios)	44
Figura 15 - Modelos de Regressão	46
Figura 16 - Erro absoluto entre modelos	46

LISTA DE TABELAS

Tabela 1 - Identificação dos Equipamentos	18
Tabela 2 - Parâmetros comerciais do Hipoclorito de Sódio	20
Tabela 3 - Plano de Amostragem	29
Tabela 4 - Reagentes e vidrarias utilizados	30
Tabela 5 - Dados amostrais NaClO	37
Tabela 6 - Métricas dos modelos de regressão	45
Tabela 7 - Valores estatísticos	48

LISTA DE ABREVIATURAS E SIGLAS

ABICLOR	Associação Brasileira de Cloro e Derivados
Cl ₂	Gás Cloro
ICMC	Instituto de Ciências Matemáticas e Computação
g.kg ⁻¹	Grama por Quilo
g.L ⁻¹	Grama por Litro
HCl	Ácido Clorídrico
HClO	Ácido Hipocloroso
kg.L ⁻¹	Quilograma Por Litro
MSE	Erro Quadrático Médio
mg.kg ⁻¹	Miligrama por Quilo
mL	Mililitro
mV	Milivolts
NaCl	Cloreto De Sódio
NaClO	Hipoclorito de Sódio
NaOH	Hidróxido De Sódio
ORP	Potencial de Oxirredução
pH	Potencial Hidrogeniônico
PVC	Policloreto de Vinila
R ²	Coefficiente de Regressão
RMSE	Raiz do Erro Quadrático Médio
RNA	Redes Neurais Artificiais
t	Valor Estatístico

SUMÁRIO

1	INTRODUÇÃO	13
2	OBJETIVOS	15
2.1	Geral	15
2.2	Específico	15
3	REVISÃO BIBLIOGRÁFICA	16
3.1	Obtenção do Cloro por Processo de Eletrolise	16
3.2	Produção de Hipoclorito de Sódio	17
3.2.1	Controle Reacional	19
3.2.2	Especificação do Hipoclorito de Sódio	19
3.3	Métodos Potenciométricos	20
3.3.1	Potenciometria	20
3.3.2	Potencial de Oxirredução na Produção de NaClO	20
3.4	Métodos Estatísticos	21
3.4.1	Regressão Linear Simples	21
3.4.2	Regressão Linear Múltipla	22
3.4.3	Teste de Hipótese	22
3.5	Ferramentas Computacionais	25
3.5.1	Redes Neurais	25
3.6	Fatores de Interferência	27
3.6.1	Potencial Hidrogeniônico (pH)	27
4	METODOLOGIA	28
4.1	Coleta de Dados	28
4.2	Amostragem	28
4.2.1	Plano de Amostragem	28
4.2.2	Volumetria do Cl ₂ e Medição ORP	30
4.3	Redes neural	32
4.4	Métricas	34
4.4.1	Acurácia	34
4.4.2	Coeficiente de determinação R ²	34
4.4.3	Erro Quadrático Médio	35
4.4.4	Raiz do Erro Quadrático Médio	35
4.5	Teste de hipótese	35
5	RESULTADOS E DISCUSSÕES	37
5.1	Dados amostrais	37
5.2	Modelos de Regressão	39

5.2.1	Modelo Exponencial.....	39
5.2.2	Rede Neural.....	40
5.3	Comparação Entre Modelos de Regressão	45
5.4	Validação Estatística.....	47
6	CONCLUSÃO	49
7	REFERÊNCIAS BIBLIOGRÁFICAS	50
8	ANEXOS	52
	ANEXO I - Resultados dos modelos de regressão para concentração de cloro	52

1 INTRODUÇÃO

A importância da indústria de cloro e soda no mundo é indiscutível, estes estão entre os dez produtos químicos mais produzidos no mundo e estão ligados à maioria da produção de outros produtos que são utilizados no dia-a-dia da sociedade (O'BRIEN *et al.* 2005).

As aplicações do cloro são diversas e variam desde a produção do gás lacrimogênio até o tratamento de água e produção de hipoclorito de sódio e outros diversos produtos. Os principais segmentos do mercado mundial que consomem cloro são vinílicos (33%) e produtos orgânicos (19%). Já o mercado de soda está dividido principalmente entre produtos orgânicos (18%), papel e celulose (16%) e produtos inorgânicos (15%) (JUNIOR, 2006).

Nas indústrias que produzem cloro liquefeito, a área de liquefação frequentemente apresenta a presença de gases inertes, como oxigênio, dióxido de carbono, nitrogênio e hidrogênio. Esses inertes dificultam a completa liquefação do cloro produzido. Antes de serem liberados para a atmosfera, esses gases não condensáveis devem passar por um processo de remoção de resíduos de cloro gasoso, conhecido como abatimento (WANJIH, 2024).

Durante esse procedimento, o cloro gasoso não liquefeito da etapa de liquefação é utilizado para produzir hipoclorito de sódio, transformando o resíduo em um coproduto de alto valor comercial e evitando a liberação de cloro na atmosfera (LOPES, 2006). Esse processo ocorre por meio da absorção do cloro em contracorrente em uma solução de soda cáustica (NaOH), que deve alcançar uma concentração adequada para manter ao menos 12% de cloro ativo.

O controle das concentrações de soda cáustica e cloro na reação para produzir hipoclorito de sódio é realizado por meio de titulação de amostras coletadas periodicamente. À medida que o ponto final da reação se aproxima, intensifica-se o monitoramento para garantir um excesso de soda cáustica entre 8 e 12 g.L⁻¹ e manter o cloro livre em 12%, assegurando a estabilidade e conformidade do produto (LUCCA, 2006).

O aumento intensivo de análises para manter a especificação do produto, eleva significativamente o consumo de reagentes além de maior tempo exposição por parte dos operadores de área.

Assim, demonstra-se a necessidade do desenvolvimento de técnicas que contribuam no

aprimoramento do controle das concentrações de hipoclorito de sódio. A implementação de um método de controle apropriado, resultará em um produto de maior estabilidade, um uso eficiente dos reagentes, além de maior confiabilidade ao processo de produção.

2 OBJETIVOS

2.1 Geral

Este trabalho tem como objetivo principal desenvolver modelos de regressão linear ou não linear que possibilitem a medição de concentração de hipoclorito de sódio (NaClO) e cloro ativo em solução (Cl_2), a partir do potencial Oxirredução através de analisadores inferenciais de ORP (*Oxidation Reduction Potencial*).

2.2 Específico

- Estabelecer o plano de amostragem e análise do hipoclorito, em tanques de produção, a partir de dados reais de processo.
- Utilizar modelos inferenciais baseados na estatística e em redes neurais artificiais que sejam capazes de descrever efetivamente as concentrações de hipoclorito de sódio durante sua produção;
- Analisar e comparar os diferentes modelos baseados em regressão linear e redes neurais artificiais;
- Avaliar a confiabilidade dos dados por meio de validação estatística.

3 REVISÃO BIBLIOGRÁFICA

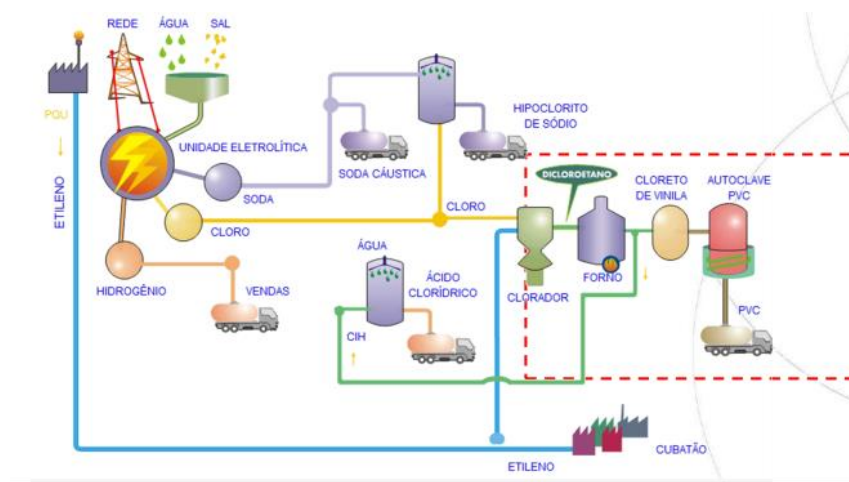
3.1 Obtenção do Cloro por Processo de Eletrolise

O desenvolvimento do cloro remonta ao século XIII, quando experimentos com a chamada "água régia" revelaram a presença de uma fumaça densa, amarelada e irritante, que chamava atenção pela sua natureza corrosiva. Em 1774, o químico Carl Wilhelm Scheele foi o primeiro a notar uma aplicação potencial para o cloro: seu poder de branqueamento em materiais vegetais. Esse feito marcou o início de várias descobertas sobre o elemento, incluindo a produção de dicloroetano em 1795 e a síntese do clorofórmio em 1831.

Em 1800, o cientista Cruikshank realizou a primeira produção de cloro via eletrólise de salmoura. Esse método se tornou a base de produção industrial e foi amplamente adotado ao longo das décadas seguintes, consolidando-se como um processo eficiente e seguro (O'Brien *et al.* 2005).

Atualmente, três principais tecnologias de produção de cloro e soda cáustica são usadas globalmente: o método de diafragma (46% de participação), o de membrana (32%) e o de mercúrio (22%) (JUNIOR, 2006). Segundo Hine (1985), o processo de produção industrial utilizando células eletrolíticas com diafragmas envolve etapas de tratamento da salmoura, eletrólise, evaporação da soda e processamento do cloro, hidrogênio e ainda a produção de subprodutos como hipoclorito de sódio. Em unidades industriais, o processo inclui também a produção de outros produtos para confecção final de policloreto de vinila (PVC), conforme ilustrado na Figura 1.

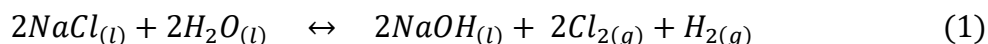
Figura 1 - Esquema simplificado do processo de uma indústria cloro e soda através da Eletrolise.



Fonte: Apresentação Industrial, Unipar, 2019.

A solução aquosa de salmoura é preparada através da mistura de sal e água até uma concentração máxima de até 320 g.L⁻¹ de NaCl, normalmente essa salmoura passa por etapas de pré-tratamento, para remoção de impurezas e posteriormente é enviada até as células eletrolíticas.

Esta salmoura tratada é injetada na célula onde ocorre a eletrólise, segundo a reação:



Gerando os principais produtos: Cloro, Hidrogênio e Hidróxido de sódio.

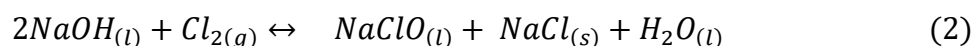
Após esta etapa o cloro (Cl₂) segue para a área de processamento com o objetivo de retirar a água e as impurezas, através do resfriamento, compressão e liquefação, para posterior envio para a área de produção de (PVC), e o cloro residual não liquefeito é utilizado para produção de coprodutos como ácido clorídrico e hipoclorito de sódio.

O hidrogênio é distribuído para as caldeiras e para a área de produção de ácido clorídrico e aquecimento da salmoura. A solução aquosa composta por NaOH (12% em massa) e NaCl (17% em massa), também conhecida como licor de células, flui por gravidade para o tanque de licor e posterior envio para a evaporação até a concentração da soda em 50%, e é utilizada para obtenção de subprodutos.

3.2 Produção de Hipoclorito de Sódio

O hipoclorito de sódio é um composto líquido de coloração amarelada e aspecto turvo à temperatura ambiente, conhecido por sua alta degradabilidade e instabilidade, especialmente em soluções sem a presença de agentes estabilizadores.

Segundo (LOPES, 2006), o processo químico envolvendo a produção de Hipoclorito de sódio em unidades industriais, ocorre a partir da cloração direta em hidróxido de sódio, a temperatura controlada em no máximo 35°C, via reação, conforme Equação 2.

