



**UNIVERSIDADE FEDERAL DO CEARÁ**  
**CENTRO DE TECNOLOGIA**  
**DEPARTAMENTO DE ENGENHARIA QUÍMICA**  
**CURSO DE ENGENHARIA QUÍMICA**

**DAYANE WENDY DUARTE DE MELO**

**PREDIÇÃO DAS VELOCIDADES ESPECÍFICAS DA PRODUÇÃO DE XILITOL A  
PARTIR DO BAGAÇO DE CAJU POR *KLUYVEROMYCES MARXIANUS* CCA 510,  
UTILIZANDO REDES NEURAS ARTIFICIAIS**

**FORTALEZA**

**2023**

DAYANE WENDY DUARTE DE MELO

PREDIÇÃO DAS VELOCIDADES ESPECÍFICAS DA PRODUÇÃO DE XILITOL A  
PARTIR DO BAGAÇO DE CAJU POR *KLUYVEROMYCES MARXIANUS* CCA 510,  
UTILIZANDO REDES NEURAS ARTIFICIAIS

Trabalho de Conclusão de Curso apresentado ao  
Curso de Graduação em Engenharia Química da  
Universidade Federal do Ceará, como requisito  
parcial à obtenção do grau de bacharel em  
Engenharia Química.

Orientadora: Profa. Dra. Andréa da Silva  
Pereira.

FORTALEZA

2023

Dados Internacionais de Catalogação na Publicação  
Universidade Federal do Ceará  
Sistema de Bibliotecas  
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

---

M485p Melo, Dayane Wendy Duarte de.  
Predição das velocidades específicas da produção de xilitol a partir do bagaço de caju por *kluveromyces marxianus* CCA 510, utilizando redes neurais artificiais / Dayane Wendy Duarte de Melo. – 2023.  
58 f. : il. color.

Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, Centro de Tecnologia,  
Curso de Engenharia Química, Fortaleza, 2023.  
Orientação: Profa. Dra. Andréa da Silva Pereira.

1. Velocidade específica. 2. Xilitol. 3. Redes neurais artificiais. I. Título.

CDD 660

---

DAYANE WENDY DUARTE DE MELO

PREDIÇÃO DAS VELOCIDADES ESPECÍFICAS DA PRODUÇÃO DE XILITOL A  
PARTIR DO BAGAÇO DE CAJU POR *KLUYVEROMYCES MARXIANUS* CCA 510,  
UTILIZANDO REDES NEURAS ARTIFICIAIS

Trabalho de Conclusão de Curso apresentado ao  
Curso de Graduação em Engenharia Química da  
Universidade Federal do Ceará, como requisito  
parcial à obtenção do título de bacharel em  
Engenharia Química.

Aprovado em: \_\_\_ / \_\_\_ / \_\_\_\_.

BANCA EXAMINADORA

---

Profª. Dra. Andréa da Silva Pereira (Orientadora)  
Universidade Federal do Ceará (UFC)

---

Prof. Dr. Daniel Vasconcelos Gonçalves  
Universidade Federal do Ceará (UFC)

---

Me. Matheus de Oliveira Barros  
Universidade Federal do Ceará (UFC)

A Deus, à Nossa Senhora e à minha família:  
Raimundo, Juraci e Marcos.

## AGRADECIMENTOS

A Deus e à Nossa Senhora, sempre presentes em minha vida.

Aos meus pais e irmão, Raimundo e Juraci e Marcos, pelo suporte e sabedoria durante todos esses anos.

Ao meu namorado, Lucas Rodrigues, pelo apoio e motivação em tantos dias apenas estudando comigo e abdicando de fins de semana em prol do nosso futuro. Grata imensamente por tudo.

