



**UNIVERSIDADE FEDERAL DE SÃO PAULO
CAMPUS DIADEMA**



ORLANDO PARDINI GARDEZANI

**SENSOR VIRTUAL PARA MONITORAMENTO DA FERMENTAÇÃO DE
HIDROMEL COM ADIÇÃO DE EXTRATO DE CASCA DE JABUTICABA
APLICANDO REDES NEURAIS ARTIFICIAIS**

DIADEMA

2025

ORLANDO PARDINI GARDEZANI

**SENSOR VIRTUAL PARA MONITORAMENTO DA FERMENTAÇÃO DE
HIDROMEL COM ADIÇÃO DE EXTRATO DE CASCA DE JABUTICABA
APLICANDO REDES NEURAIS ARTIFICIAIS**

Trabalho de Conclusão de Curso apresentado
como exigência parcial para obtenção do título
de Bacharel em Engenheira Química ao
Instituto de Ciências Ambientais, Químicas e
Farmacêuticas da Universidade Federal de São
Paulo – Campus Diadema.

Orientador: Prof. Dr. Tiago Dias Martins.

Coorientador: Prof. Dr. Rafael Ramos de
Andrade.

DIADEMA

2025

À minha avó, que sempre acreditou em mim, me incentivou a trilhar o caminho até a universidade e, com sua sabedoria e amor, me ofereceu o suporte essencial para minha vida.

Aos meus tios, que me proporcionaram oportunidades e construíram, com generosidade e cuidado, o futuro que hoje vivencio.

E à minha versão mais nova, que sonhou com determinação e consolidou tão bem esse sonho, inspirando quem sou hoje a jamais desistir, independentemente dos desafios enfrentados.

AGRADECIMENTOS

A Deus por fazer-me quem sou, conceder-me um propósito, fornecer-me sempre o que necessito e livrar-me de todo o mal.

Aos meus amigos Halex, Lucas, Jefferson, Victor, Caroline e Giovanna, meu profundo agradecimento por estarem ao meu lado ao longo dessa jornada. Vocês compartilharam comigo os pequenos e grandes triunfos que marcaram essa etapa da minha vida.

Ao meu orientador e amigo, Tiago, minha eterna gratidão pelo conhecimento compartilhado. Sua orientação foi essencial em minha vida. Mais do que um orientador, você foi um verdadeiro amigo.

A todos vocês, meu muito obrigado por fazerem parte dessa trajetória tão especial.

“O desconforto de estudar é apenas temporário, mas o arrependimento de não estudar pode durar a vida inteira”.

(Miyuki Shirogane)

RESUMO

O hidromel é uma bebida alcoólica fermentada à base de água e mel produzida pela ação de leveduras, normalmente linhagens de *Saccharomyces cerevisiae*, sobre carboidratos como glicose e frutose. O produtor, muitas vezes, se depara com empecilhos derivados da falta de conhecimento aprofundado e de controle assertivo de parâmetros importantes do processo. Sendo assim, o presente trabalho teve como objetivo a elaboração de um sensor virtual baseado em Redes Neurais Artificiais capaz de predizer as concentrações de células (X), açúcares redutores (S) e etanol (P) durante a fermentação para a produção de hidromel com adição polpa de casca de jaboticaba, utilizado para atribuir propriedades antioxidantes e sabor à bebida. Com isso, foram selecionadas como variáveis de entrada: temperatura operacional (T), pH, concentração de sólidos solúveis totais (“S.S.” medido em °Brix) e a densidade óptica (D.O.). Para a obtenção e simulação da rede com treinamento supervisionado, utilizou-se dados experimentais de fermentações realizadas previamente. As redes neurais foram treinadas e validadas em diversas combinações e a configuração de maior potencial foi retreinada utilizando o algoritmo de otimização baseado na Regularização Bayesiana. A arquitetura da RNA continha 4 neurônios na entrada, 40 e 25 nas camadas ocultas, respectivamente, e 3 neurônios na camada de saída. As funções de ativação foram *poslin softmax tansig* nas camadas ocultas e de saída, respectivamente. Essa configuração apresentou a menor média dos Erros Percentuais Médios Absolutos de 9.34% e alto coeficiente de correlação médio ($\rho=0.92$), indicando excelente ajuste aos dados.

Palavras-chave: redes neurais artificiais; fermentação alcoólica; treinamento supervisionado; modelo de regressão; otimização de processo químico.

ABSTRACT

Mead is a fermented alcoholic beverage made from water and honey, produced through the action of yeast, typically *Saccharomyces cerevisiae* strains, on carbohydrates such as glucose and fructose. Producers often encounter challenges stemming from a lack of in-depth knowledge and insufficient control over critical process parameters. Therefore, the objective of this study was to develop a virtual sensor based on Artificial Neural Networks (ANNs) capable of predicting the concentrations of cells (X), reducing sugars (S), and ethanol (P) during the fermentation process for mead production, enriched with Jabuticaba peel pulp to impart antioxidant properties and enhance the beverage's flavor. The selected input variables were operational temperature (T), pH, total soluble solids concentration (S.S., measured in °Brix), and optical density (DO). Experimental data from previously conducted fermentations were used for supervised training and simulation of the neural network. The neural networks were trained and validated across various configurations, and the most promising setup was retrained using the Bayesian Regularization optimization algorithm. The ANN architecture consisted of 4 neurons in the input layer, 40 and 25 neurons in the hidden layers, respectively, and 3 neurons in the output layer. The activation functions employed were *poslin*, *softmax*, and *tansig* in the hidden and output layers, respectively. This configuration achieved the lowest average Mean Absolute Percentage Error (MAPE) of 9.34% and a high average correlation coefficient ($\rho = 0.92$), indicating an excellent fit to the data.

Keywords: artificial neural networks; alcoholic fermentation; supervised training; regression model; chemical process optimization.

LISTA DE FIGURAS

Figura 1 – Mapa do Mel de abelha - Valor da produção (Mil Reais).....	6
Figura 2 - Processo geral para obtenção do hidromel.....	9
Figura 3 – Processo fermentativo do hidromel.....	11
Figura 4 – Pigmentos derivados de Antocianinas.....	14
Figura 5 – Neurônio Biológico.....	19
Figura 6 – Neurônio Matemático.....	20
Figura 7 – Rede com duas camadas intermediárias.....	22
Figura 8 – Funções de ativação empregadas no modelo.....	27
Figura 9 – Processo geral de obtenção de uma rede neural.....	34
Figura 10 – Rede neural que simboliza o processo	38
Figura 11 – Ajuste suave dos dados gerados para S.S. contra os dados experimentais em função do tempo.	40
Figura 12 – Ajuste suave dos dados gerados para pH contra os dados experimentais em função do tempo.....	40
Figura 13 – Ajuste suave dos dados gerados para DO. contra os dados experimentais em função do tempo.....	41
Figura 14 – Ajuste suave dos dados gerados para P contra os dados experimentais em função do tempo.....	41
Figura 15 – Ajuste suave dos dados gerados para X contra os dados experimentais em função do tempo.....	42
Figura 16 – Ajuste suave dos dados gerados para S contra os dados experimentais em função do tempo.....	42
Figura 17 – Concentrações de produto em função do tempo para todos os experimentos realizados.	49
Figura 18 – Concentrações de produto em função do tempo para todos os experimentos realizados.	50
Figura 19 – Concentrações de produto em função do tempo para todos os experimentos realizados.	50
Figura 20 – Concentrações de P em função do tempo para todos os experimentos realizados.	51
Figura 21 – Concentrações de X em função do tempo para todos os experimentos realizados.....	52
Figura 22 – Concentrações de S em função do tempo para todos os experimentos realizados.	52
Figura 23 – Teste de permutação pelo modelo de Perceptron Multicamada.	54
Figura 24 – Distribuição normal das variáveis de saída.	56
Figura 25 – Linearidade de S.S. versus P.	57
Figura 26 – Linearidade de S.S. versus X.	58
Figura 27 – Linearidade de S.S. versus S.	59
Figura 28 – Matriz de Correlação de Spearman.	60
Figura 29 – Concentrações das variáveis de saída em função do tempo: comparativo do ajuste dos dados simulados em relação aos experimentais para o experimento 1.	70
Figura 30 – Concentrações das variáveis de saída em função do tempo: comparativo do ajuste dos dados simulados em relação aos experimentais para o experimento 2.	71
Figura 31 – Concentrações das variáveis de saída em função do tempo: comparativo do ajuste dos dados simulados em relação aos experimentais para o experimento 3.Fonte: Autoria própria.....	71
Figura 32 – Concentrações das variáveis de saída em função do tempo: comparativo do ajuste dos dados simulados em relação aos experimentais para o experimento 4.	72
Figura 33 – Concentrações das variáveis de saída em função do tempo: comparativo do ajuste dos dados simulados em relação aos experimentais para o experimento 5.	72

Figura 34 – Erros relativos percentuais (ERP) de P em função do tempo de simulado.....	75
Figura 35 – Erros relativos percentuais (ERP) de X em função do tempo simulado.....	75
Figura 36 – Erros relativos percentuais (ERP) de S em função do tempo simulado.	76
Figura 37 - Ajuste dos dados de treinamento da RNA para P em relação com os dados fornecidos ao modelo... .	77
Figura 38 - Ajuste dos dados de treinamento da RNA para P em relação com os dados fornecidos ao modelo... .	77
Figura 39 - Ajuste dos dados de treinamento da RNA para P em relação com os dados fornecidos ao modelo.....	78
Figura 40 - Ajuste dos dados simulados da RNA para P em relação com os dados fornecidos ao modelo.....	79
Figura 41 - Ajuste dos dados simulados da RNA para S em relação com os dados fornecidos ao modelo.....	80
Figura 42 - Ajuste dos dados simulados da RNA para S em relação com os dados fornecidos ao modelo.....	80
Figura 43 - Resíduos da RNA para P para o experimento 2.....	82
Figura 44 - Resíduos da RNA para X para o experimento 2.	82
Figura 45 - Resíduos da RNA para S para o experimento 2.....	83
Figura 46 - Resíduos da RNA para P ao longo do tempo para o experimento 2.	84
Figura 47 - Resíduos da RNA para X ao longo do tempo para o experimento 2.....	84
Figura 48 - Resíduos da RNA para S ao longo do tempo para o experimento 2.	85

LISTA DE TABELAS

Tabela 1 – Configuração geral de treinamento.....	44
Tabela 2 – Número do experimento e suas respectivas temperaturas.....	51
Tabela 3 – Panorama geral do desempenho para as 10 melhores RNAs.....	61
Tabela 4 – Métricas de erro para as variáveis de saída.....	73

SUMÁRIO

1 INTRODUÇÃO	1
2 OBJETIVO	3
3 REVISÃO BIBLIOGRÁFICA	4
3.1 O HIDROMEL	4
3.2 DEFINIÇÃO E IMPORTÂNCIA HISTÓRICA DO HIDROMEL	4
3.3 CONSUMO ATUAL CONTEXTUALIZADO	6
3.4 OTIMIZAÇÃO E DESAFIOS EM BIOPROCESSOS.....	7
3.5 PROCESSO PRODUTIVO.....	7
3.6 PROCESSO FERMENTATIVO	9
3.6.1 Reações químicas que Envolvem a Fermentação	13
3.6.2 Adição da Casca de Jabuticaba	13
3.7 MODELAGEM E OTIMIZAÇÃO DA FERMENTAÇÃO DO HIDROMEL	16
3.8 REDES NEURAIS ARTIFICIAIS	17
3.8.1 Estrutura, Funcionamento e Treinamento	17
3.8.2 Algoritmos de Otimização	24
3.8.3 Funções de Ativação	26
3.8.4 Treinamento, Validação, Teste e Simulação	29
3.8.5 Aplicações de RNAs em Processos Fermentativos.....	30
4 MÉTODOS.....	34
4.1 FERMENTAÇÃO	35
4.2 OBTENÇÃO DOS DADOS EXPERIMENTAIS	35
4.3 OBTENÇÃO DAS RNAs.....	36
4.3.1 Seleção das variáveis de entrada e saída	37
4.3.2 Tratamento dos dados experimentais	39
4.4 TREINAMENTO E VALIDAÇÃO DAS RNAs	43
4.4.1 Configuração e detalhes do treinamento das RNAs.....	43
4.4.2 Definição dos parâmetros de eficiência da rede	45
4.4.3 Obtenção da rede	47
4.4.4 Condições de parada para o treinamento do modelo.....	47
4.4.5 Validação externa e análise dos resultados obtidos	47

5 RESULTADO E DISCUSSÃO.....	49
5.1 DADOS EXPERIMENTAIS EMPREGADOS NA OBTENÇÃO DA REDE.....	49
5.2 Análise das Importâncias por Permutação no Modelo de Regressão	54
5.3 Teste de Normalidade dos dados.....	55
5.4 Teste de Linearidade.....	57
5.5 Matriz de correlação	59
5.6 ANÁLISE DAS REDES OBTIDAS E CLASSIFICAÇÃO GERAL.....	62
5.7 SELEÇÃO DA MELHOR REDE COM BASE NA ESTRATÉGIA ABORDADA	67
5.8 ANÁLISE DAS MÉTRICAS DE ERRO	73
5.9 AVALIAÇÃO DO PROCESSO DE TREINAMENTO E VALIDAÇÃO EXTERNA.....	76
5.10 ANÁLISE DE RESÍDUOS DO MODELO	81
5.11 MODELAGEM FINAL PARA O MODELO GERADO	86
6 CONCLUSÕES E SUGESTÕES	88
6.1 CONCLUSÕES	88
6.2 SUGESTÕES PARA TRABALHOS FUTUROS	89
8 REFERÊNCIAS BIBLIOGRÁFICAS	91
9 APÊNDICE A - RESUMO GRÁFICO	100

1 INTRODUÇÃO

Atualmente, a demanda por produtos com significado histórico e impacto socioeconômico tem crescido, beneficiando pequenos produtores e a produção artesanal. O hidromel, uma das bebidas alcoólicas mais antigas, tem experimentado um renascimento, especialmente nos Estados Unidos. Esse ressurgimento é atribuído ao crescente interesse na cultura *geek*, com entusiastas de festivais medievais e da mitologia nórdica valorizando a bebida. A influência de grandes produções cinematográficas e literárias, como "*Game of Thrones*" e "*Senhor dos Anéis*", pode ter contribuído para essa tendência, embora estudos específicos sobre essa relação ainda sejam necessários (Nakada *et al.*, 2020).

Além disso, observa-se um ressurgimento do interesse pelo hidromel, que tem conquistado um espaço cativo nos eventos de cultura pop e recriações históricas, atraindo uma base de consumidores diversificada e abrindo novas perspectivas para seu mercado (Gurian, 2023). No Brasil, a produção de hidromel, embora artesanal, tem um potencial de mercado notável, especialmente nas regiões Sul e Sudeste. Cidades como São Paulo e Curitiba são reconhecidas por sediar feiras dedicadas a essa bebida (Nakada, 2020).

O hidromel é uma bebida alcoólica tradicional derivada da fermentação do mel e, possui uma composição que varia conforme o tipo de mel utilizado na produção do mosto. Esta variação influencia características como teor de ácido orgânico, pH, açúcar residual e teor alcoólico (Kawa-Rydlewská *et al.*, 2023). Os principais componentes do hidromel incluem açúcares, ácidos orgânicos e compostos fenólicos (Archilia *et al.*, 2021). A presença de compostos biologicamente ativos no hidromel destaca sua capacidade antioxidante, que é moldada por fatores como o tipo de mel, tratamento térmico do mosto, parâmetros de fermentação e ingredientes adicionais como ervas e frutas (Oganesyants *et al.*, 2015).

Estudos recentes têm explorado a adição de frutas, especialmente nativas do Brasil, ao hidromel, visando enriquecer sua composição e sabor. A jabuticaba, por exemplo, apresenta antocianinas, que são estabilizantes naturais, além de conferir cor e sabor distintos à bebida. Além disso, a escolha da levedura, especialmente do gênero *Saccharomyces*, é crucial para a fermentação, que é influenciada pela qualidade do mel, pH, temperatura e outros fatores. O mel, embora rico em açúcares fermentáveis, carece

de nitrogênio e outros nutrientes essenciais para a levedura, tornando a fermentação um processo lento. Portanto, a composição final do hidromel é determinada não apenas pelo mel, mas também pelos ingredientes adicionados ao mosto (Nyhan; Sahin; Arendt, 2023), tais como: compostos nitrogenados.

A diversidade de leveduras utilizadas na fermentação do vinho, incluindo leveduras não-*Saccharomyces*, tem mostrado impacto significativo no aroma e sabor do vinho. A aplicação dessas descobertas ao hidromel pode levar a bebidas com perfis aromáticos e gustativos mais complexos e desejáveis (Maicas, 2023).

Monitorar as concentrações de células, dos açúcares redutores e do produto é crucial para a análise e controle do processo fermentativo. No entanto, as técnicas tradicionais de medição dessas variáveis são demoradas. Assim, sensores virtuais baseados em Redes Neurais Artificiais (RNAs) são vistos como uma solução inovadora para predizer variáveis importantes do processo, de forma mais rápida e eficiente.

2 OBJETIVO

Neste estudo, a principal meta foi a concepção de um sensor virtual baseado em Redes Neurais Artificiais destinado a prever, durante a fermentação do hidromel, as concentrações de células (X), açúcares redutores (S), e, crucialmente, o produto, o etanol (P) em função de variáveis de fácil medição.

Para alcançar tal objetivo, dados experimentais foram utilizados, incluindo como variáveis de entrada: a temperatura operacional (T), a concentração de sólidos solúveis ($^{\circ}$ Brix), o pH e a densidade óptica (D.O.), selecionadas pela facilidade de medição em relação a outras variáveis, proporcionando uma abordagem prática e eficaz para o monitoramento do processo.

Adicionalmente, este trabalho teve como propósito avaliarmeticulosamente a eficácia de mais de 8 algoritmos de otimização distintos: Resilient-Backpropagation, Scaled Conjugate Gradient, Levenberg-Marquardt, Levenberg-Marquardt com Regularização Bayesiana, Gradiente Conjugado com Reinicializações de Powell-Beale, Broyden-Fletcher-Goldfarb-Shanno, Gradiente Descendente com Momentum e Taxa de Aprendizado Adaptativa, One Step Secant entre outros. Paralelamente, analisou-se o impacto de diversas funções de ativação, tais como logaritmo, tangente hiperbólica, linear, softmax e ReLu, na precisão das previsões da rede. Estas análises fornecem diretrizes valiosas para futuras investigações no campo da produção de hidromel e destacam a versatilidade das RNAs em processos fermentativos.

É imperativo enfatizar que as variáveis de predição foram selecionadas com foco na produção tradicional do hidromel, que envolve a utilização exclusiva de água, leveduras e mel com adição de um extrato de casca de jabuticaba, que teve como objetivo oferecer um caráter antioxidante. Além disso, o estudo ressalta o vasto potencial das RNAs como modelos preditivos não apenas para o hidromel, mas também para outras bebidas que passam por processos fermentativos. Esta aplicabilidade se estende a variações da bebida e a diferentes condições operacionais, incluindo fatores temporais e variáveis como temperatura.

3 REVISÃO BIBLIOGRÁFICA

3.1 O hidromel

O hidromel, uma tradicional bebida alcoólica, tem seu teor alcoólico variando entre 4% e 14%, de acordo com a legislação brasileira, sendo fruto da fermentação do mel por leveduras. Esta bebida tem sido objeto de estudos que buscam entender a cinética da fermentação e os parâmetros envolvidos no processo, especialmente quando se utiliza a levedura *Saccharomyces cerevisiae*, conhecida por sua robustez e eficácia na conversão de açúcares em etanol (Brito, 2023). A *Saccharomyces cerevisiae* é amplamente empregada devido à sua capacidade de tolerar diferentes condições de fermentação, como variações de temperatura e níveis de pH, permitindo maior controle sobre o processo fermentativo e favorecendo o desenvolvimento de aromas e sabores característicos do hidromel. A investigação dos fatores que influenciam essa fermentação, como a proporção de mel e água, a temperatura ideal e o tempo de fermentação, são fundamentais para otimizar a produção e melhorar a qualidade final do produto, especialmente considerando o perfil de mercado e o gosto do consumidor brasileiro.

Há registros históricos que indicam a longa trajetória do hidromel, sendo uma das primeiras bebidas alcoólicas consumidas pela humanidade, com menções que remontam a tempos antigos, antes mesmo do vinho (Gurian, 2023). No Brasil, o mercado tem mostrado um crescente interesse por bebidas fermentadas, com destaque para a produção e comercialização de cervejas artesanais. O hidromel, em particular, possui um grande potencial para se destacar nesse cenário, dada a riqueza e diversidade de méis produzidos no país, o que permite a criação de produtos únicos e diferenciados, capazes de atender tanto o consumidor nacional quanto o mercado de exportação (Lima, 2023).

3.2 Definição e importância histórica do hidromel

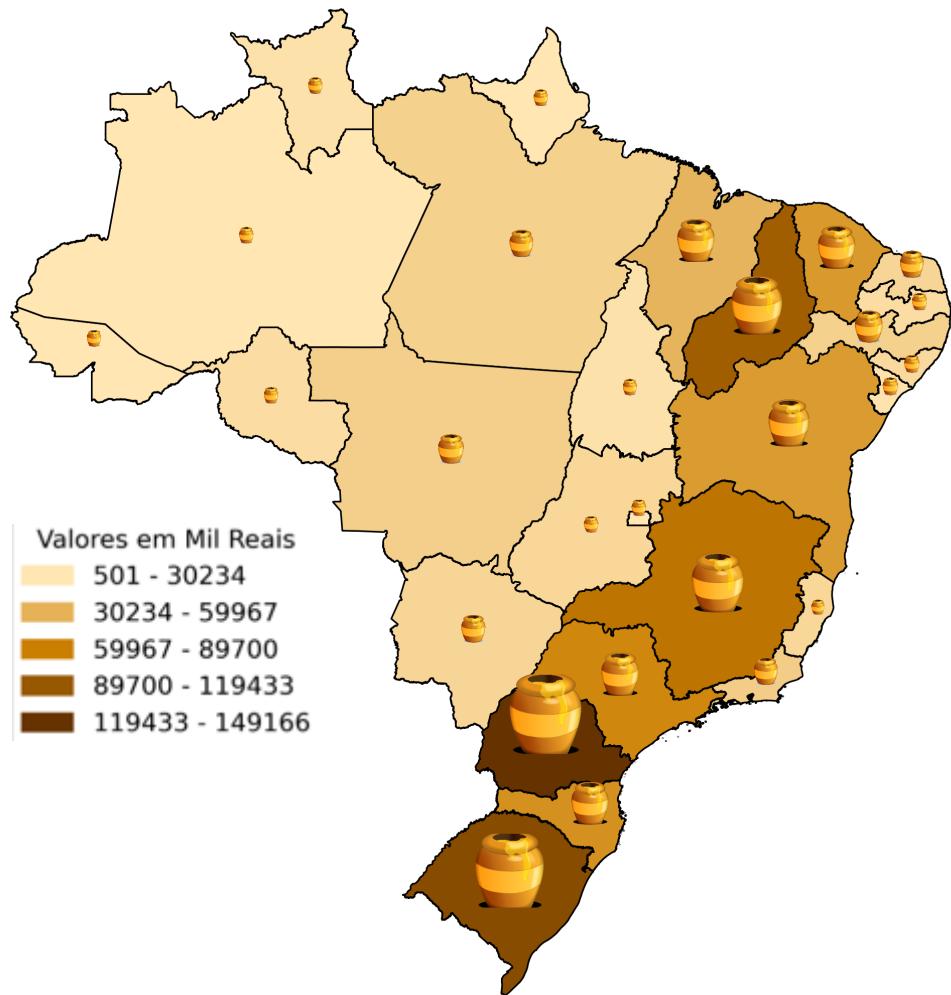
Na mitologia nórdica, o hidromel era altamente valorizado, sendo visto como uma bebida de prestígio entre civilizações antigas, como os celtas, anglo-saxões e vikings. Esta bebida desempenhava um papel central em rituais e era frequentemente associada à imortalidade, ao conhecimento e ao dom da poesia. Era também acreditado que o

hidromel possuía propriedades sobrenaturais e curativas, capazes de aumentar a força, virilidade e fertilidade (Moraes, 2023).

Nos tempos modernos, a presença do hidromel em séries de televisão, como "Vikings", "Ragnarok" e "Game of Thrones", bem como em filmes, incluindo "Harry Potter", reforçou seu reconhecimento e apreciação (Moraes, 2023). Globalmente, o hidromel é valorizado em muitos países, incluindo Inglaterra, Polônia, Alemanha e Eslovênia. No Brasil, há um potencial significativo para o crescimento do mercado de hidromel, com um público-alvo principal de homens e mulheres acima de 20 anos. As regiões Sul e Sudeste se destacam como os principais centros de produção e comercialização desta bebida, especialmente nas cidades de São Paulo e Curitiba. São Paulo, com seu mercado vibrante e consumidores em busca de novas experiências gastronômicas, consolidou-se como um polo importante para o desenvolvimento de marcas de hidromel e para a realização de eventos que celebram a cultura medieval e a produção artesanal de bebidas. Curitiba, por sua vez, é reconhecida pelo apoio a pequenos produtores e pela presença de um público interessado em bebidas artesanais, o que impulsiona o crescimento de microempresas locais e o fortalecimento da cadeia de produção artesanal do hidromel na região (Nakada, 2020).

A Figura 1 apresenta a distribuição espacial do valor da produção de mel de abelha no Brasil, em mil reais, conforme dados da Pesquisa da Pecuária Municipal (PPM) realizada pelo IBGE. O mapa evidencia uma concentração significativa da produção em regiões específicas, com destaque para os estados do Sul e Nordeste, que apresentam os maiores valores agregados, representados pelas tonalidades mais escuras. As regiões Centro-Oeste, Sudeste e Norte mostram uma menor participação relativa, com valores concentrados nas faixas inferiores. Essa distribuição reflete diferenças regionais relacionadas a fatores climáticos, socioeconômicos e estruturais que influenciam diretamente a produção e comercialização de mel no país. Além disso, a inserção de ícones de potes de mel nos estados contribui para uma visualização mais intuitiva dos principais polos produtores. O mapa, portanto, sintetiza as disparidades regionais e destaca os principais produtores, permitindo uma análise geográfica e econômica da cadeia produtiva de mel no Brasil.

Figura 1 – Mapa do Mel de abelha - Valor da produção (Mil Reais).



Fonte: IBGE - PPM - Pesquisa da Pecuária Municipal pelos parâmetros de Valor da produção, Quantidade produzida, Maior produtor.

3.3 Consumo atual contextualizado

O hidromel é uma bebida amplamente consumida e apreciada em países europeus, sendo historicamente reconhecido como a "bebida dos deuses". No entanto, no Brasil, o consumo e a produção de hidromel ainda estão nos seus estágios iniciais. A produção nacional é majoritariamente artesanal, muitas vezes realizada por apicultores que veem no hidromel uma oportunidade de agregar valor ao mel. Esta abordagem artesanal, no entanto, enfrenta desafios, como a falta de uniformidade do produto e o longo período de fermentação. Além disso, a produção de hidromel no país é limitada devido à escassa utilização de tecnologia, à falta de tradição no mercado e à demanda interna que é fortemente influenciada por fatores culturais (Ilha *et al.*, 2008).

Apesar de sua longa história, a produção de hidromel ainda enfrenta desafios, como teores alcoólicos abaixo do ideal e a ocorrência de sabores e odores indesejados. Para superar esses obstáculos, é essencial desenvolver ferramentas e técnicas que otimizem as condições de produção, visando obter um produto de alta qualidade e rendimento (Lima, 2022).

3.4 Otimização e desafios em bioprocessos

Um dos principais fatores que contribuem para a persistência da produção artesanal e a limitada disseminação do hidromel no Brasil é a lacuna existente com relação aos avanços científicos e tecnológicos voltados especificamente para a otimização e expansão desse setor. A baixa frequência de pesquisas aprofundadas sobre os processos de fermentação, controle de qualidade, e padronização de ingredientes limita o desenvolvimento de técnicas industriais eficientes. Esse cenário restringe a capacidade de produção em larga escala e dificulta a adoção de práticas que aumentem a competitividade do hidromel frente a outras bebidas fermentadas no mercado. Atualmente, o hidromel é produzido em pequena escala industrial, e o processo de fabricação ainda é pouco compreendido. Nas indústrias, o método predominante é a fermentação em batelada, porém, esse processo apresenta desafios, uma vez que resulta em baixa produtividade devido à inibição pelo etanol e pelo substrato e, isso pode levar a um tempo maior de fermentação, que poderia ser diminuído sem perda da qualidade. Porém, isso depende de estudos aprofundados (Queiroz *et al.*, 2014).

Diversos problemas estão associados à produção de hidromel artesanal, incluindo a falta de controle do pH e da temperatura, a escassez de nutrientes essenciais e substrato durante a fermentação, bem como a inibição causada pelo etanol. A dificuldade em monitorar e manter esses fatores em condições ideais durante a fabricação do hidromel pode levar a um prolongado tempo de fermentação, uma taxa de crescimento microbiano reduzida e, consequentemente, uma diminuição na produtividade (Gomes, 2010).

3.5 Processo produtivo

A produção de hidromel, uma bebida alcoólica derivada da fermentação de mel, segue etapas semelhantes à fabricação de vinho (Mattietto *et al.*, 2006). O processo se

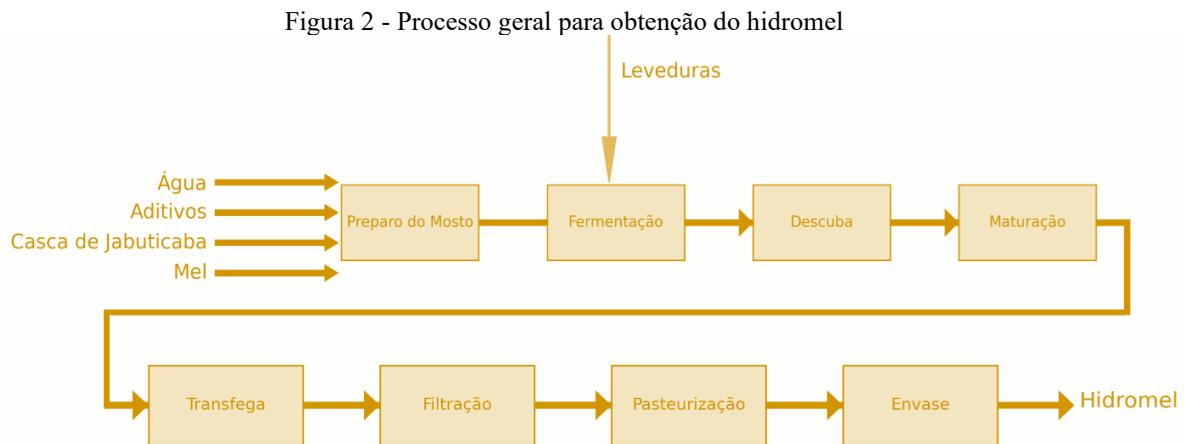
inicia com a preparação do mosto, onde o mel é diluído em água até atingir 22 °Brix de sólidos solúveis (Brunelli, 2015). Pode-se adicionar frutas, ervas e especiarias para diversificar o sabor. A correção do pH e a adição de suplementos nutricionais são essenciais, dado que o mel não fornece nutrientes suficientes para o crescimento microbiano (Gupta, 2009).

A fermentação ocorre com a adição de inóculo em um fermentador esterilizado. O método descontínuo é o mais utilizado (Brunelli, 2015), e a conclusão da fermentação é identificada pela estabilização do teor de sólidos solúveis. Posteriormente, realiza-se o processo de descuba, que consiste em separar o sedimento do líquido fermentado, evitando reações químicas indesejadas. (Brunelli, 2015).

O hidromel passa por um período de maturação em anaerobiose, onde compostos aromáticos são formados. Durante essa etapa, a concentração de acetato de etila deve ser monitorada para não exceder o valor de 2.000 mg L⁻¹, acima desse limite, o composto confere um sabor e aroma indesejáveis à bebida, tornando-a desagradável ao paladar. Além disso, níveis elevados de acetato de etila podem mascarar outros compostos aromáticos delicados formados durante a maturação, comprometendo a qualidade final do hidromel (Mileski, 2016). A maturação dura de 1 a 6 meses a temperaturas entre 10 e 12 °C (Brunelli, 2015).

Após o período de maturação, o hidromel passa por um processo fundamental conhecido como trasfega. Esse procedimento consiste na transferência cuidadosa do líquido para um novo recipiente, com o objetivo de separar o hidromel dos sedimentos acumulados no fundo do tanque. Esses sólidos, compostos principalmente por resíduos de leveduras e outras partículas resultantes da fermentação, podem impactar o sabor e a clareza do produto. A trasfega, portanto, é essencial para garantir um hidromel mais puro e de qualidade superior, preservando os aromas e sabores desejados e conferindo uma aparência límpida à bebida. A clarificação é realizada para remover sólidos insolúveis, podendo ser auxiliada por filtração e agentes clarificantes (Brunelli, 2015). A pasteurização é sugerida para prolongar a vida útil do hidromel (Brunelli, 2015).

O fluxograma do processo produtivo pode ser visto na Figura 2 demonstrando como se dá a entrada na preparação do mosto até o envase do hidromel.



Fonte: Autoria Própria.

O fluxograma do processo produtivo, ilustrado na Figura 2, descreve detalhadamente cada etapa envolvida na produção do hidromel, desde a preparação do mosto até o envase final da bebida. Inicialmente, os ingredientes principais – água, aditivos, mel e casca de jabuticaba – são combinados na etapa de preparo do mosto. Esse mosto é, então, submetido ao processo de fermentação com a adição de leveduras, responsáveis por converter os açúcares presentes em álcool e gerar os compostos característicos da bebida.

Após a fermentação, o hidromel passa pela etapa de "descuba", onde é transferido para remover resíduos sólidos. Em seguida, entra na fase de maturação, momento em que se desenvolvem os aromas e sabores mais complexos. Na sequência, o hidromel é submetido a um processo de trasfega, separando o líquido dos sedimentos que se formaram no fundo do recipiente durante a maturação. Esse líquido é, então, filtrado para garantir maior pureza e transparência. Após a filtração, ocorre a pasteurização, que tem como finalidade eliminar microrganismos indesejados, garantindo a estabilidade microbiológica do produto. Por fim, o hidromel é envasado, pronto para consumo.

Cada uma dessas etapas contribui para a qualidade final da bebida, assegurando que o hidromel atenda aos padrões desejados em termos de sabor, aparência e segurança para o consumidor.

3.6 Processo fermentativo

A produção de uma bebida alcoólica é influenciada por diversos fatores, como a qualidade das matérias-primas, o tipo de agente fermentativo, a presença de aditivos e o processo de maturação, entre outros (Gupta, 2009). No contexto da fermentação do hidromel, a linhagem de levedura utilizada, a variedade do mel, a composição do meio de cultura e o controle do pH extracelular desempenham papéis cruciais na obtenção de uma bebida de alta qualidade (Navrátil, 2001).

Conforme observado por Sroka (2007), um dos principais desafios enfrentados na produção do hidromel está relacionado à prolongada duração do processo fermentativo, o que resulta em baixa produtividade. Para superar esse obstáculo, é essencial garantir que o pH, a temperatura, a linhagem de levedura e os fatores de crescimento estejam ajustados de forma adequada. A identificação e a mitigação dos fatores que diminuem a atividade celular podem acelerar o processo de produção e torná-lo mais economicamente viável (Pereira, 2008).

Vale destacar também que a maioria das leveduras enológicas não está adaptada às condições presentes em meios de cultivo à base de mel. Esses meios se caracterizam por um baixo pH, elevados níveis de açúcares, concentrações reduzidas de nitrogênio e a presença de uma variedade de ácidos originados no mel. Tipicamente, essas condições exercem um impacto adverso na fermentação, criando um ambiente estressante que não favorece o crescimento celular (Ivorra, 1999; Melo, 2006; Pereira, 2008).

Esta seção abordará com maior detalhe os principais fatores que influenciam o processo fermentativo, sendo um deles, a temperatura, que é um parâmetro de grande influência na cinética do processo fermentativo, pois afeta as taxas de consumo de substrato (e, consequentemente, o tempo de fermentação), a produção de etanol e o crescimento microbiano. O uso de uma temperatura inadequada impacta negativamente o rendimento da fermentação, a conversão de açúcares em etanol e, consequentemente, o teor alcoólico da bebida (Andrade, 2012).