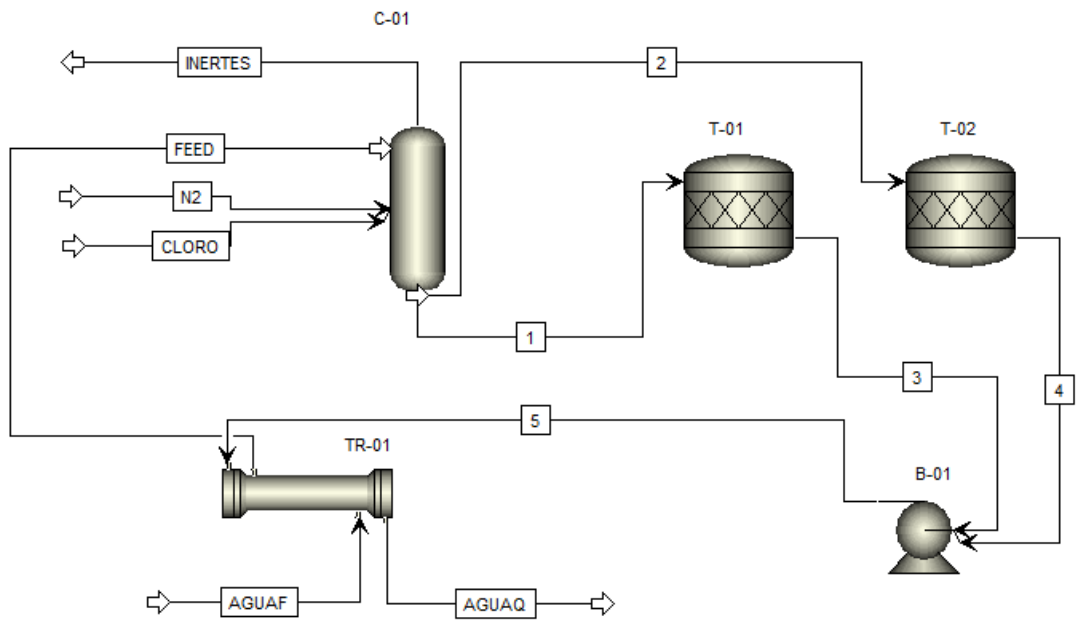


A reação ocorre em meio aquoso em pH entre 12-13, nessas condições há formação do íon Hipólito, conforme Equação 3.



As reações envolvidas no processo possuem características altamente exotérmicas, sendo necessário assim o sistema de resfriamento muito bem controlado, utilizando-se de trocadores de calor (SHREVE, 1980). A Figura 2 demonstra a configuração do sistema de produção industrial de hipoclorito de sódio.

Figura 2 - Fluxograma simplificado do processo de produção de Hipoclorito de Sódio.



Fonte: Autor, 2024.

Tabela 1 - Identificação dos Equipamentos

TAG	Equipamento
C-01	Coluna de Absorção
B-01	Bomba
T-01	Tanque de produção
T-02	Tanque de produção
TR-01	Trocador de calor

Fonte: Autor, 2024.

O processo de produção de NaClO em batelada inicia-se com a adição de uma solução de soda cáustica, inicialmente a 150 g.L⁻¹, no tanque de operação. Esta solução circula para uma coluna de absorção, onde ocorre o processo de absorção em contracorrente com cloro gasoso, que é injetado na base da coluna. O hipoclorito de sódio formado é então bombeado do tanque para um trocador de calor. Após a troca térmica, a solução é retornada ao topo da coluna,

permitindo que a concentração de NaClO aumente progressivamente até o final da batelada (SHREVE, 1980).

O hipoclorito de sódio permanece em circuito fechado até que a reação de formação se complete e o produto alcance a concentração desejada, com 12 g.L⁻¹ de soda cáustica e 12% de cloro livre em reação, (LUCCA, 2006).

3.2.1 Controle Reacional

A injeção de ar ou nitrogênio no topo da coluna é necessária para que se dilua o hidrogênio presente no cloro, e que não é absorvido no processo. Assim, o hidrogênio pode ser controlado em um nível de concentração abaixo do risco de explosividade e encaminhado com segurança para abatimento.

De acordo com o Pamphlet 089 publicado pelo *Chlorine Institute* em 2008, o controle das concentrações de soda cáustica (NaOH) e cloro (Cl₂) na solução que circula na coluna é essencial para evitar a supercloração. Quando a solução se torna saturada com cloro, ela se torna mais ácida, pois o excesso de cloro diminui o pH da reação. À medida que o pH cai abaixo de 11, ocorre a formação de ácido clorídrico (HCl). Com a redução progressiva do pH para valores abaixo de 5, o ácido hipocloroso (HClO) se decompõe em ácido clorídrico. Uma vez acumulada uma quantidade significativa de HCl, a formação de Cl₂ pode ocorrer em pH ácido, o que pode comprometer a estabilidade da solução e a eficiência do processo.

3.2.2 Especificação do Hipoclorito de Sódio

Dependendo do tipo de indústria, e da aplicabilidade do hipoclorito de sódio, ele possui uma especificação pré-determinada para suas características comerciais. Segundo a ABICLOR (2020), a média dessas características podem ser visualizadas na Tabela 2 abaixo:

Tabela 2 - Parâmetros comerciais do Hipoclorito de Sódio

Aspecto	Unidade	Quantidade
Cloro Ativo	% em Peso	12,0
Densidade (20 °C)	kg.L ⁻¹	1,2
Cloro Como Cloreto	g.kg ⁻¹	13,0
Cloro Como Clorato	g.kg ⁻¹	2,5
Hidróxido De Sódio	g.kg ⁻¹	5,0
Carbonato De Sódio	g.kg ⁻¹	8,0
Ferro	mg.kg ⁻¹	5,0

Fonte: ABICLOR, 2024.

Os parâmetros acima são determinados exclusivamente para um lote recém-produzido de hipoclorito de sódio, pois, devido à sua alta instabilidade, suas características químicas se alteram com o tempo. Essa degradação ocorre naturalmente por meio de vários tipos de mecanismos.

3.3 Métodos Potenciométricos

3.3.1 Potenciometria

A potenciometria ou método potenciométrico de análise química se baseia na medida da diferença de potencial entre dois eletrodos que estarão em contato com a solução do analito, isto é, baseia-se na medida da diferença de potencial entre um eletrodo indicador e um eletrodo de referência expostos à solução do analito. O potencial do eletrodo indicador varia com a concentração da espécie química de interesse, enquanto o potencial do eletrodo de referência permanece constante. O eletrodo de referência é um eletrodo cujo potencial é fixo e constante durante a análise, ou seja, o potencial desse eletrodo não varia com a concentração da espécie química a ser determinada. O eletrodo indicador é sensível à variação da espécie de interesse, que, por sua vez, influencia o potencial (SKOOG *et al*, 2006).

3.3.2 Potencial de Oxirredução na Produção de NaClO

A natureza das reações envolvidas na produção do hipoclorito de sódio é do tipo oxidação-redução. Nesse tipo de reação tem-se a transferência de elétrons de um reagente para o outro (SKOOG *et al*, 2006), onde a partir deste transporte, ocorre a geração de voltagem (em

mV) na solução que pode ser medida a partir de medidores de potencial de oxirredução (ORP) (PYE *et al*, 1950).

O potencial de um eletrodo de platina em uma solução cloro-álcali é representado segundo a Equação de Nest (SKOOG *et al*, 2006):

$$E = E^{\circ} - \frac{0,0591}{2} \log \frac{(\text{ClO}^-)(\text{H}^+)^2}{(\text{Cl}^-)} \quad (4)$$

Onde E° é o potencial padrão da reação. Como a medição do potencial é dependente da concentração do íon hidrônio, a neutralização da solução de soda cáustica pela adição de cloro gasoso exibe um ponto de inflexão próximo ao ponto final da reação, sendo também possível haver o controle do processo a partir da medida do pH da solução durante a produção de hipoclorito de sódio (PYE *et al*, 1950).

3.4 Métodos Estatísticos

3.4.1 Regressão Linear Simples

A Regressão é uma técnica estatística utilizada para investigar a relação entre variáveis. Um modelo de Regressão Matemática descreve a relação entre uma variável dependente (ou resposta) e uma ou mais variáveis independentes (ou explicativas), permitindo fazer previsões ou inferências sobre o comportamento da variável resposta com base nos valores das variáveis explicativas. Quando o estudo sobre duas variáveis quantitativas e a relação entre elas pode ser bem representada por uma linha reta, tem-se o caso da Regressão Linear Simples (MARTINS, 2019).

Nesse modelo, assume-se que existe uma relação linear entre a variável dependente e a variável independente, o que significa que a variação em uma está associada a uma variação proporcional na outra. A Equação da reta é expressa como:

$$Y = a + b_i x_i \quad (5)$$

Onde “Y” é a variável dependente, “ x_i ” é a variável independente, “a” representa a interceptação da reta e “b” o coeficiente angular, indicando a inclinação da reta (MARTINS, 2019).

3.4.2 Regressão Linear Múltipla

Modelos multivariados de pesquisa analisam o relacionamento entre múltiplas variáveis explicativas e, em alguns casos, múltiplas variáveis dependentes. A Regressão Múltipla é uma técnica estatística amplamente utilizada, que avalia a relação entre uma variável dependente e várias independentes.

Tabachnick e Fidell (1996), definem a regressão linear múltipla como um conjunto de técnicas úteis, já que as variáveis independentes costumam ser correlacionadas. Destacam estudos dessa área, que, a correlação entre preditores pode levar à exclusão de variáveis importantes, e a regressão linear múltipla é uma abordagem apropriada para lidar com essa redundância (DUNLAP, LANDIS, 1998).

Regressões lineares múltiplas resultam em uma equação que melhor prediz a variável dependente a partir das variáveis independentes, representando um modelo aditivo. A Equação é dada por:

$$Y = a + b_1x_1 + b_2x_2 + \dots + b_nx_n + \varepsilon \quad (6)$$

Onde “y” é a variável dependente, “a” é a constante, “b_n” são os coeficientes de regressão, “x_n” são as variáveis independentes, e “ε” é o erro associado a equação (DUNLAP, LANDIS, 1998).

A Regressão Linear Múltipla é uma abordagem estatística para descrever associações simultâneas entre várias variáveis. Um item essencial nessa modelagem inclui a estimativa, a inferência e a seleção das variáveis que constroem o modelo, bem como, a avaliação do ajuste do modelo (HAHS-VAUGHN *et al*, 2020).

3.4.3 Teste de Hipótese

Os testes de hipóteses são ferramentas estatísticas amplamente utilizadas para avaliar a igualdade ou desigualdade entre duas ou mais medidas, comparar valores esperados ou previstos com valores observados, ou ainda verificar a consistência entre estatísticas de diferentes conjuntos de dados separados no tempo e no espaço. Essas técnicas são essenciais para a validação científica, permitindo avaliar a relação entre variáveis e sustentar decisões baseadas em evidências. Por exemplo, no campo da saúde, pode-se empregar um teste de

hipótese para verificar a eficácia de uma vacina na cura da gripe aviária em galinhas poedeiras, com suporte do cálculo de probabilidades (ASSIS *et al.*, 2020).

O processo de um teste de hipótese envolve a formulação de uma hipótese nula (H_0), que representa a ausência de efeito ou diferença, e uma hipótese alternativa (H_1), que sugere a existência de um efeito ou diferença. A partir dos dados amostrais, calcula-se um valor estatístico que é comparado a um nível de significância pré-estabelecido (geralmente 0,05), para determinar se há evidências suficientes para rejeitar (H_0) em favor de (H_1) (ASSIS *et al.*, 2020).

Segundo (ASSIS *et al.*, 2020) em um conjunto de dados com tamanho amostral definido seguindo uma distribuição normal, conhecendo-se o desvio padrão, o teste de hipótese pode ser aplicado, de acordo com a Equação.

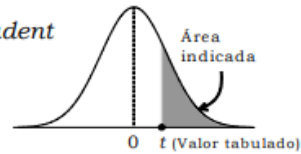
$$t = \frac{x - \mu_0}{s\sqrt{n}} \quad (7)$$

Onde:

- n – Tamanho da amostra;
- s – Desvio padrão da amostra;
- x – Média da amostra;
- μ_0 – Média populacional.

Com o valor obtido a partir da Equação 7, e definido o nível de significância (α), pode-se buscar pelo valor associado “ $t_{\text{crítico}}$ ”, de acordo com a Tabela “t-student”, referenciada na Figura 3.