Aos meus colegas de curso que se tornaram grandes amigos, por sempre acreditarem no nosso sucesso e dividirem comigo inúmeras risadas e noites acordados: Gabriella Freitas, Mateus Silas, Brenno Misquista, Catarine Alves, Larissa Castro, Henrique Aragão, Yanka Monteiro, Rebecca Holanda, Nara Ádyna, Diony Gomes, Amanda Bezerra, Erika Maria e Mauro Filho.

Às professoras doutoras Andréa Pereira e Maria Valdez, pela oportunidade, disponibilidade e pelos ensinamentos transmitidos durante esse tempo.

A todos os professores das cadeiras de Engenharia, que compartilharam seus conhecimentos e experiências e ensinaram mais do que as matérias do curso na minha trajetória.

À Universidade Federal do Ceará que proporcionou todas as atividades de extensão que participei - Centro Acadêmico, Ciclo Jr, Escola Piloto de Engenharia Química. Além da EMBRAPA, onde fui orientada pelo Matheus Oliveira no estágio, minha admiração eterna. Uma menção com carinho para quem me deu a mão e acreditou no meu potencial, a professora doutora Ana Iraidy Santa Brígida.

E a todas as pessoas que, direta ou indiretamente, contribuíram para a minha formação.

“Nada na vida deve ser temido, somente compreendido. Agora é tempo de entender mais e temer menos.” – Marie Curie

## RESUMO

A produção agrícola nordestina de caju é a maior nacionalmente. Do processo, resta uma considerável quantidade de bagaço de caju, biomassa que serve de material para geração de diversos açúcares, destacando-se glicose e xilose. Ambos, processados por via biotecnológica com micro-organismos, geram subprodutos como o etanol e xilitol, de grande valor comercial. O objetivo principal deste trabalho é propor um modelo matemático utilizando Rede Neural Artificial (RNA) aplicada no processo de obtenção biotecnológica de xilitol originado do hidrolisado de bagaço de caju a partir da levedura *Kluyveromyces marxianus* CCA 510, possibilitando verificar o comportamento dos perfis de concentração de biomassa, xilose, glicose, etanol e xilitol e as velocidades específicas de cada um. Para isso, foram utilizados os dados experimentais da tese “Produção biotecnológica de xilitol a partir de hidrolisado de bagaço de caju” (ALBUQUERQUE, 2014). Foi realizado um tratamento dos dados a partir de regressões e interpolações, buscando suavizar os ruídos das medidas experimentais e aumentar a quantidade de dados para o treinamento de redes neurais artificiais (RNAs). Neste trabalho, as velocidades específicas são previstas pelas RNAs como função da concentração instantânea de célula e substrato para cada tipo de hidrolisado hemicelulósico de bagaço de caju (HBC). Os resultados das simulações foram avaliados estatisticamente pelo erro médio absoluto (MAE), além da validação frente um novo ensaio. De acordo com os resultados, os modelos utilizados se mostraram eficazes para prever as velocidades específicas de acordo com o comportamento dos perfis de concentração avaliados para cada condição testada, com  $R^2$  para treino e validação variando de 0,99 a 0,94 para biomassa, 0,99 a 0,90 para xilose, 0,99 a 0,93 para glicose, 0,96 a 0,90 para xilitol e 0,98 a 0,94 para etanol e sem variação significativa entre as condições analisadas, indicando boa qualidade nos ajustes realizados e adequação dos modelos aos dados observados. Bons valores de acurácia foram alcançados onde se obteve um MAE inferior a 1%.