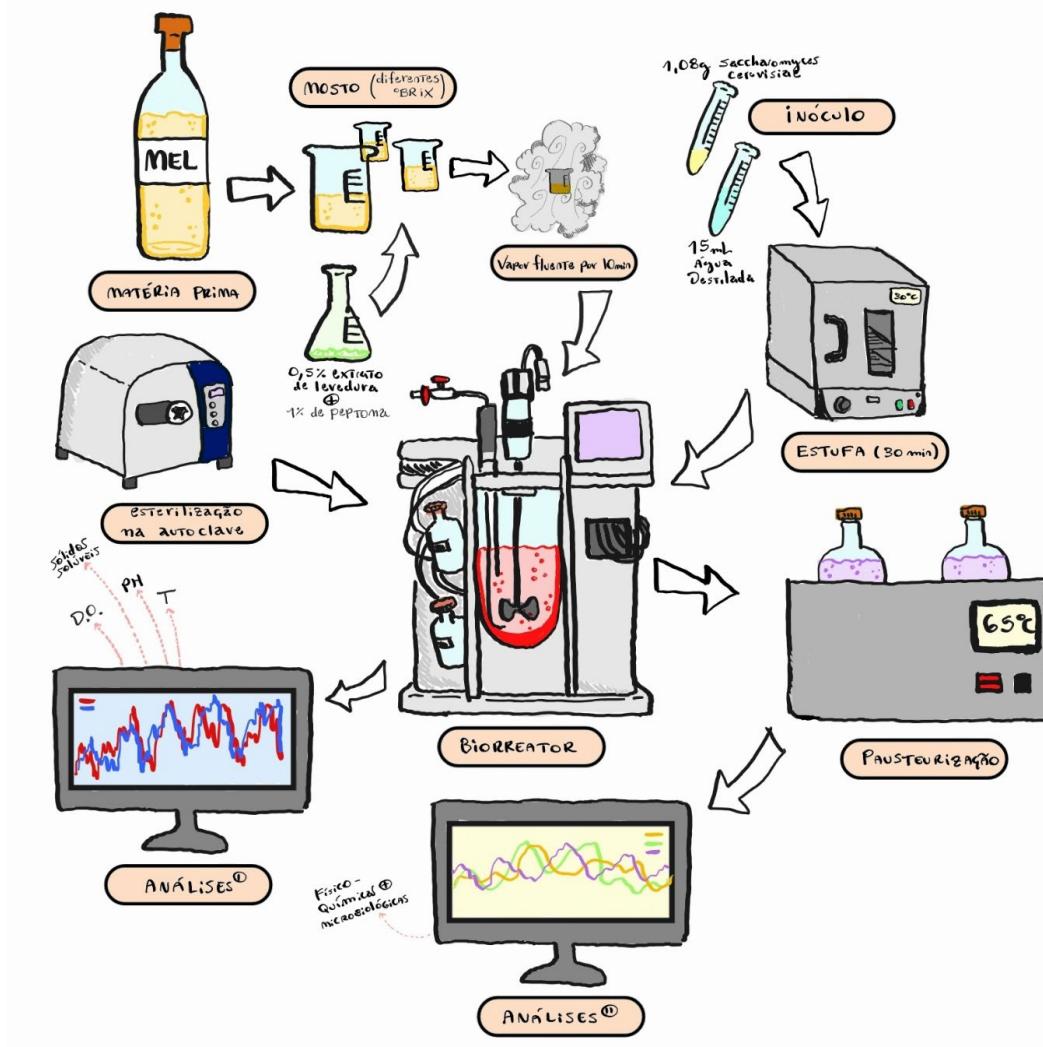
O crescimento em temperaturas elevadas é prejudicial para a *Saccharomyces cerevisiae* e leva a uma reorganização do fluxo metabólico. Inicialmente, ocorre um aumento no metabolismo em resposta ao aumento da temperatura em todo o sistema. Em seguida, ocorre a resposta metabólica ao estresse, que pode resultar em adaptação, inativação ou morte celular (Melo, 2006).

As taxas de fermentação mais eficazes são alcançadas em temperaturas situadas entre 20 e 30 °C, especialmente para leveduras do gênero *Saccharomyces*. Temperaturas muito baixas (inferiores a 15 °C) e muito altas (superiores a 30 °C) estão associadas a

fermentações mais lentas (Brunelli, 2015). É relevante destacar que temperaturas superiores ao valor ótimo diminuem a taxa de fermentação devido à desnaturação enzimática e ao aumento da toxicidade do etanol (Andrade, 2012).

Portanto, o controle preciso da temperatura é essencial para otimizar o processo de fermentação alcoólica, alcançar alto rendimento e altas taxas de produção. Conforme observado por Brunelli (2015), é crucial que a fermentação para a produção de hidromel ocorra a 24 °C, a fim de aumentar as taxas de reação (velocidade do processo) e obter um hidromel com excelentes propriedades químicas. Na Figura 3, são detalhados os passos do processo fermentativo, desde a obtenção da matéria prima, até as análises físico-químicas e microbiológicas, para a obtenção dos dados experimentais.

Figura 3 – Processo fermentativo do hidromel.



Fonte: Autoria Própria.

A figura ilustra o processo de fermentação a partir do mel como matéria-prima, destacando as etapas essenciais envolvidas. Inicialmente, o mel é transformado em

mosto com diferentes concentrações de sólidos solúveis ($^{\circ}$ Brix), seguido por sua esterilização em autoclave para garantir a assepsia do meio. Paralelamente, o inóculo é preparado com *Saccharomyces cerevisiae*, previamente hidratado em água destilada e ativado em estufa. Após a inoculação, o mosto esterilizado é transferido para o biorreator, onde ocorre a fermentação sob condições controladas de pH, temperatura (T) e oxigênio dissolvido (DO), monitoradas em tempo real. A fermentação resulta em um produto que é posteriormente pasteurizado e submetido a análises fisico-químicas e microbiológicas para avaliação da qualidade. A representação esquemática do processo destaca a integração entre controle biotecnológico e análise de resultados, fornecendo uma visão clara e didática das etapas envolvidas na produção fermentativa a partir do mel.

O pH do mosto é outro fator de grande relevância a ser considerado, uma vez que um pH baixo resulta na redução do crescimento de muitos microrganismos. A fermentação alcoólica industrial é conhecida por envolver o reciclo de células e a exposição contínua a valores baixos de pH tanto no mosto quanto no pré-inóculo (Melo, 2006). O mosto composto por mel e água possui uma capacidade tamponante reduzida, o que pode levar a uma diminuição do pH durante o processo de fermentação a ponto de inviabilizar a atividade microbiana (Brunelli, 2015; Ramalhosa *et al.*, 2011). Conforme Sroka & Tuszníki (2007), essa diminuição na atividade pode ser atribuída à produção de ácido acético, um forte inibidor, e ácido succínico pelas próprias leveduras.

Outro fator de importância nas fermentações é a concentração de açúcares no mosto, uma vez que um alto teor desse composto pode tornar o processo fermentativo extremamente lento. Türkmen & Ekşi (2011) definem o grau Brix ($^{\circ}$ Brix) como uma escala numérica baseada no índice de refração que quantifica o teor de sólidos solúveis (açúcares totais) na solução de fermentação. Um grau Brix corresponde a 1 g de sacarose em 100g de solução e expressa a concentração da solução em porcentagem em massa (Amorim *et al.*, 2018).

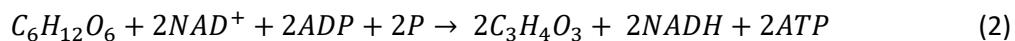
Em fermentações conduzidas em batelada, o valor de concentração de sólidos solúveis sofre alterações ao longo do processo. Esta concentração excessivamente elevada pode resultar na inibição por substrato, levando ao acúmulo de açúcares no produto. Por outro lado, um valor muito baixo dessa variável resulta em baixas concentrações de produto e baixa eficiência do processo, pois os açúcares são essenciais para o crescimento dos microrganismos e a produção. Nas fermentações em batelada, à

medida que o microrganismo consome o substrato, a concentração de sólidos solúveis diminui e, como resultado, o teor alcoólico da bebida aumenta (Amorim *et al.*, 2018).

Diante do exposto, a análise da concentração de sólidos solúveis ao longo da fermentação é fundamental para monitorar o consumo de substrato e pode ser usada para avaliar o progresso do processo. Além de indicar o término da fermentação (Brunelli, 2015), ele fornece informações sobre o consumo de substrato pelas leveduras e, se necessário, pode orientar a tomada de decisões relacionadas ao ajuste das condições do processo.

3.6.1 Reações químicas que Envolvem a Fermentação

O experimento começou com a introdução de um inóculo ao fermentador, iniciando assim, o experimento fermentativo, cujas reações podem (de forma genérica) ser observadas nas equações (1, 2 e 3):



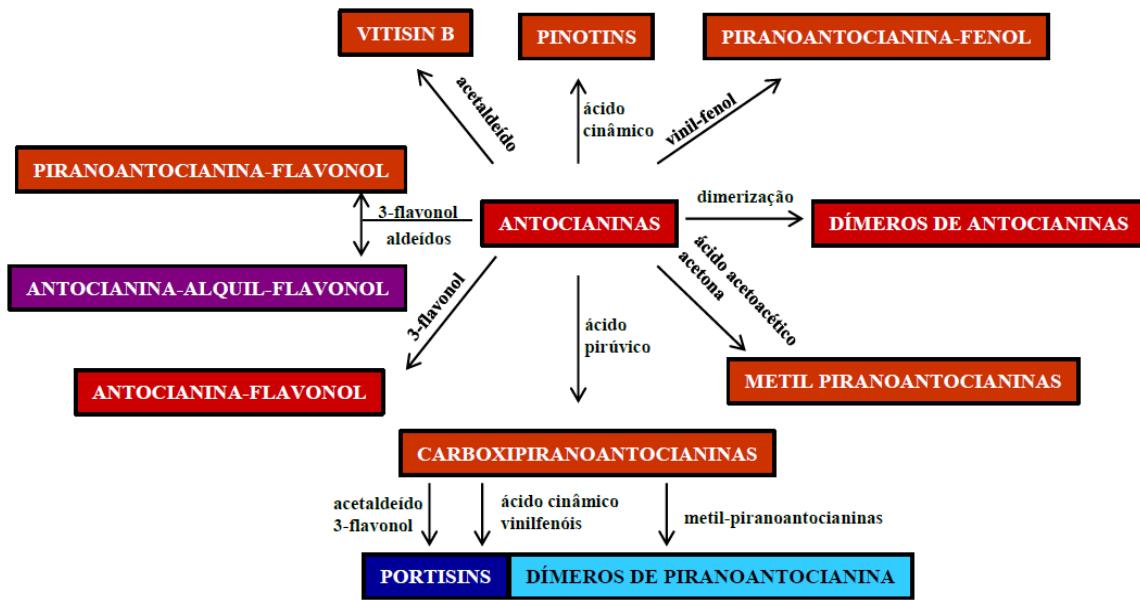
As Equações 1, 2 e 3 representam, respectivamente, os processos de quebra da sacarose, glicólise e fermentação alcoólica. Na primeira etapa, a sacarose ($C_{12}H_{22}O_{11}$) é hidrolisada em seus monômeros, glicose ($C_6H_{12}O_6$) e frutose ($C_6H_{12}O_6$), graças à ação da enzima invertase presente no microrganismo *Saccharomyces cerevisiae*. Em seguida, na Equação (2), a glicose é metabolizada na via glicolítica, onde é convertida em piruvato ($C_3H_4O_3$), com geração de ATP e NADH. Por fim, na Equação (3), ocorre a fermentação alcoólica, onde o piruvato é transformado em etanol (C_2H_5OH) e dióxido de carbono (CO_2), regenerando NAD^+ necessário para a continuidade do processo glicolítico. Em seguida, os monômeros são convertidos em piruvato, liberando a energia necessária para o funcionamento dos microrganismos (2). Por fim, o produto alcoólico é produzido como resultado desse processo (3) (Nelson, 2014).

3.6.2 Adição da Casca de Jabuticaba

No contexto das características sensoriais de bebidas fermentadas, como o hidromel e o vinho, a cor sempre foi considerada um dos principais atributos

perceptíveis pelos consumidores. Essa característica sensorial está originalmente associada à presença de um grupo significativo de pigmentos fenólicos solúveis em água, conhecidos como antocianinas, que também conferem um caráter antioxidante às bebidas. Nas jabuticabas e uvas tintas, algumas das antocianinas originais incluem cianidina, petunidina, peonidina, delfinidina e malvidina, juntamente com outros compostos, como demonstrado na Figura 4 (Burns *et al.*, 2002).

Figura 4 – Pigmentos derivados de Antocianinas.



Fonte: Freitas, 2011.

A Figura 4 apresenta as principais vias de transformação e interação química das antocianinas, compostos fenólicos amplamente conhecidos por suas propriedades antioxidantes e sua contribuição para a cor de produtos naturais, como frutas e vinhos. As antocianinas servem como núcleo central nesse esquema, sendo transformadas em diferentes derivados estruturais por meio de reações com compostos como acetaldeído, ácido pirúvico, ácidos fenólicos e flavonoides. Esses derivados incluem carboxipiranoantocianinas, piranoantocianinas-flavonol, dimetil piranoantocianinas e vitisin B, entre outros. Reações como dimerização, adição de grupos fenólicos e formação de compostos vinil-fenólicos destacam a diversidade de interações químicas que podem ocorrer, ampliando as propriedades químicas e funcionais das antocianinas. A figura também enfatiza a complexidade da formação de compostos estáveis, como portisins e dímeros de piranoantocianinas, que desempenham papéis importantes em

características sensoriais e na estabilidade da cor em produtos como vinhos envelhecidos. Essa representação esquemática ilustra a ampla reatividade das antocianinas e a formação de compostos bioativos adicionais que influenciam a qualidade e funcionalidade dos alimentos e bebidas.

As antocianinas, também conhecidas como antocianos, são pigmentos flavonoides que, dependendo do pH, e em alguns casos de agentes complexantes, conferem aos vinhos cores distintas como vermelho, púrpura ou azul. A concentração de antocianinas livres em vinhos tintos jovens é de 500 - 2000 mg L⁻¹, dependendo da variedade, sistema de cultivo, sistema de extração, práticas enológicas, entre inúmeros outros fatores (Burns *et al.*, 2002; Mazza, G. & Francis, F. J., 2005; Busse-Valverde *et al.*, 2011).

Em geral, as antocianinas se encontram localizadas, principalmente, na película da uva, sendo extraídas durante o processo de maceração que ocorre simultaneamente com a fermentação alcoólica em processos de vinificação clássica de vinhos tintos (Rizzon *et al.*, 1999).

Nos vinhos, durante a fermentação e nos anos subsequentes de maturação, as antocianinas sofrem uma série de reações e associações que determinam mudanças de cor e são essenciais na estabilidade da mesma (Brouillard *et al.*, 2003; He, F. *et al.*, 2012). Além disso, apesar de serem basicamente inodoras e insípidas, as antocianinas podem interagir com outras substâncias e influenciar as características de aroma e sabor dos vinhos e hidroméis (Dufour, 2000).

As antocianinas, além de estarem presentes na uva, também são encontradas em grande quantidade na casca da jabuticaba. Esse composto, responsável pela coloração vibrante de diversas frutas, é particularmente abundante na jabuticaba, onde contribui para seu intenso tom roxo-escuro (Rizzon *et al.*, 1999). A casca da jabuticaba, rica em antocianinas, tem atraído o interesse de pesquisadores e produtores devido às suas propriedades antioxidantes e potenciais benefícios à saúde, como a capacidade de combater radicais livres e reduzir o risco de doenças crônicas. Em processos de fermentação e produção de bebidas, a extração dessas antocianinas da casca da jabuticaba pode agregar valor ao produto, proporcionando não apenas coloração, mas também benefícios funcionais, o que a torna uma matéria-prima promissora para a indústria de bebidas fermentadas e outras aplicações alimentícias.

3.7 Modelagem e otimização da fermentação do hidromel

A produção de hidromel, especialmente no contexto da otimização do processo fermentativo, é uma área de pesquisa em expansão. Com o avanço da tecnologia e a integração de técnicas modernas, como a Inteligência Artificial, há um potencial significativo para melhorar e refinar o processo de fermentação do hidromel.

A valorização de resíduos alimentares através de processos fermentativos tem ganhado destaque nas pesquisas científicas, especialmente pelo seu potencial em promover uma economia mais sustentável e circular. No contexto do hidromel, cuja produção se baseia na fermentação do mel, esses avanços tecnológicos podem oferecer benefícios significativos, permitindo a incorporação de ingredientes adicionais derivados de resíduos alimentares, como frutas e cascas, que enriquecem o perfil sensorial e nutricional da bebida (Ortiz-Sánchez *et al.*, 2023). Além disso, a otimização do meio de cultura é fundamental para alcançar um alto rendimento do produto, com maior qualidade e redução nos custos de produção.

Esse processo de otimização vem sendo explorado por meio de abordagens algorítmicas e modelos computacionais, que permitem ajustar variáveis como nutrientes, pH e temperatura para maximizar a eficiência do processo fermentativo. Técnicas como essas, que incluem o uso de algoritmos de aprendizado de máquina e métodos estatísticos, são cada vez mais aplicadas na produção de bebidas fermentadas e podem ser adaptadas para o hidromel, potencializando sua produção em escala industrial e contribuindo para a competitividade do produto no mercado (Zhou *et al.*, 2023).

Um exemplo é o estudo de Zhang *et al.* (2014), no qual a Metodologia de Superfície de Resposta (MSR) foi empregada para otimizar a produção de hidromel. Nesse trabalho, a metodologia foi utilizada para determinar as condições ideais de temperatura, pH e concentração de fosfato diamônio no processo de fabricação da bebida. Como resultado, foi obtido um modelo de regressão empírico polinomial que se ajustou bem aos dados experimentais, permitindo a determinação das condições ideais do processo, resultando em um hidromel com teor alcoólico de 9,3% (v/v) e excelente sabor.

Vale destacar que, embora não tenha sido encontrado na literatura um uso específico dessa metodologia de superfície de resposta em conjunto com RNAs para a fermentação de hidromel, existem trabalhos que aplicaram essa abordagem com sucesso em outros

processos fermentativos. Santos *et al.* (2018), por exemplo, desenvolveram um sensor virtual neural para monitorar a produção de biossurfactante por *Bacillus subtilis* e utilizaram a MSR para determinar os valores ótimos dos parâmetros de operação.

Portanto, é evidente que há uma lacuna no que diz respeito à modelagem do processo de fermentação para a produção de hidromel com foco no uso de sensores virtuais neurais. Nesse contexto, este trabalho se destaca como inovador, com o potencial de proporcionar o monitoramento da fermentação e, consequentemente, oferecer suporte para a obtenção de um hidromel de alta qualidade.

3.8 Redes neurais artificiais

Os avanços recentes na área de Inteligência Artificial (IA) têm levado a uma crescente adoção das Redes Neurais Artificiais em diversos setores. Estas redes, inspiradas na estrutura e funcionamento do cérebro humano, têm mostrado ser ferramentas poderosas para lidar com grandes volumes de dados e extrair padrões complexos a partir deles (Costa, 2023).

As Redes Neurais Artificiais (RNAs) representam modelos computacionais inspirados no funcionamento dos neurônios biológicos, capazes de simular o processamento e a transmissão de informações de maneira semelhante ao cérebro humano. Essas redes são organizadas em camadas de neurônios artificiais interconectados, em que cada ligação possui um peso específico, ajustado continuamente durante o processo de aprendizado. Esse ajuste dos pesos permite que as RNAs adaptem seu desempenho e aumentem a precisão em tarefas de reconhecimento de padrões e previsão de resultados (Costa, 2023).

3.8.1 Estrutura, Funcionamento e Treinamento

As Redes Neurais Artificiais são inspiradas na estrutura neural de organismos vivos e consistem em unidades de processamento, semelhantes aos neurônios biológicos. A capacidade de processamento dessas redes é determinada pela informação armazenada nas conexões entre essas unidades, conhecidas como pesos. Esses pesos são ajustados através de um processo de aprendizado que utiliza padrões de treinamento específicos (Kaveh, 2022).

A eficácia de uma rede neural artificial é avaliada pela precisão com que ela pode replicar os dados de saída desejados. Para isso, o neurônio artificial precisa ser treinado para executar uma tarefa específica (Fleck *et al.*, 2016).

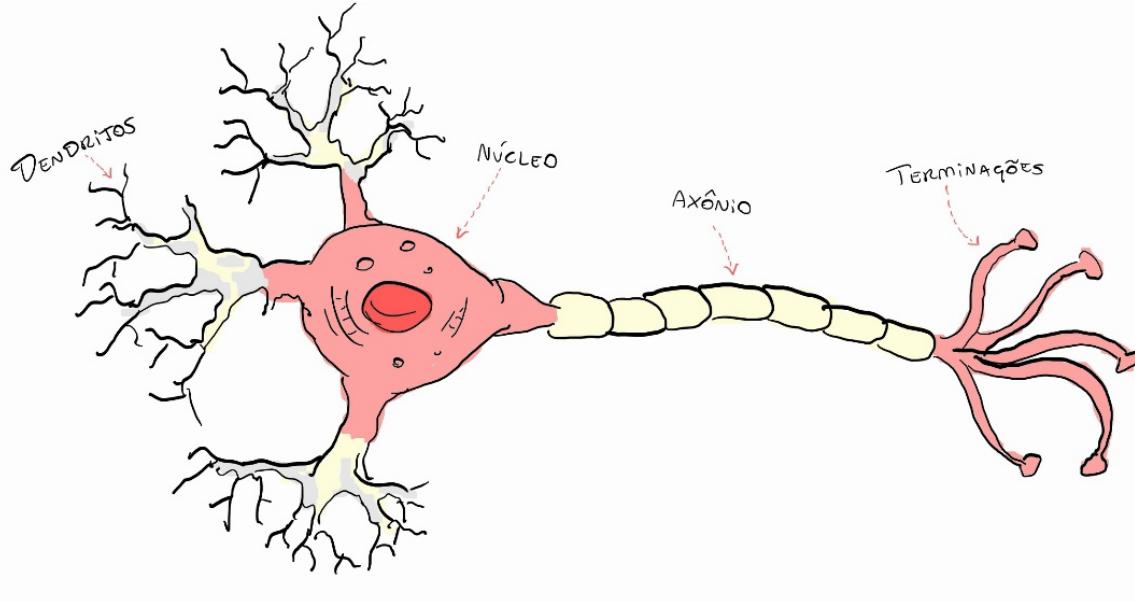
Durante o treinamento, a RNA analisa e assimila informações de padrões apresentados, formando uma compreensão única do problema em questão. Nesse processo, os pesos das conexões entre as unidades são ajustados. Assim, ao término do treinamento, a RNA possui uma compreensão do contexto em que opera, "salva" em seus parâmetros. Existem diversas técnicas para treinar RNAs, mas quando aplicadas à modelagem, o método de aprendizado supervisionado é frequentemente utilizado (Fleck *et al.*, 2016).

O termo "aprendizado supervisionado" se refere à presença de um orientador externo que fornece os dados de entrada e saída para a RNA, com o objetivo de otimizar seus parâmetros. Durante esse aprendizado, o orientador avalia e guia o desempenho da rede, comparando sua saída com a saída desejada e ajustando os pesos para reduzir erros. Esse ajuste é gradual, com pequenas modificações feitas a cada ciclo de treinamento. A diferença acumulada entre a saída desejada e a saída real é frequentemente usada como métrica de desempenho e como critério de otimização (Silva, 2016).

Existem dois modos principais de implementação do aprendizado supervisionado: "off-line" e "on-line". No modo "off-line", o conjunto de dados de treinamento é fixo; portanto, qualquer adição de novos dados exige a criação de um novo conjunto de treinamento. No modo "on-line", a RNA opera em um processo de adaptação constante, ajustando seus parâmetros de forma dinâmica à medida que novos dados são incorporados (Fleck *et al.*, 2016).

As redes são compostas por unidades de processamento chamadas neurônios artificiais que imitam os neurônios biológicos (Figura 5) em sua funcionalidade básica (Ripley, 2001).

Figura 5 – Neurônio Biológico.



Fonte: Autoria Própria.

No início do processo biológico, informações ou impulsos nervosos são captados pelos dendritos e, a partir daí, são encaminhados para o corpo celular. No corpo celular, ocorre o processamento dessas informações e a geração de novos impulsos, que são então transmitidos através do axônio até suas ramificações terminais. A partir dessas terminações, os impulsos são transmitidos a outros neurônios por meio das sinapses, que representam os pontos de contato entre a extremidade do axônio de um neurônio e o dendrito de outro (Braga *et al.*, 2000).

Para que o impulso seja transmitido à célula nervosa seguinte, é essencial que o limiar de excitação (threshold) seja alcançado. Segundo Braga *et al.* (2000), o disparo de um neurônio depende do acúmulo de uma grande quantidade de entradas, incluindo entradas inibitórias e excitatórias, avaliadas pelo corpo celular em um curto espaço de tempo. Se o percentual resultante for suficientemente elevado, o corpo celular combina os sinais recebidos e o neurônio "dispara", gerando um impulso que é então transmitido aos neurônios subsequentes. Após o disparo, o neurônio entra em um estado de repouso, preparando-se para gerar um novo impulso.

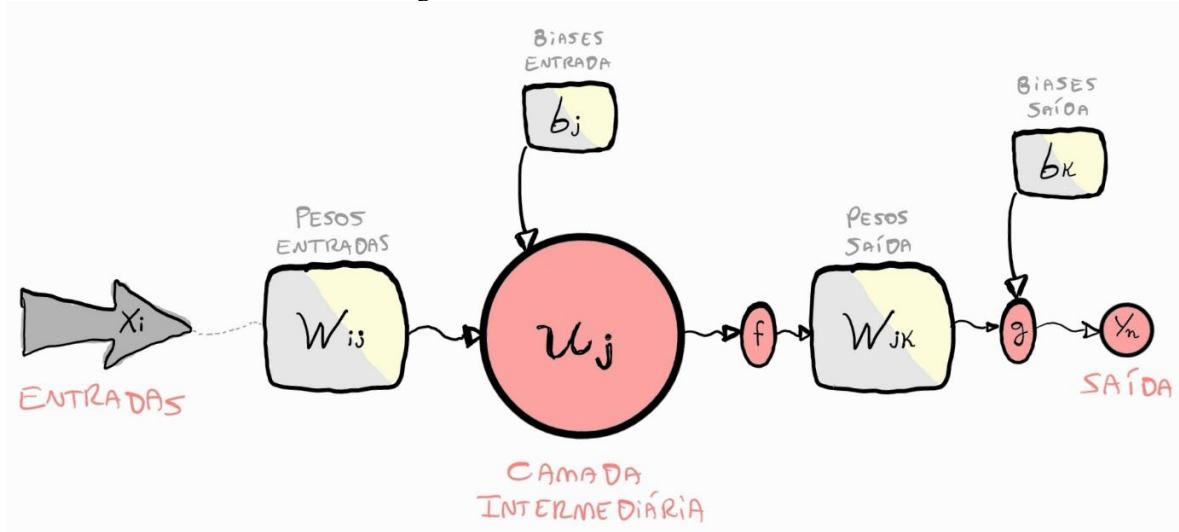
A analogia mais abrangente entre neurônios artificiais e biológicos reside no fato de que as conexões entre os nodos dos neurônios artificiais representam os axônios e dendritos, enquanto os pesos correspondem às sinapses, e o limiar se assemelha à atividade na soma. É fundamental destacar que tanto as Redes Neurais Artificiais

(RNAs) quanto as redes biológicas aprendem por meio do ajuste incremental dos pesos ou da força das sinapses (Silva, 2016).

A estrutura de uma RNA é geralmente organizada em camadas: uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída. A informação flui da camada de entrada, através das camadas ocultas, até a camada de saída. As conexões entre os neurônios, conhecidas como sinapses, são onde os pesos são armazenados e ajustados durante o treinamento (Sperduti, 2001).

A Figura 6 apresenta o modelo matemático de uma rede neural artificial de uma camada oculta (podendo existir diversas dessas camadas intermediárias dependendo da complexidade do modelo).

Figura 6 – Neurônio Matemático.



Fonte: Autoria Própria.

De forma implícita, j representa um neurônio presentes na camada oculta (ou de saída) unitária adotada como u_j . Nesse modelo, o neurônio recebe um número i de entradas (inputs), identificadas como x_i . Cada entrada está associada a um respectivo peso, representado por w_{ij} , sendo i de onde está vindo (no caso a entrada), até j , seu respectivo neurônio na camada oculta. Assim como w_{ik} , tendo j como o referencial do respectivo neurônio da camada oculta até k , a saída do modelo. O limiar (viés) da camada oculta é representado por b_j e saída, por b_k . A função de ativação da camada oculta é representada por f , e a função de ativação de saída é g , enquanto a saída é denotada por y_k , com k representando o número de saídas.

Matematicamente, o modelo pode ser representado pelas Equações 4 e 5. É importante destacar que essas equações permitem determinar o valor de ativação de cada neurônio (Braga, 2000).

$$u_j = f \left(\sum_{i=1}^X w_{ij} x_i + b_j \right)_j \quad (4)$$

$$y_k = g \left(\sum_{j=1}^Z w_{jk} u_j + b_k \right)_k \quad (5)$$

A junção somatória é indicada por Σ , que representa a soma ponderada das entradas multiplicadas pelos seus pesos sinápticos (coeficiente de ativação) acrescidas de seus respectivos vieses.

Cada neurônio em uma RNA recebe uma série de entradas, as processa e produz uma saída ou diversas. As entradas são ponderadas com base em pesos, que são ajustados durante o processo de treinamento da rede, orquestradas pelo algoritmo de otimização que sempre compara o valor produzido no final com o esperado, nisso ajusta os pesos conforme o erro e a tolerância imposta, recomeçando o processo até que haja um valor que sacie o erro ou as condições de parada. A soma ponderada das entradas passa por uma função de ativação de saída, que determina a saída do neurônio sendo aqui, introduzida para conferir não-linearidade ao modelo. Essa adição é crucial nas redes neurais artificiais para permitir saídas com valores distintos, além de proporcionar ao modelo a capacidade de representar relações não-lineares. Diversas funções de ativação podem ser empregadas, sendo as mais comuns as funções: linear, tangente hiperbólica e sigmoide (Dayhoff, 2001).

Unindo as equações (4) e (5), tem-se a modelagem que fornece a equação 6 para uma rede de uma camada genérica (podendo ser expandida para mais camadas intermediárias:

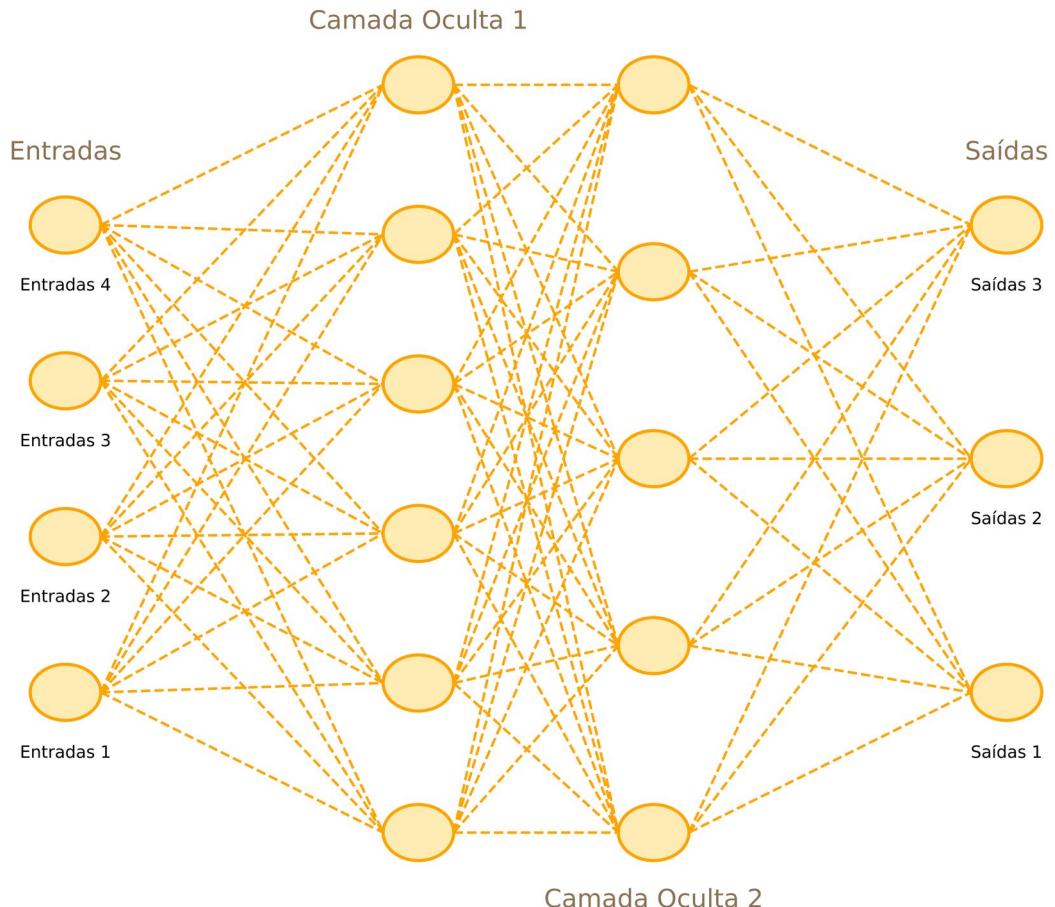
$$y_k = g \left(\sum_{j=1}^Z w_{jk} f \left(\sum_{i=1}^X w_{ij} x_i + b_j \right)_j + b_k \right)_k \quad (6)$$

As Redes Neurais Artificiais (RNAs) podem assumir uma variedade de estruturas, e a escolha da configuração adequada está intimamente relacionada com o algoritmo de aprendizado utilizado para treinar a rede. Esse algoritmo influencia o

número de neurônios e camadas presentes na rede, bem como a organização dos neurônios. As RNAs podem ser essencialmente divididas em três categorias distintas: redes *feedforward* com apenas uma camada oculta, redes *feedforward* de múltiplas camadas e redes recorrentes (Haykin, 2000).

A Figura 7, ilustra uma rede genérica de múltiplas camadas de 4 entradas e 3 saídas, interligadas por 2 camadas ocultas de 6 e 5 neurônios, respectivamente. Tal processo genérico ilustrado pela Figura 6, poderia (por exemplo), representar uma gama de eventos, tais como: temperatura (T), pressão (P), calor (Q) e vazão volumétrica (\dot{V}) como entradas, prevendo volume (V) densidade (ρ) e energia interna (U) como saída, ou em processos mais complexos, como a aplicação em um sistema de previsão de vendas para uma loja de varejo, onde as quatro variáveis de entrada seriam: o preço do produto, os gastos em publicidade, a temporada do ano e o volume de estoque disponível. As três saídas desejadas poderiam incluir: o número de unidades vendidas de um produto específico, a receita total obtida e a margem de lucro diária.

Figura 7 – Rede com duas camadas intermediárias.



Fonte: Autoria Própria.

Nesse contexto, a rede neural seria treinada com dados históricos para aprender as relações entre variações no preço, investimentos em publicidade, sazonalidade e estoque, e como essas variáveis impactam os resultados da loja. Com base nas previsões geradas, a loja poderia ajustar suas estratégias de preços, alocação de orçamento em marketing e gestão de estoque, visando maximizar as vendas, otimizar a receita e melhorar a margem de lucro. Esse exemplo demonstra como redes neurais com múltiplas variáveis de entrada e saída podem ser aplicadas a problemas reais e complexos de previsão.

O processo de treinamento de uma RNA envolve a apresentação de um conjunto de dados de entrada e saída desejada à rede. A rede faz previsões com base nas entradas e ajusta seus pesos para minimizar a diferença entre suas previsões e as saídas desejadas. Esse processo é repetido várias vezes até que a rede alcance um nível aceitável de precisão. Existem vários algoritmos de treinamento disponíveis, sendo o algoritmo de retropropagação um dos mais populares, esse processo é crucial para que as RNAs sejam eficazes na solução de uma ampla gama de problemas, incluindo aqueles relacionados ao monitoramento e otimização de processos, como na fermentação de hidromel (Michaelides, 2001).

Dois aspectos críticos a serem cuidadosamente definidos são o número de camadas e de neurônios na configuração da rede. Geralmente, adicionar duas ou mais camadas intermediárias pode facilitar o processo de treinamento. No entanto, a incorporação de um grande número de camadas não é recomendada, pois pode resultar em uma rede superdimensionada e menos precisa, uma vez que erros se propagam para camadas anteriores (sendo este, o intuito deste presente trabalho, tentando trazer o menor esforço computacional possível, ou seja, uma rede mais “simples”), porém, é válido salientar que em modelos mais complexos, o número de camadas é proporcionalmente aumentado, tornando o parâmetro de esforço computacional incontrolável. A quantidade de neurônios em uma rede neural é influenciada por diversos fatores, como o tamanho do conjunto de treinamento, a presença de ruído nos dados, a complexidade da função a ser aprendida e a distribuição estatística dos dados. Portanto, é válido salientar que, sua definição geralmente envolve uma abordagem empírica (Braga *et al.*, 2000).

Um modelo empírico é uma representação baseada em observações e experimentações diretas, em vez de derivada exclusivamente de teorias ou princípios

matemáticos. Em vez de partir de leis ou fórmulas predefinidas, ele é desenvolvido a partir de dados coletados, buscando padrões e relações que permitam fazer previsões ou descrições sobre determinado fenômeno. Esse tipo de modelo é especialmente útil quando o sistema em estudo é complexo e difícil de ser descrito apenas por equações matemáticas.