Figura 3 - Distribuição T

Tabela 5 Distribuição *t* de Student

gl	Área na cauda superior								
	0,25	0,10	0,05	0,025	0,01	0,005	0,0025	0,001	0,0005
1	1,000	3,078	6,314	12,71	31,82	63,66	127,3	318,3	636,6
2	0,816	1,886	2,920	4,303	6,965	9,925	14,09	22,33	31,60
3	0,765	1,638	2,353	3,182	4,541	5,841	7,453	10,21	12,92
4	0,741	1,533	2,132	2,776	3,747	4,604	5,598	7,173	8,610
5	0,727	1,476	2,015	2,571	3,365	4,032	4,773	5,894	6,869
6	0,718	1,440	1,943	2,447	3,143	3,707	4,317	5,208	5,959
7	0,711	1,415	1,895	2,365	2,998	3,499	4,029	4,785	5,408
8	0,706	1,397	1,860	2,306	2,896	3,355	3,833	4,501	5,041
9	0,703	1,383	1,833	2,262	2,821	3,250	3,690	4,297	4,781
10	0,700	1,372	1,812	2,228	2,764	3,169	3,581	4,144	4,587
11	0,697	1,363	1,796	2,201	2,718	3,106	3,497	4,025	4,437
12	0,695	1,356	1,782	2,179	2,681	3,055	3,428	3,930	4,318
13	0,694	1,350	1,771	2,160	2,650	3,012	3,372	3,852	4,221
14	0,692	1,345	1,761	2,145	2,624	2,977	3,326	3,787	4,140
15	0,691	1,341	1,753	2,131	2,602	2,947	3,286	3,733	4,073
16	0,690	1,337	1,746	2,120	2,583	2,921	3,252	3,686	4,015
17	0,689	1,333	1,740	2,110	2,567	2,898	3,222	3,646	3,965
18	0,688	1,330	1,734	2,101	2,552	2,878	3,197	3,610	3,922
19	0,688	1,328	1,729	2,093	2,539	2,861	3,174	3,579	3,883
20	0,687	1,325	1,725	2,086	2,528	2,845	3,153	3,552	3,850
21	0,686	1,323	1,721	2,080	2,518	2,831	3,135	3,527	3,819
22	0,686	1,321	1,717	2,074	2,508	2,819	3,119	3,505	3,792
23	0,685	1,319	1,714	2,069	2,500	2,807	3,104	3,485	3,768
24	0,685	1,318	1,711	2,064	2,492	2,797	3,091	3,467	3,745
25	0,684	1,316	1,708	2,060	2,485	2,787	3,078	3,450	3,725
26	0,684	1,315	1,706	2,056	2,479	2,779	3,067	3,435	3,707
27	0,684	1,314	1,703	2,052	2,473	2,771	3,057	3,421	3,689
28	0,683	1,313	1,701	2,048	2,467	2,763	3,047	3,408	3,674
29	0,683	1,311	1,699	2,045	2,462	2,756	3,038	3,396	3,660
30	0,683	1,310	1,697	2,042	2,457	2,750	3,030	3,385	3,646
35	0,682	1,306	1,690	2,030	2,438	2,724	2,996	3,340	3,591
40	0,681	1,303	1,684	2,021	2,423	2,704	2,971	3,307	3,551
45	0,680	1,301	1,679	2,014	2,412	2,690	2,952	3,281	3,520
50	0,679	1,299	1,676	2,009	2,403	2,678	2,937	3,261	3,496
z	0,674	1,282	1,645	1,960	2,326	2,576	2,807	3,090	3,291

Nota: A coluna em destaque é a mais usada.

Fonte: BARBETTA, 2010

Se o valor obtido do teste estatístico (*t*) for menor, em valor absoluto, do que o valor crítico ($t_{\text{crítico}}$), então não há evidências estatisticamente significativas para rejeitar a hipótese nula (H_0). Isso indica que os dados amostrais são consistentes com H_0 , dentro do nível de significância estabelecido.

Uma abordagem alternativa para avaliar a hipótese nula baseia-se no valor-p, que representa a probabilidade de observar um valor do teste estatístico tão extremo quanto (ou mais extremo que) o valor observado, sob a suposição de que H_0 é verdadeira.

A região crítica para rejeição de H_0 é definida em função do nível de significância (α) e do tipo de teste conduzido:

- Teste unilateral (esquerda ou direita): A região crítica está em uma única extremidade da curva de distribuição.

- Teste bilateral: A região crítica está em ambas as extremidades da curva, considerando desvios significativos tanto para valores maiores quanto menores que o esperado sob H_0

Para se calcular o valor-p, seguimos a seguinte equação:

$$\text{Valor Alpha} = 2 * \text{valor} - p \mid t > \text{Tabela Estatística "t"} \quad (8)$$

A decisão de rejeitar ou não H_0 é baseada na comparação entre o valor-p e o nível de significância estabelecido (α):

- Se $\text{valor-p} < \alpha$: Há evidências estatisticamente significativas para rejeitar H_0 .
- Se $\text{valor-p} \geq \alpha$: Não há evidências suficientes para rejeitar H_0 .

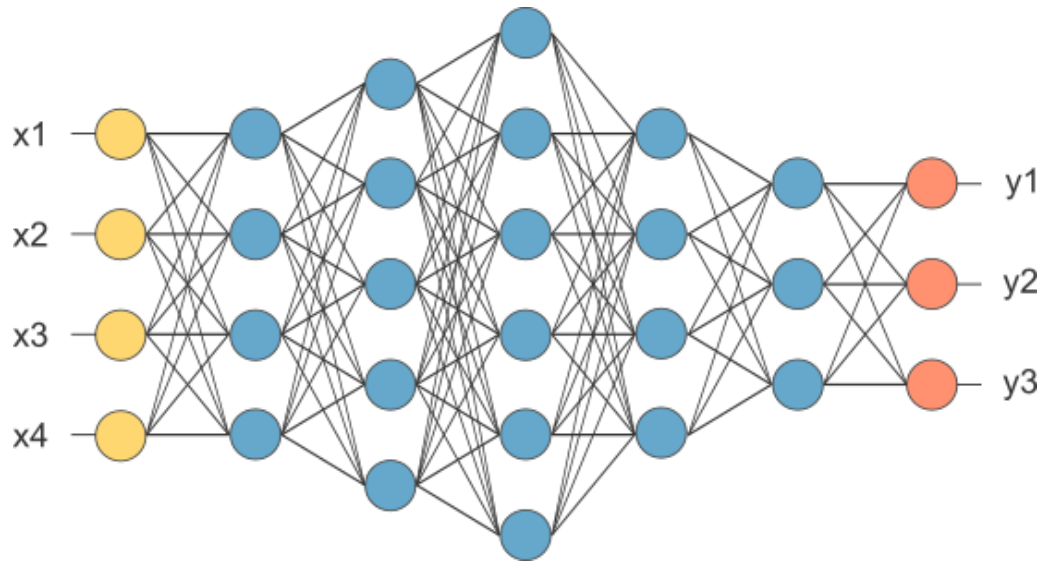
3.5 Ferramentas Computacionais

3.5.1 Redes Neurais

A aplicação de redes neurais na análise de dados é um campo vasto e em constante evolução, integrando conceitos de inteligência artificial com métodos estatísticos para resolver problemas complexos. Redes neurais são modelos computacionais inspirados no funcionamento do cérebro humano, compostos por unidades de processamento (neurônios) que se conectam e ajustam seus pesos sinápticos através de processos de aprendizado, (FERNEDA, 2006).

Redes Neurais possuem funções iterativas que permitem a elas serem “treinadas” para resolver determinados tipos de problemas. As primeiras funções de treinamento foram idealizadas por cientistas associados a biólogos, na Figura 4 é possível visualizar uma representação de Rede neural com 4 entradas e 3 saídas (COELHO, 2024).

Figura 4 - Exemplo de Rede Neural Artificial

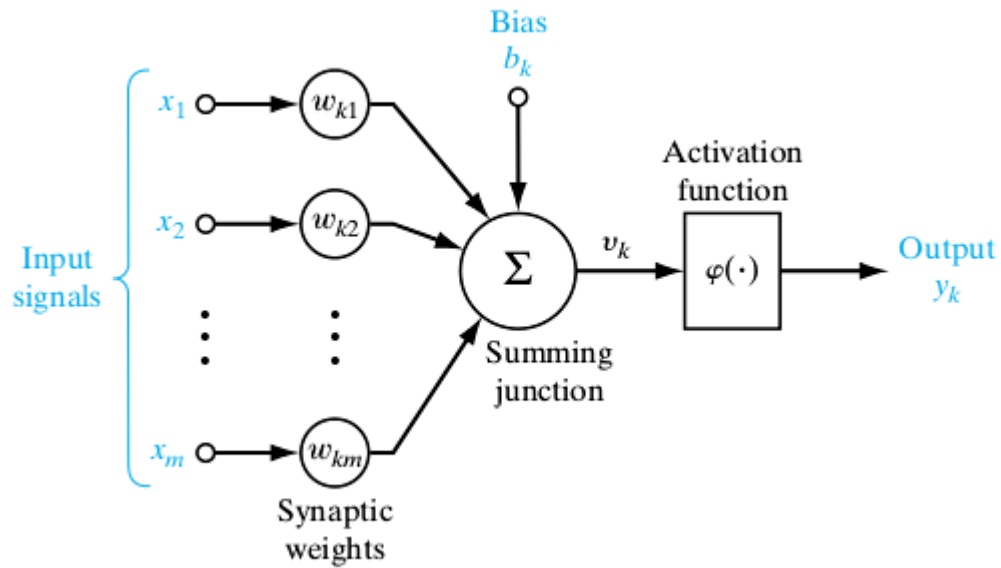


Fonte: COELHO, 2024.

As (RNAs) possuem como unidade fundamental o neurônio artificial, uma estrutura inspirada no neurônio biológico. Esses neurônios artificiais recebem um conjunto de entradas, comparáveis aos "axônios" dos neurônios biológicos, e produzem uma saída, equivalente ao "dendrito". Assim como os neurônios presentes em sistemas biológicos, os neurônios artificiais também possuem a capacidade de aprender, ajustando seus parâmetros durante um processo de treinamento para resolver problemas específicos. No contexto computacional, o aprendizado dos neurônios artificiais é realizado com base em dados de entrada e saída. Durante o treinamento, os parâmetros dos neurônios são ajustados para minimizar erros e melhorar a precisão nas previsões. Em aplicações modernas, utiliza-se uma abordagem conhecida como *Deep Learning*, que treina redes neurais profundas compostas por várias camadas de neurônios interligados (COELHO, 2024).

Os neurônios artificiais são estruturados com base em uma arquitetura inspirada nos biológicos, mas adaptada para atender às demandas computacionais. As entradas de um neurônio são ponderadas por pesos, que determinam a importância de cada entrada, e combinadas de acordo com uma função de ativação. Essa função é responsável por processar os valores ponderados das entradas, realizando uma combinação linear (que pode incluir um valor adicional chamados bias) e produzindo um valor de saída ativado. O funcionamento básico de um neurônio genérico está representado na Figura 5, destacando suas principais componentes (COELHO, 2024).

Figura 5 - Representação de um Neurônio artificial



Fonte: COELHO, 2024.

Tais estruturas formam a base de todas as redes neurais artificiais, sendo amplamente utilizadas em aplicações que envolvem aprendizado supervisionado, não supervisionado ou por reforço.

Uma das vantagens das redes neurais é a capacidade de aprender e melhorar seu desempenho à medida que são expostas a novos dados. Elas podem identificar padrões ocultos e realizar inferências baseadas em dados passados, o que as torna ideais para análises preditivas. (FERNEDA, 2006).

3.6 Fatores de Interferência

3.6.1 Potencial Hidrogeniônico (pH)

O potencial de oxirredução, é uma medida completa, que varia a depender da concentração dos componentes em solução. A interferência do pH na medição do potencial (ORP), se dá ao aumento da concentração de cloro livre em solução, diminuindo o pH que favorece a formação de ácidos, que por sua vez, interferem diretamente na medição (ORP), tendo em vista a relação logarítmica linear entre concentração e tensão da solução (GONGORA *et al*, 2024).

4 METODOLOGIA

A metodologia deste trabalho baseou-se em dados reais de processo, utilizando uma solução concentrada de hipoclorito de sódio. Para a análise dos dados, foram empregadas ferramentas computacionais como Google Colaboratory e Microsoft Excel, que auxiliaram na organização, tratamento e interpretação dos resultados experimentais.

4.1 Coleta de Dados

Em dezembro de 2023, foram realizadas a coleta de amostras de hipoclorito de sódio e realizados o experimento em laboratório, para obtenção do potencial de oxirredução e concentração de cloro ativo em solução. Essa coleta ocorreu com o NaClO estável e especificado para venda.

4.2 Amostragem

4.2.1 Plano de Amostragem

Inicialmente, foi desenvolvido um plano de amostragem para coleta dos dados, no qual foi definido a volumetria da amostra inicial e de todos os pontos de diluição, incluindo concentração de cloro esperada a cada ponto, com o objetivo de orientar o experimento, conforme referenciado na Tabela 3.