**Palavras-chave:** Velocidade específica. Xilitol. Redes neurais artificiais.

## ABSTRACT

The northeastern region of Brazil has the highest national production of cashews. From this process, a substantial amount of cashew bagasse remains, which serves as a material for generating various sugars, notably glucose and xylose. Both, processed through biotechnological methods with microorganisms, generate by-products such as ethanol and xylitol, which have significant commercial value. The main objective of this study is to propose a mathematical model using Artificial Neural Networks (ANN) applied to the biotechnological process of obtaining xylitol from cashew bagasse hydrolysate using the yeast *Kluyveromyces marxianus* CCA 510. This allows for examining the behavior of concentration profiles of biomass, xylose, glucose, ethanol, and xylitol, as well as the specific rates of each. Experimental data from the thesis "Biotechnological Production of Xylitol from Cashew Bagasse Hydrolysate" (ALBUQUERQUE, 2014), were used for this purpose. The data underwent treatment through regressions and interpolations to smooth experimental measurement noise and increase the dataset for training Artificial Neural Networks (ANNs). In this study, specific rates are predicted by ANNs as a function of the instantaneous concentration of cells and substrate for each type of cashew bagasse hemicellulosic hydrolysate (HBC). Simulation results were statistically evaluated using the Mean Absolute Error (MAE), in addition to validation against a new experiment. According to the results, the models used proved effective in predicting specific rates according to the behavior of the evaluated concentration profiles for each tested condition, with  $R^2$  for training and validation from 0.99 to 0.94 for biomass, 0.99 to 0.90 for xylose, 0.99 to 0.93 for glucose, 0.96 to 0.90 for xylitol and 0.98 to 0.94 for ethanol, showing no significant variation between the analyzed conditions. This indicates good quality in the model adjustments and the suitability of the models to the observed data. High accuracy values were achieved where MAE was less than 1%.

**Keywords:** Specific rate. Xylitol. Artificial Neural Networks.

## LISTA DE FIGURAS

Figura 1	– Pré-tratamento de biomassa lignocelulósica .....	16
Figura 2	– Principais produtos formados a partir do material lignocelulósico .....	19
Figura 3	– Esquema geral da fermentação de D-xilose .....	20
Figura 4	– Classes de modelos .....	22
Figura 5	– Processamento do sinal em um neurônio de uma RNA .....	23
Figura 6	– Redes feedforward .....	24
Figura 7	– Arquitetura de RNA do tipo retropropagação .....	26
Figura 8	– Fluxograma dos experimentos realizados visando avaliar a influência de destoxificação do HBA na produção do xilitol .....	30
Figura 9	– Fermentação, por <i>Kluyveromyces marxianus</i> CCA 510, de meios formados por HBC com diferentes tratamentos .....	32
Figura 10	– Simulação de RNAs: comparativo entre as previsões das RNAs e os dados pseudo-experimentais .....	44

## LISTA DE TABELAS E QUADROS

Tabela 1 – Composição dos meios formulados a partir de hidrolisado do bagaço de caju .....	31
Tabela 2 – Matriz das condições operacionais iniciais para Biomassa .....	34
Tabela 3 – Interpolação utilizando os dados experimentais da biomassa .....	37
Tabela 4 – Limites de interpolação .....	39
Tabela 5 – R <sup>2</sup> para validação dos ajustes de curva de concentração dos compostos .....	40
Tabela 6 – R <sup>2</sup> dos modelos preditivos usando RNAs .....	44
Tabela 7 – Validação estatística – MAE .....	45
Quadro 1 – Ajustes de regressão não linear aplicados em Origin .....	36
Quadro 2 – Arquitetura das RNAs .....	42

## LISTA DE ABREVIATURAS E SIGLAS

GPBIO	Grupo de Pesquisa e Desenvolvimento de Processos Biotecnológicos
HBC	Hidrolisado hemicelulósico de bagaço de caju
HMF	5-Hidroximetilfurfural
LACAM	Laboratório de Microbiologia Agrícola e Molecular
MAE	Erro Absoluto Médio
MAPE	Erro Percentual Absoluto Médio
MSE	Erro Quadrático Médio
RNA	Rede Neural Artificial
UFC	Universidade Federal do Ceará
UFSCar	Universidade Federal de São Carlos
XDH	Xilitol Desidrogenase
XR	Xilose Redutase