A abordagem empírica, por sua vez, é o método de desenvolver modelos ou tomar decisões baseado na experimentação e observação. Nesse caso, diferentes configurações ou variáveis são testadas e os resultados são avaliados para verificar qual delas é a mais eficaz para o problema em questão. Em redes neurais, modelo implementado nesse trabalho, uma abordagem empírica envolve ajustar parâmetros como a quantidade de neurônios, taxas de aprendizado e estruturas de camadas até encontrar a configuração que oferece o melhor desempenho com base nos dados disponíveis.

É essencial evitar o uso excessivo de neurônios, o que pode impedir a rede de extrair características gerais que permitam o reconhecimento de padrões não encontrados durante o treinamento, um problema conhecido como sobreajuste (*sobreajuste*). Por outro lado, utilizar um número muito reduzido de neurônios pode levar a redes que levam um tempo excessivo para encontrar uma representação ideal, ou seja, subajustado. Quando a quantidade de dados de treinamento é muito maior que o número de conexões, o sobreajuste é improvável, mas pode fazer com que a rede não alcance a convergência durante o treinamento (Braga *et al.*, 2000).

3.8.2 Algoritmos de Otimização

O entendimento profundo dos algoritmos de treinamento começa com uma análise detalhada do modelo de aprendizado de máquina segundo Goodfellow (2015), o sistema absorve informações do ambiente, que são processadas e integradas à sua base de conhecimento. Esta base é, subsequentemente, utilizada para realizar tarefas específicas. Dada a natureza frequentemente incompleta das informações do ambiente, a máquina não tem uma compreensão total a priori de como preencher as lacunas em seus dados. Assim, ela opera com base em inferências e, após a execução, recebe *feedback*, possibilitando a avaliação e revisão de suas suposições (Goodfellow, 2015).

De acordo com LeCun, Bengio e Hinton (2015), o aprendizado é o mecanismo pelo qual os parâmetros de uma rede neural são otimizados. Existem múltiplos

algoritmos de treinamento, cada um com suas características distintas. A principal distinção entre eles, conforme discutido por Kingma e Ba (2015), é a metodologia de ajuste dos pesos (Lecun, 2015; Kingma; Ba, 2015)

O processo de treinamento da rede é guiado pelas Equações 7 e 8, onde $w(i)$ simboliza os pesos sinápticos, $b(i)$ denota os vieses, e i indica a iteração corrente. Durante este processo, ambos os parâmetros são refinados para melhorar a eficácia da rede.

$$w_{i+1} = w(i) + \Delta w(i) \quad (7)$$

$$b_{i+1} = b(i) + \Delta b(i) \quad (8)$$

Dada a diversidade de algoritmos de treinamento existentes, este estudo focará nos que têm mostrado excelência em pesquisas correlatas e que serão utilizados na configuração da Rede Neural Artificial para este projeto. Os algoritmos mais pertinentes para esta análise são: *Resilient Backpropagation*, *Scaled Conjugate Gradient*, Levenberg-Marquardt, Levenberg-Marquardt com Regularização Bayesiana e *Conjugate Gradient* com reinicializações de Beale entre outros.

No contexto de treinamento de Redes Neurais Artificiais (RNAs) utilizando a plataforma MATLAB, diversos algoritmos são disponibilizados para otimizar o ajuste dos pesos e viés da rede. Cada algoritmo apresenta características intrínsecas que o tornam mais adequado para determinados cenários e tipos de problemas (Tekin, 2021).

O algoritmo identificado pela *trainrp*, denominado Gradiente Resiliente, adota uma metodologia de atualização de peso menos influenciada pela escala das derivadas. Esta particularidade o torna uma opção viável para situações em que os pesos e viés da rede necessitam de ajustes em escalas variadas, aumentando a variação bruscamente, tentando sair das “amarras” de um possível mínimo local (Riedmiller, 1993).

O algoritmo *trainscg*, ou Gradiente Conjugado Escalonado, fundamenta-se no método do gradiente conjugado. Este é notório por sua eficiência e rápida convergência, mostrando-se especialmente benéfico para RNAs com uma vasta quantidade de parâmetros (Karhe; Kale, 2020). Dada sua eficácia, é frequentemente indicado como uma escolha versátil para o treinamento de RNAs (Møller, 1993). Por sua vez, o algoritmo *trainlm*, baseado no método de Levenberg-Marquardt, destaca-se pela acelerada convergência, sendo altamente efetivo em problemas relacionados à regressão

e aproximação de funções. Este método é particularmente recomendado para redes neurais que contam com uma única camada oculta (Marquardt, 1963).

O *trainbr*, que associa o método de Levenberg-Marquardt à regularização Bayesiana, incorpora mecanismos de regularização ao treinamento, visando prevenir o fenômeno de sobreajuste. Esta estratégia é crucial quando se identifica um risco elevado de ajuste excessivo aos dados de treinamento (MacKay, 1992; Awang *et al.*, 2018).

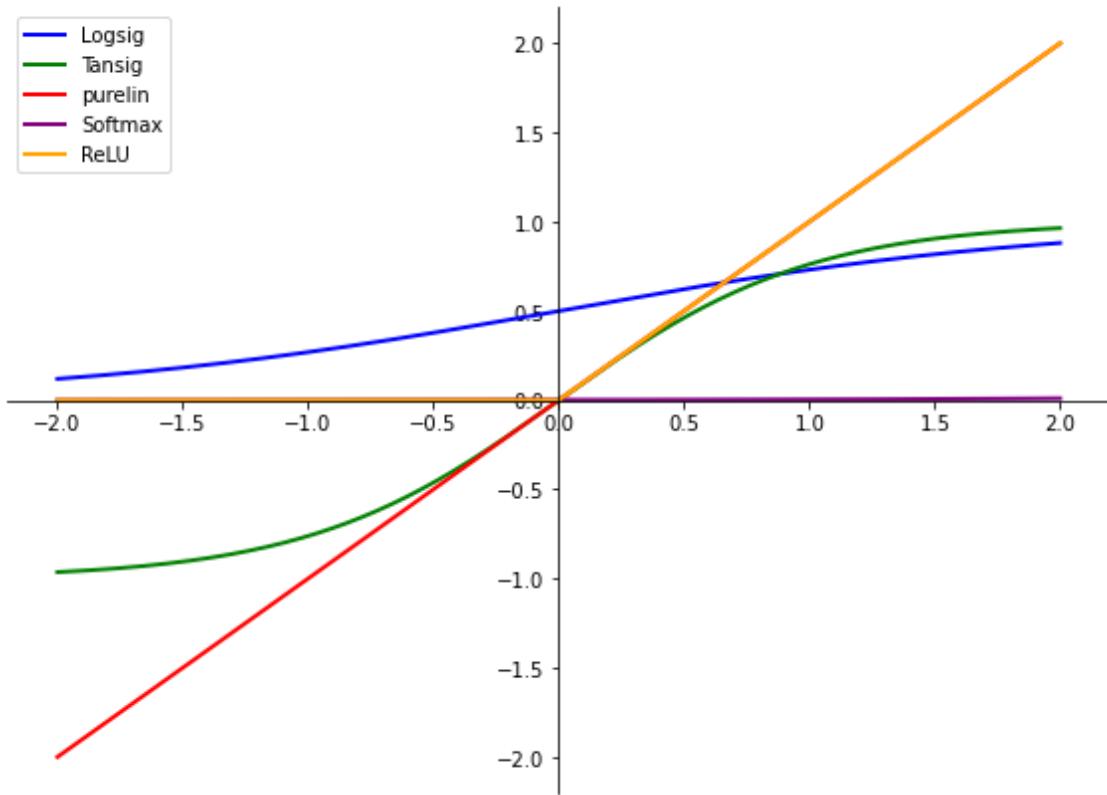
Finalmente, o algoritmo *traincgb*, conhecido como Gradiente Conjugado com Reinicializações de Powell-Beale, integra o método do gradiente conjugado a reinicializações específicas, que atuam como uma estratégia para evitar armadilhas comuns de convergência em mínimos locais. Essas reinicializações são aplicadas de forma seletiva durante o processo de treinamento, permitindo que o algoritmo retome o cálculo a partir de uma nova direção quando a convergência se torna lenta ou o avanço em direção ao mínimo global é comprometido. Essa característica torna o *traincgb* uma alternativa valiosa em cenários onde outros métodos encontram dificuldades para superar barreiras de otimização e alcançar uma solução eficaz (Sherali, 1990).

É imperativo enfatizar que a seleção do algoritmo de treinamento deve sermeticulosamente realizada, considerando as peculiaridades do problema e dos dados em análise. Processos de experimentação e validação cruzada são essenciais para determinar o algoritmo mais propício para cada contexto (Tekin; Can, 2021).

3.8.3 Funções de Ativação

Conforme observado por Braga *et al.* (2000), o modelo original de Rede Neural Artificial (RNA), proposto por McCulloch e Pitts em 1943, passou por inúmeras variações ao longo do tempo. Essas variações possibilitam que a RNA gere não apenas saídas binárias (zero e um), mas também saídas de natureza diversa, utilizando diferentes funções de ativação. Dal'Boit (2008) enfatiza que a função de ativação pode ser tanto uma função linear quanto não linear de u_i , e algumas das funções de ativação mais comuns incluem a função logaritmo-sigmoide (*logsig*), a tangente hiperbólica-sigmoide (*tansig*), a função linear (*purelin*) e duas outras funções de ativação comuns como a linear retificada positiva (*poslin*) e a função exponencial normalizada (*softmax*). Representações gráficas dessas funções estão ilustradas na Figura 8.

Figura 8 – Funções de ativação empregadas no modelo.



Fonte: Autoria Própria.

Um ponto relevante a ser destacado é que o único requisito essencial para uma função de ativação é que ela seja diferenciável, ou seja, que seja uma função contínua (Haykin, 2000). As Equações de 9 a 13 representam, respectivamente, a função logaritmo-sigmoide, a tangente hiperbólica-sigmoide e a função linear (Representadas pela Figuras 8). Essas funções desempenham um papel fundamental na capacidade da RNA de aprender e modelar informações complexas em diversos tipos de problemas.

$$f(x) = \left(\frac{1}{1 + e^{-x}} \right) \quad (9)$$

$$f(x) = \left(\frac{2}{1 + e^{-2x}} \right) - 1 = \left(\frac{e^x - e^{-x}}{e^x + e^{-x}} \right) \quad (10)$$

$$f(x) = x \quad (11)$$

$$f(x)_i = \frac{e^{x_i}}{\sum_{j=1}^k e^{x_j}} \quad (12)$$

$$f(x) = \max(0, x) \quad (13)$$

A Equação (9) também chamada de função sigmoide logística, essa função mapeia a entrada para um intervalo entre 0 e 1, também sendo reconhecida em algumas literaturas comumente na literatura de $\sigma(x)$. É amplamente utilizada em camadas ocultas de redes neurais mais antigas e em modelos onde se requer uma saída probabilística binária. A *logsig* é particularmente sensível a valores extremos e pode sofrer saturação, onde gradientes se tornam muito pequenos e prejudicam o aprendizado (problema do *vanishing gradient*).

Conhecida também como tanh a Equação (10), essa função transforma a entrada em um intervalo entre -1 e 1, também representada como $\tanh(x) = 2\sigma(2x) - 1$. Assim como a *logsig*, a *tansig* é utilizada para camadas ocultas, mas possui uma amplitude maior (entre -1 e 1), o que pode facilitar o aprendizado, reduzindo problemas de saturação. A função linear pura (Equação (11)) é simples, mapeando a entrada diretamente para a saída. É comum em camadas de saída para redes de regressão, pois não comprime a saída e permite que ela represente uma faixa contínua de valores, em vez de valores restritos, como em funções de ativação não lineares.

Também conhecida como *ReLU* (Rectified Linear Unit), a função *poslin*, Equação (12), mapeia entradas negativas para zero e mantém as entradas positivas inalteradas. É uma função amplamente usada em camadas ocultas por sua simplicidade e eficiência em evitar o problema do *vanishing gradient*, além de introduzir não linearidade sem a saturação das funções sigmoides. O *vanishing gradient* é um problema que ocorre durante o treinamento de redes neurais profundas, especialmente ao usar funções de ativação como a sigmoide ou a tangente hiperbólica. Em termos simples, o problema acontece quando os gradientes (as mudanças necessárias para ajustar os pesos da rede) ficam muito pequenos à medida que vão sendo propagados para as camadas iniciais da rede.

A função *softmax*, ou exponencial normalizada, representada pela Equação (13), é usada em camadas de saída para problemas de classificação multiclasse. Ela transforma um vetor de saídas em uma distribuição de probabilidade, com cada valor entre 0 e 1 e a soma total igual a 1. O *softmax* facilita a interpretação probabilística e a escolha da classe mais provável, essencial em tarefas de classificação.

3.8.4 Treinamento, Validação, Teste e Simulação

O treinamento de uma Rede Neural Artificial (RNA) é essencialmente um problema de otimização não linear sem restrições, no qual os pesos sinápticos são ajustados iterativamente com o objetivo de minimizar o erro quadrático médio entre a saída desejada e a saída real da rede. Do ponto de vista estatístico, essa fase envolve a estimativa dos parâmetros do modelo com base em um conjunto de dados disponíveis (Hastie, 2009). Vários fatores desempenham um papel crucial nesta etapa, incluindo a inicialização da rede, o tempo de treinamento, o número de camadas intermediárias, a quantidade de neurônios em cada camada e o volume de dados disponíveis para o treinamento (Silva, 2016).

O tempo de treinamento pode ser afetado por diversos fatores, e é fundamental estabelecer um critério de parada apropriado. Geralmente, o treinamento é concluído quando a rede demonstra uma boa capacidade de generalização e a taxa de erro atinge um valor aceitável. Encontrar esse ponto ideal de parada, onde o erro é mínimo e a capacidade de generalização é máxima, é uma tarefa crítica. O número máximo de ciclos de iteração (épocas) também influencia o tempo de treinamento (Dal'boit, 2008).

Em qualquer problema que envolva Redes Neurais Artificiais com treinamento supervisionado, é necessário ter acesso aos pares de entrada e saída conhecidos. Os dados utilizados para modelagem devem ser divididos em conjuntos distintos, permitindo que as etapas de treinamento, validação, teste e simulação sejam realizadas sem a reutilização dos mesmos dados. Embora diferentes proporções possam ser empregadas, geralmente a maior parte dos dados é reservada para o treinamento (Hastie, 2009). Após o treinamento, a avaliação do desempenho da rede é conduzida utilizando um conjunto de teste independente. Nesse estágio, a rede é testada em pares de entrada-saída que não foram usados durante o treinamento, permitindo uma avaliação objetiva de sua capacidade de generalização (Hastie, 2009).

Por último, pode existir uma etapa adicional chamada de simulação (ou validação externa), na qual a RNA é testada sob outras condições diferentes das quais o modelo foi treinado. Essa etapa também serve para evitar o sobreajuste, no qual a RNA se ajusta excessivamente aos dados de treinamento, perdendo sua capacidade de generalização. Vale ressaltar que a maioria dos estudos existentes na literatura relacionados a Redes Neurais Artificiais se concentra nas etapas de treinamento, validação interna e teste, ignorando a etapa de simulação. No entanto, a simulação desempenha um papel crucial

na confirmação da capacidade preditiva da RNA e pode fornecer resultados que contradizem as conclusões obtidas nas etapas anteriores (Melo, 2020).

Um exemplo notável da importância da simulação foi evidenciado por Ottaiano *et al.* (2021). Neste estudo, a etapa de validação inicial sugeriu que uma configuração de rede era a melhor opção, mas a simulação subsequente revelou que essa rede não era capaz de prever adequadamente as variáveis de saída. A decisão final sobre a melhor RNA foi, portanto, baseada na etapa de simulação.

Ademais, outro trabalho que ratifica a importância da parte da simulação, é o de Valera (2021). Neste trabalho, mostra que um dos métodos mais comuns para controlar o SO₂ é o processo de dessulfurização usando torre de pulverização. Devido ao grande número de parâmetros a serem avaliados e à complexidade do processo, sendo observadas dificuldades na proposição de modelos matemáticos para prever a eficiência de remoção e o coeficiente volumétrico de transferência de massa em fase gasosa (k_{ga}). O objetivo deste estudo foi obter uma rede neural artificial (RNA) para prever a eficiência de remoção e o k_{ga} em uma torre de pulverização para remoção de SO₂. Os resultados mostraram que a escolha do melhor modelo a partir das etapas de treinamento e validação não gerou resultados confiáveis. A melhor estrutura foi definida analisando os resultados de uma etapa de simulação, que utilizou dados independentes. O melhor modelo foi obtido com a estrutura 5-9-2, treinada usando o algoritmo Levenberg-Marquardt com Regulação Bayesiana (*trainbr*) e tendo as funções de transferência softmax e linear nas camadas ocultas e de saída, respectivamente. Essa rede apresentou um erro médio de 8,44% para a concentração de SO₂ na saída e 4,53% para o k_{ga} . Este trabalho demonstrou que o uso de redes neurais é promissor na previsão de variáveis importantes nos processos de remoção de poluentes do ar em torres de pulverização.

3.8.5 Aplicações de RNAs em Processos Fermentativos

Há uma extensa quantidade de estudos relacionados ao uso de Redes Neurais Artificiais (RNAs) aplicadas à fermentação alcoólica disponível na literatura. Esses estudos geralmente visam a criação de modelos capazes de descrever as características mais significativas do processo de fermentação ou a utilização desses modelos para fins de controle e otimização. Abordaremos alguns desses trabalhos em ordem cronológica.

Zhang *et al.* (1994) desenvolveram um protótipo de sistema baseado em RNAs para monitorar as fermentações em batelada de *Bacillus thuringiensis*. Neste estudo, vários

parâmetros foram utilizados como entradas para a RNA, incluindo o tipo de inóculo, a temperatura de operação, o valor de pH, o tempo de processo acumulado, a densidade óptica do mosto e a variação na Densidade Óptica (D.O.). A RNA, neste caso, previa como saída a densidade óptica para o próximo período de amostragem.

O sistema foi implementado tanto em simulações computacionais quanto em experimentos de fermentação em escala laboratorial. Os resultados demonstraram que a RNA do tipo *feedforward Backpropagation* melhorou a supervisão do processo fermentativo e manteve a operação em condições ótimas durante o processo. O sucesso deste estudo foi em parte atribuído à preparação adequada dos dados de treinamento, incluindo a normalização dos dados, e à configuração apropriada (Zhang *et al.*, 1994).

Um estudo abrangente sobre o tema foi conduzido por Radke (2002), que investigou técnicas de otimização para o processo de fermentação alcoólica com o objetivo de determinar faixas operacionais adequadas e desenvolver um modelo do processo. Para isso, Radke combinou equações de balanço de massa (partes de um modelo fenomenológico) com RNAs (técnica empírica) para descrever a cinética desconhecida do processo (se caracterizando como um modelo semiempírico). O estudo também considerou o uso de medidas secundárias, como turbidez, pH e concentração de sólidos solúveis ($^{\circ}$ Brix), como variáveis de entrada para a RNA. Foram elaborados três modelos de RNA, cada um com quatro variáveis de entrada e uma variável de saída correspondente a uma taxa cinética específica de crescimento microbiano (Radke, 2002).

Os resultados demonstraram um bom desempenho das RNAs na descrição da cinética do processo. Além disso, o estudo ressaltou que a escolha da função de ativação, particularmente a função tansig nas camadas intermediárias, resultou em melhores resultados de validação em comparação com a função logsig. Radke também desenvolveu dois modelos híbridos, um dos quais empregou variáveis primárias (concentração de: X - células, S – Substrato, P - etanol) como entradas da RNA e obteve resultados satisfatórios na aproximação de dados desconhecidos de concentração celular, de açúcares redutores e produto (etanol). O segundo modelo híbrido usou variáveis secundárias (pH, turbidez e concentração de sólidos solúveis ($^{\circ}$ Brix)) como entradas e alcançou resultados ainda melhores que o primeiro. Este estudo destacou a importância de modelos elaborados com variáveis facilmente mensuráveis para melhorar o monitoramento do processo e sugeriu que essas variáveis poderiam ser incorporadas ao modelo em tempo real (Radke, 2002).

Uma das bases deste trabalho, estabelecida no estudo de Amgarten em 2012, utilizou redes neurais artificiais para desenvolver um modelo matemático preditivo capaz de estimar o rendimento do processo de fermentação. Esse modelo foi construído com entradas selecionadas de forma criteriosa, empregando o método estatístico de correlação linear para identificar as variáveis mais relevantes entre uma série de parâmetros monitorados em usinas de produção de etanol e açúcar. A seleção cuidadosa das variáveis de entrada foi essencial para aprimorar a precisão do modelo, permitindo que ele refletisse de maneira mais fiel as condições reais de produção industrial e, assim, oferecesse uma ferramenta eficiente para a previsão de rendimentos. O estudo de Amgarten destaca a importância de métodos estatísticos na escolha das variáveis em modelos preditivos, reforçando o valor da integração entre inteligência artificial e análise estatística para otimizar processos industriais (Amgarten, 2012).

Santos, em 2018, desenvolveram um sensor virtual neural em planilha eletrônica para monitorar a produção de biossurfactante por *Bacillus subtilis* usando apenas casca de beterraba como meio de cultura. Embora este estudo envolvesse um cultivo aeróbico, sua aplicação é semelhante ao desenvolvimento de RNAs. Os autores basearam-se em uma rede multicamada, no algoritmo de treinamento de Levenberg-Marquardt com regularização Bayesiana. As variáveis de entrada incluíam concentração celular, concentração de glicose, tensão superficial, concentração de oxigênio dissolvido e tensão superficial diluída em dez ou cem vezes. A saída da RNA era a concentração de biossurfactante bruto. O uso da Densidade Óptica (D.O.) como entrada da RNA foi adotado devido à rapidez desse método em comparação com a gravimetria. Os resultados revelaram que a RNA obteve um coeficiente de determinação (R^2) de 0,9955, o que permitiu o desenvolvimento de um sensor virtual neural (Santos *et al.*, 2018).

Silva, em 2021, empregou na criação do sensor virtual, duas estratégias de abordagens distintas. Na primeira estratégia, as concentrações de células e açúcares redutores foram previstas separadamente por duas RNAs. Na segunda estratégia, uma única RNA previu ambas as variáveis simultaneamente. Tendo um sensor virtual desenvolvido com potencial para monitorar e otimizar a produção de bebidas, visando maximizar o rendimento e a produtividade, também empregando como dados de entrada variáveis como pH, D.O. e concentração de sólidos solúveis ($^{\circ}$ Brix).

E mais recentemente, em um estudo envolvendo a fermentação de mel com adição de extrato de casca de jaboticaba, foi obtida uma rede que fornece previsões simultâneas de X, S e P com soma dos erros MAPE (Erro médio absoluto percentual) de $10^{-3}\%$ com

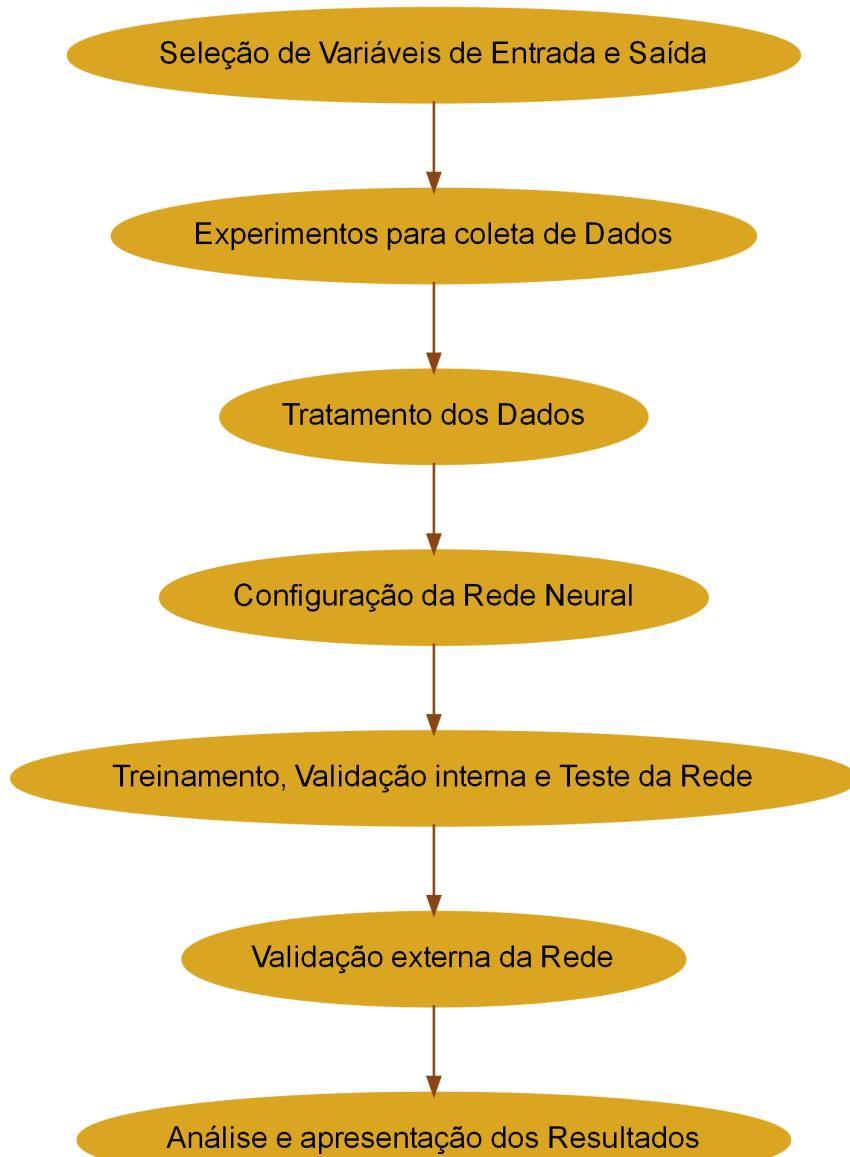
somente uma camada intermediária de configuração 3-15-3 utilizando também pH, D.O. e concentração de sólidos solúveis ($^{\circ}$ Brix) como variáveis de entrada (Gomes, 2023). Apesar do bom resultado, o trabalho considerou apenas a temperatura de 25 $^{\circ}$ C.

Em resumo, as RNAs são ferramentas poderosas que podem aprender padrões complexos e fazer previsões com base em dados. Elas são amplamente utilizadas em diversas áreas, desde a detecção de padrões em dados climáticos até aplicações médicas, como diagnóstico, prognóstico de doenças, processos químicos desde fermentações e tratamentos de efluentes entre outras diversas áreas.

4 MÉTODOS

A abordagem adotada abrange desde a seleção de variáveis até a análise dos resultados. O processo inicia-se com a coleta de dados por meio de experimentos de fermentação, seguida pelo tratamento desses dados e pela configuração de uma rede neural. Após o treinamento e validação da rede, realizam-se simulações para, finalmente, equacionar e analisar os resultados obtidos, tentando garantir que o sensor virtual tenha sido obtido adequadamente. A Figura 9 apresenta o panorama geral para a construção de uma rede neural.

Figura 9 – Processo geral de obtenção de uma rede neural.



Fonte: Autoria Própria.

4.1 Fermentação

O panorama geral dos métodos experimentais é apresentado de forma abrangente e fundamenta-se integralmente no trabalho de iniciação científica de Costa, realizado em 2021.

Para o preparo do extrato de casca de jabuticaba, a ser utilizado nas fermentações, foi adicionado 50 % (m/m) de casca de jabuticaba, e 50 % de água mineral. Após, a mistura foi triturada em liquidificador por 5 minutos, peneirada usando peneiras de 60 mesh. A fração líquida, denominada de extrato, foi armazenada.

O meio de cultivo utilizado foi uma mistura de 2/3 (vol) de mel (adquirido de produtores rurais na região de São Miguel Arcanjo, SP), diluído com água mineral até atingir uma concentração de sólidos solúveis de 30 °Brix, e 1/3 de extrato de casca de jabuticaba adicionalmente, o meio foi enriquecido com 0,5% de extrato de levedura e 1% de peptona.

Após o preparo, o meio foi submetido a vapor fluente por 10 minutos. Previamente, o fermentador vazio fora esterilizado em autoclave por 20 minutos.

Antes da inoculação no fermentador, uma etapa crucial de preparação do inóculo foi executada. Essa etapa consistiu na hidratação do inóculo, para a qual 1,08 g de *Saccharomyces cerevisiae* da marca comercial Premier Blanc, da Fermentis, foram adicionados a 15 mL de água destilada. Essa quantidade foi utilizada para obter uma concentração inicial de inóculo no fermentador de 0,36 g/L, de acordo com recomendações do fabricante. Em seguida, o inóculo foi mantido em uma estufa a 30 °C por um período de 30 minutos.

As fermentações foram conduzidas em um fermentador do modelo Labfors 5, fabricado pela Infors HT na Suíça, em temperaturas de 20, 25 e 30°. A agitação foi realizada com o auxílio de duas turbinas de pás planas, operando a 200 rpm, conforme descrito por Ferraz (2015). O volume total do meio de cultivo utilizado foi de 3,0 L, (Costa, 2021). No total, foram realizados 5 experimentos de fermentação.

4.2 Obtenção dos dados experimentais

No estudo de Costa (2021) foram adotadas diversas técnicas para a determinação e análise de diferentes parâmetros relacionados à fermentação e composição de amostras. A concentração celular foi determinada utilizando a técnica gravimétrica. Para isso, as amostras foram submetidas à centrifugação a uma velocidade de 4000 rpm durante 15

minutos, utilizando-se uma centrífuga modelo 5702 R da Eppendorf. O precipitado obtido após a centrifugação foi lavado em duas etapas consecutivas e, em seguida, seco em estufa a 70°C, conforme metodologia descrita por Andrade *et al.* (2013).

Para a quantificação da concentração de açúcares redutores, que compreendem a glicose e frutose, e etanol, o sobrenadante das amostras foi analisado utilizando cromatografia líquida de alta eficiência (HPLC), equipado com a coluna Aminex HPX-87 H (Biorad Laboratories), e detector de índice de refração. O volume de injeção foi de 20 µL, temperatura do forno de 30°C, e vazão de fase móvel (solução de H₂SO₄ em pH 1,4) de 0,6 mL min⁻¹ (Andrade *et al.*, 2013).

O monitoramento do pH durante o processo de fermentação foi realizado em tempo real, utilizando um sensor autoclavável adquirido a partir da empresa Mettler Toledo.

A análise da absorbância das amostras (densidade óptica), que foram previamente diluídas em uma proporção de 1:15, foi conduzida em um espectrofotômetro modelo Genesys 10S da marca Thermo Scientific, com um comprimento de onda fixado em 600 nm. Adicionalmente, a quantificação de sólidos solúveis (°Brix), que serve como uma predição da concentração de sacarose, no sobrenadante, foi realizada utilizando um refratômetro digital modelo HI96801 da marca Hanna Instruments.

4.3 Obtenção das RNAs

O desenvolvimento de Redes Neurais Artificiais (RNAs) é uma área de pesquisa em rápido crescimento, com aplicações significativas em diversos campos industriais e científicos. A metodologia que será empregada na construção e treinamento dessas redes foi essencial para assegurar sua eficácia e precisão. Este estudo propôs uma abordagem metodológica para o desenvolvimento de RNAs, cobrindo etapas críticas desde a preparação dos dados até a avaliação dos resultados.

A fase inicial consistiu na preparação e tratamento dos dados experimentais. Foi essencial que os dados coletados estivessem organizados e rigorosamente analisados para identificar e remover possíveis *outliers* reportados, assegurando que estivessem corretamente dispostos e adequados para análises subsequentes. Esse cuidado é fundamental para preservar a integridade dos resultados e maximizar a relevância das etapas posteriores (Goodfellow *et al.*, 2015).

A seleção de variáveis de entrada e saída foi a próxima etapa crítica. A escolha apropriada dessas variáveis foi relevante para a capacidade da RNA de aprender e prever as características essenciais do processo (Goodfellow *et al.*, 2015).

Definir parâmetros para avaliar a eficiência da rede foi outra etapa crucial. Métricas como erro quadrático médio e coeficiente de determinação (R^2) serão frequentemente usadas para essa finalidade (Bishop, 2006).

A configuração das RNAs, incluindo a seleção do número de camadas, neurônios, algoritmo de otimização e funções de transferência, foi fundamental para o desempenho da rede (Lecun *et al.*, 2015).

O treinamento, validação interna e teste das RNAs foram etapas vitais. Durante o treinamento, a RNA ajustou seus pesos com base nos dados fornecidos. A validação interna ajustou parâmetros para evitar sobreajuste, enquanto os dados de teste avaliaram a capacidade de generalização da rede (Schmidhuber, 2015).

Após o treinamento, a RNA passou por uma validação externa e os resultados foram avaliados para determinar sua precisão em prever as variáveis de interesse (Hinton *et al.*, 2012).

4.3.1 Seleção das variáveis de entrada e saída

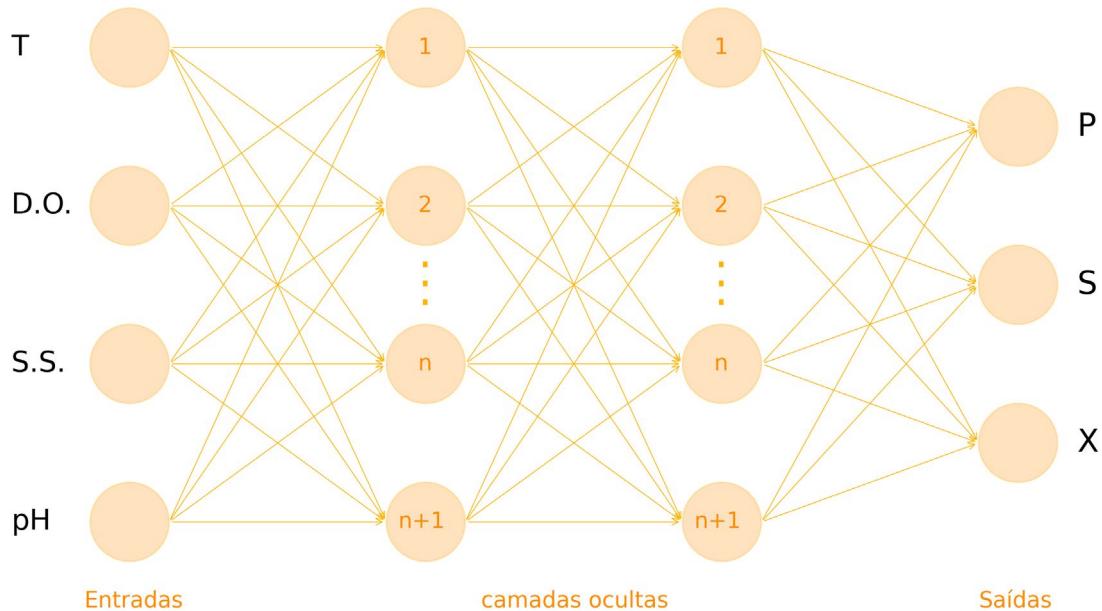
A abordagem envolveu o desenvolvimento de uma única RNA capaz de prever simultaneamente três variáveis essenciais no processo de fermentação: X (concentração de células), S (concentração dos açúcares redutores) e P (concentração de produto), todas em g L^{-1} . Para alcançar esse objetivo, as variáveis de entrada foram cuidadosamente selecionadas, considerando diversos aspectos cruciais para o sucesso da modelagem e visando maximizar a precisão nas previsões da RNA.

Na escolha das variáveis de entrada, foram considerados diversos fatores importantes. O pH, por exemplo, foi incluído devido à sua relevância como uma variável crítica indicadora de produção no processo. A temperatura também foi selecionada, uma vez que regula o processo e possui uma relação diretamente proporcional com a conversão dos reagentes e as velocidades de reação. A densidade óptica (D.O.) foi outra variável de entrada escolhida, por sua relação com o crescimento microbiano, desempenhando um papel essencial na caracterização do processo. Além disso, incorporou-se a concentração de sólidos solúveis (S.S.), pois esta indica o consumo do substrato ao longo da fermentação.

Outro aspecto relevante na escolha dessas variáveis de entrada foi sua facilidade de medição, uma vez que era fundamental que as variáveis fossem práticas, rápidas e acessíveis de serem obtidas durante o processo de fermentação.

A Figura 10 ilustra uma RNA com duas camadas intermediárias, projetada para prever as variáveis P, X e S simultaneamente.

Figura 10 – Rede neural que simboliza o processo.



Fonte: Autoria Própria.

Na Figura 10, pode-se observar a confirmação da interação entre as variáveis de entrada, os *inputs* (T, pH, Sólidos Solúveis (S.S.) e densidade óptica (D.O.)) com a camada intermediária. Cada seta que conecta essas camadas representa um peso que indica a intensidade da comunicação entre elas, ou seja, o impacto que cada variável tem sobre o neurônio intermediário. Da camada intermediária para a camada de saída, os *outputs* (neurônios de saída), tem-se as concentrações das três variáveis: X (células), S (Açúcares redutores) e P (etanol).