O volume inicial da amostra utilizada foi de 50 mL, com um incremento fixo de diluição de 5 mL, dessa forma foram estabelecidos 30 pontos de diluição, incluindo a previsão das concentrações de cloro esperadas, baseando-se na concentração da amostra inicial conforme a Equação 9.

$$C_1V_1 = C_2V_2 \quad (9)$$

Onde “ C_1 e V_1 ” é a concentração e o volume inicial da amostra, “ V_2 ” é o volume final e “ C_2 ” é concentração desconhecida. A Equação 9, dessa forma poder ser manipulada com objetivo de ser determinar “ C_2 ”.

$$\frac{C_1V_1}{V_2} = C_2 \quad (10)$$

Tabela 3 - Plano de Amostragem

Pontos	Volume da Amostra (mL)	Concentração de cloro (g.L ⁻¹)
1	50,0	165,0
2	55,0	150,0
3	60,0	137,5
4	65,0	126,9
5	70,0	117,9
6	75,0	110,0
7	80,0	103,1
8	85,0	97,1
9	90,0	91,7
10	95,0	86,8
11	100,0	82,5
12	105,0	78,6
13	110,0	75,0
14	115,0	71,7
15	120,0	68,8
16	125,0	66,0
17	130,0	63,5
18	135,0	61,1
19	140,0	58,9
20	145,0	56,9
21	150,0	55,0
22	155,0	53,2

23	160,0	51,6
24	165,0	50,0
25	170,0	48,5
26	175,0	47,1
27	180,0	45,8
28	185,0	44,6
29	190,0	43,4
30	195,0	42,3

Fonte: Autor, 2024.

4.2.2 Volumetria do Cl_2 e Medição ORP

Os reagentes e equipamentos usados para a prática da análise do teor de cloro ativo e tensão da solução estão descritos encontrados na Tabela 4.

Tabela 4 - Reagentes e vidrarias utilizados

Reagentes	Vidrarias
Solução de iodeto de potássio (KI) a 5% (m/m)	Pipetas volumétricas (1 mL)
Solução de ácido acético 1:3 (v/v)	Proveta graduada (50 mL)
Solução de tiosulfato de sódio 0,1 N padronizada	Bureta de (50 mL)
Indicador de solução de amido em frasco conta gotas	Frasco de Erlenmeyer (250 mL)
Água destilada	Analizador ORP
Solução de calibração	

Fonte: Autor, 2024.

Para determinação do cloro ativo, foram realizados os seguintes procedimentos:

1. Pipetou-se 1 mL de amostra para um Erlenmeyer de 250 mL, a completar o volume com 30 mL de água destilada e homogeneizou-se a solução;
2. Pipetou-se 10 mL de solução de iodeto de potássio a 5% (m/m) para um frasco de Erlenmeyer de 250 mL;
3. Adicionou-se 10 mL da solução de ácido acético a 33% e titulou-se, rapidamente, com a solução de tiosulfato de sódio 0,1 N contida na bureta até atingir a cor

amarelo claro;

4. Colocou-se 3 gotas de indicador de amido e prosseguiu-se a titulação até que desapareça a cor azul.
5. Foi anotado o volume gasto de solução de tiosulfato de sódio.

Este procedimento foi realizado para cada ponto de diluição. A ordem dos reagentes no frasco Erlenmeyer não altera o resultado das análises desde que a amostra seja inserida por último. Para determinar o teor de cloro ativo em solução a cada ponto de diluição, utilizou-se da Equação 11.

$$\text{Teor de cloro ativo (g.L}^{-1}\text{)} = \text{Volume de Tiosulfato de sódio} * 3,546 \quad (11)$$

Para determinação da tensão em solução (potencial de oxirredução), seguiu-se o procedimento conforme descrito abaixo:

1. Conectou-se o eletrodo de ORP ao medidor HI9126 e ligou-se o dispositivo.
2. Configurou-se o medidor no modo "mV", pressionando a tecla "RANGE".
3. Após a configuração, o eletrodo de ORP foi imerso na amostra até uma profundidade de 4 centímetros.
4. Com o eletrodo estabilizado na solução, registrou-se o valor correspondente ao potencial de oxirredução (em mV) para cada ponto de diluição preparado.

Esse processo foi repetido para todas as amostras, garantindo a precisão e a reprodutibilidade das medições.

4.3 Redes neural

Para o desenvolvimento do código utilizado na análise e construção do modelo preditivo, foram adotadas ferramentas computacionais utilizadas no processamento de dados e aprendizado de máquina. O código foi implementado em Python, uma linguagem de programação versátil e aplicada em contextos científicos e acadêmicos, utilizando-se dos dados obtidos experimentalmente.

Para garantir a eficiência e a precisão das análises, as seguintes bibliotecas foram utilizadas:

- **NumPy**: Manipulação de arrays e operações matemáticas.
- **Matplotlib**: Visualização gráfica dos dados e resultados.
- **scikit-learn**: Redimensionamento de dados, divisão em conjuntos de treinamento e teste, e cálculo do coeficiente de determinação R^2 .
- **TensorFlow**: Construção e treinamento da rede neural.
- **Joblib**: Persistência dos escaladores de normalização para futura reutilização.

A estruturação do código foi realizada em partes, afim de garantir que estejam adequados para serem utilizados.

1. Carregamento de Dados; os dados de entrada são carregados diretamente em formato de *array NumPy*, contendo dois atributos:
 - ORP (mV): Variável independente.
 - Concentração de Cloro (g.L^{-1}): Variável dependente.
2. Pré-Processamento dos Dados
 - Escalamento: Utilizou-se a classe *MinMaxScaler* da biblioteca *scikit-learn* para normalizar os valores entre 0 e 1, melhorando o desempenho da rede neural.
 - Divisão em Conjuntos de Treinamento e Teste: O conjunto foi dividido em 50% para treinamento e 50% para teste, garantindo validação cruzada.
3. Construção da Rede Neural
 - Foi utilizada uma rede sequencial da biblioteca *TensorFlow/Keras*, que permite a inserção de camada por camada de neurônios na RNA.

- Camada de entrada: 1 neurônio (correspondente ao ORP).
- Duas camadas ocultas: Cada camada com um número de neurônios a serem definidos e função de ativação ReLU.
- Camada de saída: 1 neurônio com função de ativação linear, para prever a concentração de cloro.

Para compilação do modelo, utilizou-se um otimizador “*Adam*”, com taxa de aprendizado de 0,001 e uma função de perda, *Mean Squared Error* (Erro Quadrático Médio).

Durante o treinamento do modelo, foi definido o número de épocas em 500, com possibilidade de parada antecipada, à medida que o erro médio dos dados de treinamento fosse menor que 0,01 em comparação com os dados amostrais reais, para cada ciclo de treinamento, o tamanho do lote (*batch*) a ser testado foi de 8 amostras por vez, incluindo uma validação cruzada de dados para monitorar o desempenho do modelo a cada época testada.

No modelo desenvolvido, a função de ativação ReLU (*Rectified Linear Unit*) foi utilizada devido à sua eficiência em redes profundas, pois ajuda a mitigar problemas associados ao gradiente desaparecido, proporcionando uma convergência mais rápida e estável durante o treinamento.

Além disso, a normalização dos dados foi um passo para garantir que todas as variáveis estivessem na mesma escala. Isso não apenas melhora a estabilidade numérica do modelo, mas também acelera o processo de convergência ao facilitar a otimização do algoritmo de aprendizado.

4.4 Métricas

Para avaliação da RNA foi incluindo métricas, já que o modelo é avaliado no conjunto de teste utilizando a função de perda. O cálculo de métricas em RNAs é fundamental para avaliar o desempenho e orientá-lo durante o treinamento. As métricas de erro medem a discrepância entre as previsões feitas pela rede e os valores reais, ajudando a entender como o modelo está se saindo em relação aos dados de treinamento e validação.

4.4.1 Acurácia

A Acurácia é uma métrica simples e comum para avaliar o desempenho de modelos de classificação. Ela calcula a proporção de previsões corretas sobre o total de previsões feitas. Seguindo uma equação simples.

$$Acurácia = \frac{\text{Número de previsões corretas}}{\text{Número total de exemplos}} \quad (12)$$

Além da acurácia, a métrica de “*Loss*”, foi implementada no modelo, onde determina a perda da rede neural artificial em comparação com os dados reais que estão sendo avaliados

4.4.2 Coeficiente de determinação R^2

O R^2 , também conhecido como “coeficiente de determinação” é uma métrica que indica o quão bem o modelo se ajusta aos dados. Ele mede a proporção da variação nos dados de saída que é explicada pelo modelo, que é determinado pela Equação.

$$R^2 = 1 - \frac{\sum_{i=1}^N (\gamma_I - \gamma_{I2})^2}{\sum_{i=1}^N (\gamma_I - \gamma_{I2})^2} \quad (13)$$

Onde:

- γ_I é o valor da real da amostra;
- γ_{I2} é o valor previsto pela rede neural;
- N é o tamanho da amostra.

4.4.3 Erro Quadrático Médio

O Erro Quadrático Médio (MSE) é uma das métricas de erro mais comuns, especialmente em problemas de regressão. Ele calcula a média dos quadrados das diferenças entre os valores previstos e os valores reais, que pode ser calculado seguindo a Equação.

$$MSE = \frac{1}{N} \sum_{I=1}^N (\gamma_I - \gamma_{I2})^2 \quad (14)$$

4.4.4 Raiz do Erro Quadrático Médio

O RMSE é simplesmente a raiz quadrada do MSE, e seu valor tem as mesmas unidades dos dados originais, o que pode facilitar a interpretação dos resultados. Calculado pela Equação.

$$RMSE = \sqrt{\frac{1}{N} \sum_{I=1}^N (\gamma_I - \gamma_{I2})^2} \quad (15)$$

4.5 Teste de hipótese

Para realizar a análise de significância da relação entre a concentração de cloro e o potencial de oxidação e redução (ORP), uma metodologia estatística foi adotada com o objetivo de validar a existência de uma possível dependência entre essas variáveis. A validação dos dados tem como objetivo determinar se a concentração de cloro pode ser considerada um fator preditivo confiável para o comportamento do ORP, com implicações práticas para processos químicos e industriais que dependem do equilíbrio redox. Onde as hipóteses estabelecidas foram.

- Hipótese nula H_0 : Não há relação significativa entre a concentração de cloro e o potencial de oxidação e redução, onde seu coeficiente de regressão é próximo de 0;
- Hipótese alternativa H_1 : Existe uma relação significativa entre a concentração de cloro e o potencial de oxidação (ORP). Ou seja, o coeficiente de regressão é diferente de zero.

Para testar as hipóteses formuladas, foi empregada uma análise de regressão linear simples, considerando a concentração de cloro como variável independente e o ORP como

variável dependente.

O coeficiente de regressão foi submetido a um teste “t”, avaliando a probabilidade de que ele seja diferente de zero. A hipótese nula foi rejeitada se o valor-p obtido fosse menor que o nível de significância adotado ($\alpha = 0,05$).

A proporção da variabilidade total do ORP explicada pela concentração de cloro foi avaliada através do R^2 elevado reforça a existência de uma relação linear significativa. Os resultados da análise foram interpretados considerando a magnitude e a significância estatística do coeficiente de regressão, o valor-p correspondente ao modelo.

5 RESULTADOS E DISCUSSÕES

Nesta seção são apresentados os resultados obtidos através da aplicação de redes neurais artificiais e modelos de regressão lineares no excel, utilizado como estudo, bem como a validação estatística dos dados amostrais. Além disso, é discutido a comparação entre modelos de regressão, justamente com as métricas associadas a cada modelo de estudo.

5.1 Dados amostrais

Os dados obtidos na volumetria das amostras são apresentados nesta seção, com eles é possível calcular a concentração de cloro em cada ponto de diluição utilizando-se da Equação 11, bem como mensurar o potencial de oxirredução utilizando o analisador mencionado na Tabela 4, os dados estão apresentados na Tabela 5.