## SUMÁRIO

1	<b>INTRODUÇÃO</b> .....	14
2	<b>REFERENCIAL TEÓRICO</b> .....	15
2.1	<b>Biomassa lignocelulósica: Bagaço de caju</b> .....	15
2.2	<b>Pré-tratamento com hidrólise ácida</b> .....	16
2.3	<b>Fatores que influenciam a produção biotecnológica</b> .....	17
2.3.1	<i>Concentração inicial de substrato e micro-organismo</i> .....	17
2.3.2	<i>Tipo de substrato</i> .....	18
2.3.3	<i>Inibidores</i> .....	18
2.4	<b>Produção biotecnológica de xilitol</b> .....	19
3	<b>MODELO MATEMÁTICO DE PROCESSOS FERMENTATIVOS</b> .....	21
3.1	<b>Redes neurais artificiais</b> .....	22
3.1.1	<i>Estrutura da RNA</i> .....	24
3.1.2	<i>Algoritmo de aprendizagem</i> .....	25
3.1.3	<i>Função de ativação</i> .....	26
3.1.4	<i>Arquitetura da RNA</i> .....	26
3.2	<b>Cinética do processo</b> .....	27
3.3	<b>Discriminação entre modelos</b> .....	28
4	<b>MATERIAIS E MÉTODOS</b> .....	30
4.1	<b>Metodologia para obtenção dos dados experimentais</b> .....	30
4.1.1	<i>Base de dados</i> .....	33
4.2	<b>Tratamento dos dados experimentais</b> .....	35
4.2.1	<i>Regressão não linear e Interpolação</i> .....	35
4.2.2	<i>Normalização dos dados pseudo-experimentais</i> .....	40
4.2.3	<i>Metodologia</i> .....	40
5	<b>RESULTADOS E DISCUSSÃO</b> .....	44
5.1	<b>Resultados</b> .....	44
5.2	<b>Perspectivas para trabalhos futuros</b> .....	47
6	<b>CONCLUSÃO</b> .....	48
	<b>REFERÊNCIAS</b> .....	49
	<b>APÊNDICE A – CONDIÇÕES OPERACIONAIS INICIAIS PARA XILOSE, GLICOSE, XILITOL E ETANOL</b> .....	53

<b>APÊNDICE B – DADOS INTERPOLADOS: PSEUDO- EXPERIMENTAIS .....</b>	<b>54</b>
<b>APÊNDICE C – ERROS POR ÉPOCA E BISSETRIZES DA PREDIÇÃO COM REDES NEURAS PARA CADA SUBSTÂNCIA .....</b>	<b>55</b>
<b>APÊNDICE D – GRÁFICOS DE REGRESSÃO NÃO-LINEAR PARA CADA SUBSTÂNCIA .....</b>	<b>57</b>

## 1 INTRODUÇÃO

A região nordeste do Brasil abrange aproximadamente 700 mil hectares dedicados à cultura do caju, sendo o estado do Ceará o principal produtor. Apesar do avanço na industrialização do pedúnculo do caju, com o suco integral como produto principal e uma produção anual em torno de 70 mil toneladas, a utilização efetiva do pedúnculo na cultura do caju ainda é relativamente baixa (SILVA NETO, 2000).

Assim, por conta da quantidade significativa de resíduos gerados e para encontrar uma proposta que integre a preservação ambiental como um tópico e ainda traga benefícios econômicos ao processo produtivo da castanha e do caju, tem-se observado um interesse cada vez maior na exploração de fontes alternativas de energia para a produção de bioprodutos de alto valor agregado, tais como o etanol e xilitol (ROCHA, 2010).

Nesse sentido, a utilização da modelagem matemática no processo fermentativo do bagaço do caju, possibilita a avaliação de condições não previamente testadas experimentalmente. Isso permite a previsão do comportamento dos perfis de concentração inicial de células, substratos e produtos diante de variações nas condições da reação. Como resultado, é possível identificar um ponto ótimo de operação do sistema por meio de simulações (NIELSEN; VILLADSEN, 1994).