As setas que partem dos neurônios da camada intermediária em direção aos neurônios de saída representam os pesos sinápticos, os quais desempenham um papel crucial na facilitação da comunicação entre esses níveis. Esses pesos modulam a intensidade com que os sinais são transmitidos, determinando a contribuição de cada neurônio intermediário para o valor final das variáveis de saída. Em essência, eles

ajustam a influência de cada unidade na predição, permitindo que a rede neural capture padrões complexos e interações específicas entre as variáveis, o que é fundamental para a precisão do modelo preditivo (Braga *et al.*, 2000).

A Figura representa uma arquitetura com uma camada intermediária adicional (rede de duas camadas), o que eleva a complexidade da rede e amplia o número de parâmetros a serem ajustados. Essa configuração aumenta a capacidade da rede de modelar relações não lineares e padrões mais complexos nos dados, permitindo uma captura mais detalhada das interações entre variáveis. Contudo, o aumento de camadas também requer um ajuste cuidadoso para evitar o risco de sobreajuste e garantir a generalização adequada do modelo.

4.3.2 Tratamento dos dados experimentais

Para realizar a modelagem da fermentação, foram utilizados os dados experimentais disponíveis de Costa (2021). No entanto, devido à limitação na quantidade de dados coletados (em torno de 10 a 12 pontos por fermentação), adotou-se uma estratégia comum em estudos envolvendo Redes Neurais Artificiais (RNAs), que consistiu na geração de dados adicionais por meio de ajuste de equações suaves (Becker *et al.*, 2002; Silva; Fileti, 2016). Essa abordagem foi aplicada com êxito em estudos recentes conduzidos pelo grupo Choji *et al.* (2021) e Irizawa, Martins e Veggi (2021).

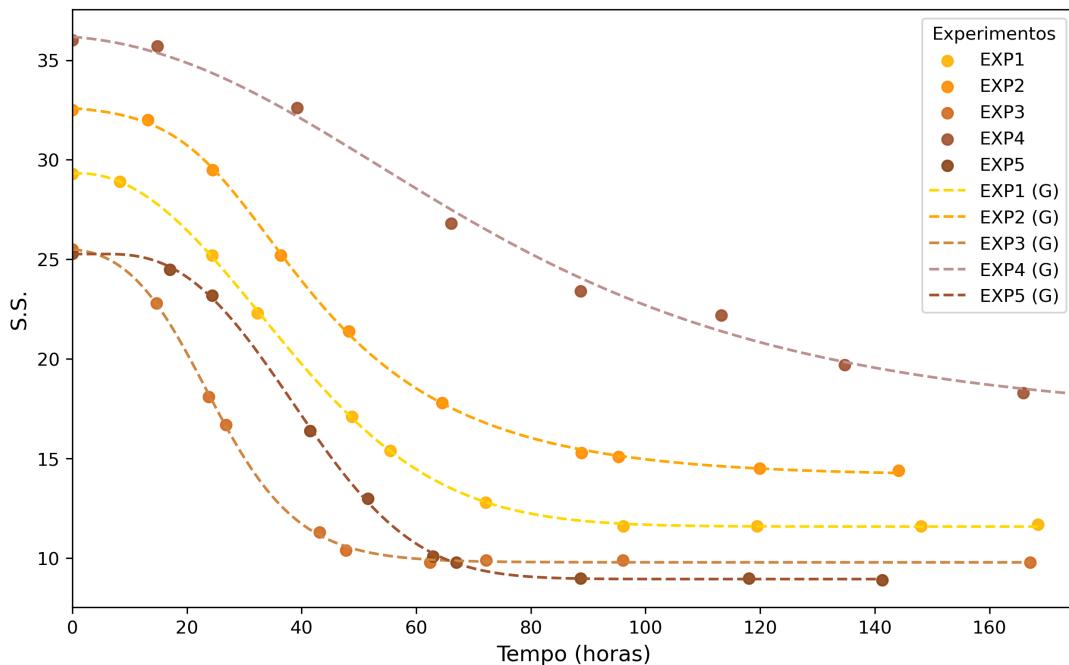
O procedimento de geração de dados, foi realizado no software TableCurve2D, do qual contempla uma interface para imputar as variáveis de entrada e saída juntamente do seu respectivo tempo. Com isso, é gerado um gráfico com uma ampla gama de funções para serem selecionadas. É então, escolhida a função com o ajuste mais suave, mostrando um comportamento que se enquadra à respectiva variável (devendo ser estudada anteriormente) e, então, seleciona-se a função, gerando os dados em formato separado por vírgulas (CSV).

Por meio desse procedimento, foram gerados 200 pontos para cada uma das variáveis de entrada e saída, abrangendo os 5 experimentos de fermentação realizados e mantendo um intervalo de tempo constante. Isso resultou em uma base de dados expandida que melhorou significativamente o processo de treinamento da RNA.

Essa estratégia de geração de dados adicionais foi fundamental para garantir que a RNA fosse devidamente treinada e pudesse aprender com maior precisão as complexas interações do processo de fermentação ao longo do tempo.

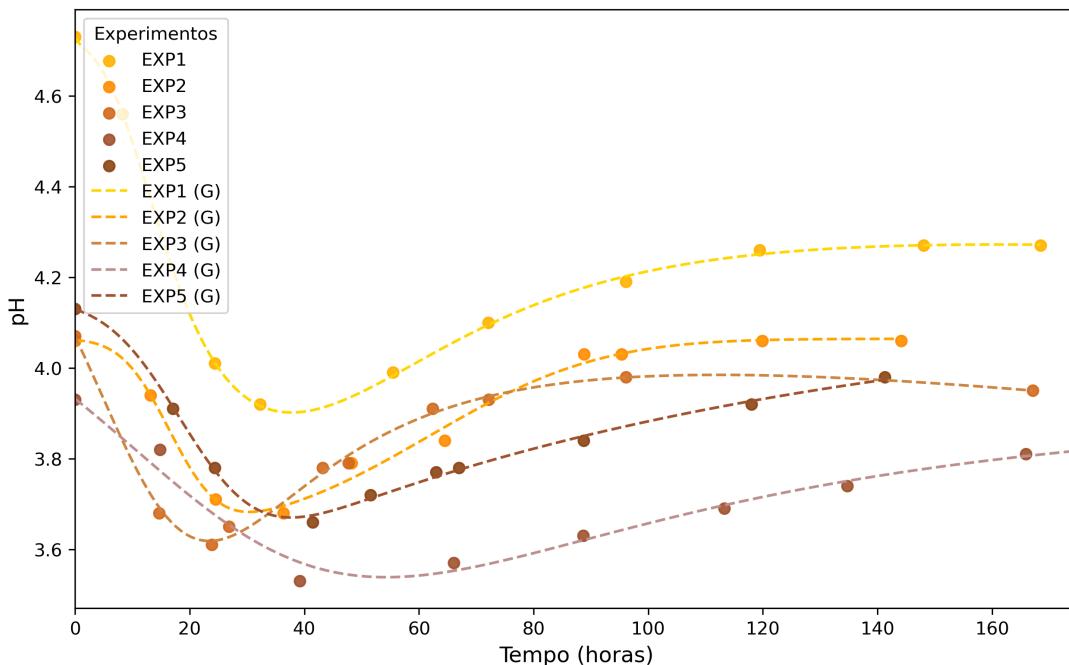
As Figuras 11 a 16 mostram o emprego dos ajustes aplicados todas as variáveis com o intuito de aumentar a representatividade dos dados.

Figura 11 – Ajuste suave dos dados gerados para S.S. contra os dados experimentais em função do tempo.



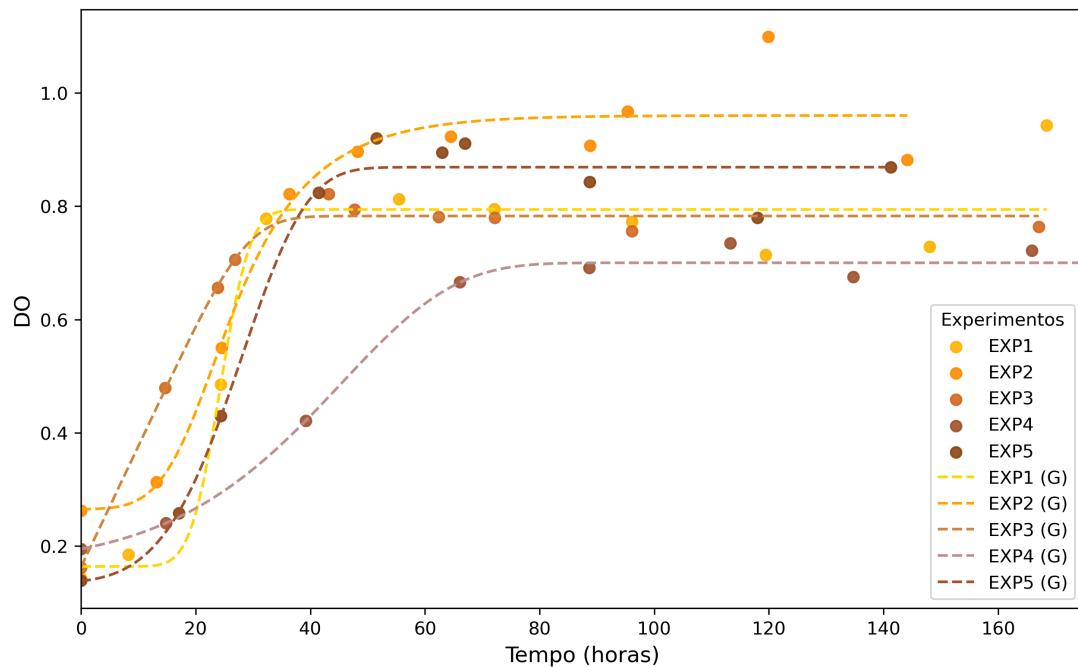
Fonte: Autoria Própria.

Figura 12 – Ajuste suave dos dados gerados para pH contra os dados experimentais em função do tempo.



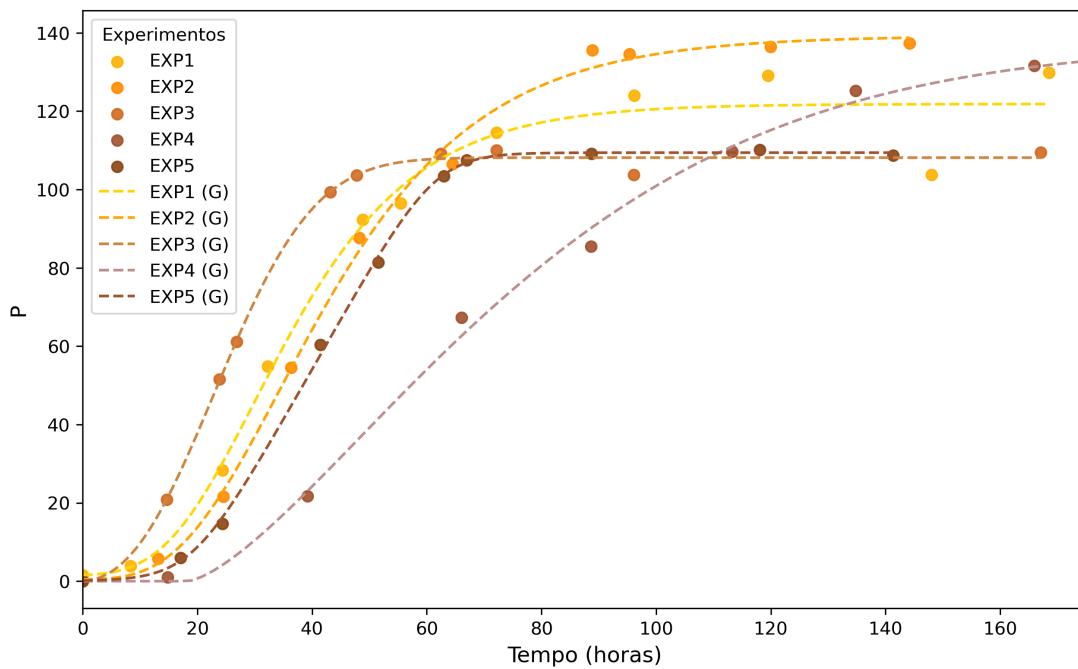
Fonte: Autoria Própria.

Figura 13 – Ajuste suave dos dados gerados para DO. contra os dados experimentais em função do tempo.



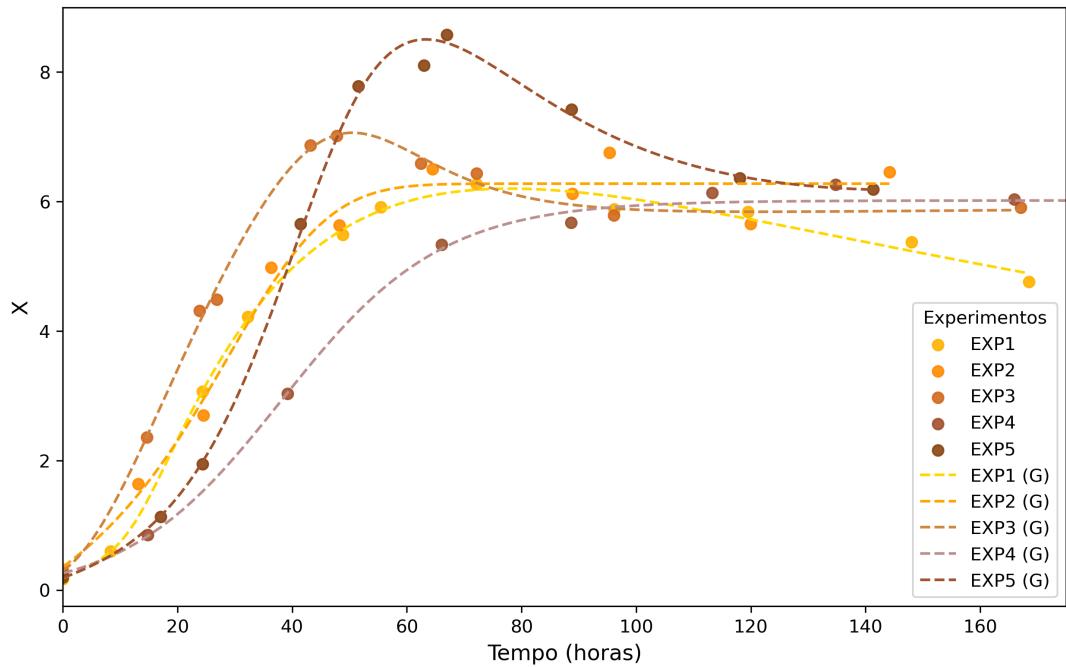
Fonte: Autoria Própria.

Figura 14 – Ajuste suave dos dados gerados para P contra os dados experimentais em função do tempo.



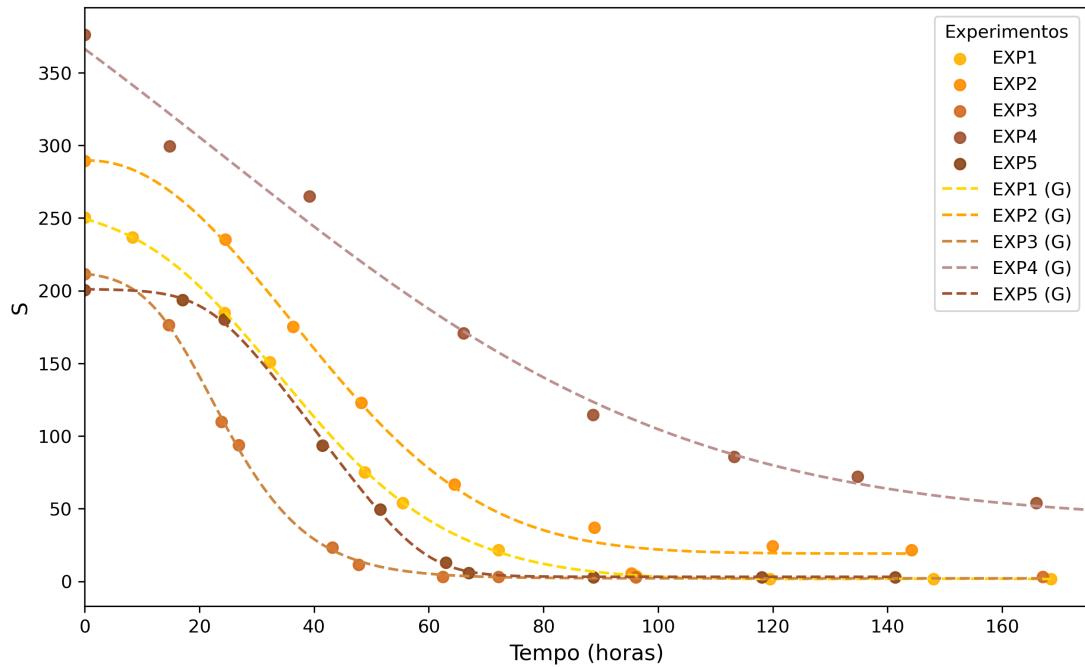
Fonte: Autoria Própria.

Figura 15 – Ajuste suave dos dados gerados para X contra os dados experimentais em função do tempo.



Fonte: Autoria Própria.

Figura 16 – Ajuste suave dos dados gerados para S contra os dados experimentais em função do tempo.



Fonte: Autoria Própria.

Para todas as demais funções, os novos dados foram gerados a partir de funções suaves aplicadas a todas as variáveis. As capturas de tela dos gráficos que geraram essas

funções estão disponíveis no apêndice. Para acessar os dados, o leitor pode fazer uma solicitação no endereço: tdmartins@unifesp.br.

Além disso, também foram realizadas análises estatísticas prévias nas variáveis de entrada: (i) método de importância por permutação para estimar a relevância de cada variável na previsão dos valores de interesse; (ii) teste de normalidade de Shapiro-Wilk; (iii) análise das relações entre as variáveis de saída para verificação de diferentes graus de linearidade; (iv) a matriz de correlação de Spearman para verificação de interrelações entre as variáveis monitoradas durante o processo fermentativo.

4.4 Treinamento e validação das RNAs

4.4.1 Configuração e detalhes do treinamento das RNAs

As fases de obtenção (treinamento, teste e validação) e simulação foram conduzidas no ambiente de desenvolvimento Mathworks Matlab® R2016b. Durante o processo, foram criadas e treinadas redes neurais artificiais (RNAs) com diferentes arquiteturas, incluindo aquelas com uma ou duas camadas intermediárias (também conhecidas como camadas ocultas). Além disso, variou-se o número de neurônios em cada camada intermediária para avaliar o impacto dessas configurações na performance da RNA.

Para a otimização do treinamento das RNAs, foram explorados diversos algoritmos de otimização: o Scaled Conjugate Gradient (*trainscg*), o Levenberg-Marquardt (*trainlm*), o Levenberg-Marquardt com Regularização Bayesiana (*trainbr*), o Resilient Backpropagation (*trainrp*), o Gradiente Conjugado com Reinicializações de Powell-Beale (*traincgb*), o Broyden-Fletcher-Goldfarb-Shanno (*trainbfg*), o Gradiente Descendente com Momentum e Taxa de Aprendizado Adaptativa (*traingdx*), o One Step Secant (*trainoss*) e Gradiente descendente simples (*traingd*) entre outros considerados relevantes para o estudo. Cada um desses algoritmos possui características específicas que influenciaram a convergência e a eficiência do treinamento.

Em relação às funções de ativação utilizadas nas camadas intermediárias, foram testadas cinco opções: a função logaritmo-sigmoide (*logsig*), a tangente hiperbólica-sigmoide (*tansig*), a linear (*purelin*), a linear retificada (ReLU ou *poslin*) e a exponencial normalizada (*softmax*). Essas funções desempenharam um papel fundamental na modelagem das relações entre as camadas da RNA, e sua escolha afetou significativamente o desempenho da rede.

Para as camadas de saída das RNAs, todas as cinco funções de ativação também foram testadas e avaliadas em configurações de uma única camada intermediária, além da segunda camada. Essas escolhas foram feitas com o objetivo de analisar como as diferentes combinações de funções de ativação influenciaram a capacidade da RNA de representar e generalizar os dados de forma eficaz.

Para cada um dos algoritmos de otimização, foram testados de 5 a 40 neurônios na primeira e na segunda (variando de 5 em 5), com todas as 5 possíveis funções de ativação nas duas camadas intermediárias e 3 na saída, sendo cada combinação, retreinada 5 vezes, totalizando 24.000 redes por algoritmo. Portanto, a estratégia principal previu em torno de 216.000 possíveis redes neurais testadas, proporcionando uma avaliação criteriosa para identificar a configuração ideal antes do critério final.

Vale ressaltar que, caso a configuração com a menor média de erros do processo de simulação para cada variável de saída apresentasse um erro inferior ao limite de 10%, essa configuração foi submetida a um retreinamento intensivo, variando de 1.000 a 3.000 repetições, com o objetivo de minimizar o erro ao máximo. Totalizando, em torno de 300 mil redes diferentes ao longo de todo o processo.

A Tabela 1 resume as diversas configurações testadas durante o estudo, abrangendo as variações nos algoritmos de treinamento, nas funções de transferência e na estrutura geral das RNAs. Essa análise permitiu uma avaliação abrangente do desempenho das RNAs em relação a diferentes parâmetros e configurações, proporcionando insights valiosos para a modelagem e simulação do processo em questão.

Tabela 1 – Configuração geral de treinamento.

Algoritmos de otimização	Funções de ativação (1 ^a camada intermediária)	Funções de ativação (2 ^a camada intermediária)	Funções de ativação (saída)	Neurônios (1 ^a camada)	Neurônios (2 ^a camada)
<i>trainscg, trainlm, traincgb, trainbr, trainbfg, traingdx, trainrp, trainoss, traingd</i>	<i>tansig, purelin, softmax, logsig, poslin</i>	<i>tansig, purelin, softmax, logsig, poslin</i>	<i>tansig, purelin, logsig</i>	5 - 40	5 - 40

Fonte: Autoria Própria.

Conforme apresentado na Tabela 1, optou-se por utilizar um número reduzido de neurônios na segunda camada intermediária das redes neurais. Essa decisão foi motivada pela intenção de evitar que as RNAs com duas camadas intermediárias se tornassem excessivamente complexas, com um número elevado de parâmetros a serem

ajustados. Assim, o intervalo de neurônios nessa camada intermediária foi limitado a uma faixa entre 5 e 40 unidades.

Essa abordagem foi adotada com o objetivo de mitigar possíveis impactos adversos no desempenho computacional e no tempo de resposta do processo de treinamento das redes. Evitar uma complexidade excessiva nas RNAs mostrou-se fundamental para garantir que o treinamento fosse eficiente e que a RNA pudesse aprender as relações entre as variáveis de entrada e saída de maneira adequada.

Dessa forma, quando se equilibrar o número de neurônios nas camadas intermediárias, a busca será para encontrar um ponto ótimo que permita que as RNAs com duas camadas continuem eficazes na modelagem do processo, ao mesmo tempo em que mantenham um desempenho computacional gerenciável e um tempo de resposta aceitável. Esse equilíbrio é essencial para garantir que as RNAs sejam uma ferramenta viável e eficiente na abordagem do problema em questão.

Por fim, cada configuração de rede foi retreinada, pois devido à natureza não convexa da função de perda, somado ao fato de que os parâmetros iniciais sempre possuem chutes iniciais aleatórios providos pelo Matlab®, o modelo pode convergir para mínimos locais em vez do mínimo global, levando a soluções subótimas. Para evitar isso, adota-se o retreinamento com diferentes pesos iniciais, ampliando a chance de alcançar o mínimo global ou regiões próximas. Neste estudo, o modelo foi retreinado cinco vezes com inícios aleatórios, aumentando a robustez e a representação fiel do sistema. O monitoramento em tempo real do erro permitiu a interrupção do treinamento em caso de aumento do erro, otimizando recursos e garantindo a generalização do modelo final.

4.4.2 Definição dos parâmetros de eficiência da rede

Para avaliar a eficácia da rede neural, foram utilizados três parâmetros fundamentais: a função objetivo (F_{OBJ}), o erro médio percentual absoluto (MAPE) e o coeficiente de correlação de Pearson (ρ). Ao final de cada uma das quatro etapas (treinamento, validação, teste e simulação), os valores desses parâmetros serão armazenados para uma análise abrangente do desempenho da rede (Silva, 2021).

A função objetivo escolhida será a média da diferença quadrática entre os valores calculados pela RNA e os valores experimentais reais, conhecida como Mean Squared Error (MSE), conforme será expresso na Equação 14:

$$F_{OBJ} = MSE = \frac{1}{n} \sum_{j=1}^n (h_{j,exp} - h_{j,rede})^2 \quad (14)$$

Na Equação 14, n representa o número de pontos experimentais, $h_{j,rede}$ é o valor que será previsto pela RNA, e $h_{j,exp}$ é o valor real para o ponto experimental j . É crucial destacar que quanto mais próxima de zero for a função objetivo (F_{OBJ}), menor será a discrepância entre os dados experimentais e as previsões da RNA, indicando, em geral, um melhor desempenho da rede.

O erro médio percentual absoluto (MAPE) foi calculado com base na Equação 15 a seguir:

$$MAPE = \frac{1}{n} \sum_{j=1}^n \frac{|h_{j,exp} - h_{j,rede}|}{h_{j,exp}} \cdot 100\% \quad (15)$$

O coeficiente de correlação de Pearson, por sua vez, é um parâmetro que avalia a intensidade e a direção da correlação entre duas variáveis, variando de -1,0 a 1,0 (Silva, 2021). Dado o contexto fornecido, a equação para o coeficiente de correlação de Pearson (ρ) é representado pela Equação 16:

$$\rho = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (16)$$

Onde: x_i é o valor da variável x para o ponto experimental i ; y_i é o valor previsto pela RNA (Rede Neural Artificial) para o ponto experimental i ; \bar{x} é a média aritmética dos valores de x ; \bar{y} é a média aritmética dos valores de y .

O numerador da fração representa a covariância entre x e y , enquanto o denominador é o produto dos desvios padrão de x e y . Quando o coeficiente de Pearson é calculado desta forma, ele fornece uma medida da força e direção da relação linear entre as duas variáveis.

Esses parâmetros fornecerão uma visão abrangente da precisão da RNA, permitindo avaliar a concordância entre os resultados a serem previstos e os dados experimentais reais. Quanto menor o MSE e o MAPE, e quanto mais próximo de 1 for ρ , melhor será o desempenho da rede neural no processo de modelagem e simulação (Awang, 2018).

4.4.3 Obtenção da rede

O conceito de “obtenção da rede” foi compreendido pela estruturação das seguintes etapas: treinamento, teste e, por fim, validação, para a qual foram fornecidos 4 experimentos. Foram disponibilizados 800 dados (200 para cada experimento) para a obtenção da rede, utilizando 70% dos dados para treinamento (560 dados), 15% para validação (120 dados) e 15% para teste (120 dados). Um experimento completo (200 dados, escolhido aleatoriamente) foi reservado exclusivamente para a validação externa da rede. Nesse caso, o Experimento 2, escolhido aleatoriamente.

A rede foi então submetida à análise da influência de uma gama de fatores de predição, que regiam as variáveis de saída, como funções de ativação, quantidade de camadas ocultas, número total de neurônios por camada e algoritmos de otimização.

4.4.4 Condições de parada para o treinamento do modelo

A primeira condição de parada, consiste no acompanhamento em tempo real do erro ao longo do processo de validação. Caso o erro comece a aumentar, o processo é interrompido após 20 aumentos consecutivos, pois esse comportamento indica uma possível divergência do modelo, saindo de um ponto ótimo global. Embora o modelo possa estar em um mínimo local, associa-se esse aumento do erro à perda de convergência em direção ao máximo global esperado.

A segunda condição de parada, igualmente relevante, está relacionada à tolerância imposta ao modelo, definida como 1.10^{-8} para a Função Objetivo do Treinamento (Equação 14). A terceira condição estabelece um limite de 1000 iterações como máximo permitido. Por fim, a quarta condição de parada determina que o tempo total para o treinamento individual não deve exceder 1800 segundos.

4.4.5 Validação externa e análise dos resultados obtidos

A fase de validação externa das redes neurais foi executada utilizando os 200 dados exclusivos escolhidos aleatoriamente de um único experimento (neste trabalho, denominado de experimento 2). Essa etapa desempenhou um papel crucial na identificação do modelo mais apropriado para o processo em análise, além de mitigar potenciais problemas de superajuste. Durante essa simulação, novos conjuntos de dados foram introduzidos como variáveis de entrada, permitindo avaliar a capacidade de

generalização da rede e, consequentemente, determinar qual modelo melhor descrevia o processo de fermentação para a produção de hidromel.

Após a obtenção dos resultados, a seleção da melhor RNA baseou-se na análise do erro médio percentual absoluto (MAPE) das simulações atreladas das variáveis de saída e do coeficiente de correlação de Pearson obtidos na etapa de simulação. Além desses critérios, também foi considerado o número de parâmetros presentes na RNA.

Outro método inovador adotado para analisar os resultados foi o estabelecimento de um limiar de 25%. Esse critério garante que os erros da validação externa sejam homogêneos e distribuídos igualmente entre as variáveis de saída, respeitando o limiar definido. A mesma abordagem foi aplicada durante as fases de treino, validação e teste, assegurando uma distribuição equilibrada dos erros ao longo de todo o processo de treinamento. Esse procedimento visa garantir a homogeneidade no treinamento e na obtenção das variáveis, resultando em um modelo mais estável e consistente.

Assim, a escolha do modelo ideal levou em consideração uma combinação de fatores, incluindo a capacidade de generalização da rede, precisão na predição dos resultados e a complexidade do modelo. Essa abordagem visou garantir que a RNA selecionada não apenas se adaptasse bem aos dados de treinamento, mas também fosse capaz de generalizar efetivamente para novos conjuntos de dados, tornando-se uma ferramenta eficaz na modelagem do processo de produção de hidromel.

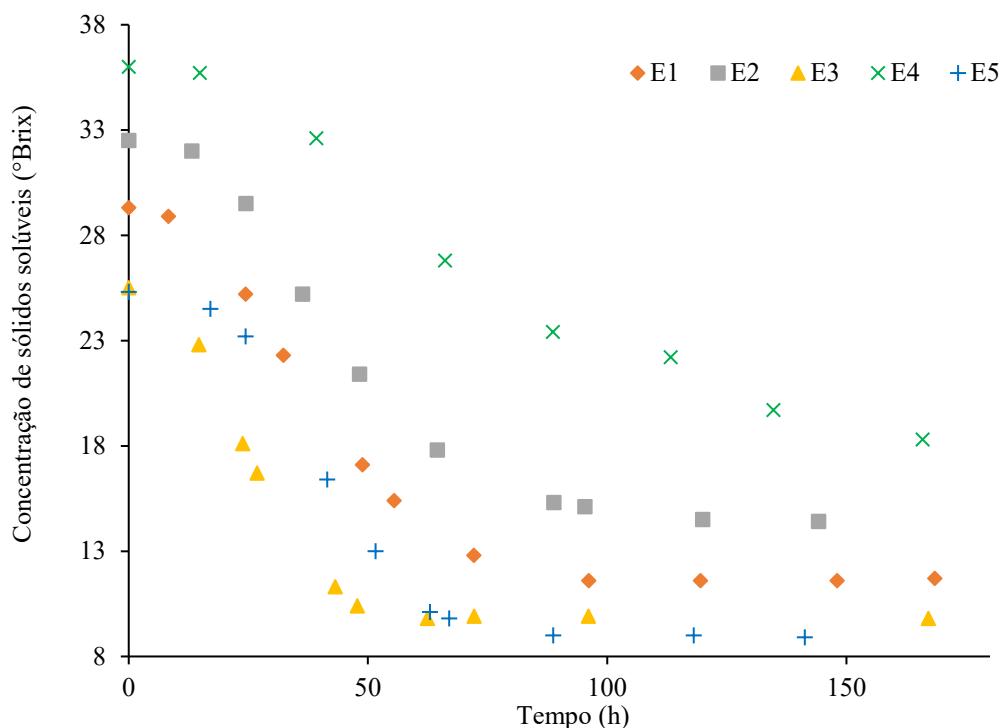
5 RESULTADO E DISCUSSÃO

No presente estudo, diversas estratégias foram meticulosamente avaliadas em relação ao desempenho de redes neurais artificiais. Para cada estratégia considerada, foram exploradas configurações com uma e duas camadas intermediárias. Durante essa avaliação, houve uma variação sistemática dos algoritmos de otimização, da quantidade de neurônios presentes em cada camada e das respectivas funções de ativação utilizadas. As variáveis de entrada selecionadas para este estudo foram o potencial de hidrogênio (pH), Sólidos Solúveis (S.S.), densidade óptica (D.O.) e a Temperatura (T). O processo de treinamento foi realizado e validando, obtendo-se um total de 300 mil RNAs.

5.1 dados experimentais empregados na obtenção da rede

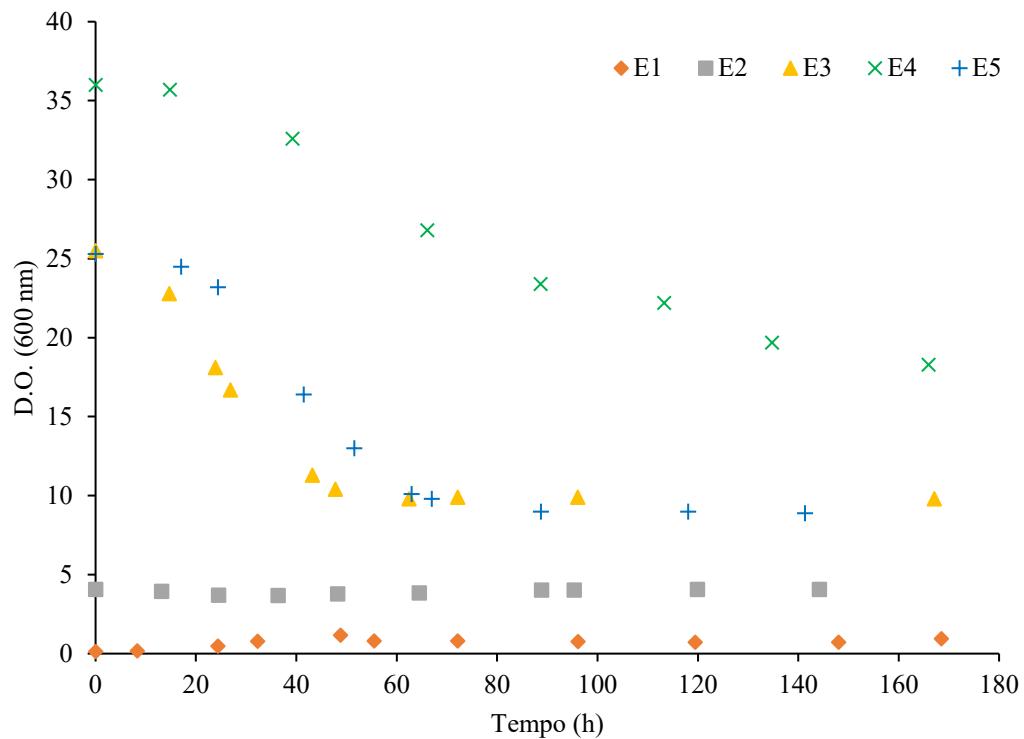
As Figuras 17, 18 e 19 ilustram as variáveis de entrada utilizadas na rede neural em função do tempo, enquanto a Tabela 2 apresenta as respectivas temperaturas e sólidos solúveis a que cada experimento foi submetido.

Figura 17 – Concentrações de produto em função do tempo para todos os experimentos realizados.



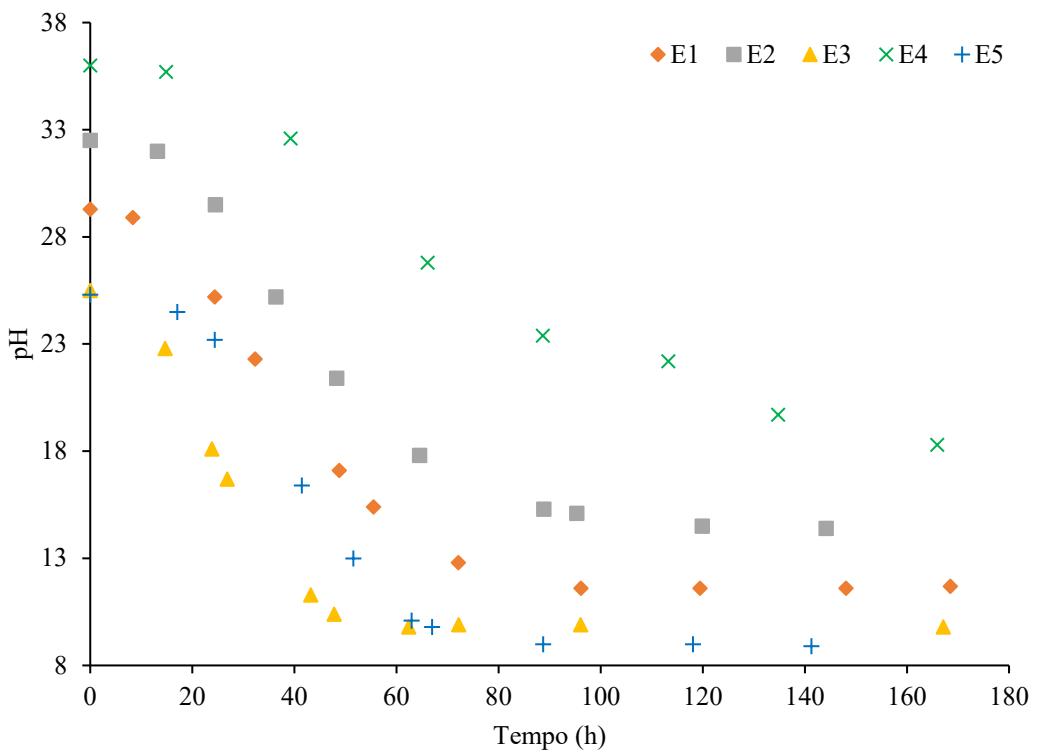
Fonte: Autoria própria.