Tabela 5 - Dados amostrais NaClO

Volume de Titulante	Cloro Livre (g.L ⁻¹)	ORP (mV)
46,59	165,20	499,90
45,09	159,90	500,20
41,99	148,90	502,70
40,47	143,50	503,30
39,00	138,30	504,50
38,24	135,60	504,90
37,17	131,80	505,30
35,50	125,90	506,90
33,36	118,30	507,10
32,01	113,50	508,70
30,01	106,40	511,50
28,37	100,60	512,30
27,44	97,30	513,50

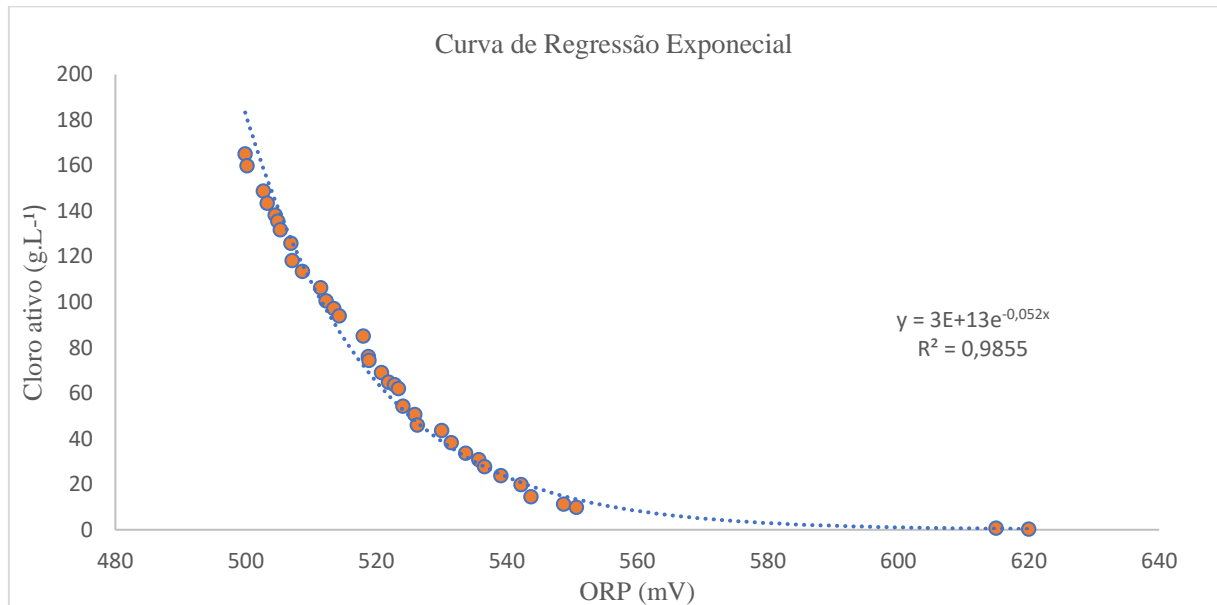
26,51	94,00	514,30
24,00	85,10	518,00
21,49	76,20	518,80
21,01	74,50	518,90
19,49	69,10	520,80
18,30	64,90	521,90
17,96	63,70	522,80
17,51	62,10	523,40
15,31	54,30	524,10
14,30	50,70	525,90
13,00	46,10	526,30
12,30	43,60	530,00
10,80	38,30	531,50
9,50	33,70	533,70
8,71	30,90	535,70
7,81	27,70	536,60
6,71	23,80	539,10
5,61	19,90	542,20
4,09	14,50	543,70
3,19	11,30	548,70
2,79	9,90	550,70
0,20	0,70	615,00
0,08	0,30	620,00

5.2 Modelos de Regressão

5.2.1 Modelo Exponencial

Com base nos dados apresentados na tabela 5, é possível então estabelecer um modelo de regressão simples, linear ou não, utilizando o software Excel, conforme Figura 6.

Figura 6 - Modelo de Regressão (Excel)



Fonte: Autor, 2024.

Na Figura 6, pode-se observar que o melhor ajuste pra o conjunto de dados é do tipo exponencial, representado pela Equação 14.

$$y = 3 \cdot 10^{13} e^{-0,052x} \quad (14)$$

O coeficiente de determinação ($R^2 = 0,985$), indica que o modelo exponencial ajusta os dados com alta precisão mesmo utilizando-se de uma ferramenta menos robusta, explicando 98,55% da variabilidade observada no cloro ativo em função do ORP. O comportamento do gráfico sugere uma relação inversa entre o ORP e a concentração de cloro ativo (g.L⁻¹). À medida que o ORP aumenta, a quantidade de cloro ativo diminui de maneira exponencial, indicando que em condições de maior potencial oxidativo, há uma redução mais acentuada na concentração de cloro ativo.

5.2.2 Rede Neural

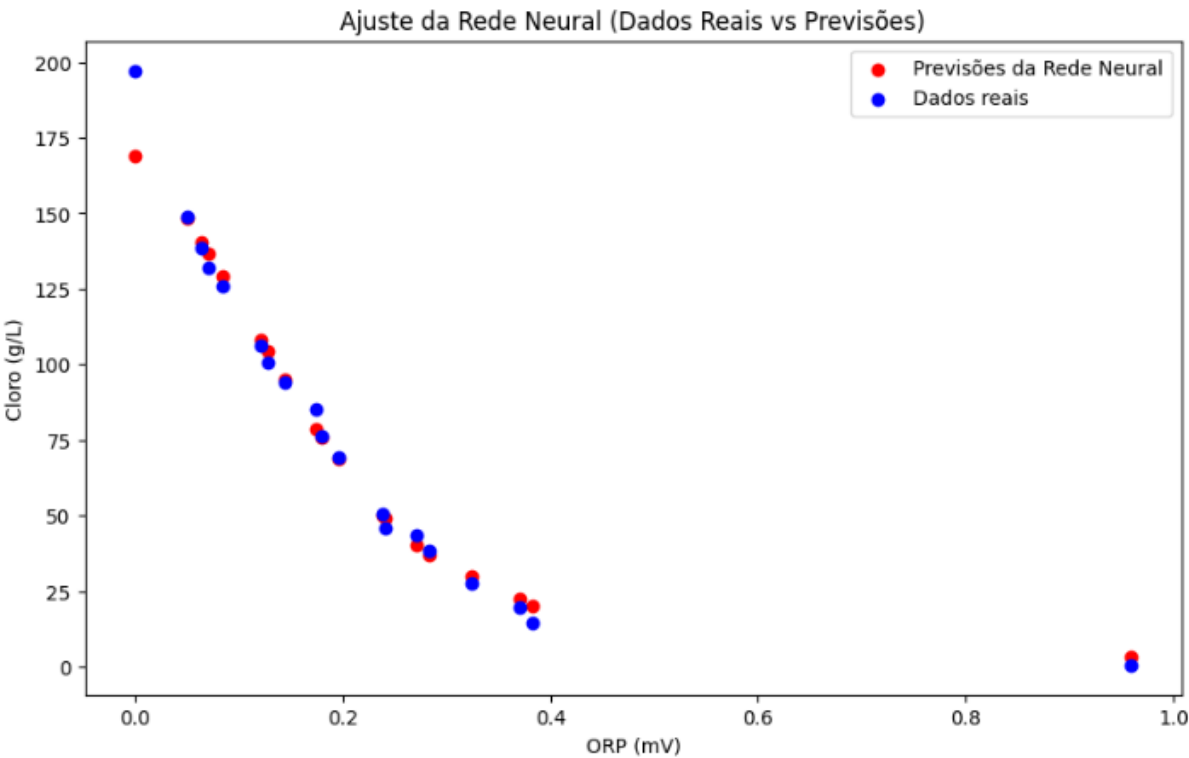
O modelo de regressão foi testado utilizando redes neurais artificiais em quatro condições distintas, com o objetivo de avaliar o impacto do número de neurônios na camada oculta no desempenho do modelo. Para determinar os parâmetros do modelo, considerou-se o tamanho do conjunto de dados disponível, que possui 36 amostras, sendo este o número máximo de neurônios recomendado para evitar o sobreajuste e manter a generalização do modelo.

As condições testadas consistiram em variar o número de neurônios na camada oculta entre 10, 20 e 36, além de uma quarta condição que visava simular um cenário de sobreajuste. Nesta última condição, o número de neurônios foi ampliado significativamente, utilizando 64 neurônios em ambas as camadas ocultas. Esse cenário foi implementado para investigar como o modelo se comporta em situações de complexidade excessiva, onde o número de parâmetros excede significativamente o tamanho do conjunto de dados, o que pode levar ao sobreajuste e à perda de capacidade preditiva em dados não vistos.

A escolha dos números de neurônios testados foi baseada em boas práticas de modelagem de redes neurais, que sugerem manter a complexidade do modelo alinhada ao tamanho do conjunto de dados e à natureza do problema (ICMC, 2024).

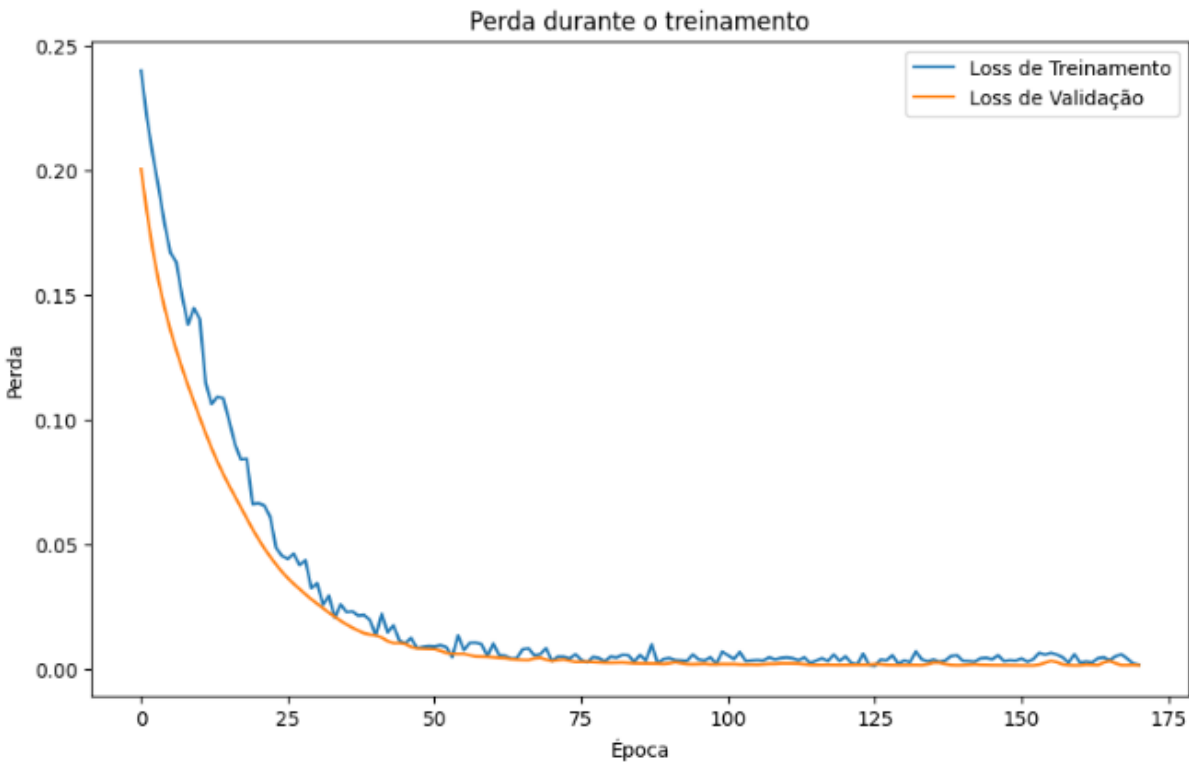
Os resultados obtidos em cada modelo estão apresentados nas figuras a seguir.

Figura 7 – Curva de regressão (10 neurônios)



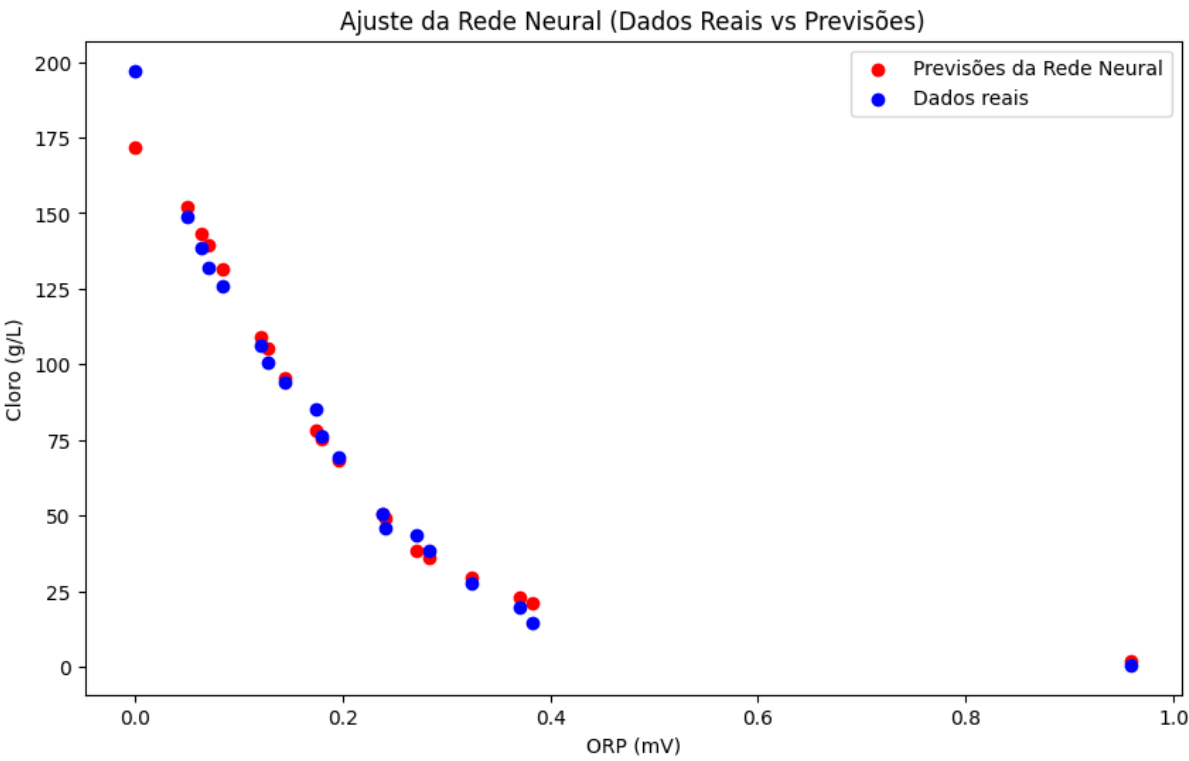
Fonte: Autor, 2024.