O presente trabalho tem como objetivo geral apresentar uma proposta de utilização de uma Rede Neural Artificial (RNA) que sirva como modelo matemático satisfatório para descrever a predição das velocidades específicas de crescimento celular, consumo de substrato e geração de produtos para biomassa, xilose, glicose, etanol e xilitol. Todos foram listados da base de dados experimentais oriundos da dissertação de Tiago Lima de Albuquerque: Produção biotecnológica de xilitol a partir de hidrolisado de bagaço de caju (2014), na qual baseia-se a análise preditiva demonstrada nos capítulos seguintes. Acrescento ao final, sugestões para trabalhos futuros, com passos a serem seguidos e estudados a partir dos dados obtidos nas conclusões deste trabalho.

## 2 REFERENCIAL TEÓRICO

### 2.1 Biomassa lignocelulósica: Bagaço de caju

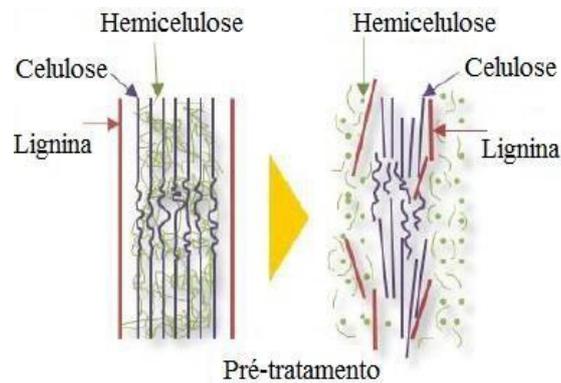
Os materiais lignocelulósicos representam a fração mais significativa da biomassa vegetal, sendo uma fonte preponderante de compostos orgânicos e matéria-prima renovável na biosfera, caracterizando-se por seu custo acessível e baixo impacto ambiental (MULAKHUDAIR, 2016). Dentre esses materiais, destaca-se o bagaço de caju, subproduto da extração do suco do pseudofruto, correspondendo a aproximadamente 20% do peso total do pedúnculo e figurando como uma das principais fontes de resíduos na indústria agrônômica (SANTOS et al., 2007).

O bagaço de caju apresenta elevado teor de açúcares fermentáveis em sua composição, originando-se de uma cultura adaptada ao clima do nordeste brasileiro e caracterizada por baixas exigências (BRIONES, SERRANO e LABIDI, 2012). Além disso, destaca-se por sua relevância na produção de diversos produtos de alto valor de mercado, como etanol e xilitol.

Essa biomassa é composta por três tipos de polímeros - celulose, hemicelulose e lignina - os quais estão intrincadamente entrelaçados e conectados quimicamente por meio de ligações covalentes e de hidrogênio (PÉREZ et al., 2002). No bagaço de caju, esses polímeros são os principais constituintes, com proporções relatadas por Rocha (2010) e Correia et al. (2013) de 20,54% e 20,56% de celulose, 16,33% e 10,17% de hemicelulose, e 33,62% e 35,26% de lignina, respectivamente.

A transformação do bagaço de caju para a produção de substâncias de valor agregado está condicionada à sua hidrólise. Apenas por meio desse processo, os açúcares presentes, como glicose e xilose, podem ser liberados e, posteriormente, convertidos por micro-organismos específicos por meio de fermentação. Um exemplo de aplicação biotecnológica é a conversão da xilose em xilitol. Por sua vez, o sucesso dessa hidrólise dependerá de um pré-tratamento adequado e eficaz.

Figura 1 – Pré-tratamento de biomassa lignocelulósica



Fonte: Adaptada de Mosier et al. (2005).