Figura 18 – Concentrações de produto em função do tempo para todos os experimentos realizados.



Fonte: Autoria própria.

Figura 19 – Concentrações de produto em função do tempo para todos os experimentos realizados.



Fonte: Autoria própria.

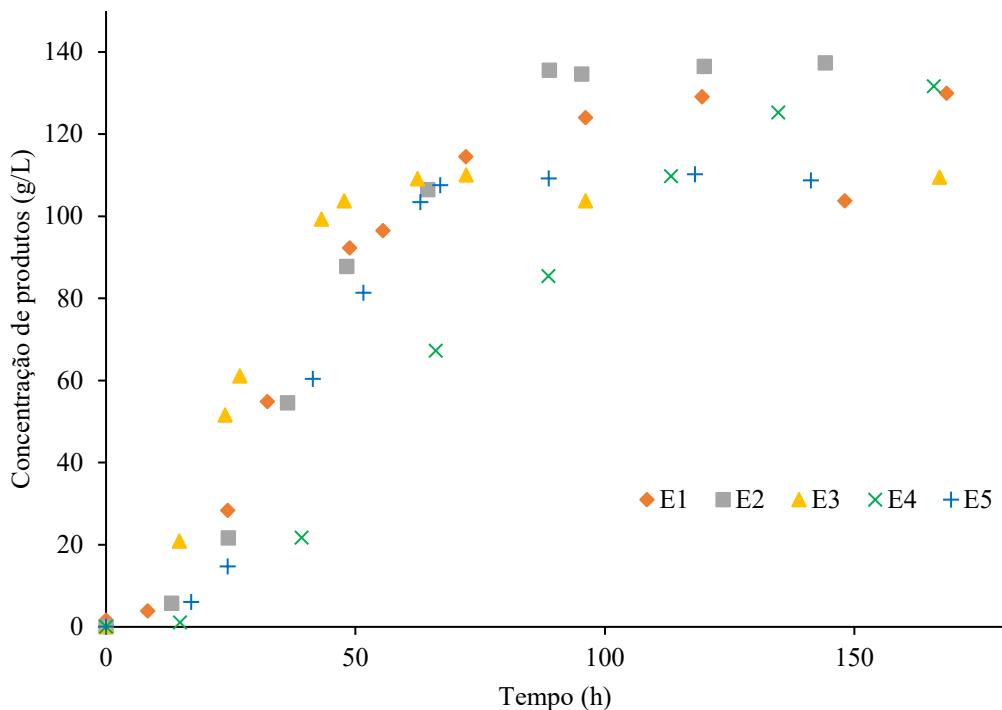
Tabela 2 – Número do experimento e suas respectivas temperaturas.

n° EXP.	Sólidos Solúveis (°Brix)	Temperatura (°C)
1	30	25,0
2	30	25,0
3	25	30,0
4	35	20,0
5	25	20,0

Fonte: Autoria Própria.

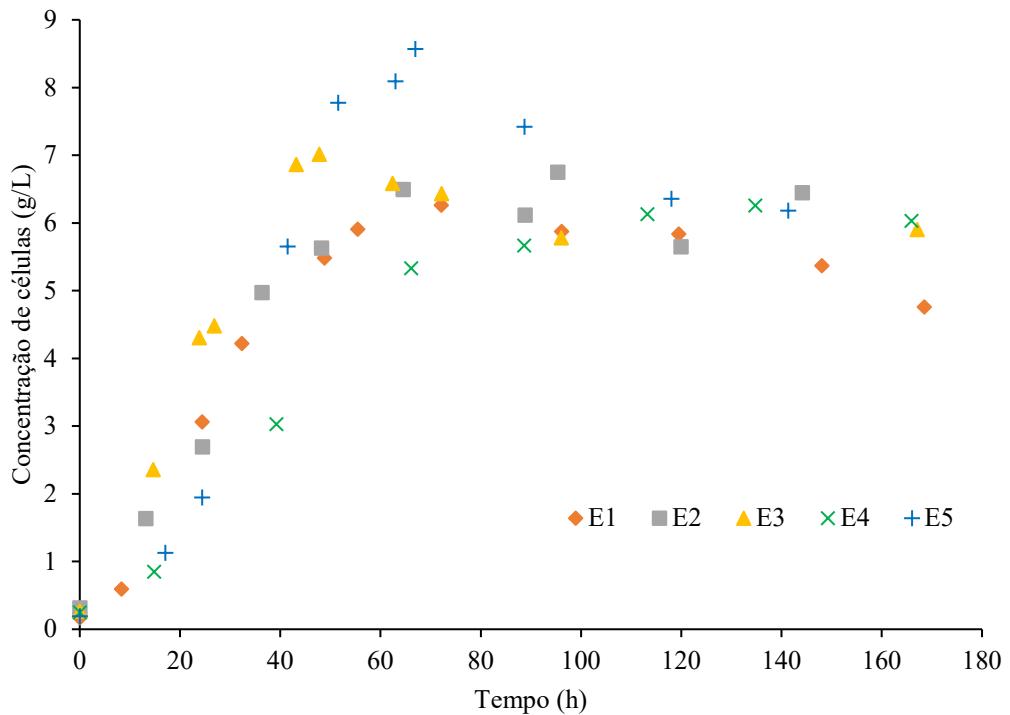
Por outra via, as Figuras 20, 21 e 22 mostram as concentrações das variáveis de saída em função do tempo:

Figura 20 – Concentrações de P em função do tempo para todos os experimentos realizados.



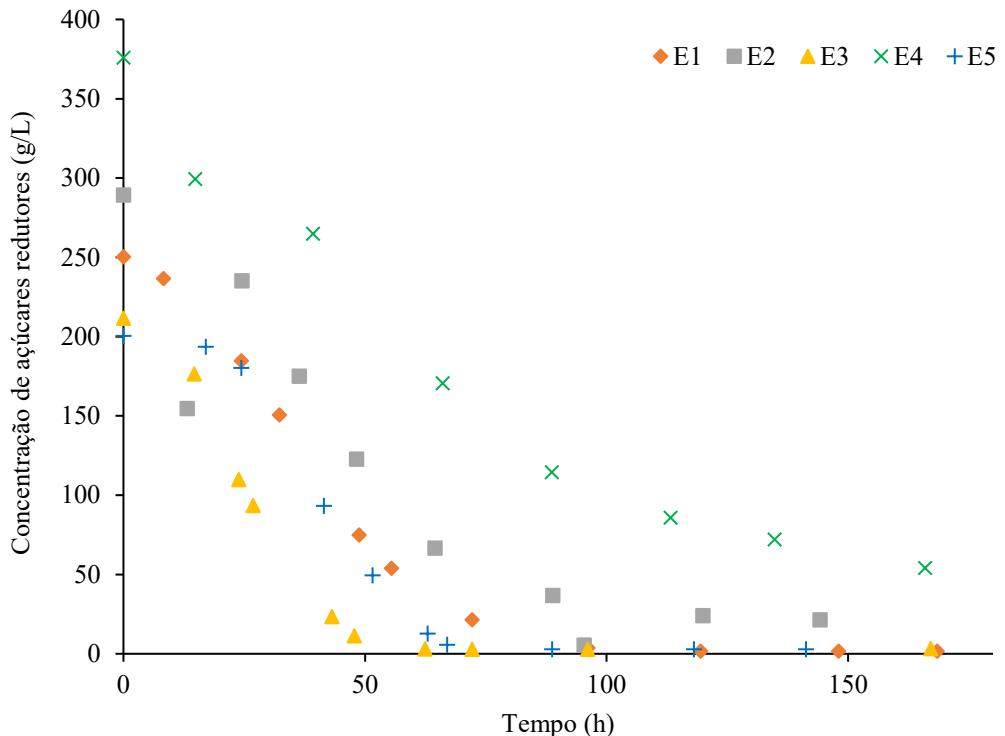
Fonte: Autoria própria.

Figura 21 – Concentrações de X em função do tempo para todos os experimentos realizados.



Fonte: Autoria própria.

Figura 22 – Concentrações de S em função do tempo para todos os experimentos realizados.



Fonte: Autoria própria.

A análise da Figura 17 revela que a concentração inicial de açúcar em cada processo fermentativo está alinhada com os valores de sólidos solúveis empregados. Observa-se que, ao término das fermentações, mesmo com a presença de 8 a 18, os açúcares residuais apresentam-se substancialmente inferiores a 60 g/L, com a exceção do processo E5, que registra 70 g/L. Este último apresenta uma concentração inicial de substrato significativamente alta, resultando em maior inibição por substrato durante a fermentação.

Conforme ilustrado na Figura 20, o perfil de crescimento celular nas fermentações segue um padrão específico. Inicialmente, ocorre um aumento rápido, indicativo da fase exponencial, seguido por uma fase estacionária e de declínio celular. Notadamente, a fermentação do Experimento 6 (E6), conduzida a 20°C com menor concentração de sólidos solúveis (25 °Brix), demonstrou um desenvolvimento maior da levedura, o que corrobora com as indicações do fabricante sobre a preferência da levedura por temperaturas mais baixas e menor concentração de açúcares, condições que minimizam o efeito inibidor sobre seu crescimento. Além disso, a taxa de crescimento celular tende a ser mais elevada em temperaturas reduzidas (Andrade, 2012).

A utilização da rede oferece diversos benefícios. Se a concentração de células estiver muito baixa, o sensor pode orientar ao operador, a mudança de alguma condição operacional, como abaixar a temperatura para favorecer o crescimento microbiano. Ela também auxilia na predição do teor alcoólico do produto, assegurando sua conformidade com a legislação nacional. Uma alta concentração de etanol pode causar inibição, e a rede possibilita o monitoramento e controle dessa concentração durante a fermentação. Além disso, a rede proporciona uma visão atual do processo, orientando futuras decisões e ajustes (Costa, 2021).

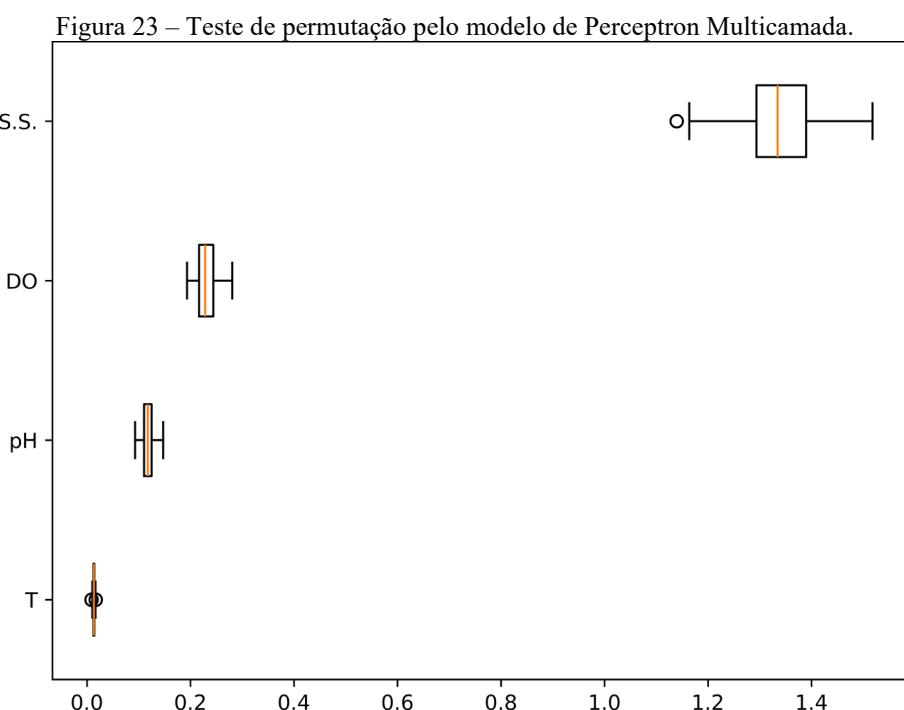
A RNA pode ser adaptada para outras variações de hidromel e bebidas fermentadas, mas requer retreinamento. Espera-se que este trabalho gere impactos sociais, econômicos e ambientais positivos, incentivando a produção da bebida. A ferramenta pode ser útil para pequenos agricultores na otimização e monitoramento da produção artesanal de hidromel, potencialmente aumentando sua renda, valorizando o mel nacional e fortalecendo a apicultura no país. Isso pode também favorecer a preservação das abelhas, promovendo sustentabilidade. A previsão de X, S e P é específica para hidromel com adição de polpa de casca de jabuticaba, mas a rede neural pode ser

adaptada para outras bebidas fermentadas e variações de hidromel, desde que seja retreinada.

5.2 Análise das Importâncias por Permutação no Modelo de Regressão

A avaliação das importâncias das variáveis preditoras é uma etapa fundamental para interpretar modelos de aprendizado de máquina, especialmente em problemas de regressão. Neste estudo, utilizou-se o método de importância por permutação para estimar a relevância de cada variável na previsão dos valores de interesse. Esse método consiste em calcular a diferença na métrica de desempenho do modelo após permutar aleatoriamente os valores de uma variável preditora, mantendo as demais constantes. A magnitude da degradação no desempenho é interpretada como a importância da variável (Breiman, 2001).

Os resultados da análise de permutação, apresentados na Figura 23, indicam que a variável S.S. possui a maior influência nas previsões realizadas pelo modelo, conforme evidenciado pelo seu valor médio de importância substancialmente superior às demais variáveis. Esse resultado sugere que os dados relacionados à concentração de sólidos suspensos (S.S.) fornecem informações críticas para o ajuste do modelo e contribuem significativamente para a explicação da variabilidade nos dados.



Fonte: Autoria Própria.

As variáveis DO e pH, embora apresentem importância menor em comparação a S.S., demonstraram impacto considerável e consistente na capacidade preditiva do modelo, indicando que elas possuem informações complementares que auxiliam na explicação do fenômeno estudado. Por outro lado, a temperatura apresentou relevância reduzida, com valores de importância próximos a zero. Isso pode indicar que a temperatura não está fortemente associada às variáveis de saída no contexto do presente estudo. Resultados similares foram reportados por Goldstein *et al.* (2015), que destacaram a importância da análise de permutação para identificar variáveis menos informativas, auxiliando na simplificação e otimização de modelos.

A escolha do método de permutação mostrou-se particularmente relevante por ser agnóstica ao modelo e robusta em cenários com múltiplas variáveis correlacionadas. No entanto, deve-se reconhecer que, em casos de alta colinearidade entre preditores, a importância atribuída a uma variável pode ser reduzida pela redundância com outras variáveis do conjunto (Altemeyer *et al.*, 2021). Isso sugere que investigações futuras podem incluir abordagens complementares, como o *SHAP* (SHapley Additive exPlanations), para capturar interações mais complexas e fornecer uma interpretação mais granular dos impactos das variáveis.

Portanto, os resultados da análise de permutação fornecem uma visão clara sobre a hierarquia de relevância das variáveis no modelo, destacando S.S. como a variável-chave para as previsões realizadas, enquanto identificam variáveis menos informativas que poderiam ser consideradas para exclusão em versões futuras do modelo, dependendo do objetivo da análise.

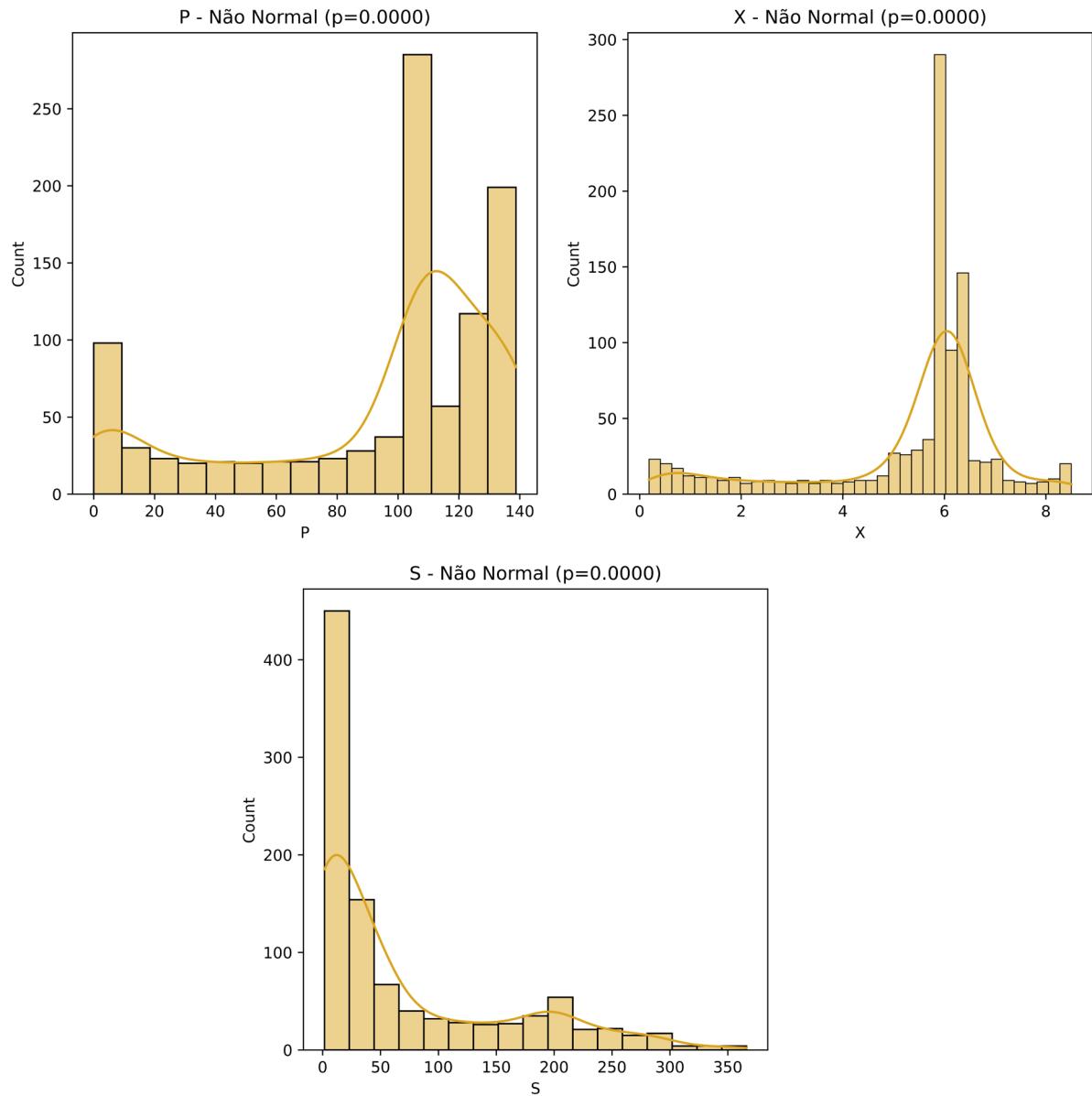
5.3 Teste de Normalidade dos dados

O teste de normalidade de Shapiro-Wilk foi aplicado para verificar a adequação das distribuições das variáveis P, X e S a uma distribuição normal. Os resultados indicaram que todas as variáveis apresentaram p-valores extremamente baixos ($p < 0,05$), confirmando que não seguem uma distribuição normal.

A Figura 24 ilustra a normalidade das variáveis de saída por meio do teste de Shapiro-Wilk, sendo este, um teste robusto para amostras pequenas e médias, sensível a desvios de normalidade, como assimetria e curtose. No caso das variáveis analisadas, a não normalidade pode ser atribuída às características do processo de fermentação de hidromel, como variações nas concentrações ao longo do tempo, oscilações no

metabolismo celular e a presença de diferentes fases do processo (crescimento celular, produção de etanol e consumo de substrato).

Figura 24 – Distribuição normal das variáveis de saída.



Fonte: Autoria Própria.

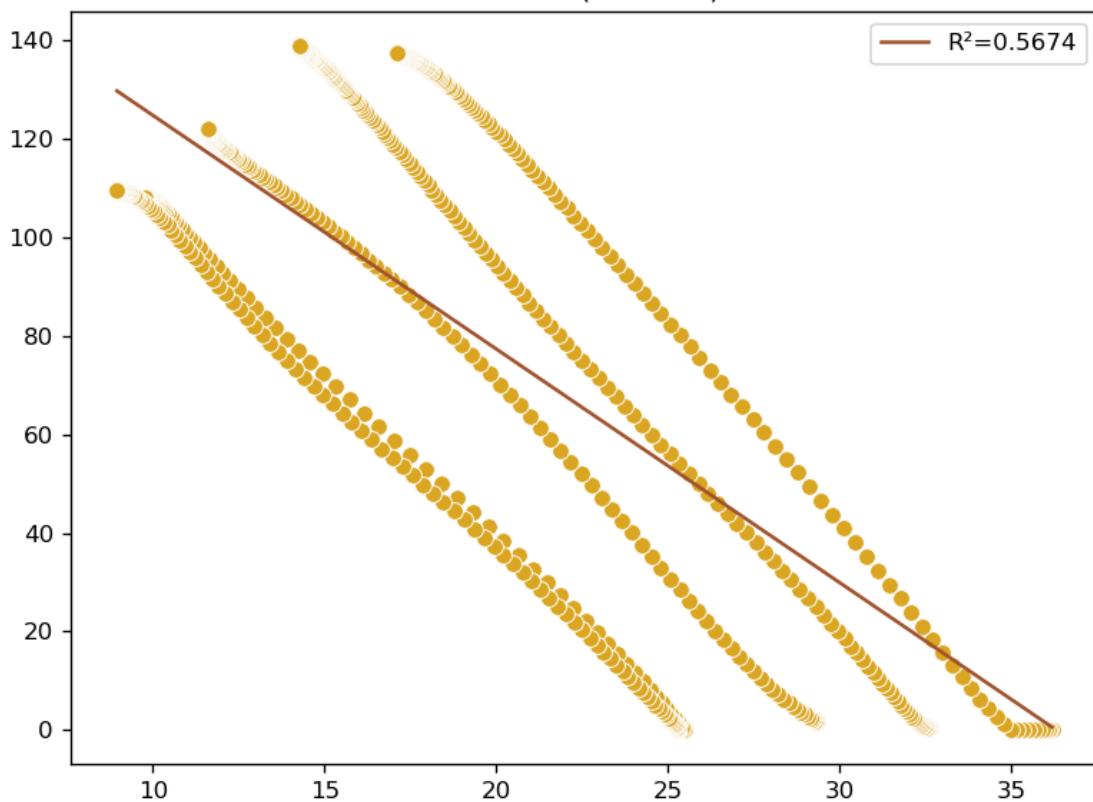
Esses resultados reforçam a importância de utilizar métodos estatísticos e computacionais que não assumam normalidade, especialmente na modelagem por redes neurais, garantindo maior precisão nas previsões e análises do processo fermentativo. Sendo também, válido salientar que, o mesmo teste foi empregado para todas as variáveis e, de forma unanime, todas apresentaram não-normalidade.

5.4 Teste de Linearidade

A análise das relações entre as variáveis de saída revela diferentes graus de linearidade, evidenciados pelos coeficientes de determinação R^2 obtidos nas regressões lineares. Essas relações são cruciais para entender a dinâmica do processo fermentativo de hidromel e avaliar a adequação dos modelos utilizados. Dado o teste de permutação, que destacou a maior representatividade de S.S. no modelo, em consonância com os resultados do teste de normalidade, apenas os detalhes das variáveis de saída foram aprofundados. Isso se justifica pelo fato de que a maioria das relações binárias apresentou baixa linearidade, limitando sua contribuição para uma análise mais robusta no contexto do modelo.

A relação entre a soma dos quadrados S.S. e o etanol P é demonstrada pela Figura 25, ela apresenta um coeficiente de determinação intermediário $R^2 = 0,57$, indicando uma correlação moderada. A dispersão dos pontos ao redor da linha de regressão sugere que essa relação não é estritamente linear, sendo possivelmente influenciada por outros fatores inerentes ao processo, como limitações na cinética enzimática ou interações com outras variáveis.

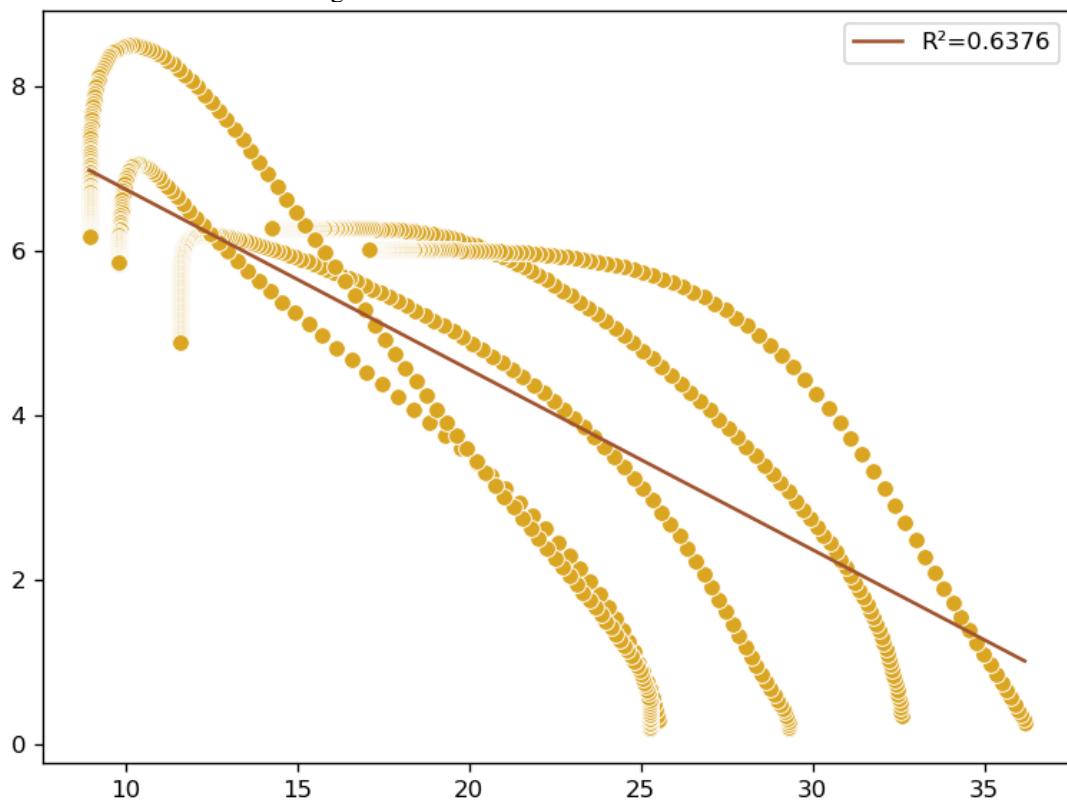
Figura 25 – Linearidade de S.S. versus P.



Fonte: Autoria Própria.

No caso da relação entre S.S. e o crescimento celular X (Figura 26), o coeficiente de determinação é ligeiramente maior $R^2 = 0,64$, indicando uma correlação moderada a forte. No entanto, a presença de curvaturas nos dados sugere que interações não lineares podem desempenhar um papel importante, especialmente considerando as diferentes fases do metabolismo celular ao longo da fermentação.

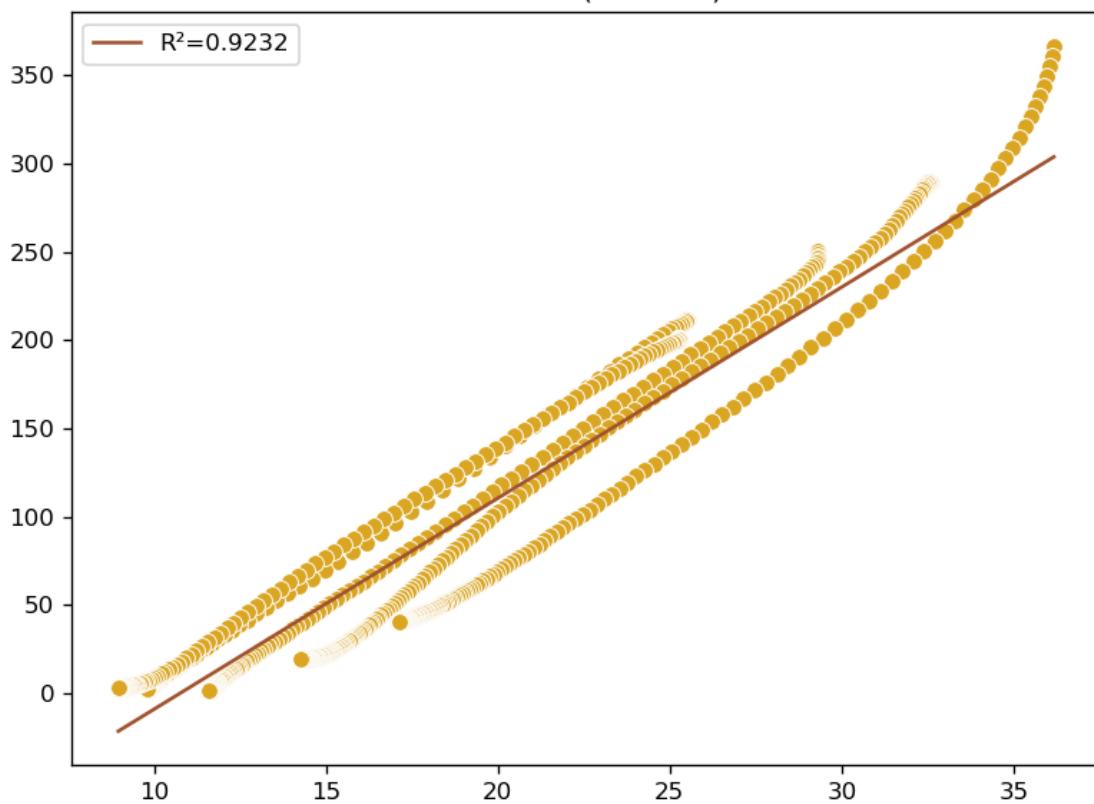
Figura 26 – Linearidade de S.S. versus X.



Fonte: Autoria Própria.

Por outro lado, a relação entre S.S. e o substrato S (Figura 27) apresenta alta linearidade, com um coeficiente de determinação de $R^2 = 0,92$. Esse resultado reflete uma forte correlação direta, coerente com o comportamento esperado, dado que o consumo de substrato é diretamente proporcional ao progresso do processo fermentativo. Essa linearidade torna S uma variável mais previsível no modelo.

Figura 27 – Linearidade de S.S. versus S.



Fonte: Autoria Própria.

Essas análises destacam que, enquanto algumas variáveis exibem alta linearidade S, outras P e X apresentam comportamentos mais complexos. Isso reforça a necessidade de abordagens avançadas para capturar de forma mais completa a dinâmica do sistema, especialmente em contextos em que as interações não lineares são relevantes para a modelagem e otimização da fermentação.

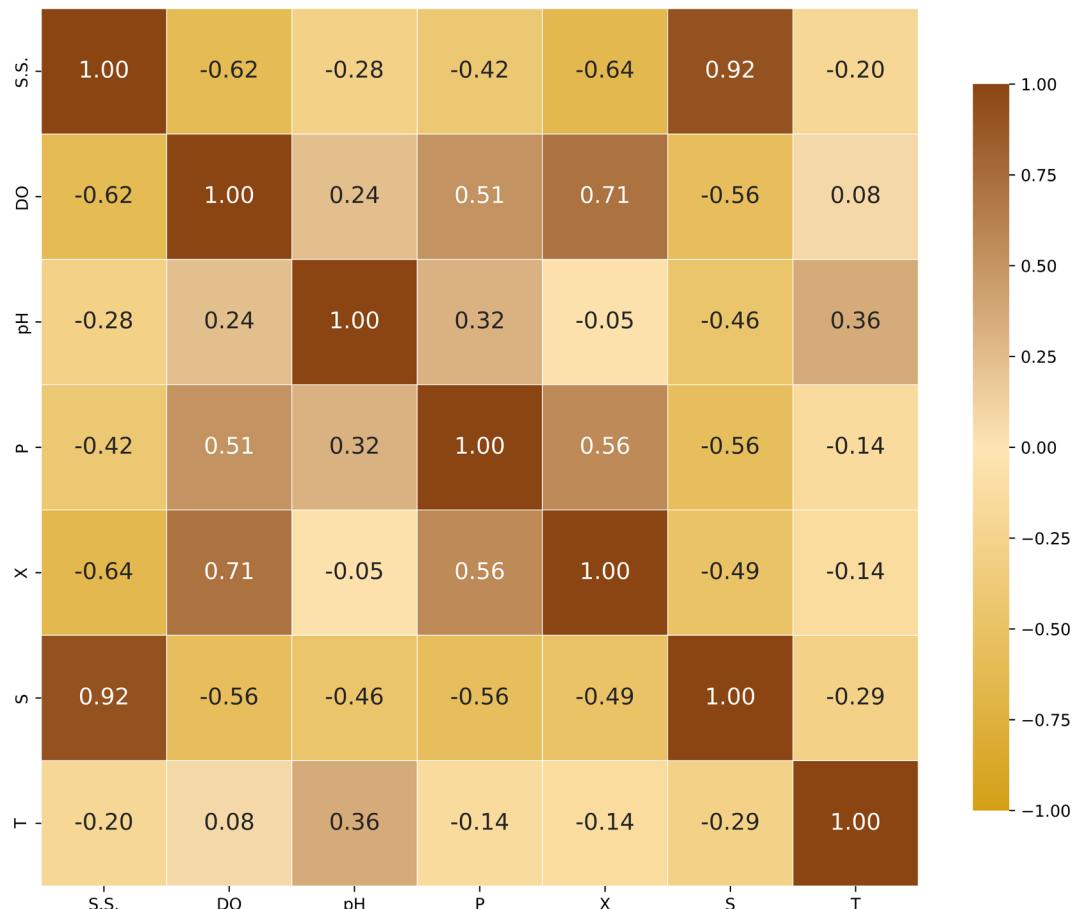
5.5 Matriz de correlação

Para investigar as interrelações entre as variáveis monitoradas durante o processo fermentativo, foi aplicada a matriz de correlação de Spearman. Essa técnica foi selecionada por sua capacidade de detectar relações monotônicas, independentemente da linearidade intrínseca dos dados, uma característica essencial em sistemas biológicos complexos (Frost, 2024). A decisão da matriz de Spearman é não a de Pearson, por exemplo, advém dos resultados dos testes de linearidade e normalidade dos dados, prevendo não linearidade e não normalidade. O coeficiente de Spearman varia de -1 (correlação negativa perfeita, inversamente proporcionais) a +1 (correlação positiva perfeita, aumento ou diminuição proporcional), fornecendo uma análise robusta das

interações entre variáveis experimentais em processos fermentativos, onde há significativas não-linearidades e variações ao longo do tempo. As variáveis avaliadas incluíram tanto parâmetros de entrada quanto de saída.

Ao analisar a Figura 28, foram evidenciadas relações críticas e fenomenologicamente explicáveis entre os parâmetros do processo de fermentação. A concentração de sólidos solúveis (S.S.) apresentou correlação positiva com os açúcares totais (S), com um coeficiente de $\rho = 0.92$, evidenciando a relação direta entre a quantidade inicial de substrato (açúcares) disponível e sua medição em °Brix. Durante a fermentação, a diminuição do valor de S.S. reflete o consumo de açúcares pelas células, resultando na produção de etanol e biomassa. Por outro lado, S.S. mostrou correlação negativa com a concentração de células (X) e etanol (P), com valores de -0.82 e -0.42, respectivamente. Esse comportamento indica o consumo progressivo de substrato com a concomitante proliferação celular e produção de etanol, fenômeno esperado na fase de crescimento exponencial da fermentação.

Figura 28 – Matriz de Correlação de Spearman.



Fonte: Autoria Própria.

A densidade óptica (DO) apresentou correlação positiva moderada com a concentração de células (X), com um valor de $\rho = 0.6457$, reforçando seu uso como um indicador indireto de biomassa. Isso ocorre porque a densidade óptica reflete a turbidez do meio, que está correlacionada com a presença e o aumento de células viáveis. Em contrapartida, a correlação negativa de -0.5728 com os açúcares (S) sugere que a redução da turbidez acompanha o consumo de substrato, associado a uma possível saturação celular ou alterações morfológicas na fase estacionária. O pH do meio demonstrou correlação moderada positiva com a produção de etanol ($\rho = 0.3155$) e negativa com os açúcares ($\rho = -0.4627$). Esse comportamento é explicado pela acidificação do meio durante a fermentação alcoólica devido à produção de subprodutos ácidos, como ácidos orgânicos leves, os quais reduzem gradualmente o pH. O controle inadequado do pH pode interferir no desempenho das enzimas responsáveis pela metabolização dos açúcares e, consequentemente, na eficiência da produção de etanol.