Figura 8 – Loss (10 neurônios)



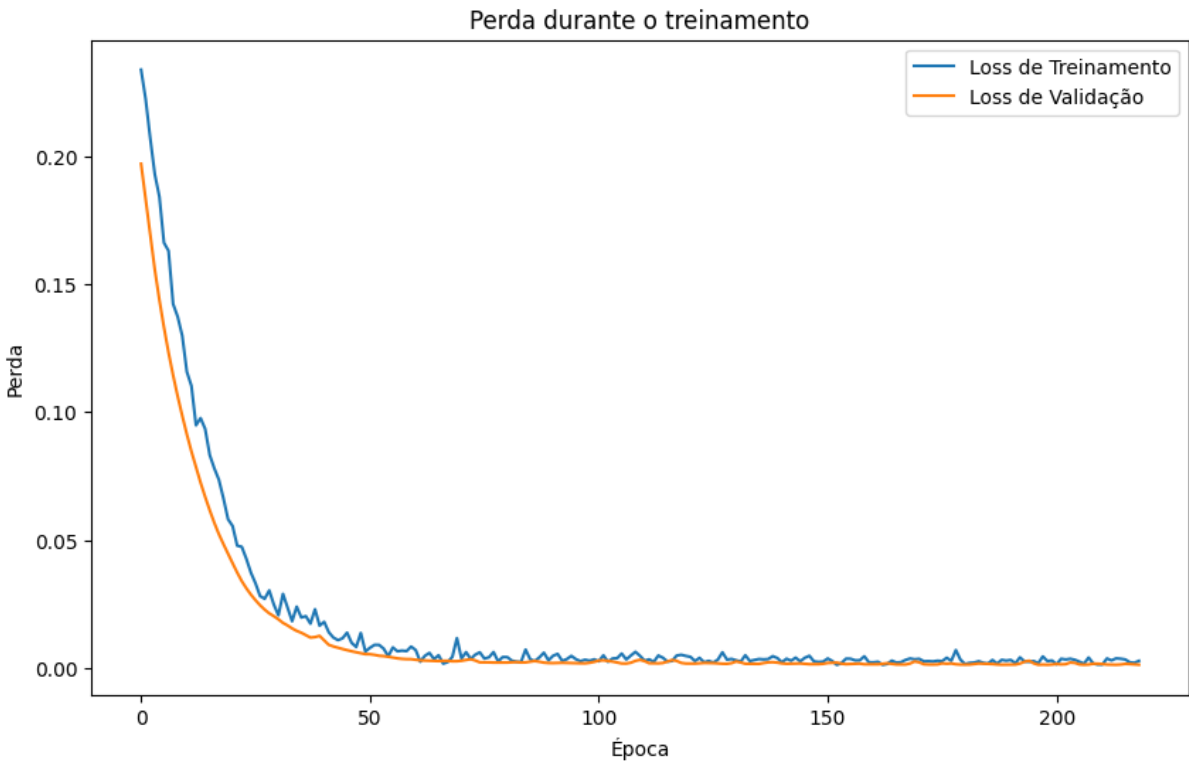
Fonte: Autor, 2024.

Figura 9 - Curva de regressão (20 neurônios)



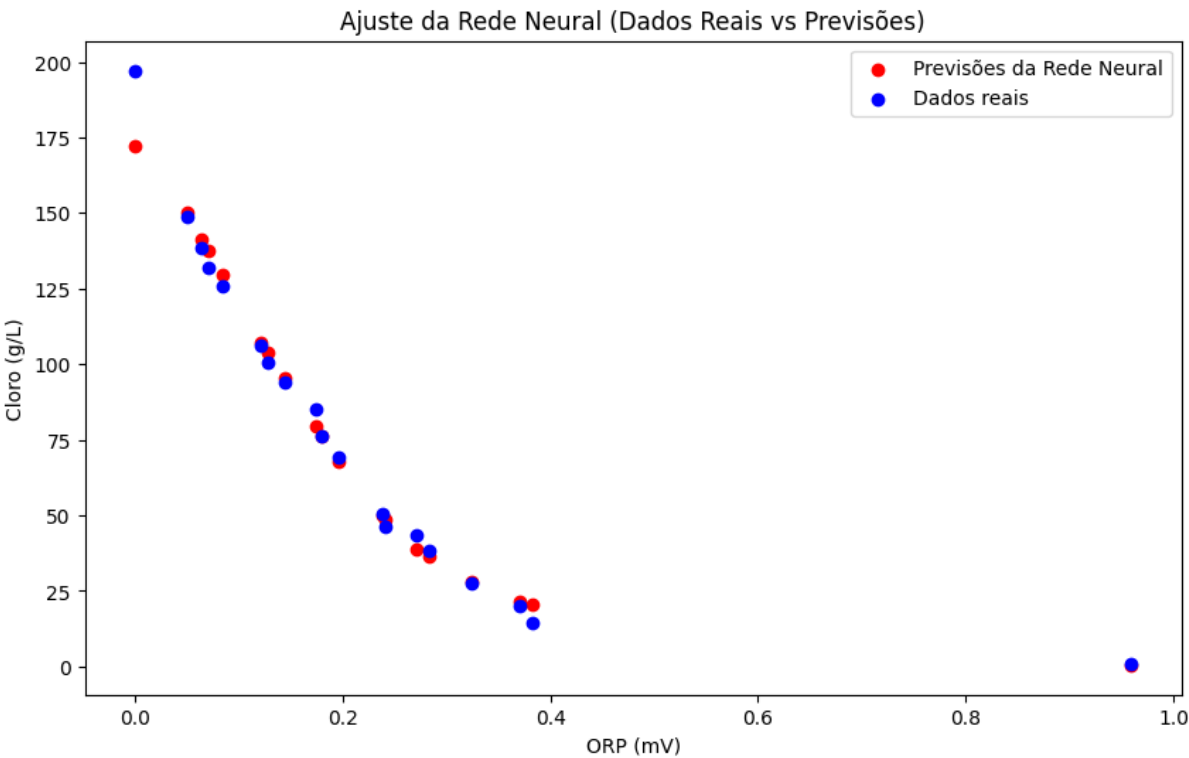
Fonte: Autor, 2024.

Figura 10 - Loss (20 neurônios)



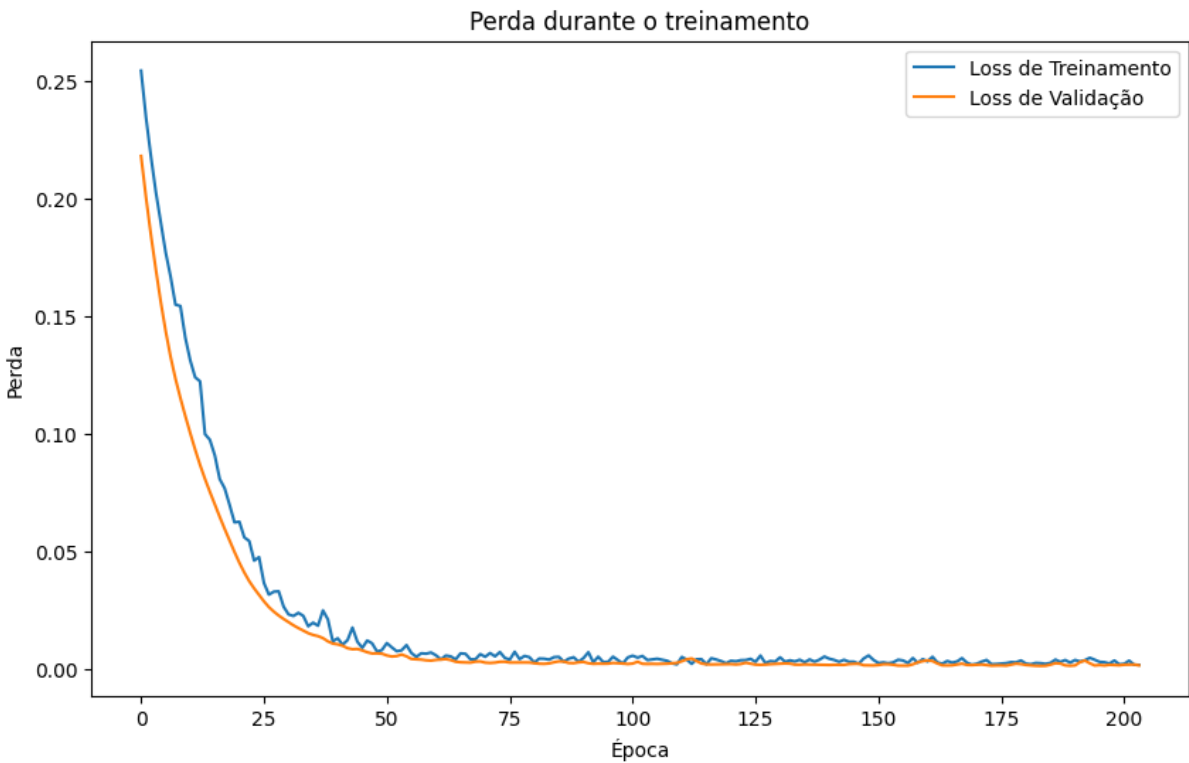
Fonte: Autor, 2024.

Figura 11 - Curva de regressão (36 neurônios)



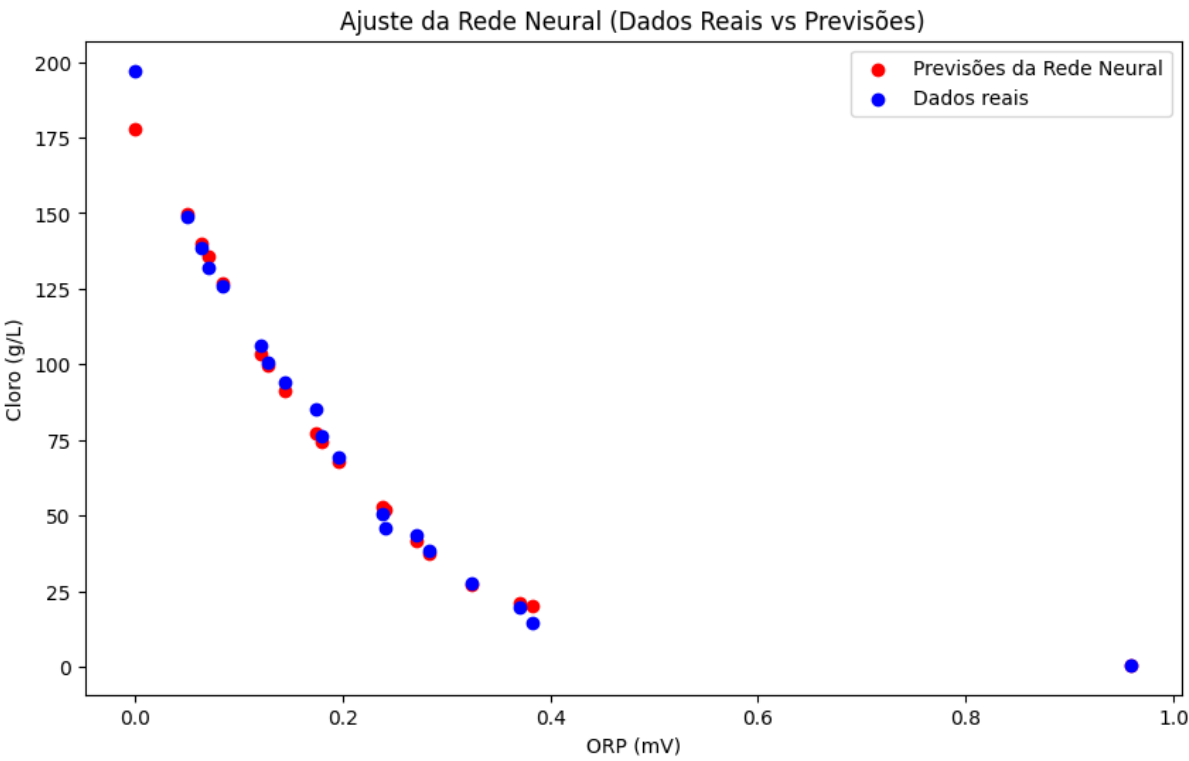
Fonte: Autor, 2024.