## 2.2 Pré-tratamento com hidrólise ácida

Empregam-se etapas distintas de preparação para processar materiais lignocelulósicos: o pré-tratamento, que visa a deslignificação para expor a celulose e hemicelulose de maneira mais eficaz; a hidrólise da celulose e hemicelulose, com o propósito de gerar açúcares fermentáveis como xilose e glicose; e, por último, a fermentação desses açúcares hidrolisados (DEMIRBAS, 2008). Ou seja, a realização de um pré-tratamento e hidrólise, a princípio, têm a finalidade de solubilizar os açúcares monoméricos presentes na biomassa lignocelulósica, para que em uma etapa subsequente, o licor produzido venha a ser utilizado como uma fonte de carbono por diferentes microrganismos. Dessa forma, permitindo que matérias-primas lignocelulósicas possam ser susceptíveis a utilização biotecnológica.

O pré-tratamento ácido é considerado um dos mais importantes, gerando uma grande quantidade de açúcares a partir de biomassas distintas. Além de garantir a formação e preservação de pentoses e hexoses na biomassa, o pré-tratamento deve minimizar a geração de subprodutos que possam causar inibição dos micro-organismos durante o processo de fermentação (MOSIER, 2005). A hidrólise com ácido diluído é uma das mais estudadas por ser eficaz e barata (RAFIQUL e SAKINAH, 2012).

As temperaturas durante o tratamento podem variar de 100 °C até 160 °C, ocorrendo dentro de um reator e usando ácidos como o ácido sulfúrico, concentrado ou diluído, mas também são empregados ácidos nítrico, fosfórico e clorídrico.

Existem alguns pontos desfavoráveis associados a essa tecnologia, como a necessidade de neutralização prévia à fermentação para evitar a inibição, o risco potencial envolvido na operação com substâncias perigosas e os custos significativos de instalação para

garantir que os materiais resistam às condições agressivas do fluido reagente (SARKAR, 2012). Uma desvantagem adicional desse método está relacionada à produção de inibidores indesejáveis para a fermentação microbiológica, como ácido acético, furfural e 5-hidroximetilfurfural. Esses inibidores resultam da degradação de açúcares pela ação do ácido, especialmente xilose e glicose. A formação desses compostos está associada a fatores como concentrações elevadas de ácidos e altas temperaturas (MOOD, 2013).

## **2.3 Fatores que influenciam a produção biotecnológica**

A xilose destaca-se como um dos principais açúcares resultantes da hidrólise de biomassa lignocelulósica. Juntamente com esse açúcar, obtêm-se outros monômeros, como glicose, arabinose, manose, entre outros. Classificada como o segundo monômero de açúcar mais abundante na biosfera, a xilose é utilizada no processo de conversão para xilitol, seja de maneira industrial ou microbiológica. Conseqüentemente, observou-se um aumento considerável nos estudos voltados para o desenvolvimento da obtenção de xilose a partir de fontes renováveis (ALBUQUERQUE, 2015).

Durante um processo fermentativo existem vários fatores que podem influenciar diretamente o rendimento e a produtividade de xilitol, entre eles podemos destacar a concentração inicial de células do inóculo, tipo de substrato e compostos inibidores como relevantes para entendermos durante este trabalho.

### ***2.3.1 Concentração inicial de substrato e micro-organismo***

Em um processo descontínuo (batelada), a quantidade de biomassa alcançada está diretamente relacionada à concentração inicial do substrato. Essa concentração inicial limita o crescimento e a eficiência do microrganismo na conversão do substrato no produto desejado. O crescimento celular apresenta um comportamento de saturação à medida que a concentração do substrato aumenta, mas, em alguns casos, a concentração de açúcar pode atingir níveis suficientemente altos para ultrapassar a região de saturação, resultando em inibição do crescimento pelo substrato (FAKRUDDIN et al., 2012).