A temperatura operacional (T) apresentou correlações fracas ou não significativas com a maioria das variáveis, com coeficientes variando entre -0.2856 e 0.3610 . Esse resultado indica uma influência limitada da temperatura no experimento, sugerindo que o processo fermentativo foi conduzido em condições térmicas controladas, próximas ao ideal fisiológico dos microrganismos utilizados. Essa estabilidade térmica garante a manutenção da atividade enzimática e a eficiência metabólica, sem inibição significativa devida a variações no processo.

Os açúcares totais (S) apresentaram forte correlação negativa com a concentração de células ($\rho = -0.6830$) e com a produção de etanol. Tal relação reflete o consumo progressivo do substrato como fonte primária de energia, favorecendo simultaneamente a produção de biomassa e metabólitos. Essa tendência é característica dos processos fermentativos em que a conversão dos açúcares ocorre de forma direta e proporcional ao crescimento celular e à formação do produto de interesse.

Essas correlações fornecem insights claros sobre as relações metabólicas subjacentes. A redução dos sólidos solúveis e dos açúcares totais ao longo do tempo reflete a conversão metabólica típica das fases *lag*, exponencial e estacionária da fermentação. A densidade óptica, como indicador indireto da biomassa, está alinhada com a cinética de crescimento microbiano, sugerindo que o aumento da concentração celular ocorre paralelamente à redução do substrato. A ausência de correlação significativa entre a temperatura operacional e os demais parâmetros evidencia que o

controle térmico mitigou possíveis efeitos de inibição ou redução da atividade enzimática.

5.6 Análise das redes obtidas e classificação geral

O objetivo primordial dessa abordagem era desenvolver uma única Rede Neural Artificial (RNA) capaz de efetuar previsões simultâneas para três variáveis críticas: concentração de produto, células e substrato. Para construir essa RNA abrangente, empregou-se dados experimentais que englobavam informações vitais de Potencial de Hidrogênio (pH), Sólidos Solúveis ($^{\circ}$ Brix), Densidade Óptica (D.O.) e Temperatura (T). Os resultados e conclusivos são apresentados na Tabela 3, que compila os dez conjuntos de parâmetros mais promissores para cada um dos nove algoritmos de otimização utilizados.

Tabela 3 – Panorama geral do desempenho para as 10 melhores RNAs.

ID	Estrutura [0]	Algoritmo de Otimização [1]	Funções de Ativação [3]	Treino F_{obj} [4]	MAPE	MAPE	MAPE	$\sum \text{MAPE}$	NP	ρ
					(%)	(%)	(%)	Simu. (%) [5]	[6]	[7]
1	4-40-25-3	<i>trainbr</i>	<i>poslin-softmax-tansig</i>	1.32E-01	13.05	8.75	6.23	9.34	1303	0.92
2	4-40-25-3	<i>trainbr</i>	<i>poslin-softmax-purelin</i>	5.36E-02	11.14	11.04	8.55	10.24	1303	0.92
3	4-30-20-3	<i>trainlm</i>	<i>softmax-poslin-purelin</i>	1.31E-01	9.91	12.34	11.83	11.36	833	0.92
4	4-25-20-3	<i>trainrp</i>	<i>softmax-softmax-purelin</i>	2.86E-01	14.57	6.35	13.71	11.55	708	1.00
5	4-40-35-3	<i>traincfg</i>	<i>tansig-softmax-purelin</i>	3.28E-01	10.73	17.49	12.99	13.74	1743	0.94
6	4-30-35-3	<i>trainbfg</i>	<i>softmax-softmax-purelin</i>	4.00E-02	21.73	20.35	15.65	19.24	1343	1.00
7	4-25-20-3	<i>traincgp</i>	<i>tansig-softmax-purelin</i>	1.17E+00	15.38	36.37	12.66	21.47	708	0.93
8	4-35-25-3	<i>trainoss</i>	<i>tansig-logsig-purelin</i>	5.94E-01	29.66	33.78	9.91	24.45	1153	0.79
9	4-35-30-3	<i>traingdx</i>	<i>poslin-softmax-purelin</i>	3.35E+01	35.75	25.69	33.24	31.56	1348	0.92
10	4-35-25-3	<i>traingd</i>	<i>softmax-softmax-purelin</i>	5.79E+03	182.69	101.50	228.01	170.73	1153	1.00

Fonte: Autoria própria.

Legenda:

[0] Código de identificação da RNA.

[1] Arquitetura da RNA: neurônios na camada de entrada - na camada intermediária 1 - na camada intermediária 2 - na camada de saída.

[2] Método de otimização empregado durante o treinamento da RNA.

[3] Função de ativação na camada intermediária 1 - na camada intermediária 2 - na camada de saída.

[4] Critério de desempenho durante a fase de treinamento.

[5] Somatório dos erros médios percentuais absolutos para X, S e P.

[6] Número de parâmetros para a rede - ilustra os pesos, por exemplo, na ID 1: $(4*40 + 40*25 + 25*3)$ e depois a soma dos vieses $(40+25+3) = 1303$.

[7] Coeficiente médio de correlação de Pearson.

A Tabela 3 proporciona uma visão abrangente dos resultados, incluindo os algoritmos de otimização empregados, as funções de objetivo utilizadas nas fases de treinamento, validação e teste, bem como o Erro Médio Percentual Absoluto (MAPE) para cada uma das variáveis. As RNAs presentes na Tabela 3 estão classificadas em ordem decrescente de acordo com a soma total do MAPE das variáveis P, X e S.

Os resultados apresentados referem-se a testes realizados com Redes Neurais Artificiais (RNAs) no MATLAB, utilizando diferentes algoritmos de otimização, funções de ativação e arquiteturas de rede. O objetivo principal foi identificar a melhor configuração para modelar variáveis relacionadas à produção de etanol (P), células (X) e açúcares redutores (S). A métrica de avaliação utilizada foi o Erro Percentual Médio Absoluto (MAPE), calculado para cada variável de saída. Também foram analisados o coeficiente de correlação médio de Pearson (ρ), o número de parâmetros da rede (NP) e o critério objetivo (F_{obj}) e a relação de erro para cada variável de saída e função objetivo obedecendo o limiar de 25% de proximidade entre os erros para cada variável relacionados à fase de treinamento. A Tabela 3 detalha os valores correspondentes a estas métricas.

Nove algoritmos de otimização foram avaliados. O algoritmo *trainbr*, baseado em regularização Bayesiana, apresentou o melhor desempenho geral. Na configuração 4-40-25-3, este algoritmo resultou em um $\Sigma MAPE$ de 10.24% e coeficiente de correlação ρ de 0.92. Esse desempenho foi consistente entre as variáveis analisadas, com destaque para o erro de açúcares redutores, que foi de apenas 8.55%. Este resultado reflete a robustez do algoritmo em cenários de alta complexidade e sua capacidade de minimizar os erros absolutos em diferentes variáveis de saída.

O algoritmo *trainlm*, baseado em Levenberg-Marquardt, apresentou bons resultados, com $\Sigma MAPE$ de 11.36% e menor número de parâmetros (NP de 833), sendo eficiente para arquiteturas menores. No entanto, esse algoritmo tem limitações em redes maiores devido ao alto custo computacional. Já o algoritmo *trainrp*, baseado em propagação resiliente, teve um desempenho intermediário, com $\Sigma MAPE$ de 11.55%, destacando-se por sua simplicidade e baixa dependência de hiperparâmetros, sendo mais estável em redes menos complexas.

Os algoritmos baseados em gradiente conjugado, *traincgv* e *traincgb*, apresentaram desempenho mediano, com $\Sigma MAPE$ de 13.74% e 21.47%, respectivamente. Embora eficientes no consumo de memória, esses algoritmos tiveram dificuldades em alcançar baixos erros, especialmente para variáveis de maior variabilidade. O algoritmo *trainbfg*, baseado em Broyden-Fletcher-Goldfarb-Shanno, teve um desempenho inferior, com $\Sigma MAPE$ de 19.24%, indicando que sua aplicação pode não ser ideal para este tipo de problema.

O algoritmo *trainoss*, baseado no método One Step Secant, e o algoritmo *traingdx*, que utiliza descida de gradiente com taxa de aprendizado adaptativa, tiveram desempenhos insatisfatórios, com $\Sigma MAPE$ de 24.45% e 31.56%, respectivamente. Ambos demonstraram incompatibilidade com a arquitetura e os dados analisados. O algoritmo *traingd*, baseado no gradiente descendente de retropropagação (Gradient Descent Backpropagation), apresentou os piores resultados, com $\Sigma MAPE$ de 170.73%, refletindo baixa capacidade de generalização, por ser simples para um problema altamente complexo e, sendo o único entre todos os algoritmos que não conseguiu atender aos critérios de possuir erros entre as variáveis e funções objetivos de no máximo 25%.

O coeficiente de correlação médio variou entre 0.92 e 1.00 nos modelos mais eficientes, indicando alta capacidade de predição em todas as variáveis. Essa métrica reforça a confiabilidade dos modelos treinados e sua aplicabilidade prática em processos complexos. A análise também destacou a importância da escolha de funções de ativação, sendo que combinações não-lineares, como *tansig* e *softmax*, foram mais eficazes para representar a variabilidade dos dados.

Os resultados obtidos evidenciam que o desempenho das RNAs está diretamente relacionado à seleção adequada de algoritmos de otimização, funções de ativação e arquitetura de rede. O algoritmo *trainbr* demonstrou superioridade ao minimizar os erros e maximizar a capacidade de generalização, sendo ideal para redes mais complexas, além de ser o mais demorado em questão de tempo de processamento e obtenção das RNAs, levando em média o quíntuplo do tempo necessário em relação aos outros algoritmos. Por outro lado, algoritmos como *traingdx* e *traingd* apresentaram limitações significativas, evidenciando a necessidade de uma abordagem criteriosa na escolha do método de treinamento.

5.7 Seleção da melhor rede com base na estratégia abordada

Após a identificação do melhor algoritmo, a configuração final foi retreinada utilizando *trainbr* com a arquitetura 4-40-25-3. Essa configuração apresentou o menor $\Sigma MAPE$ (de 9.34%) e alto coeficiente de correlação médio ($\rho=0.92$), indicando excelente ajuste aos dados. A arquitetura inclui 40 e 25 neurônios nas camadas intermediárias, totalizando 1303 parâmetros, o que reflete uma complexidade adequada para capturar padrões dos dados sem superajuste altamente problemáticos. A combinação de funções de ativação *poslin-softmax-tansig* demonstrou ser eficiente para modelar relações não-lineares nas camadas intermediárias e de saída.

A função *poslin*, utilizada na camada inicial, é amplamente conhecida por sua robustez ao evitar problemas de saturação, especialmente em cenários onde é necessário preservar informações relevantes de valores positivos (Inoe *et al.*, 2022). A função *softmax*, na segunda camada, desempenha um papel crucial na normalização probabilística das ativações, promovendo uma distribuição suave dos dados que facilita a aprendizagem hierárquica. Por fim, a função *tansig*, na camada de saída, é especialmente eficiente para modelar relações não lineares complexas, proporcionando um ajuste eficaz para dados contínuos ou altamente dinâmicos.

A determinação do número ideal de parâmetros (*NP*) em Redes Neurais Artificiais (RNAs) é um tópico amplamente debatido na literatura científica. O consenso é orientado para a configuração de uma rede com o mínimo de parâmetros, sem comprometer significativamente a precisão, conforme indicado pelo $\Sigma MAPE$. No contexto de monitoramento e controle de processos industriais, a otimização do número de parâmetros é de suma importância. Uma RNA com menos parâmetros não apenas garante uma eficiência computacional superior, mas também assegura uma resposta mais ágil, essencial para aplicações em tempo real, mesmo que neste caso não tenha sido possível, uma vez que o processo possui sua intrínseca complexidade.

A combinação das funções de ativação *poslin*, *softmax* e *tansig* nas respectivas camadas da rede resultante, foi escolhida mediante a todos os critérios de classificação e serve para maximizar o desempenho em problemas de modelagem complexa.

A função *poslin* (positive linear transfer function), é usada na primeira camada com 40 neurônios. Ela funciona como uma variação do ReLU, ativando apenas valores positivos e definindo os negativos como zero. Essa propriedade reduz problemas de saturação de gradientes, comuns em funções como sigmoid (*logsig*) ou tangente hiperbólica (*tansig*), e permite à rede modelar relações não lineares com eficiência. Com 40 neurônios, a camada inicial foi projetada para capturar um alto número de características básicas dos dados, fornecendo uma base ampla e robusta para o aprendizado das camadas subsequentes.

Na segunda camada, a função *softmax* é utilizada com 25 neurônios. Essa função transforma os valores de entrada em probabilidades normalizadas, somando 1, o que é útil para reduzir e organizar a dimensionalidade dos dados de forma probabilística. Essa camada age como um filtro que refina as representações aprendidas na camada anterior, reduzindo gradualmente a complexidade e preservando informações importantes. A escolha de 25 neurônios permite um equilíbrio entre a manutenção de informações relevantes e a redução de

dimensionalidade, evitando perdas significativas de dados e prevenindo gargalos no fluxo de informação.

Outrossim, é relevante destacar a predominância da função de ativação *softmax* nas camadas intermediárias e de saída dos 10 melhores resultados pode ser explicada por sua capacidade de normalizar as ativações em valores probabilísticos, garantindo estabilidade durante o treinamento e eficiência na propagação do gradiente. Essa função suaviza as ativações, evitando saturação e explosões, o que contribui para um aprendizado hierárquico mais organizado, permitindo que a rede priorize informações mais relevantes em problemas multivariáveis. Além disso, a *softmax* é particularmente eficiente em contextos em que as variáveis possuem relações interdependentes ou precisam ser tratadas de forma relativa, como em sistemas complexos que demandam interpretações probabilísticas. Sua capacidade de generalizar bem, reduzindo o risco de sobreajuste, e de promover ajustes eficientes explica sua recorrência nas melhores configurações, evidenciando sua versatilidade em redes neurais aplicadas a problemas não lineares (Nguyen *et al.*, 2022).

A eficiência dos algoritmos de otimização em redes neurais é um tópico de pesquisa contínuo. A seleção do algoritmo de otimização mais adequado para um modelo de Rede Neural é uma das etapas mais críticas no campo do Aprendizado Profundo e em todas as classes de Redes Neurais. A escolha do algoritmo pode influenciar significativamente a precisão, eficiência e desempenho da rede. Em um estudo realizado por Fatima (2020), foi feita uma análise comparativa de vários algoritmos de otimização populares, incluindo *trainbr*, *trainlm*, *trainrp*, *trainscg* e *traincgb*, em diferentes conjuntos de dados. O estudo concluiu que a escolha do otimizador pode ter um impacto significativo no desempenho da rede neural profunda (Fatima, 2020).

Além disso, o *trainbr* é especialmente eficaz em cenários onde os dados são limitados ou apresentam ruídos, pois consegue priorizar padrões gerais ao invés de ajustar variações específicas. Sua habilidade em lidar com funções de ativação como *softmax* e *tansig* e em arquiteturas mais complexas explica seu desempenho superior nos 10 melhores resultados, destacando-se como uma solução robusta e confiável para redes neurais que demandam alta precisão e generalização.

A camada de saída utiliza a função tangente hiperbólica, que mapeia os valores de entrada para o intervalo [-1, 1]. Essa função é ideal para problemas de regressão ou previsão contínua, onde é necessário modelar relações não lineares mais suaves. A *tansig* apresenta uma simetria em torno de zero, o que ajuda a estabilizar o treinamento e a equilibrar gradientes positivos e negativos. Essa característica a torna eficaz para ajustar saídas

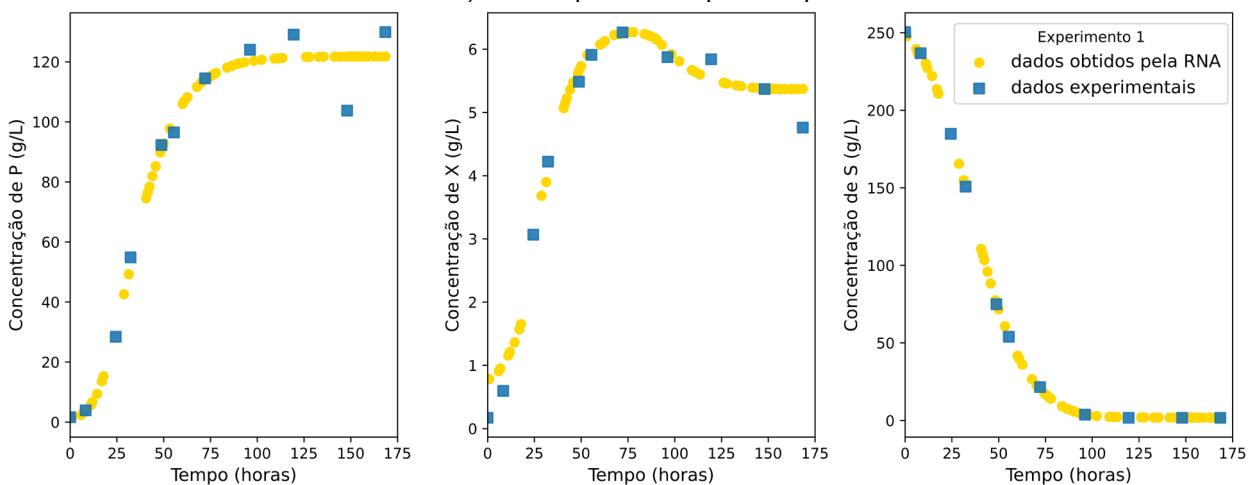
contínuas com precisão, especialmente quando as variáveis foram normalizadas previamente no mesmo intervalo.

Essa abordagem é particularmente útil em problemas onde as variáveis apresentam alta variabilidade e relações não lineares complexas, garantindo um equilíbrio entre capacidade de aprendizado e generalização.

Essa configuração demonstra ser eficaz na modelagem de sistemas complexos, combinando flexibilidade e eficiência para lidar com diferentes características dos dados. A escolha cuidadosa do número de neurônios em cada camada e das funções de ativação reflete um compromisso entre complexidade e desempenho, prevenindo sobreajustes e garantindo resultados robustos.

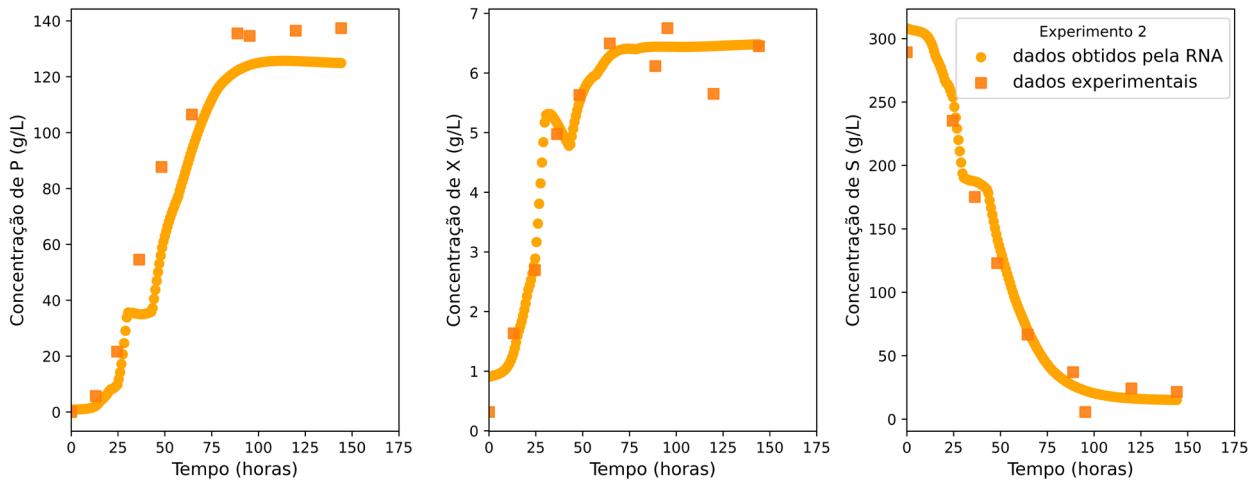
A Figura 29, ilustrando o experimento 1 mostra que a rede neural apresentou alta precisão na simulação das variáveis, com os dados simulados alinhados de forma consistente aos dados experimentais, especialmente para as concentrações de P (g/L) e S (g/L). Pequenos desvios foram observados em concentrações finais de X (g/L), indicando áreas para possíveis melhorias na modelagem, assim como nos experimentos 3, 4 e 5 (Figuras 31, 32 e 33), apresentando desvios e problemas na representatividade na predição das células (X). No geral, os resultados destacam a capacidade da RNA de capturar padrões complexos e modelar os dados com desempenho comparável ao ajuste tradicional, com potencial para otimização adicional. A Figura 30 traz também uma visão geral do Experimento 2.

Figura 29 – Concentrações das variáveis de saída em função do tempo: comparativo do ajuste dos dados simulados em relação aos experimentais para o experimento 1.



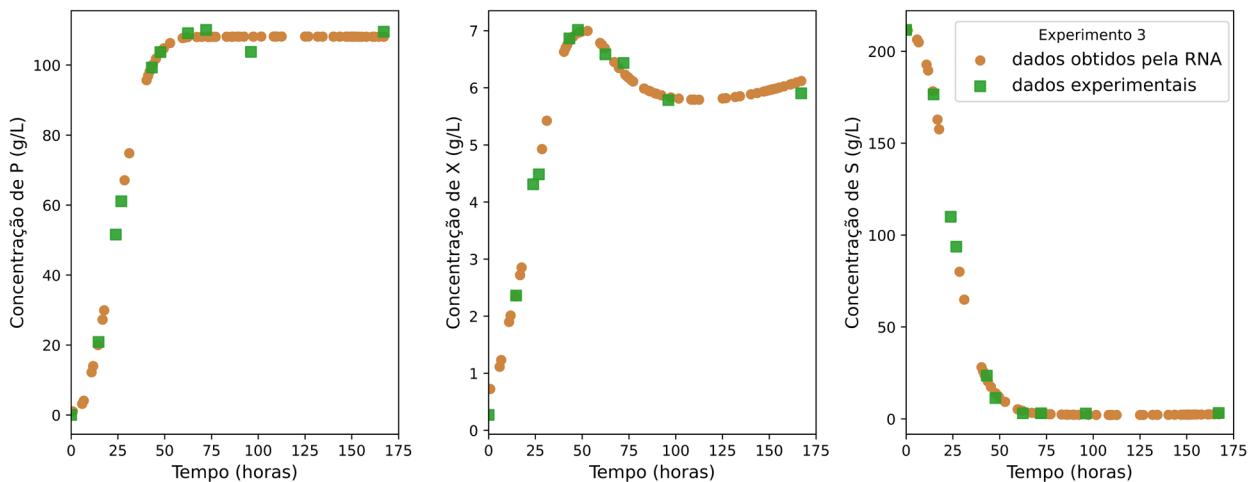
Fonte: Autoria própria.

Figura 30 – Concentrações das variáveis de saída em função do tempo: comparativo do ajuste dos dados simulados em relação aos experimentais para o experimento 2.



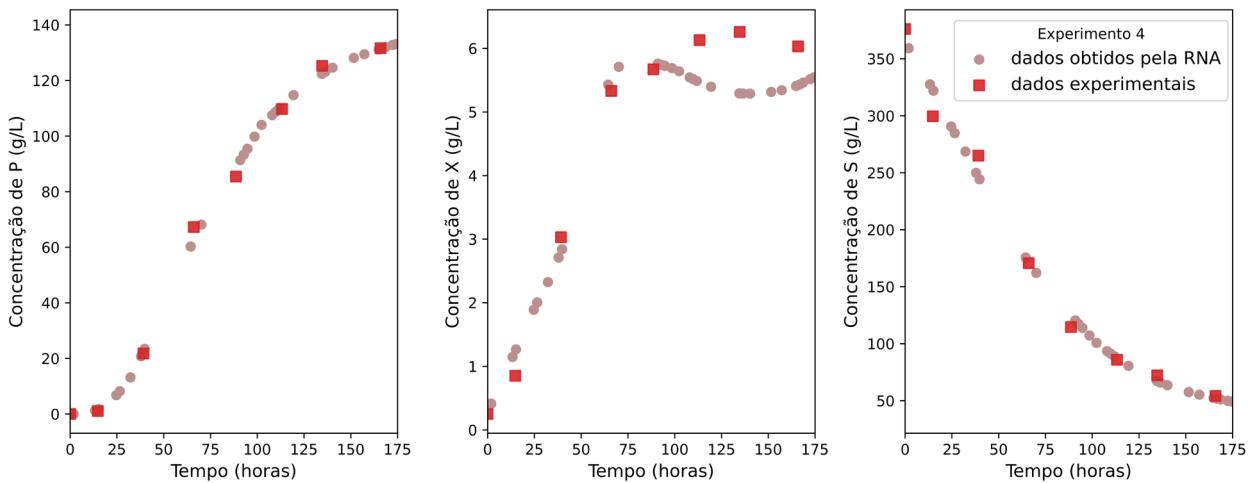
Fonte: Autoria própria.

Figura 31 – Concentrações das variáveis de saída em função do tempo: comparativo do ajuste dos dados simulados em relação aos experimentais para o experimento 3.



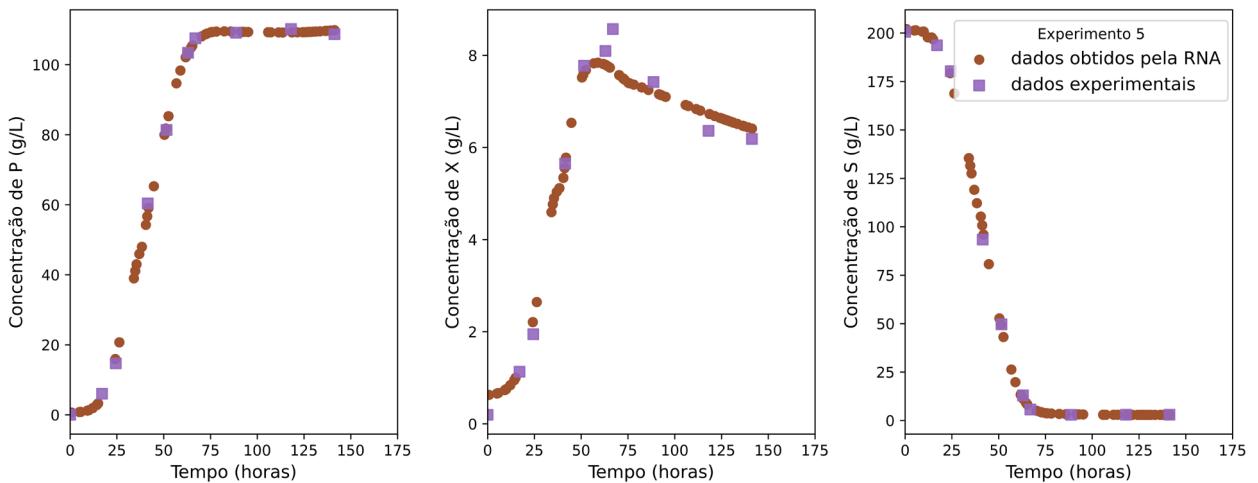
Fonte: Autoria própria.

Figura 32 – Concentrações das variáveis de saída em função do tempo: comparativo do ajuste dos dados simulados em relação aos experimentais para o experimento 4.



Fonte: Autoria própria.

Figura 33 – Concentrações das variáveis de saída em função do tempo: comparativo do ajuste dos dados simulados em relação aos experimentais para o experimento 5.



Fonte: Autoria própria.

A dificuldade em ajustar as concentrações de X pode ser explicada pela natureza biológica e dinâmica dessa variável, que apresenta comportamentos não lineares e transições abruptas entre fases de crescimento, estacionária e declínio, tornando seu padrão mais complexo para modelagem, como pode ser observado através da Figura 30, do qual continha o experimento 2, usado de forma integral no processo de validação externa. Além disso, os microrganismos interagem diretamente com o produto gerado e o substrato consumido, resultando em relações não lineares que amplificam erros no aprendizado isolado da RNA. Outro fator relevante são os possíveis outliers ou variações não modeladas nos dados experimentais de X, que podem desestabilizar o ajuste da rede, mesmo em configurações

robustas, destacando a necessidade de dados mais representativos e controle de ruídos para melhorar o desempenho.

5.8 Análise das métricas de erro

Para avaliar melhor a situação da previsão da concentração das células é visualizar os erros resultantes do processo como um todo, para isso, é necessário ver todos os erros do processo através das seguintes métricas: Erro absoluto médio (MAE), Erro quadrático médio (MSE), Raiz do erro quadrático médio (RMSE) e o Coeficiente de determinação (R^2) do qual podem ser vistos nas equações de 17 a 20.

$$MAE = \frac{1}{n} \sum_{j=1}^n |h_{j,exp} - h_{j,rede}| \cdot 100\% \quad (17)$$

$$MSE = \frac{1}{n} \sum_{j=1}^n (h_{j,exp} - h_{j,rede})^2 \cdot 100\% \quad (18)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{j=1}^n (h_{j,exp} - h_{j,rede})^2} \cdot 100\% \quad (19)$$

$$R^2 = 1 - \frac{\sum_{j=1}^n (h_{j,exp} - h_{j,rede})^2}{\sum_{j=1}^n (h_{j,exp} - \bar{h}_{j,exp})^2} \quad (20)$$

Na Equação 20, $\bar{h}_{j,exp}$ é a média dos valores experimentais.

A Tabela 4 traz o resultado de todas as métricas:

Tabela 4 – Métricas de erro para as variáveis de saída.

Erros	P	X	S	Média
MAE	5.92	0.24	4.18	3.45
MSE	102.27	0.11	71.01	57.80
RMSE	10.11	0.33	8.43	6.29
R^2	0.95	0.97	0.99	0.97

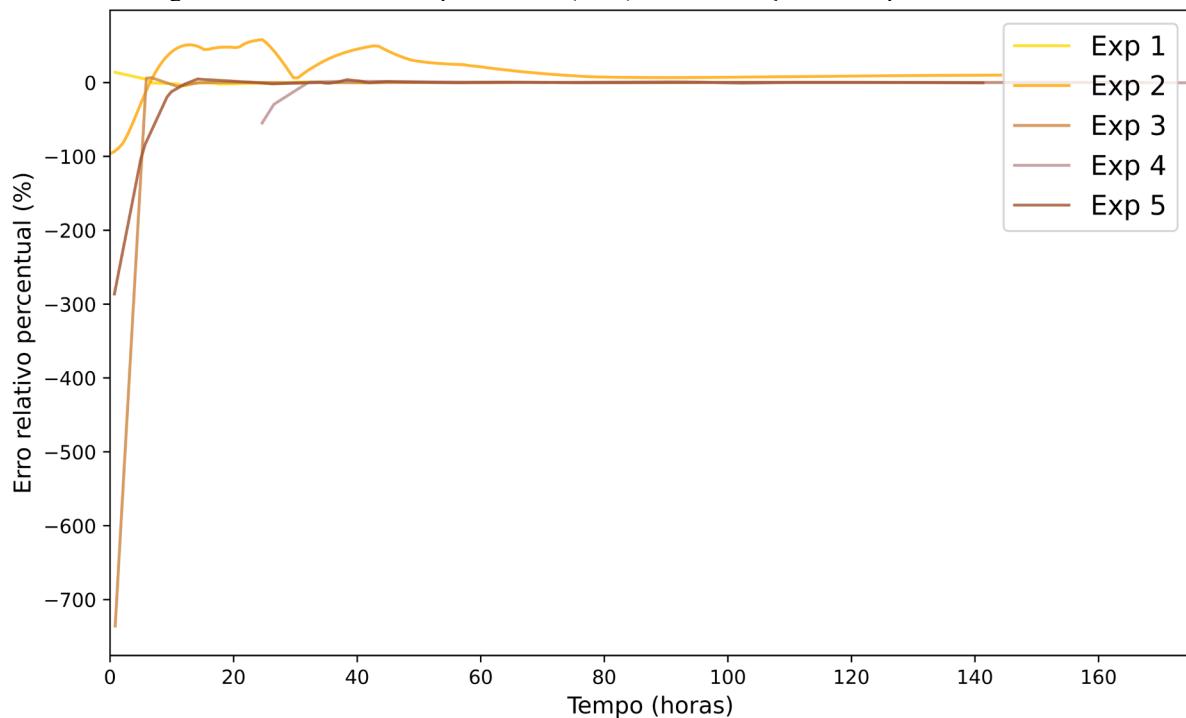
Fonte: Autoria Própria.

A variável X apresenta os menores valores de erro absoluto médio ($MAE = 0.24$), erro quadrático médio ($MSE = 0.11$) e erro médio quadrático raiz ($RMSE = 0.33$), o que poderia refletir uma alta precisão preditiva, mas isso não foi observado. O coeficiente $R^2=0.97$ confirma que a RNA modelou adequadamente a variabilidade dos dados experimentais de X, mesmo que os gráficos tenham indicado dificuldades em capturar transições dinâmicas. Isso sugere que, embora os ajustes possam parecer visualmente imprecisos, a escala relativa de X minimiza o impacto de erros pequenos em relação às outras variáveis.

Para as variáveis P e S, os erros são maiores, especialmente em P, onde o MSE atinge 102.27% e o RMSE é 10.11%. Isso indica que a RNA teve dificuldades em capturar adequadamente os padrões de P, possivelmente devido à maior variabilidade ou ruídos nos dados experimentais. No entanto, o $R^2=0.95$ demonstra que a rede conseguiu explicar a maior parte da variabilidade, indicando um ajuste geral satisfatório. Já S apresenta erros intermediários, com $MAE = 4.18\%$ e $RMSE = 8.43\%$, mas um R^2 elevado (0.99), sugerindo que os erros absolutos são relativamente pequenos quando comparados à amplitude da variável. Em resumo, as métricas médias reforçam que a RNA apresentou bom desempenho geral (R^2 médio = 0.97), mas destacam a necessidade de melhorias específicas para minimizar os erros em variáveis com maior complexidade, como P.

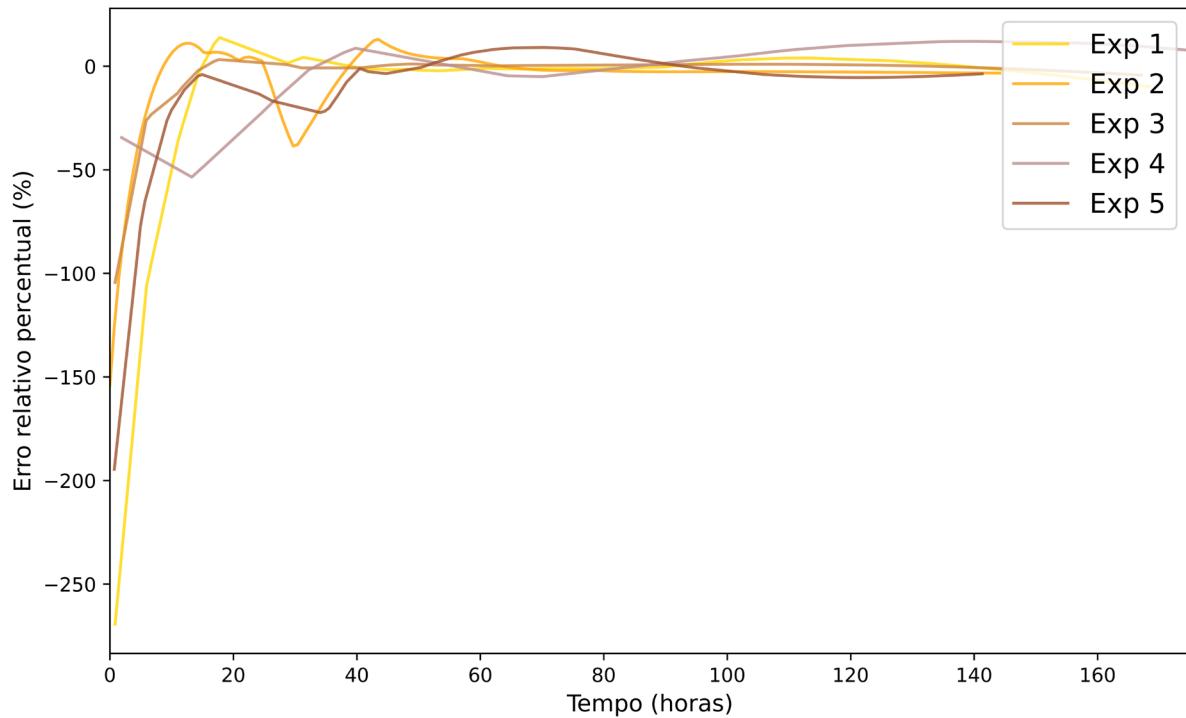
Pelas Figuras 34, 35 e 36, é evidente o maior esforço da rede em prever as concentrações iniciais até que haja um comportamento controlado no caráter preditivo. Os erros relativos percentuais para os experimentos mostram diferentes padrões de desempenho da rede neural para as variáveis de saída, os erros são mais significativos nas fases iniciais, com subestimações de aproximadamente -300% no experimento 5, evidente o maior esforço da rede em prever as concentrações iniciais até que haja um comportamento controlado no caráter preditivo, indicando que a RNA ajustou bem os padrões de crescimento após superar as discrepâncias iniciais. Em X (Figura 35), os erros apresentam maior variabilidade, com picos positivos e negativos refletindo dificuldades na modelagem de transições dinâmicas, como o pico de crescimento e o início do declínio. Já para S (Figura 36), os erros oscilam moderadamente ao longo de toda previsão, apresentando erros com limiares superior e inferior de 20 a -20%, sugerindo que a RNA modelou bem o consumo de substrato em condições mais previsíveis. Esses padrões destacam que as maiores dificuldades da RNA ocorrem nas fases iniciais e em transições não lineares, enquanto as fases mais estáveis são bem ajustadas.