Figura 12 - Loss (36 neurônios)



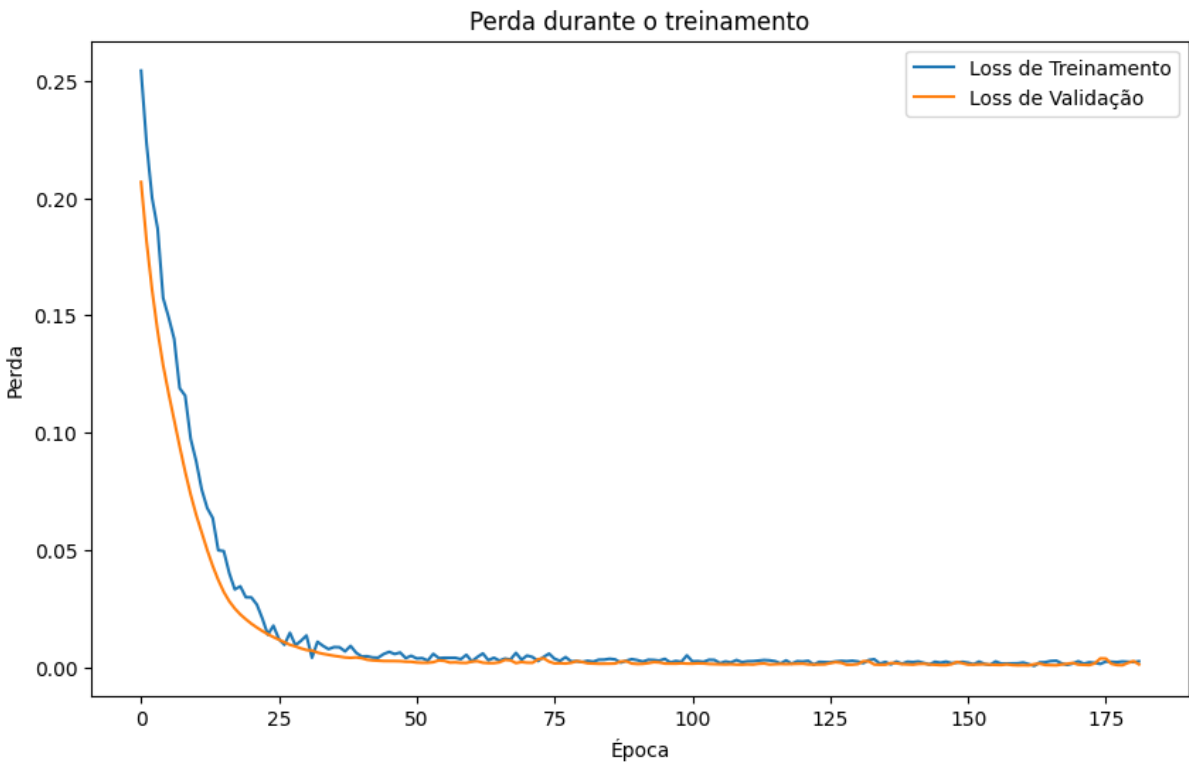
Fonte: Autor, 2024.

Figura 13 - Curva de regressão (64 neurônios)



Fonte: Autor, 2024.

Figura 14 - Loss (64 neurônios)



Fonte: Autor, 2024.

Entre o conjunto de figuras, alinhadas com os resultados exposto no Anexo I, é possível afirmar que o aumento progressivo de neurônios na camada oculta aumentam a precisão de previsão da RNA. As Figuras 13 e 14, que possuem os resultados na configuração de 64 Neurônios, quando comparada com as demais Figuras (7 a 12), apresentam nitidamente uma menor perda em validação com os dados reais.

5.3 Comparação Entre Modelos de Regressão

Nesta análise, são avaliadas métricas como acurácia, coeficiente de determinação (R^2), erro médio absoluto (MSE) e erro quadrático médio (RMSE) para determinar o desempenho dos modelos de regressão, apresentados na Tabela 6.

Tabela 6 - Métricas dos modelos de regressão

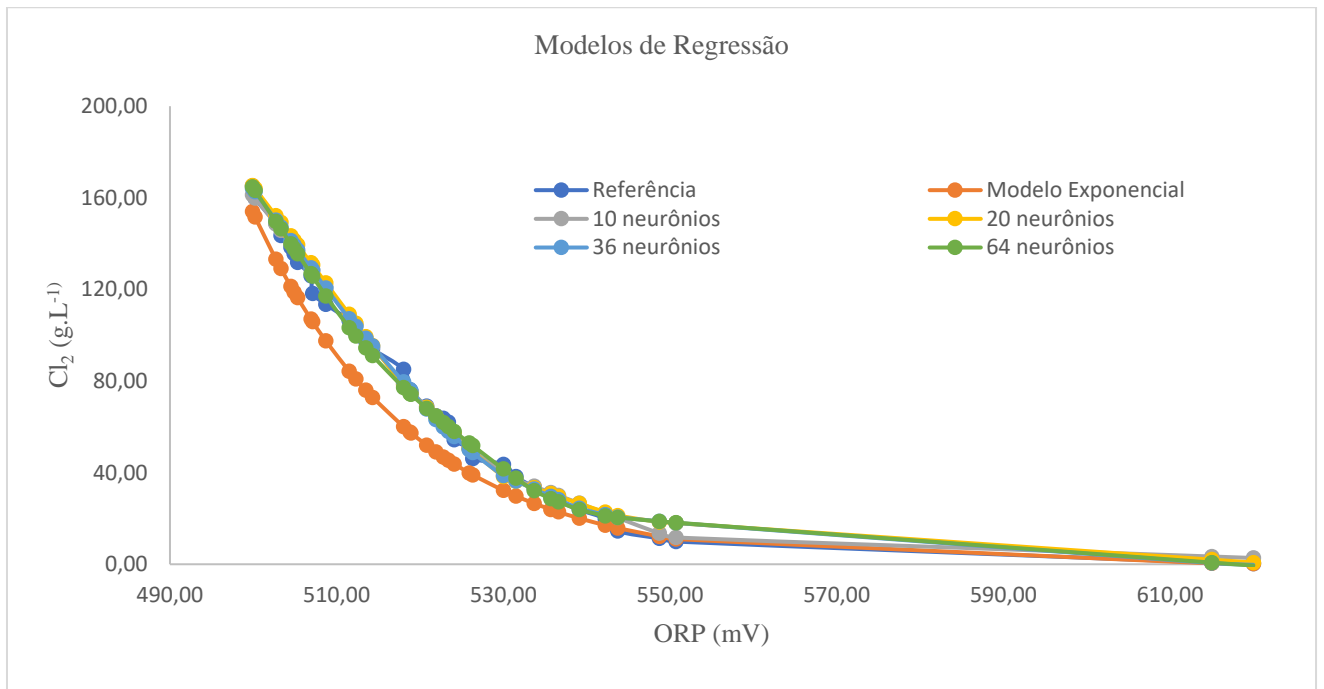
Métricas	Modelo Exponencial	RNA 10 Neurônios	RNA 20 Neurônios	RNA 36 Neurônios	RNA 64 Neurônios
ACURÁCIA	97,22%	94,74%	95,12%	96,35%	99,99%
R^2	0,985	0,981	0,982	0,985	0,998
MSE	0,073	0,020	0,019	0,018	0,016
RMSE	0,071	0,034	0,033	0,032	0,027

Fonte: Autor, 2024.

Os resultados expostos na Tabela 6, indicam que a capacidade preditiva não cresce linearmente com o número de neurônios, mas sim atinge um equilíbrio ideal, nesse caso por exemplo o ajuste no Excel é não linear, por se tratar de uma curva exponencial, demonstra uma boa acurácia de 97,22%.

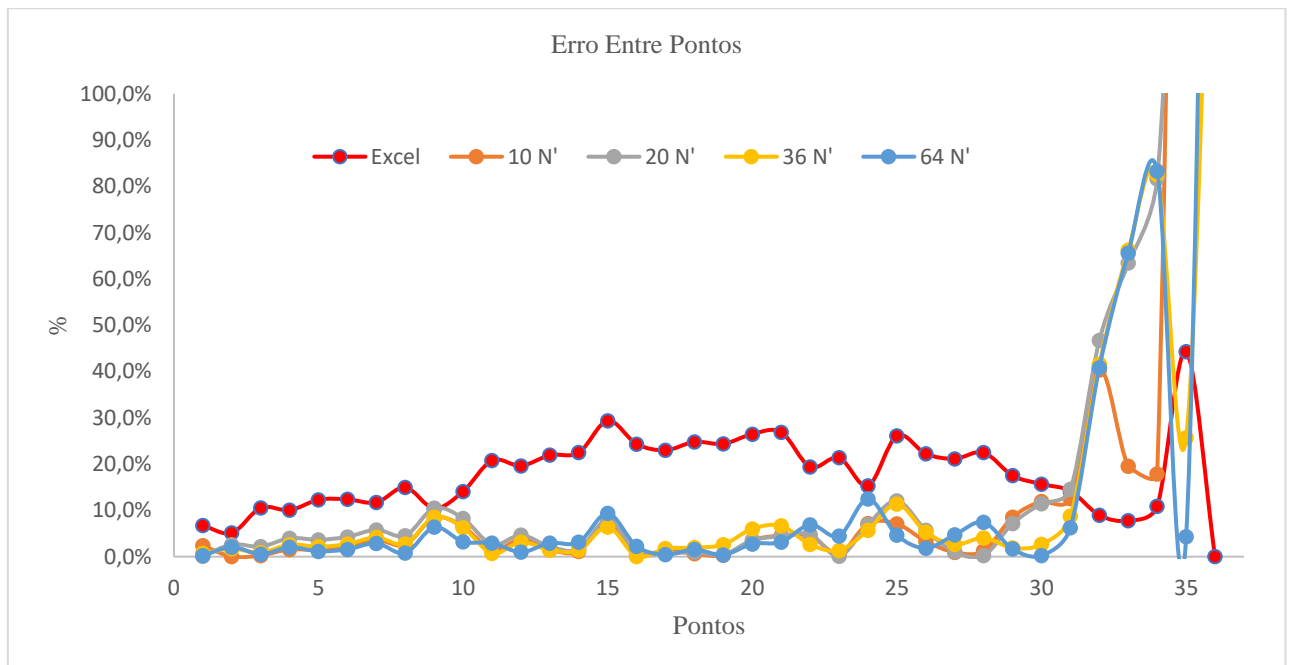
O Anexo I, demonstra os dados de concentração de Cl_2 gerados por todos os modelos e ORP associado, nas Figura 15 e 16 pode-se visualizar a distribuição de todos os modelos em comparação com a referência.

Figura 15 - Modelos de Regressão



Fonte: Autor, 2024.

Figura 16 - Erro absoluto entre modelos



Fonte: Autor, 2024

Nas Figura 15 e 16 observa-se que as previsões geradas pelos modelos de redes neurais apresentam menores variações em relação aos valores de referência.

Com 10, 20 e 36 neurônios o modelo apresenta um desempenho relativamente consistente com os valores de referência, mas sua capacidade de capturar padrões mais complexos pode ser limitada, como evidenciado um erro maior nos pontos em que se tem concentrações mais baixas de Cl_2 , porém 20 e 36 neurônios, percebe-se previsões razoavelmente próximas da referência no geral.

À medida que o número de neurônios aumenta, o modelo captura menos variações nos dados, mas também há indícios de que ele começa a ajustar padrões específicos dos dados de treinamento, reduzindo sua capacidade de generalização. Utilizando 64 neurônios, existe desvios mais pronunciados nos pontos finais (30 a 35) quando se observa o erro associado na curva em azul exposta na Figura 16. Isso é um reflexo do sobreajuste, onde o excesso de parâmetros compromete a performance em dados fora do conjunto de treinamento.