Figura 34 – Erros relativos percentuais (ERP) de P em função do tempo de simulado.



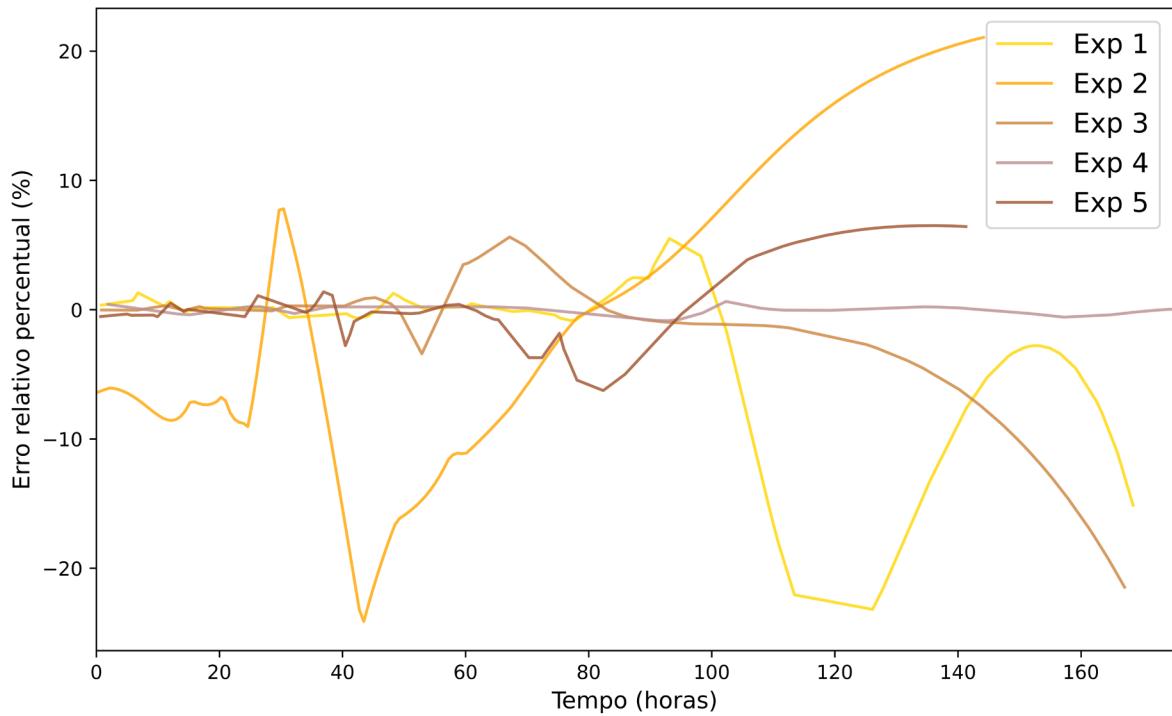
Fonte: Autoria própria.

Figura 35 – Erros relativos percentuais (ERP) de X em função do tempo simulado.



Fonte: Autoria própria.

Figura 36 – Erros relativos percentuais (ERP) de S em função do tempo simulado.



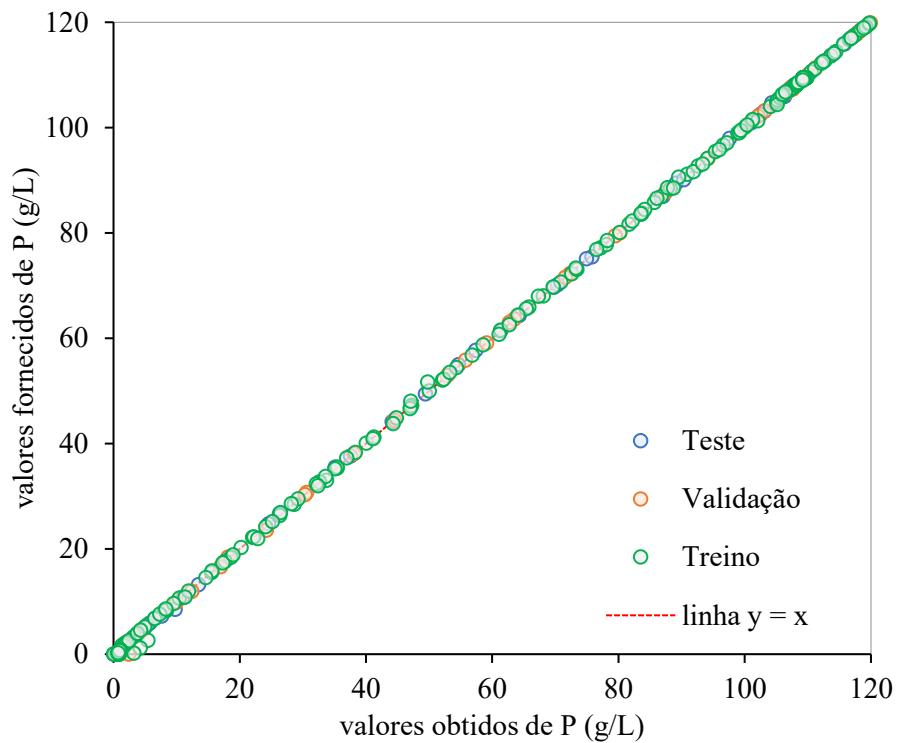
Fonte: Autoria própria.

5.9 Avaliação do processo de treinamento e validação externa

Analizar somente os resultados (ou erros) de saída do processo de simulação não demonstra o todo, é necessário ver o processo de treinamento como um todo e acompanhar o ajuste dos dados ao longo deste processo.

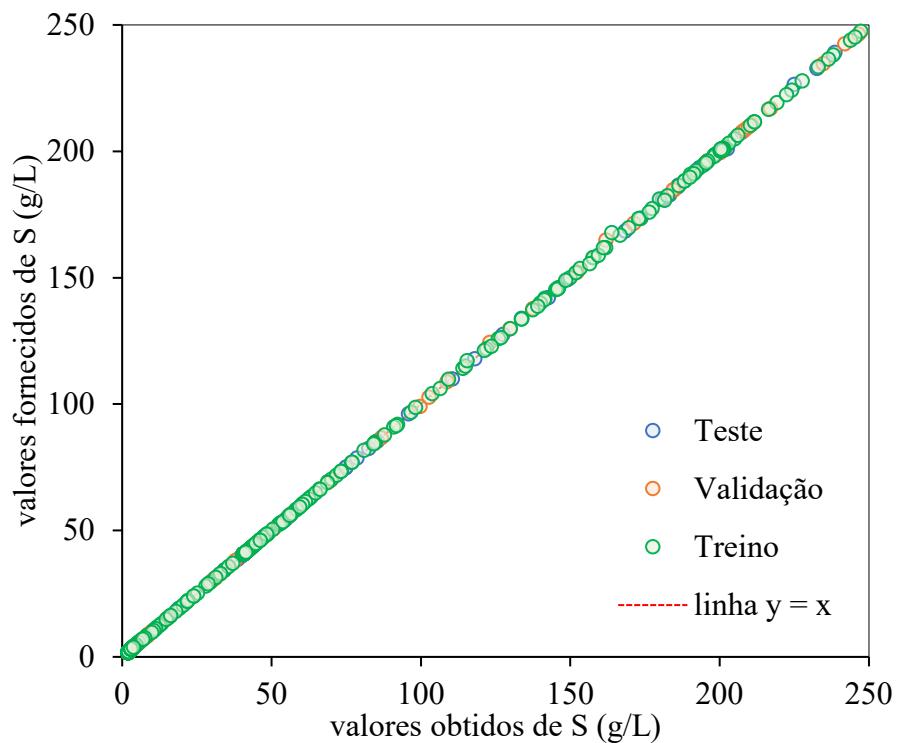
Em um contexto geral, analisando também o mesmo comportamento gráfico nas Figuras 37 e 38 da qual representam as concentrações de P e S ao longo do processo de treinamento, do qual contempla os passos de treino, validação e teste, respectivamente, infere que, a capacidade do modelo de RNA em se ajustar tão estreitamente aos dados gerados pelos ajustes suaves foi de grande sucesso. Isso não apenas válida a abordagem adotada na modelagem e treinamento da rede, mas também sugere que a RNA pode ser uma ferramenta valiosa para a previsão e otimização de processos relacionados à concentração do substrato em estudos futuros. A consistência observada entre os valores experimentais e os previstos pela RNA reforça a aplicabilidade prática deste modelo em ambientes industriais e de pesquisa, onde decisões baseadas em previsões precisas são cruciais para a eficiência e eficácia do processo.

Figura 37 - Ajuste dos dados de treinamento da RNA para P em relação com os dados fornecidos ao modelo.



Fonte: Autoria própria.

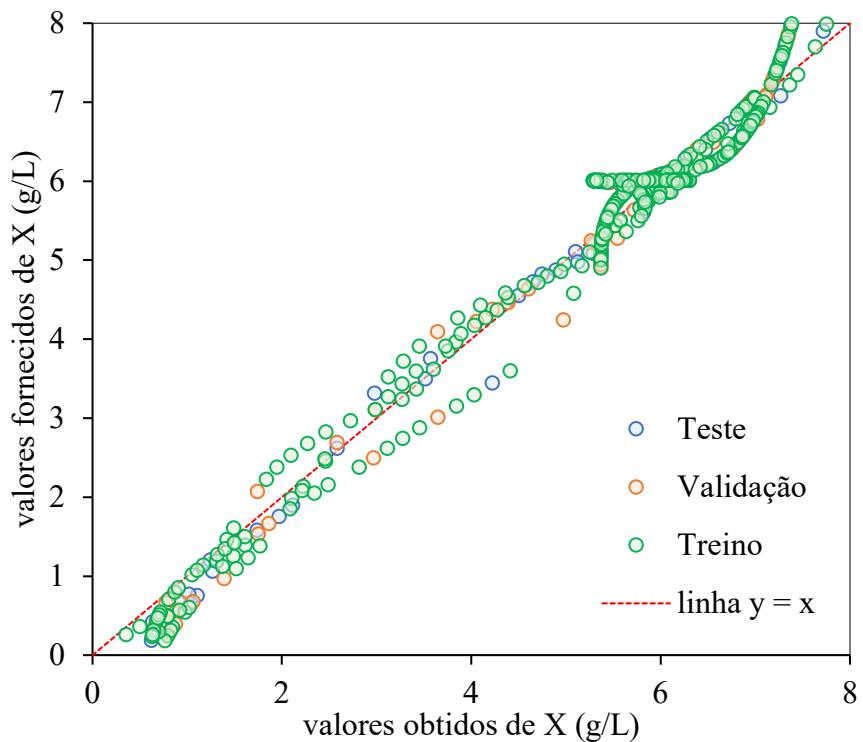
Figura 38 - Ajuste dos dados de treinamento da RNA para S em relação com os dados fornecidos ao modelo.



Fonte: Autoria própria.

Para a variável X (Figura 39), é notório uma problemática dispersão dos dados durante os processos de treinamento, mais uma vez evidenciando a complexidade de predição no comportamento da variável. Quando se observa os gráficos de dispersão, onde o eixo x representa as concentrações previstas pela RNA e o eixo y as concentrações geradas, é evidente que os pontos referentes aos conjuntos de treino, validação e teste se alinham de forma satisfatória à linha $y=x$. Esta linha representa uma correspondência ideal entre os valores previstos e os observados, indicando que o modelo de RNA tem uma capacidade excepcional de replicar os dados experimentais.

Figura 39 - Ajuste dos dados de treinamento da RNA para P em relação com os dados fornecidos ao modelo.



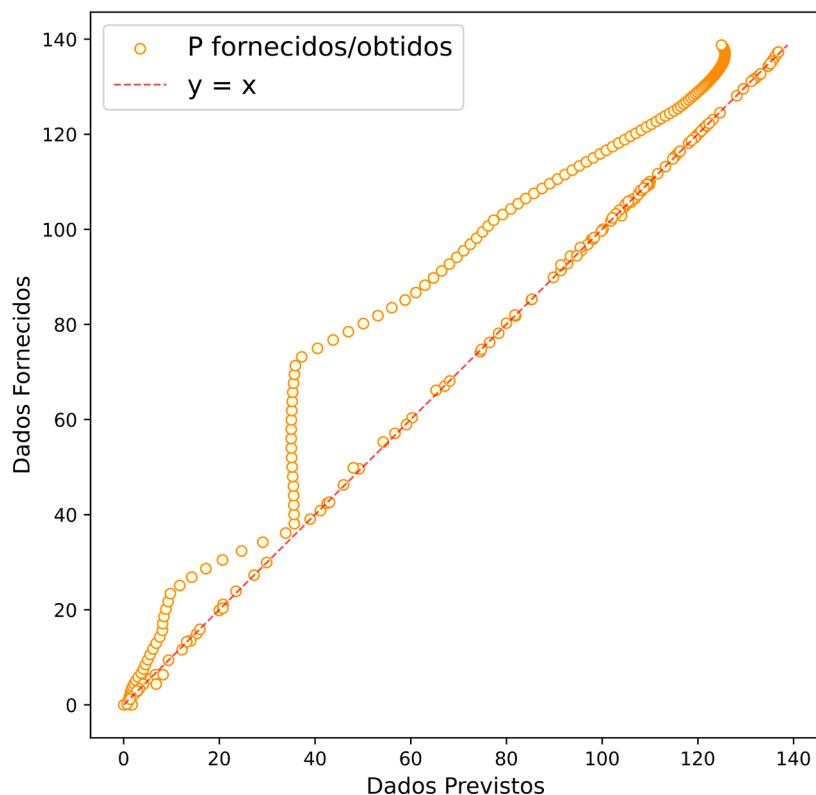
Fonte: Autoria própria.

A proximidade dos pontos de teste, validação e treino à linha $y=x$ sugere que o modelo de RNA foi bem treinado e validado, minimizando o erro entre as previsões e os valores reais (para P e S). Esta concordância reforça a robustez e a confiabilidade do modelo, demonstrando sua capacidade de generalizar bem para novos dados, além de replicar com precisão os dados nos quais foi treinado.

Outra análise criteriosa a ser observada, são os ajustes dos dados simulados (validação externa), realizando uma plotagem dos dados fornecidos (gerados pelos ajustes suaves) contra os retornados pela rede para os mesmos pontos, a fim de visualizar os ajustes desses dados.

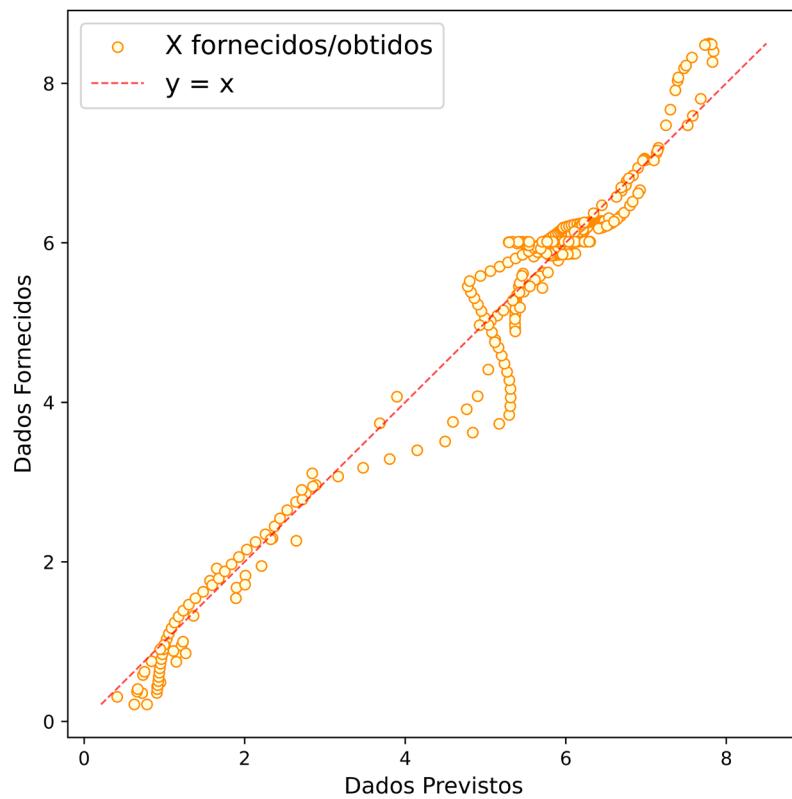
Nas Figuras 40, 41 e 42 é notório um ajuste problemático ao longo do meio do processo, embora já justificado pelas análises comportamentais dos erros relativos percentuais. Para P , há maior dispersão em valores intermediários, sugerindo que a RNA enfrenta dificuldades em ajustar essa faixa específica. Em X , os dados apresentam maior proximidade da linha de identidade, evidenciando menor dispersão e maior precisão nos ajustes, embora pequenas inconsistências ocorram em pontos de transição. Já em S , o alinhamento é mais consistente, com a dispersão concentrada em faixas específicas, provavelmente devido a variações dinâmicas na fase inicial ou final. Essa análise reforça que, apesar de boa capacidade de ajuste, há espaço para melhorias na captura de padrões mais complexos ou em transições dinâmicas. Pelas dinâmicas do processo químico aqui apresentado.

Figura 40 - Ajuste dos dados simulados da RNA para P em relação com os dados fornecidos ao modelo.



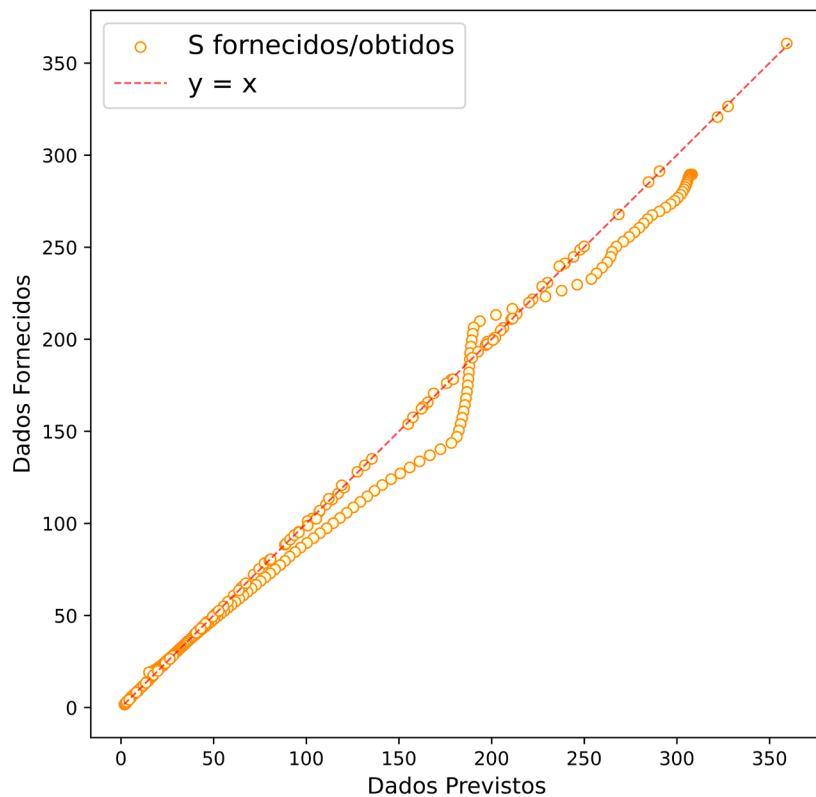
Fonte: Autoria própria.

Figura 41 - Ajuste dos dados simulados da RNA para S em relação com os dados fornecidos ao modelo.



Fonte: Autoria própria.

Figura 42 - Ajuste dos dados simulados da RNA para S em relação com os dados fornecidos ao modelo.



Fonte: Autoria própria.

5.10 Análise de resíduos do modelo

A análise dos resíduos é uma etapa fundamental na avaliação de modelos preditivos, incluindo redes neurais. Resíduos são definidos como a diferença entre os valores observados e os valores preditos pelo modelo. A investigação sistemática desses resíduos permite uma compreensão mais aprofundada sobre a capacidade do modelo em capturar os padrões subjacentes dos dados e identificar possíveis limitações em seu desempenho (Shmueli, 2010).

Se formos expressar isso de forma generalizada, temos a Equação 21:

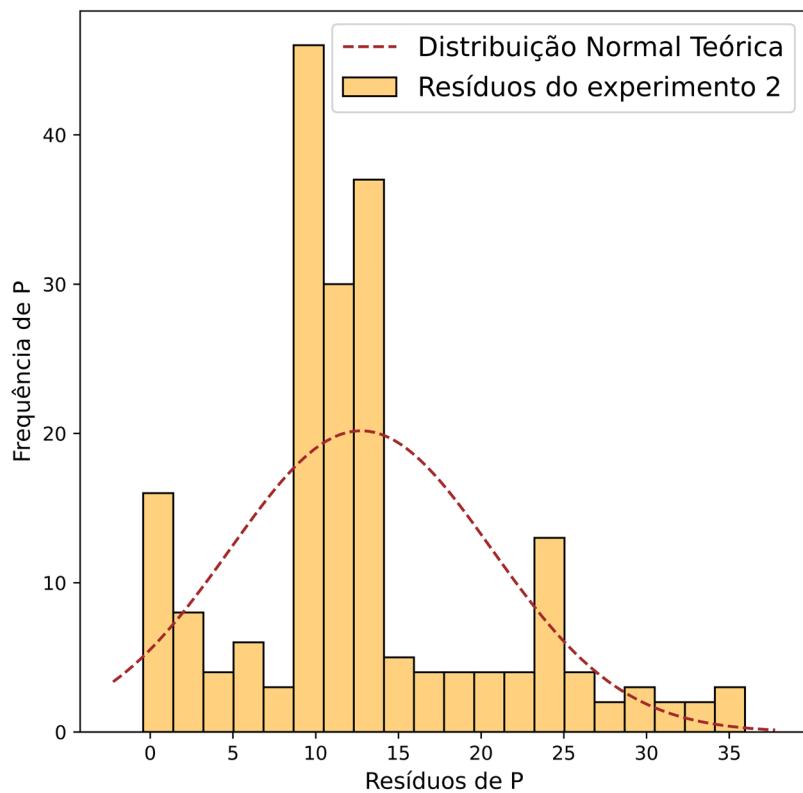
$$e_i = y_i - \hat{y}_i \quad (21)$$

Sendo, e_i é o resíduo para o ponto i , y_i o valor observado no ponto i e \hat{y}_i o valor predito pelo modelo no ponto i .

Um dos principais objetivos da análise de resíduos é verificar a adequação do modelo. Idealmente, os resíduos devem ser distribuídos de forma aleatória e homogênea em torno de zero, sem apresentar padrões estruturados. Resíduos sistemáticos, como tendências crescentes ou decrescentes, podem indicar que o modelo não está capturando relações importantes entre as variáveis. Isso pode ser reflexo de uma modelagem inadequada, insuficiência de complexidade na arquitetura da rede ou mesmo a presença de relações não lineares que o modelo não foi treinado para interpretar (Shmueli, 2010).

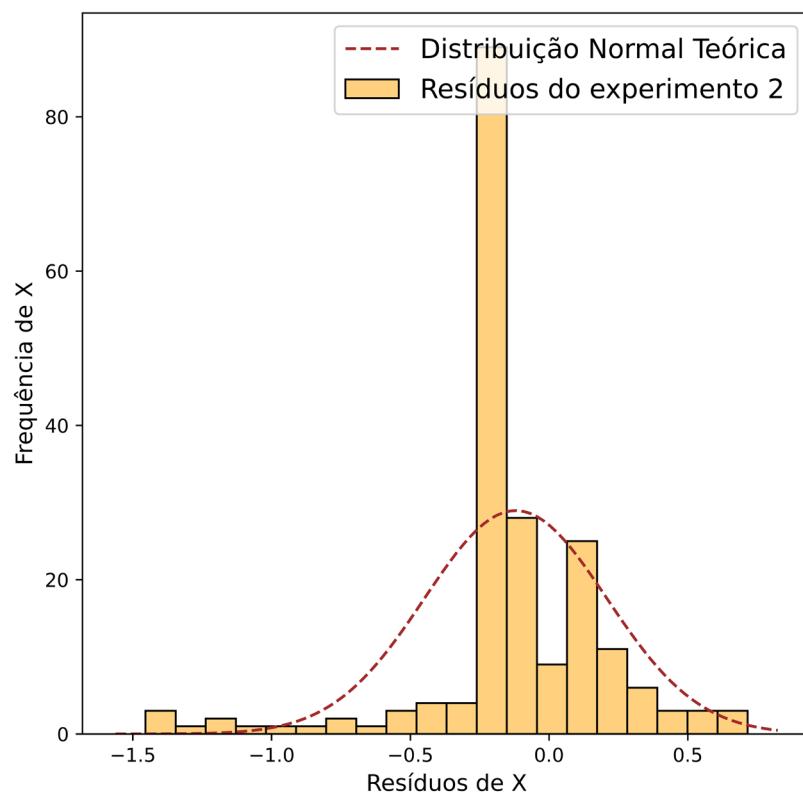
Além disso, a análise de resíduos desempenha um papel essencial na detecção de outliers e heterocedasticidade. Outliers, representados por resíduos extremamente altos ou baixos, podem ser indicativos de erros nos dados ou de cenários que o modelo não consegue generalizar. Já a heterocedasticidade, caracterizada pela variação não uniforme dos resíduos ao longo do conjunto de dados, pode sinalizar problemas na estabilidade do modelo e na confiança de suas previsões (Hastie, 2009). Os resíduos das variáveis (P, X e S) do experimento 2 (do qual foi destinado exclusivamente para a validação externa) são destacados nas Figuras 43, 44 e 45 apresentam suas respectivas distribuições e desvios da normalidade.

Figura 43 - Resíduos da RNA para P para o experimento 2.



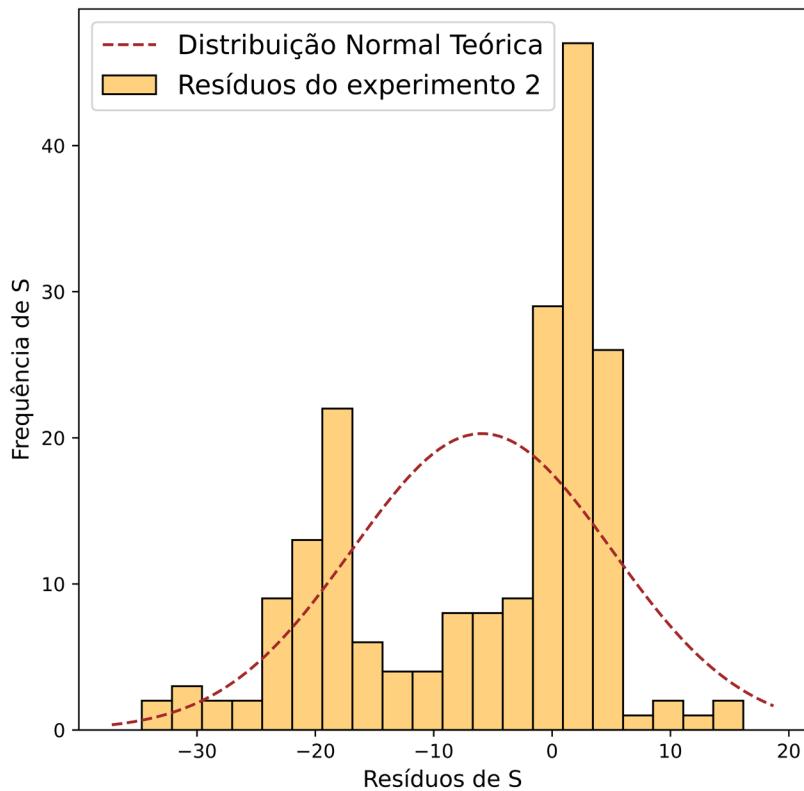
Fonte: Autoria própria.

Figura 44 - Resíduos da RNA para X para o experimento 2.



Fonte: Autoria própria.

Figura 45 - Resíduos da RNA para S para o experimento 2.

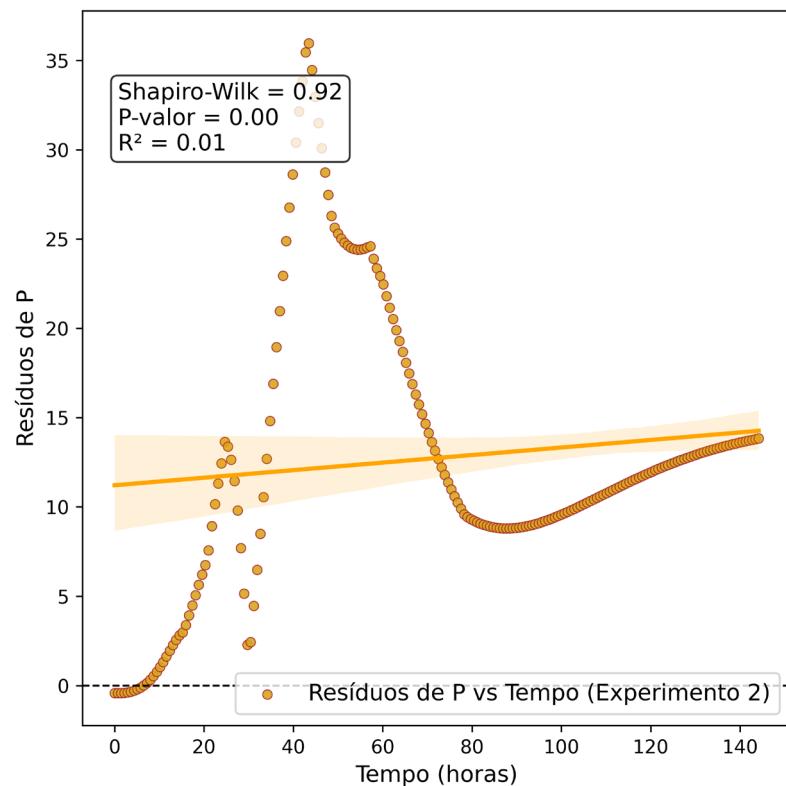


Fonte: Autoria própria.

Todas as Figuras apresentam distribuição assimétrica e desvios da normalidade, confirmados pelos baixos valores de Shapiro-Wilk e p-valores iguais a zero. Embora os resíduos não sigam uma distribuição normal teórica, uma vez que a maioria dos dados deveria ter frequências maiores em valores próximos de zero e, mais próximas deste valor, descendo a frequência de aparições (soma das quantidades que esse valor aparece) e sem *outliers*. Apesar das problemáticas destacadas, o comportamento geral sugere que o modelo capturou padrões importantes, mas ainda apresenta limitações em algumas áreas específicas, especialmente para a variável S (Figura 45), que exibe maior variabilidade. Esses resultados podem indicar a necessidade de ajustes adicionais para melhorar a capacidade do modelo em representar a variabilidade dos dados. Mesmo assim, a maioria das variáveis tiveram média de 0.78 nos valores de Shapiro-Wilk, refletindo um bom ajuste geral, por mais que precise de ajustes futuros.

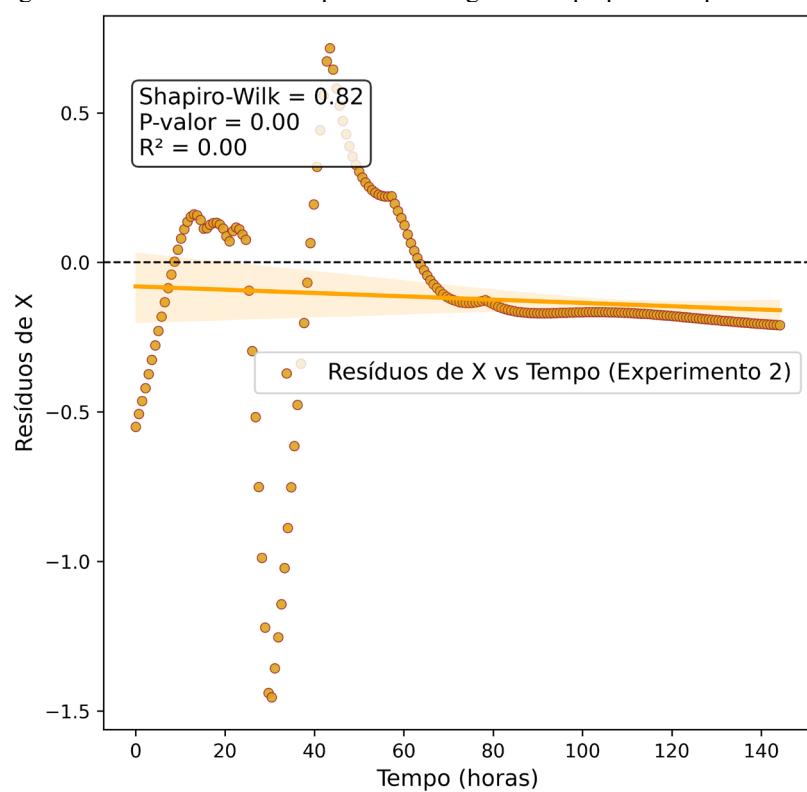
Outra análise interessante é a observação dos resíduos ao longo do tempo de validação externa, onde é possível observar o quanto os resultados apresentaram erros maiores ao longo do começo do processo, solidificando as análises dos erros. As Figuras 46 a 48 apresentam os resíduos das variáveis de saída ao longo do tempo para o experimento 2.

Figura 46 - Resíduos da RNA para P ao longo do tempo para o experimento 2.



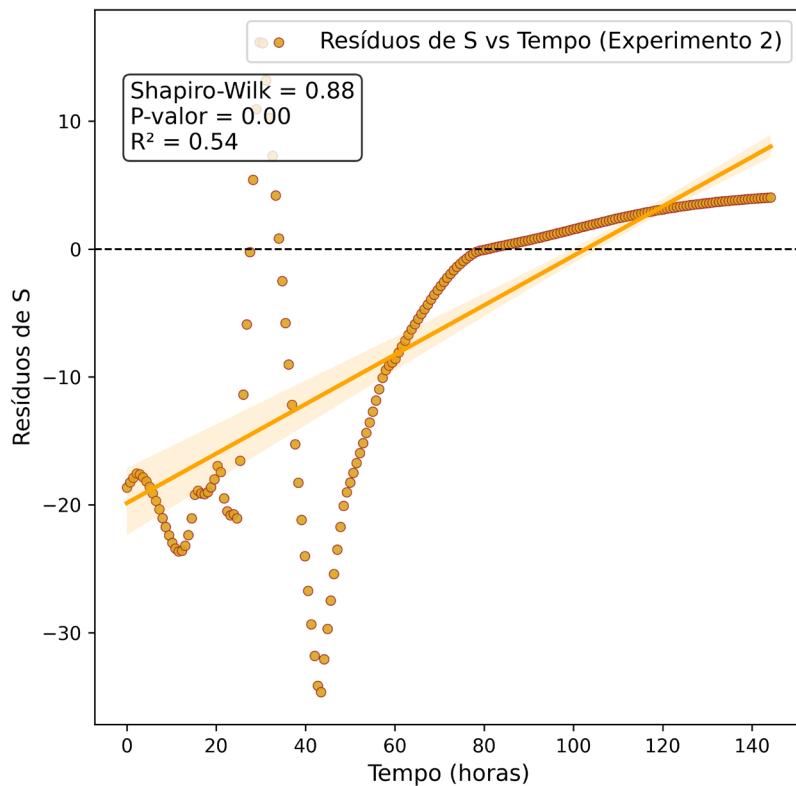
Fonte: Autoria própria.

Figura 47 - Resíduos da RNA para X ao longo do tempo para o experimento 2.



Fonte: Autoria própria.

Figura 48 - Resíduos da RNA para S ao longo do tempo para o experimento 2.



Fonte: Autoria própria.

Nas Figuras 46, 47 e 48, é possível observar um comportamento mais disperso dos erros no começo do processo. Ao longo da validação, mais além do meio do processo, é possível ver os dados agrupando-se, influenciando uma linearização do comportamento. É notório salientar que, em um processo teoricamente viável, a linha de regressão deveria acompanhar a linha de X em zero, P e S (Figuras 46 e 48) se distanciam vastamente, por mais que possuam coeficiente de determinação (R^2) mais altos pelo fato da inclinação (que aqui é um problema serem muito inclinadas), sendo a Figura 47, para a variável X, os melhores ajustes nesse experimento 2 em relação aos resíduos.

Outro aspecto relevante é a validação de suposições estatísticas subjacentes ao modelo. Embora redes neurais sejam modelos de caixa-preta, muitas aplicações assumem que os resíduos seguem uma distribuição normal, especialmente em tarefas de regressão. Testes de normalidade, como o de Shapiro-Wilk, podem ser utilizados para verificar se essa suposição é válida. Resíduos que não seguem uma distribuição normal podem indicar a necessidade de transformações nos dados ou de ajustes na arquitetura do modelo (Shmueli, 2010).

Por fim, a análise de resíduos contribui para o diagnóstico de problemas como sobreajuste ou subajuste. Um modelo em sobreajuste pode apresentar resíduos muito pequenos no

conjunto de treinamento, mas grandes no conjunto de validação ou teste, indicando que o modelo memorizou os dados de treinamento em vez de aprender padrões generalizáveis. Por outro lado, resíduos elevados e sistemáticos tanto no treinamento quanto na validação sugerem que o modelo está em subajuste, ou seja, não possui capacidade suficiente para capturar a complexidade dos dados (Hastie, 2009).

5.11 Modelagem final para o modelo gerado

Conforme elucidado por Martins (2018), uma RNA pode ser conceituada como uma equação matemática complexa e não linear, composta por uma multiplicidade de parâmetros, incluindo os pesos e os vieses associados a cada neurônio artificial. Durante a fase de treinamento, esses parâmetros sãometiculosamente ajustados. A representação matemática para a variável de saída de uma RNA que possui duas camadas intermediárias é delineada pela Equação (22).

$$y_k = g \left(\sum_{j=1}^Z w_{jk} f_2 \left(\sum_{h=1}^Y w_{hj} f_1 \left(\sum_{i=1}^X w_{ih} x_i + b_h \right)_h + b_j \right)_j + b_k \right)_k \quad (22)$$

Nesta equação, k denota o número total de saídas, y_k representa a saída específica da RNA, g é a função de ativação da camada de saída, Z é a contagem de neurônios na camada de saída, f_1 é a função de ativação pertinente à camada intermediária primária e f_2 é a camada intermediária secundária, X é o número de neurônios na primeira camada e Y os neurônios da segunda camada. Os termos b_h , b_j e b_k referem-se aos vieses, enquanto w_{ih} , w_{hj} e w_{jk} são os pesos sinápticos das camadas intermediária primária e secundária e saída, respectivamente.

Com base nessa formulação, a RNA ótima para monitorar o processo de fermentação do hidromel é descrita pela Equação (23).

$$\begin{aligned}
y_k = & \text{tansig} \left(\sum_{j=1}^3 \begin{bmatrix} 0.58 \\ \vdots \\ -1.67 \end{bmatrix} \text{softmax} \left(\sum_{h=1}^{25} \begin{bmatrix} -0.26 & \cdots & -0.43 \\ \vdots & \ddots & \vdots \\ -0.28 & \cdots & 0.43 \end{bmatrix} \text{poslin} \left(\sum_{i=1}^{40} \begin{bmatrix} -1.15 & \cdots & -0.46 \\ \vdots & \ddots & \vdots \\ -0.20 & \cdots & -0.08 \end{bmatrix} \begin{bmatrix} S.S. \\ pH \\ D.O. \\ T \end{bmatrix} \right) \right. \right. \\
& \left. \left. + \begin{bmatrix} -1.01 \\ \vdots \\ -0.74 \end{bmatrix}_h \right)_h + \begin{bmatrix} -0.28 \\ \vdots \\ -0.34 \end{bmatrix}_j \right)_j + \begin{bmatrix} -0.37 \\ -0.13 \\ -0.98 \end{bmatrix}_k \right)_k
\end{aligned} \tag{23}$$

No entanto, é imperativo ressaltar que, devido à extensiva quantidade de dados e à necessidade de precisão decimal exigida pelo software Matlab®, este estudo não detalha os pesos e vieses da RNA. Para informações adicionais ou detalhes específicos, os interessados são encorajados a entrar em contato diretamente com o orientador através do e-mail: tdmartins@unifesp.br.

6 CONCLUSÕES E SUGESTÕES

6.1 Conclusões

Neste estudo, foi concebido um sensor virtual fundamentado em Redes Neurais Artificiais (RNAs) com o propósito de prever, de maneira eficiente, as concentrações de células (X), açúcares redutores (S) e do produto (P) ao longo do processo fermentativo na produção de hidromel. A pesquisa também se dedicou a discernir os algoritmos de treinamento mais eficazes e as funções de ativação mais propícias para tal aplicação complexa que leva em consideração a temperatura. Estes insights são cruciais para orientar investigações subsequentes no domínio da produção de hidromel, bem como para ampliar o escopo da aplicação de técnicas de aprendizado de máquina e inteligência artificial na indústria.

Os resultados obtidos, em termos de eficiência, demonstraram consistência entre a estratégia adotada. Dada essa congruência, foi possível identificar uma única rede neural artificial para a previsão simultânea de X, S e P.

Na abordagem que contemplou duas camadas intermediárias, foram geradas em torno de 300.000 RNAs, com cada rede encarregada de prever as três variáveis em simultâneo. Notavelmente, a RNA 4-40-25-3 (entrada-intermediária 1, 2-saída) e algoritmo de otimização baseado em regularização bayesiana (*trainbr*) contendo as funções de ativação intermediárias *poslin* e *softmax* e saída *tansig* alcançou previsões para as concentrações de células (X), substrato (S) e produto (P) com MAPEs de 13.05, 8.75 e 6.23%, respectivamente. Esta RNA apresentou um coeficiente de correlação de Pearson (ρ) de 0.92, composta por um total de 1303 de parâmetros.

Ao avaliar o desempenho dos algoritmos de otimização, observou-se que a função *softmax* mostrou-se mais apropriada para ambas as camadas intermediárias, enquanto a função linear (*purelin*) foi predominante para a camada de saída, refletindo sua recorrência nos resultados mais promissores a estratégia, embora o resultado tenha收敛ido melhor com a função tangente hiperbólica (*tansig*).

Em relação à seleção de funções de ativação, a combinação de *poslin*, *softmax* e *tansig* apresenta características distintas que potencializam o desempenho da rede em problemas não lineares. Essa combinação tem sido amplamente estudada, mostrando resultados promissores em aplicações que demandam flexibilidade e precisão.

O algoritmo *trainbr* apresentou os melhores resultados devido à sua capacidade de combinar a eficiência no treinamento com uma excelente generalização dos modelos. Baseado

em regularização bayesiana, o *trainbr* ajusta os pesos da rede enquanto controla a complexidade do modelo, evitando sobreajuste, mesmo em problemas com alta variabilidade e não linearidades. Esse algoritmo otimiza a função de erro incluindo um termo de penalização para pesos excessivamente grandes, o que favorece modelos mais equilibrados e robustos.

É essencial que os pesquisadores continuem explorando diferentes combinações de algoritmos de otimização e funções de ativação para determinar as melhores configurações para suas aplicações específicas.

6.2 Sugestões para trabalhos futuros

Embora este estudo tenha demonstrado a aplicabilidade das Redes Neurais Artificiais (RNAs) no monitoramento de variáveis do processo de fermentação, trabalhos futuros podem explorar a eficácia de diferentes modelos de aprendizado supervisionado. Modelos como Random Forest (Floresta Aleatória), SVM (Máquina de Vetores de Suporte), XGBoost (Extreme Gradient Boosting), KNN Regressor (Regressor de Vizinhos Mais Próximos) e Gaussian Process (Processo Gaussiano) representam abordagens promissoras, cada qual com características específicas que merecem investigação detalhada.

O Gaussian Process (Processo Gaussiano), por exemplo, é uma abordagem probabilística poderosa que utiliza funções de kernel para modelar relações complexas entre as variáveis de entrada e saída. Sua principal vantagem está na capacidade de fornecer previsões precisas acompanhadas por uma estimativa de incerteza, tornando-o ideal para cenários em que a confiabilidade dos resultados é crítica, como no monitoramento de processos sensíveis. Essa abordagem permite capturar nuances nos dados e pode adaptar-se bem a problemas não lineares (Williams & Rasmussen, 2006).

Por outro lado, o Random Forest (Floresta Aleatória) é uma abordagem baseada em um conjunto de árvores de decisão que utiliza o método de agregação. Essa técnica constrói múltiplas árvores de decisão, cada uma treinada em diferentes subconjuntos aleatórios do conjunto de dados, e combina suas previsões por meio de votações (no caso de classificação) ou médias (para regressão). Esse método confere ao modelo robustez, tornando-o capaz de lidar eficientemente com dados ruidosos e de evitar o sobreajuste comum em árvores de decisão individuais (Breiman, 2001).

Além disso, um caminho interessante para pesquisas futuras seria a criação de uma *pipeline* automatizada para treinamento e validação desses modelos. A pipeline poderia

incluir a busca automatizada de hiperparâmetros utilizando métodos como grid search ou otimização bayesiana, além de comparações sistemáticas entre os modelos para selecionar a abordagem mais eficaz em diferentes cenários.

Por fim, a competição entre modelos, com base em métricas como RMSE, MAE, R², MBE (Mean Bias Error), RSE (Relative Squared Error), Log-Cosh e Hamming Loss, permitiria identificar os algoritmos mais adequados para prever variáveis críticas de processos fermentativos. Essas métricas oferecem uma avaliação abrangente, considerando desde a precisão das previsões (como RMSE e MAE) até a análise de viés (MBE) e a robustez a outliers com o Log-Cosh, por exemplo (Willmott, 2005; Chai, 2014; Hamming, 1950).

Essas direções futuras certamente ampliariam o escopo das aplicações de aprendizado de máquina em processos industriais, especialmente no monitoramento e controle de fermentações, consolidando métodos mais eficientes e precisos.

8 REFERÊNCIAS BIBLIOGRÁFICAS

- ALTEMEYER, R.; BIRNBAUM, B.; TROYANOVSKI, D. **Variable importance in machine learning: Practical applications and limitations.** Data Science Journal, v. 19, p.1-14, 2021. [Link para o Artigo.](#)↔
- AMGARTEN, E. L. **Estudo Da Viabilidade de Uso de Redes Neurais Artificiais Para a Predição Do Rendimento Da Fermentação Alcoólica.** Repositório Institucional UP Campus sede – Ecoville Dissertação, Programa de Pós-Graduação em Biotecnologia Industrial, 2012. [Link para a Dissertação.](#)↔
- AMORIM, T. S.; LOPES, S. B.; BISPO, J. A. C.; BONAFE, C. F. S.; CARVALHO, G. B. M. de; MARTÍNEZ, E. A. **Influence of acerola pulp concentration on mead production by *Saccharomyces cerevisiae*.** LWT, v. 97, p. 561-569, ISSN 0023-6438, 2018. [Link para o Artigo.](#)↔
- ANDRADE, R. R; COSTA, A. **Modelagem cinética do processo de produção de etanol a partir de hidrolisado enzimático de bagaço de cana-de-açúcar concentrado com melaço considerando reciclo de células.** 262 p. Tese (doutorado) - Universidade Estadual de Campinas, Faculdade de Engenharia Química, Campinas, SP, 2012. [Link para a Dissertação.](#)↔
- ANDRADE, R. R.; MAUGERI FILHO, F.; MACIEL FILHO, R.; COSTA, A. C. **Kinetics of ethanol production from sugarcane bagasse enzymatic hydrolysate concentrated with molasses under cell recycle.** Bioresource Technology, v.130, p. 351-359, 2013. [Link para o Artigo.](#)↔
- ANDRADE, R. R.; MARTINS, T.D.; SILVA, R.M.R. **Sensor virtual neural para monitoramento da fermentação de hidromel.** Revista de Engenharia e Tecnologia, v. 14, n. 1, 2021. [Link para o Artigo.](#)↔
- ARCHILIA, M. D.; NETO, A. A. L.; MARCUCCI, M. C.; ALONSO, R. C. B.; CAMARGO, T. C.; CAMARGO, R. C.; SAWAYA, A. C. H. F. **Characterization of Brazilian monofloral and polyfloral honey by UHPLC-MS and classic physical-chemical analyses.** Journal of Apicultural Research, v. 62, n. 3, p. 578-589, 2023. [Link para o Artigo.](#)↔
- AWANG, M. K.; ISMAIL, M.; MAKHTAR, M.; RAHMAN, M. A.; MAMAT, A. R. **Performance Comparison of Neural Network Training Algorithms for Modeling Customer Churn Prediction.** International Journal of Engineering & Technology, v. 7, n. 2.15, p. 35-37, 2018. [Link para o Artigo.](#)↔
- BISHOP, C. M. **Pattern Recognition and Machine Learning.** Springer New York, NY, Ed. 1, 2006. [Link para o Livro.](#)↔
- BEALE, E. M. L. **A derivation of conjugate gradients In Numerical Methods for Non-linear Optimization.** Academic Press. p. 39-43, 1972.
- BRAGA, A. P.; LUDEMIR, T. B.; CARVALHO, A. P. L. F. **Redes neurais artificiais: teoria e aplicações.** Rio de Janeiro: LTC. 2 Ed. [Link para o Livro.](#)↔

BREIMAN, L. Random forests. *Machine Learning*, v.45, n.1, p.5–32, 2001. DOI: 10.1023/A:1010933404324. [Link para o Artigo.](#)↔

BRITO, A. F. S. **Análise cinética e estudos dos parâmetros fermentativos para a produção de hidromel.** 67f. (Trabalho de Conclusão de Curso - Monografia), Curso de Engenharia de Biotecnologia e Bioprocessos, Centro de Desenvolvimento Sustentável do Semiárido, Universidade Federal de Campina Grande, Sumé - PA, Brasil, 2015. [Link para a Dissertação.](#)↔

BRUNELLI, L. T. **Caracterização físico-química, energética e sensorial de hidromel.** 85f. Tese (doutorado) - Universidade Estadual Paulista, Faculdade de Ciências Agronômicas de Botucatu, 2015. [Link para a Dissertação.](#)↔

BURNS, J.; MULLEN, W.; LANDRAULT, N.; TEISSEDRE, P. L.; LEAN, M. E. J.; CROZIER, A. **Variations in the profile and content of anthocyanins in wines made from Cabernet sauvignon and hybrid grapes.** Journal of Agriculture and Food Chemistry, v. 50, n. 14, p. 4096–4102, 2002. [Link para o Artigo.](#)↔

BUSSE-VALVERDE, N.; GÓMEZ-PLAZA, E.; LÓPEZ-ROC, J. M.; GIL-MUNOZ, R.; BAUTISTA-ORTIN, A. B. **The extraction of anthocyanins and proanthocyanidins from grapes to wine during fermentative maceration is affected by the enological technique.** Journal of Agriculture and Food Chemistry, v. 59, n. 10, p. 5450–5455, 2011. [Link para o Artigo.](#)↔

CARNEIRO JÚNIOR, J. B. A.; SILVA, M. C.; NAGAYAMA, R. R.; VONA, T. S. P. **Aplicação de redes neurais artificiais na análise financeira das empresas do setor de petróleo.** Revista Brasileira de Administração Científica, v.13 n.2, 2023. [Link para o Artigo.](#)↔

CHAI, T., DRAXLER, R. R. **Root mean square error (RMSE) or mean absolute error (MAE)? Arguments against avoiding RMSE in the literature.** *Geoscientific Model Development*, v.7, p.1247–1250, 2014. [Link para o Artigo.](#)↔

CHEN, T.; GUESTRIN, C. **XGBoost: A scalable tree boosting system.** *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16)*, p785–794, 2016. [Link para o Artigo.](#)↔

CHOJI, T. T.; OTTAIANO, G. Y.; SEOLATTO, A. A.; SILVA, E. A.; MARTINS, T. D. **Cinética da biossorção da mistura Ni²⁺-Cr³⁺ via redes neurais artificiais.** REVISTA DE ENGENHARIA E TECNOLOGIA, v. 13, p. 200-210, 2021. [Link para o Artigo.](#)↔

COSTA, J. S.; SILVA, G. M. A.; SCÁRDUA, L. A. **Uso de Redes Neurais Artificiais para classificação de propostas dependentes e independentes em licitações no ramo de óleo e gás.** Peer Review, v. 5, n. 7, p. 38–57, 2023. [Link para o Artigo.](#)↔

COSTA, B. S. G. **Produção de Hidromel com adição de casca de jabuticaba: Influência dos parâmetros fermentativos, qualidade da bebida e concentração de antocianinas.** 2022. 59 f. Trabalho de conclusão de curso (Bacharelado em Química) - Instituto de Ciências

Ambientais, Química e Farmacêuticas, Universidade Federal de São Paulo, Diadema, 2022. [Link para a Dissertação.](#)↔

DAL'BOIT, S. **Predição do equilíbrio líquido-vapor através de redes neurais artificiais.** 70f. Dissertação (mestrado) - Universidade Estadual de Campinas, Faculdade de Engenharia Química, Campinas, SP, 2008. [Link para o Artigo.](#)↔

DAYHOFF, J.; DELEO, J. **Artificial neural networks. Conference on Prognostic Factors and Staging in Cancer Management: Contributions of Artificial Neural Networks and Other Statistical Methods Supplement to Cancer.** American Cancer Society, v. 3, n. 3, p. 1615-1635, 2001. [Link para o Artigo.](#)↔

DE QUEIROZ, J. C. F.; DE FARIAS RAMOS, D.; ALVES, A. S. S.; RODRIGUES, J. S. L.; DE LIMA SOUZA, J. W. **Produção de Hidromel de Forma Artesanal e Avaliação dos Parâmetros durante o Processo Fermentativo.** Revista Saúde Ciência Online. v. 3, n. 3, p. 321-328, 2014. [Link para o Artigo.](#)↔

DUFOUR, C.; SAUVAITRE, I. **Interactions between anthocyanins and aroma substances in a model system. Effect on the flavor of grape-derived beverages.** Journal of Agriculture and Food Chemistry, v. 48, p. 1784–1788, 2000. [Link para o Artigo.](#)↔

FERRAZ, F. D. O. **Estudo dos parâmetros fermentativos, características físico-químicas e sensoriais de hidromel.** 129 f. Tese (Doutorado) – Escola de Engenharia de Lorena, Universidade de São Paulo, Lorena, 2015. [Link para a Dissertação.](#)↔

FLECK, L.; TAVARES, M. H. F.; EYNG, E.; HELMANN, A.; ANDRADE, M. **Redes neurais artificiais: princípios básicos.** Revista Eletrônica Científica Inovação e Tecnologia, v. 1, n. 13, p. 47-57, 2016. [Link para o Artigo.](#)↔

FREITAS, V. De; MATEUS, N. **Formation of pyranoanthocyanins in red wines: a new and diverse class of anthocyanin derivatives.** Analytical and Bioanalytical Chemistry, v.401, p.1463-1473, 2011. [Link para o Artigo.](#)↔

FROST, JIM. **Spearman's Correlation Explained.** Acesso em: 15 dez. 2024. [Link para o Material.](#)↔

FU, Y.; SHI, X.; LI, F.; YAN, X.; LI, B.; LUO, Y.; JIANG, G.; LIU, X.; WANG, L. **Fermentation of mead using *Saccharomyces cerevisiae* and *Lactobacillus paracasei*: Strain growth, aroma components and antioxidant capacity.** Food Bioscience, v. 52, p. 102402, 2023. ISSN 2212-4292. [Link para o Artigo.](#)↔

GARDEZANI, O. P. Project 1: **Neural virtual sensor for monitoring the fermentation of mead with Jabuticaba extract.** [linguagem em Python]. GitHub, 2023. [Link para o Projeto.](#)↔

GOMES, T. M. C., **Produção de Hidromel: efeito das condições de fermentação.** Instituto Politécnico de Bragança, Escola Superior Agrária. Dissertação de Mestrado em Biotecnologia, 2010. [Link para o Artigo.](#)↔

GOLDSTEIN, A.; KAPELNER, A.; BLEVINS, T.; LEHNHARDT, T. **Peeking inside the black box: Visualizing statistical learning with plots of individual conditional expectation.** Journal of Computational and Graphical Statistics, v. 24, n. 1, p. 44-65, 2015. [Link para o Artigo.](#)↔

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning.** MIT Press. Nature, v.521, n.7553, p.436-444, 2015. [Link para o artigo.](#)↔

GUPTA, J. K.; SHARMA, R. **Production technology and quality characteristics of mead and fruit-honey wines: A review.** Natural Product Radiance, New Delhi, v.8, p.345-355, 2009. [Link para o Artigo.](#)↔

GURIAN, G. F. **Bebidas fermentadas brasileiras: considerações histórico-semânticas sobre aluás.** Boletim do Museu Paraense Emílio Goeldi. Ciências Humanas, Vol. 18, no. 2, 2023. [Link para o Artigo.](#)↔

HAGAN, M. T., & MENHAJ, M. B. (1994). **Training feedforward networks with the Marquardt algorithm.** IEEE transactions on Neural Networks, v.5, n.6, p.989-993. [Link para o Artigo.](#)↔

HAMMING, R. W. **Error detecting and error correcting codes.** *The Bell System Technical Journal*, v.29, n.2, P.147–160, 1950. [Link para o Artigo.](#)↔

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. **The Elements of Statistical Learning: Data Mining, Inference, and Prediction.** 2nd ed. New York: Springer, 2009. [Link para o Livro.](#)↔

HAYKIN, S. **Redes Neurais: princípios e prática.** Trad. Paulo Martins Engel. 2.ed. Porto Alegre: Bookman, 2000. Consultado em biblioteca em: 14 set. 2023.

HE, F.; LIANG, N. N.; MU, L.; PAN, Q. H.; WANG, J.; REEVERS, M. J.; DUAN, C. Q. **Anthocyanins and their variation in red wines I. Monomeric anthocyanins and their color expression.** Molecules, v. 17, n. 2 p. 1571-1601, 2012. [Link para o Artigo.](#)↔

HINTON, G.E.; SRIVASTAVA, N.; KRIZHEVSKY, A.; SUTSKEVER, I.; SALAKHUTDINOV, R.R. **Improving neural networks by preventing co-adaptation of feature detectors.** Journal of Machine Learning Research v. 15, p. 1929-1958 arXiv, 2012. [Link para o Artigo.](#)↔

ILHA, E. C.; BERTOLDI, F. C.; DOS REIS, V. D. A.; SANT'ANNA, E. **Rendimento e eficiência da fermentação alcoólica na produção de hidromel.** Embrapa Pantanal-Boletim de Pesquisa e Desenvolvimento, 2008. [Link para o Documento.](#)↔

IRIZAWA, I. A.; MARTINS, T. D.; VEGGI, P. C. **Modelagem da cinética de extração supercrítica de artemisinina via redes neurais artificiais.** REVISTA DE ENGENHARIA E TECNOLOGIA, v.13, n.2, p.112-124, 2021. [Link para o Artigo.](#)↔

IVORRA, C.; PÉREZ-ORTÍN, J. E.; DEL OLMO, M. **An inverse correlation between stress resistance and stuck fermentations in wine yeasts. A molecular study.** Biotechnol Bioeng, v.64, n.6, p.698-708, 1999. [Link para o Artigo.](#)↔

- KARHE, R. R.; KALE, S. N. **Classification of Cardiac Arrhythmias Using Feed Forward Neural Network.** Helix, v.10 n.5, p.15-20, 2020. [Link para o Artigo.](#) ↵
- KAVEH, M.; MESGARI, M. S. **Application of Meta-Heuristic Algorithms for Training Neural Networks and Deep Learning Architectures: A Comprehensive Review.** Springer - Neural Processing Letters, v.55, p.4519-4622, 2022. [Link para o Artigo.](#) ↵
- KAWA-RYDLEWSKA, J.; KUCHARSKA, A. Z.; WASZKIW, M. D.; SROKA, P. **How Do Different Ingredients and Additives Affect the Production Steps and the Bioactive Potential of Mead?** National library of Medicine - Food Technol Biotechnol, v.61, n.2, p.179-190, 2023. [Link para o Artigo.](#) ↵
- KINGMA, D. P., BA, J. Adam: **A Method for Stochastic Optimization.** *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*. v.9, 2015. [Link para o Artigo.](#) ↵
- LECUN, Y.; BENGIO, Y.; HINTON, G. **Deep learning.** Nature, v.521, n.7553, p.436-444, 2015. [Link para o Artigo.](#) ↵
- LIMA, S. S. **Modelagem e simulação da produção de hidromel em biorreator em batelada utilizando a levedura *Saccharomyces Cerevisiae*.** 59f. Trabalho de Conclusão de Curso – Monografia), Curso de Engenharia de Biotecnologia e Bioprocessos, Centro de Desenvolvimento Sustentável do Semiárido, Universidade Federal de Campina Grande, Sumé – Paraíba – Brasil, 2022. p.12-25, 2022. [Link para a Dissertação.](#) ↵
- MACKAY, D. J. C. **Bayesian Interpolation, Neural Compute.** v. 4, p. 415–447, 1992. [Link para o Artigo.](#) ↵
- MAICAS, S.; MATEO, J. J. J. **The Life of Non-*Saccharomyces* Yeasts in Wine.** Preprints – microorganisms, v.11, p.1178, 2023. [Link para o Artigo.](#) ↵
- MAPA. MINISTÉRIO DA AGRICULTURA, PECUÁRIA E ABASTECIMENTO. Instrução Normativa nº34. 2012. [Link para a Norma.](#) ↵
- MARQUARDT, D. W. **An Algorithm for Least-Squares Estimation of Nonlinear Parameters,** J. Soc. Ind. Appl. Math. v. 11, n. 2, p. 431–441, 1963. [Link para o Artigo.](#) ↵
- MARTINS, T. D. **Predição da recorrência de tromboembolismo venoso via Redes Neurais Artificiais.** 2018. 128 p. Tese (doutorado) - Universidade Estadual de Campinas, Faculdade de Engenharia Química, Campinas, SP. [Link para a Tese.](#) ↵
- MATTIETTO, R. A.; LIMA, F.C. C.; VENTURIERI, G. C.; ARAÚJO, A. A. **Tecnologia para obtenção artesanal de Hidromel do tipo doce.** Embrapa. Comunicado Técnico 170, p.1-5, 2006. [Link para o Comunicado Técnico.](#) ↵
- MAZZA, G.; FRANCIS, F. J. **Anthocyanins in grapes and grape products.** Critical Reviews in Food Science and Nutrition, v.35, p.341–371, 2005. [Link para o Artigo.](#) ↵

MCCULLOCH, W., PITTS, W. **A logical calculus of the ideas immanent in nervous activity.** Bulletin of Math. Bio., n.5, p.115-133, 1943. [Link para o Livro.](#)↔

MELO, E. B.; OLIVEIRA, E. T.; MARTINS, T. D. **A neural network correlation for molar density and specific heat of water: Predictions at pressures up to 100 MPa,** Fluid Phase Equilibria, v.506, p.112411, 2020. [Link para o Artigo.](#)↔

MELO, H. F. D. **Resposta ao estresse ácido em leveduras da fermentação alcoólica industrial.** 110p. Tese (Doutorado) - Universidade Federal de Pernambuco, Recife/PE, 2006. [Link para a Dissertação.](#)↔

MORAES, G. L. de; LIMA, V. U. B. C. de. **Ragnarök: uma análise da Völuspá e sua influência nas sociedades escandinavas da era viking.** Ed. 5, v.2, p.27-48, 2023. [Link para o Artigo.](#)↔

MICHAELIDES, S.; PATTICHIS, C.; KLEOVOULOU, G. **Classification of rainfall variability by using artificial neural networks.** INTERNATIONAL JOURNAL OF CLIMATOLOGY Int. J. Climatol. v.21, p.1401-1414, 2001. [Link para o Artigo.](#)↔

MILESKI, J. P. F. **Produção e caracterização de hidromel utilizando diferentes cepas de leveduras *Saccharomyces*.** Dissertação (Mestrado em Tecnologia de Alimentos) - Universidade Tecnológica Federal do Paraná, Londrina, 2016. [Link para a Dissertação.](#)↔

MILLER, G. L. **Use of dinitrosalicylic acid reagent for determination of reducing sugar.** Analytical Chemistry, v.31, n.3, p.426-428, 1959. [Link para o Artigo.](#)↔

MØLLER, M. F. **A scaled conjugate gradient algorithm for fast supervised learning.** Science Direct: Neural Networks, v.6, n.4, p. 525-533, 1993. [Link para o Artigo.](#)↔

NAKADA, P. E.; CACIATORI, R.; PANDOLFI, J. R. **O renascimento do hidromel e sua relação com a cultura geek.** Revista de Bebidas e Cultura, v.5, n.2, p.45-60, 2020.

NAVRÁTIL, M.; ŠTURDÍK, E.; GEMEINER, P. **Batch and continuous mead production with pectate immobilised, ethanol-tolerant yeast.** Biotechnology Letters, v.23, p.977-982, 2001. [Link para o Artigo.](#)↔

NELSON, D. L.; COX, M. M. **Princípios de Bioquímica de Lehninger.** 6. ed. Porto Alegre: Armed, Cap.14-14.1, p.543-557, ISBN: 9788582710722, 2014.

NYHAN, L.; SAHIN, A. W.; ARENDT, E. K. **Co-fermentation of non-*Saccharomyces* yeasts with *Lactiplantibacillus plantarum* FST 1.7 for the production of non-alcoholic beer,** European Food Research and Technology, v.249, n.1, p.167-181, 2022. [Link para o Artigo.](#)↔

OGANESYANTS, L. A.; PANASYUK, A. L.; KUZMINA, E. I.; SVIRIDOV, D. A. **Modern analysis methods used in order to establish the geographic origin of food products.** Published as a conference paper at ICLR, arXiv:1412.6980, v.9, 2015. [Link para o Artigo.](#)

ORTIZ-SÁNCHEZ, M., INOCENCIO-GARCÍA, P.-J., ALZATE-RAMÍREZ, A. F., & ÁLZATE, C. A. C. **Potential and Restrictions of Food-Waste Valorization through Fermentation Processes**. Alessia Tropea. v.9, n.3, p.274, 2023. [Link para o Artigo.](#)↔

PATI, S.; DE, S.; CHOWDHURY, R. **Exploring the hybrid route of bio-ethanol production via biomass co-gasification and syngas fermentation from wheat straw and sugarcane bagasse: Model development and multi-objective optimization**. Science Direct: Journal of Cleaner Production, v.395, p.136441, 2023. ISSN 0959-6526. [Link para o Artigo.](#)↔

PEREIRA, A. P. **Caracterização de mel com vista à produção de hidromel**. Instituto Politécnico de Bragança, Escola Superior Agrária, Dissertação de Mestrado em Qualidade e Segurança Alimentar, 2008. [Link para a Dissertação.](#)↔

POWELL, M. J. D. **Restart procedures for the conjugate gradient method. Mathematical Programming**, Springer - Mathematical Programming, v.12, n.1, p.241-254, 1977. [Link para o Artigo.](#)↔

RADKE, E. **Desenvolvimento de modelos híbridos-neurais para fermentação alcoólica e estudo de técnicas de otimização do processo**, p133. Dissertação (mestrado) – Universidade Estadual de Campinas, Faculdade de Engenharia Química, Campinas, SP, 2002. [Link para a Dissertação.](#)↔

RIBÉREAU-GAYON, J.; PEYNAUD, E.; RIBÉREAU-GAYON, P.; SUDRAUD, P. **Effect of three processes of vinification on chemical composition and quality of Cabernet Franc wine**. Sciences et techniques du vin. Paris: Dunod, v.3, p.719, 1976.

RIZZON, L. A.; MIELE, A.; MENEGUZZO, J.; ZANUZ, M. C. **Efeito de três processos de vinificação sobre a composição química e a qualidade do vinho Cabernet Franc**. Pesq. agropec. bras., Brasília, v.34, n.7, p.1285-1293, 1999. [Link para o Artigo.](#)↔

RIEDMILLER, M. & BRAUN, H. **A direct adaptive method for faster backpropagation learning: The RPROP algorithm**. In IEEE International Conference on Neural Networks, vol. 1, p.586-591, 1993. [Link para o Artigo.](#)↔

RIEDMILLER, M. **Advanced supervised learning in multi-layer perceptrons — From backpropagation to adaptive learning algorithms**. Computer Standards & Interfaces, v.16, n.3, p.265-278, 1994 [Link para o Artigo.](#)↔

RIPLEY, B.; RIPLEY, R. **Neural networks as statistical methods in survival analysis in Clinical applications of artificial neural networks**, edited by DYBOWSKI, R. and GANT, V., published by the press syndicate of the university of Cambridge, v.11, p.237-255, 2001. [Link para o Artigo.](#)↔

SANTOS, B.; SIMIQUELI, A.; PONEZI, A.; PASTORE, G.; FILETI, A. **Monitoring of biosurfactant production by *bacillus subtilis* using beet peel as culture medium via the development of a neural soft sensor in an electronic spreadsheet**. Brazilian Journal of Chemical Engineering, v.35, p.1355-1367, 2018. [Link para o Artigo.](#)↔

SANTOS, I.; BOSMAN, G.; DU TOIT, W.; ALEXANDRE-TUDO, J. L. **The use of non-invasive fluorescence spectroscopy to quantify phenolic content under red wine real-time fermentation conditions.** Science Direct: Food Control, v.147, p.109616, 2023. [Link para o Artigo.](#)↔

SCHMIDHUBER, J. **Deep learning in neural networks: An overview.** Science Direct: Neural Networks, v.61, p.85-117, 2015. [Link para o Artigo.](#)↔

SHMUELI, G. **To Explain or to Predict?** Statistical Science, v.25, n.3, 2010. [Link para o Artigo.](#)↔

SHERALI H. D., ULULAR, O. **Conjugate gradient methods using quasi-Newton updates with inexact line searches,** Journal of Mathematical Analysis and Applications, v.150, n.2, p.359-377, ISSN 0022-247X, 1990. [Link para o Artigo.](#)↔

SILVA, I. N; SPATTI, D. H. & FLAUZINO, R. A. **Redes neurais artificiais para engenharia e ciências aplicadas.** 2^a ed. São Paulo: Artliber Editora. 2016. Consultado em biblioteca em: 15 set. 2023.

SROKA, P.; TUSZYŃSKI, T. **Changes in organic acid contents during mead wort fermentation.** Science Direct: Food Chemistry - FOOD CHEM, v.104, p.1250-1257, 2007. [Link para o Artigo.](#)↔

SPERDUTI, A. **Neural Networks for Adaptive Processing of Structured Data.** In: DORFFNER, G., BISCHOF, H., HORNIK, K. (eds) Artificial Neural Networks, Springer, Berlin, Heidelberg - ICANN 2001. Lecture Notes in Computer Science, v.2130. p.5-12, 2001. [Link para o Artigo.](#)↔

TEKIN, S.; CAN, T. **Slide type landslide susceptibility assessment of the Büyük Menderes watershed using artificial neural network method.** Environmental Science and Pollution Research, Environ, v.29, p.47174–47188, 2021. [Link para o Artigo.](#)↔

VALERA, V. Y.; CODOLO, M. C.; MARTINS, T. D. **Artificial neural network for prediction of SO₂ removal and volumetric mass transfer coefficient in spray tower.** Chemical Engineering Research and Design, v. 170, p. 1-12, 2021. [Link para o Artigo.](#)↔

VASCONCELOS, N. M.; PINTO, G. A. S.; ARAGAO, F. A. S. de. **Determinação de açúcares redutores pelo ácido 3,5-dinitrosalicílico: histórico do desenvolvimento do método e estabelecimento de um protocolo para o laboratório de bioprocessos.** Fortaleza: Embrapa Agroindústria Tropical, 2013. (Embrapa Agroindústria Tropical. Boletim de Pesquisa e Desenvolvimento, v.87). [Link para o Boletim.](#)↔

WILLIAMS, C. K. I., RASMUSSEN, C. E. **Gaussian Processes for Machine Learning.** MIT Press. ISBN: 9780262182539, 2006. [Link para o Artigo.](#)↔

WILLMOTT, C. J., MATSUURA, K. **Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance.** Climate Research, v.30, n.1, p79-82, 2005. [Link para o Artigo.](#)↔

ZHANG, Q.; REID, J. F.; BRUCE LITCHFIELD, J.; REN, J.; CHANG, S. W. **A prototype neural network supervised control system for *Bacillus thuringiensis* fermentations.** Biotechnology and bioengineering. v.43, n.6, p.483-489, 1994. [Link para o Artigo.](#)↔

ZHANG, Y.; AI, Y. Q.; WU, Q.; LI, C. F.; CHEN, W. X. **Optimization of Fermentation Conditions of Mead by Response Surface Methodology.** Advanced Materials Research, v.881-883, p.789-792, 2014. [Link para o Artigo.](#)↔

ZHOU, T.; ZHOU, T.; REJI, R.; KAIRON, R. S.; CHIAM, K. H. **A review of algorithmic approaches for cell culture media optimization.** Frontiers in Bioengineering and Biotechnology, Frontiers in Bioengineering and Biotechnology, v.11, 2023. [Link para o Artigo.](#)↔

APÊNDICE A - Resumo Gráfico