Os resultados obtidos pelo ajuste realizado Excel fornecem um ponto de referência simples e direto, menos sofisticado e apresenta um erro significativamente maior. Os modelos de redes neurais, especialmente nas configurações intermediárias 20 e 36 neurônios, demonstram um potencial significativo para superar as previsões não-lineares do Excel, mas com a vantagem de capturar relações não lineares entre as variáveis. As maiores discrepâncias entre os valores previstos e os de referência ocorrem em regiões de baixas concentrações de Cl_2 , como evidenciado pelas últimas linhas da tabela 7. Essa discrepância pode estar associada à dificuldade dos modelos em aprender adequadamente em regiões de menor densidade de dados. Técnicas como a ponderação de amostras ou a adição de dados específicos nessas regiões poderiam melhorar o desempenho do modelo.

5.4 Validação Estatística

A análise estatística da relação entre a concentração de cloro e o potencial de oxidação e redução (ORP) foi conduzida com o objetivo de validar a existência de uma dependência significativa entre essas variáveis, permitindo avaliar se o cloro pode ser considerada um fator preditivo confiável para o comportamento do potencial de redox. Dessa forma, foram formuladas as seguintes hipóteses:

- Hipótese nula (H_0): Não há relação significativa entre a concentração de cloro e o ORP, ou seja, o coeficiente de regressão é próximo de zero;
- Hipótese alternativa (H_1): Existe uma relação significativa entre a concentração de cloro

e o ORP, com coeficiente de regressão diferente de zero.

Para testar essas hipóteses, foi aplicada uma análise de regressão linear simples, em que a concentração de cloro foi considerada a variável independente e o ORP a variável dependente.

O coeficiente de regressão foi submetido ao teste t, que avalia a probabilidade de o coeficiente ser significativamente diferente de zero, adotando-se um nível de significância ($\alpha = 0,05$.), os resultados estatísticos, estão expostos na Tabela 7.

Tabela 7 - Valores estatísticos

Estatística	Valor
Coeficiente de Regressão	-0,434
Erro Padrão	0,060
Valor “t”	-7,010
Valor-p	0,001
Coeficiente de determinação	0,975

Fonte: Autor, 2024.

Os resultados obtidos, incluindo a magnitude e a significância estatística do coeficiente de regressão e o valor-p associado ao modelo, foram analisados para determinar a influência direta da concentração de cloro sobre o comportamento do ORP.

Na Tabela 7, o valor-p observado foi inferior ao nível de significância adotado ($\text{valor-p} < \alpha = 0,05$). Dessa forma, a hipótese nula (H_0) é rejeitada, e a hipótese alternativa (H_1) é aceita, indicando a existência de uma relação estatisticamente significativa entre a concentração de cloro e o potencial de oxidação e redução (ORP).

Esse resultado sugere que o pH da solução exerce uma influência relevante no comportamento do potencial de oxirredução, conforme discutido por Góngora em 2024. De acordo com suas abordagens, a relação entre essas variáveis apresenta um comportamento logarítmico, o que reforça a complexidade e a interdependência entre a concentração de cloro, o pH e o ORP em soluções aquosas.

6 CONCLUSÃO

O presente estudo abordou a relação entre a concentração de cloro ativo e o potencial de oxidação e redução (ORP), utilizando técnicas de regressão disponíveis no Excel, redes neurais artificiais (RNA) e validação estatística. Os resultados obtidos confirmaram a existência de uma relação significativa entre as variáveis, reforçando que a concentração de cloro pode ser considerada um fator preditivo relevante para o comportamento do ORP.

A partir dos modelos desenvolvidos, verificou-se que o modelo de regressão exponencial com Excel apresentou um coeficiente de determinação ($R^2 = 0,985$), explicando 98,55% da variabilidade nos dados. Demonstrando ser uma solução simples e eficaz, mas limitada ao capturar relações precisas entre ORP e concentração de cloro.

Para os Modelos de Redes Neurais Artificiais, a performance dos modelos com 20 e 36 neurônios mostrou-se superior em capturar as relações não lineares entre as variáveis, apesar da alta acurácia (até 99,99% com 64 neurônios), observou-se sobreajuste em cenários mais complexos, comprometendo a generalização do modelo em regiões de baixa concentração de cloro, o equilíbrio ideal entre precisão e generalização foi observado com 20 e 36 neurônios, com menor erro quadrático médio ($MSE = 0,018$).

Os resultados estatísticos validaram a significância da relação entre as variáveis: o teste t do coeficiente de regressão apresentou valor-p = 0,001, inferior ao nível de significância adotado ($\alpha = 0,05$), levando à rejeição da hipótese nula (H_0). Esse resultado corrobora a existência de uma dependência significativa entre a concentração de cloro e o ORP. Além disso, a análise reforçou o impacto do pH da solução sobre a estabilidade do ORP, conforme apontado por Góngora em 2024, evidenciando um comportamento logarítmico entre essas variáveis.

Por fim, os modelos propostos mostraram-se promissores como alternativas viáveis e econômicas para substituir métodos tradicionais de análise titulométrica, especialmente em processos industriais que exigem controle rigoroso e tempo-resposta ágil. A implementação de técnicas baseadas em RNAs, aliada ao controle preciso do pH e ao uso de equipamentos calibrados, pode elevar significativamente a confiabilidade e a qualidade na produção de hipoclorito de sódio.

7 REFERÊNCIAS BIBLIOGRÁFICAS

ABICLOR. **Manual de Hipoclorito de Sódio**. Associação Brasileira da Indústria de Álcalis, Cloro e Derivados. Disponível em: <http://www.abiclor.com.br>. Acesso em 05/11/2024.

ASSIS, J. P., SOUDA, R. P., LINHARES, P. F. **Teste de Hipótese estatística**. Universidade Federal Rural do Semi-Árido, 2020.

BARBETTA, P. A. **Estatística aplicada às Ciências Sociais**. 7 ed. Universidade Federal de Santa Catarina - UFSC, 2010.

CHLORINE SCRUBBING SYSTEMS, ed. 3; **Pamphlet 089**; The Chlorine Institute: Arlington, VA, 2008.

COELHO, M. **Fundamentos de Redes Neurais**. Disponível em: <https://www2.decom.ufop.br/imobilis/fundamentos-de-redes-neurais/>. Acesso em 20/11/2024.

DUNLAP, W. P.; LANDIS, R. S. **Interpretativos of multiple regression biorrede from factor analysis and canonical correlativo**. The Jornal of General Psychology, 125, p. 397-407, 1998.

ICMC. **Instituto de Ciências Matemáticas e Computação**. Redes Neurais Artificiais. Disponível em: <https://sites.icmc.usp.br/andre/research/neural/>. Acesso em 20/12/2024.

FERNEDA, E. **Redes neurais e sua aplicação em sistemas de recuperação de informação**. Scielo Brasil, p. 25–30, Abril. 2006.

GONGORA, K. et al. The influence of pH on the efficacy of oxidation-reduction potential (ORP) to predict chlorine disinfection of surrogate bacteria, *Escherichia coli* O157: H7 and *Listeria monocytogenes* in oxidant demand free conditions and fresh produce wash water. **ELSEVIER**, v. 121, 2024.

HAHS-VAUGHN, D. L. **Multiple Linear Regression. Statistical Concepts**, v. 404, p. 527–599, 2020.

HINE, F. **Electrode Processes and Electrochemical Engineering**, Plenum Press. New York, 1985.

JUNIOR, R. N. J. **Modelagem Matemática de um Processo Industrial de Produção de Cloro e Soda por Eletrólise de Salmoura Visando a sua Otimização**. Dissertação de Mestrado. Escola Politécnica da Universidade de São Paulo, Departamento de Engenharia Química. São Paulo, 2006.

LOPES, A. **Estudo da degradação do hipoclorito de sódio com o níquel metálico**. 2006. 58. Dissertação (Mestrado) - Curso de Engenharia Química, Centro de Ciências e Tecnologia, Universidade Federal de Campina Grande, Campina Grande, 2006. Cap. 2

LUCCA, L. **Controle de qualidade do hipoclorito de sódio no processo de produção**. TCC (graduação) - Universidade Federal de Santa Catarina. Centro de Ciências Físicas e Matemáticas. Curso de Química - UFSC, 2006.

MARTINS, E.G.M. **Regressão linear simples**, Rev. Ciência Elem., V7(3):045, 2019.

O'BRIEN, T. F., BOMMARAJU, T. V., HINE, F. **Handbook of Chlor-Alkali Technology**. Springer, New York, 2005.

PYE, D. J. **The control of bleach manufacture by the oxidation potential**. Journal of Electrochemical Society, vol. 97, nº 8, 1950.

SHREVE, R.N; BRINK Jr, J.A. **Indústrias de Processos Químicos** - 4ª ed. Guanabara Koogan, S.A. Rio de Janeiro, 1980

SKOOG, D. A. **Fundamentos de Química Analítica**. 8ª ed. São Paulo: Pioneira Thomson Learning, 2006.

TABACHNICK, B.; FIDELL, L. S. **Using multivariate statistics** (3a ed.). New York: Harper Collins, 1996.

UNIPAR, **Apresentação industrial**. Disponível em: <https://unipar.com>. Acesso em 11/11/2024.

WANJIH, L. **Processo de liquefação secundária de cloro residual em liquefação de cloro e método de configuração de equipamento**. GRUPO JIANGSU SUHUA CO., LTD. Disponível em: <https://patents.google.com/patent/CN1410730A/en>. Acesso em 08 de novembro de 2024.

8 ANEXOS

ANEXO I - Resultados dos modelos de regressão para concentração de cloro

ORP (mV)	Cloro Livre (g.L ⁻¹)	Cl ₂ (g.L ⁻¹)	Cl ₂ (g.L ⁻¹)	Cl ₂ (g.L ⁻¹)	Cl ₂ (g.L ⁻¹)	Cl ₂ (g.L ⁻¹)
Referência	Referência	Modelo Exponencial	10 neurônios	20 neurônios	36 neurônios	64 neurônios
499,90	165,20	154,07	161,28	165,37	164,03	164,82
500,20	159,90	151,69	159,93	164,08	162,61	163,20
502,70	148,90	133,20	148,52	152,14	150,21	149,66
503,30	143,50	129,10	145,76	149,21	147,23	146,41
504,50	138,30	121,29	140,23	143,35	141,27	139,91
504,90	135,60	118,80	138,39	141,40	139,29	137,74
505,30	131,80	116,35	136,54	139,45	137,30	135,58
506,90	125,90	107,06	129,17	131,63	129,36	126,91
507,10	118,30	105,96	128,25	130,66	128,37	125,83
508,70	113,50	97,50	120,88	122,85	120,61	117,17
511,50	106,40	84,29	107,98	109,18	107,17	103,24
512,30	100,60	80,85	104,29	105,27	103,78	99,61
513,50	97,30	75,96	98,76	99,41	98,69	94,43
514,30	94,00	72,87	95,07	95,51	95,30	91,07
518,00	85,10	60,11	78,78	78,34	79,63	77,14
518,80	76,20	57,66	75,89	75,41	76,24	74,50
518,90	74,50	57,36	75,53	75,06	75,81	74,17
520,80	69,10	51,97	68,67	68,43	67,76	68,04
521,90	64,90	49,08	64,69	64,60	63,20	64,64

522,80	63,70	46,83	61,44	61,46	59,88	61,94
523,40	62,10	45,40	59,27	59,36	57,97	60,14
524,10	54,30	43,77	56,73	56,92	55,73	58,05
525,90	50,70	39,86	50,35	50,64	50,00	52,97
526,30	46,10	39,04	49,39	49,29	48,72	51,85
530,00	43,60	32,21	40,51	38,33	38,64	41,58
531,50	38,30	29,79	37,03	36,12	36,28	37,59
533,70	33,70	26,57	34,02	33,35	32,81	32,11
535,70	30,90	23,95	31,29	30,82	29,65	28,60
536,60	27,70	22,85	30,06	29,68	28,23	27,24
539,10	23,80	20,07	26,64	26,53	24,44	23,86
542,20	19,90	17,08	22,40	22,80	21,64	21,14
543,70	14,50	15,80	20,35	21,27	20,53	20,42
548,70	11,30	12,18	13,51	18,47	18,78	18,71
550,70	9,90	10,98	11,67	17,98	18,07	18,15
615,00	0,70	0,39	3,38	2,01	0,52	0,73
620,00	0,30	0,30	2,72	0,75	-0,25	-0,61

Fonte: Autor, 2024.