Com o objetivo de aplicação industrial, os microrganismos devem demonstrar alta eficiência na conversão do substrato em produto, evitar a produção de substâncias incompatíveis com o produto, não serem patogênicos e possibilitar a rápida liberação do produto no meio ambiente (LIMA et al., 2001). Nesse contexto, a levedura *K. marxianus* tem

despertado interesse na biotecnologia em comparação com outras espécies de leveduras. Isso se deve à sua ampla diversidade metabólica, capacidade de assimilar diversos açúcares, crescimento mais rápido e amplo espectro de termotolerância (FONSECA et al., 2008; LANE & MORRISEY, 2010; LANE et al., 2011; FERREIRA et al., 2015; NACHAIWIENG et al., 2015; FASOLI et al., 2016). Neste trabalho, os dados de produção de xilitol via microorganismo *K. marxianus* CCA 510 pertencem ao Grupo de Pesquisa e Desenvolvimento de Processos Biotecnológicos (GPBIO) da Universidade Federal do Ceará e a cepa *Kluyveromyces marxianus* CCA 510 é uma levedura oriunda do banco de cultura do Laboratório de Microbiologia Agrícola e Molecular - LACAM, da Universidade Federal de São Carlos - UFSCar, São Carlos, Brasil.

### ***2.3.2 Tipo de substrato***

Os hidrolisados hemicelulósicos, neste trabalho o hidrolisado do bagaço de caju, contêm várias hexoses e pentoses, e apesar de haver variação na composição desses hidrolisados, xilose, glicose e arabinose são os três principais açúcares presentes. Uma estratégia adotada na produção biotecnológica de xilitol envolve o uso de glicose e arabinose para fornecer a energia necessária ao crescimento microbiano, enquanto a xilose é assimilada para a produção de xilitol.

### ***2.3.3 Inibidores***

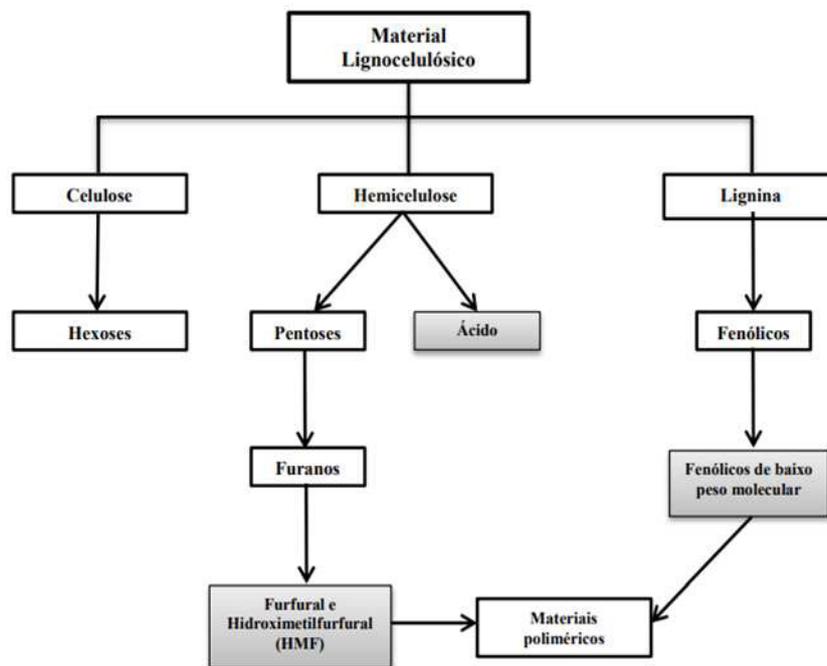
Como é comum a presença de compostos que podem inibir o metabolismo dos microrganismos nos licores provenientes de hidrolisados hemicelulósicos, resultando na morte microbiana ou na redução da produtividade do xilitol. Esses produtos são gerados pela decomposição de açúcares presentes na hemicelulose e de compostos da lignina. A concentração tolerável de inibidores varia de acordo com o microrganismo utilizado e sua adaptação ao meio de fermentação específico (RAO, 2016).

Os principais inibidores incluem o ácido acético, 5-hidroximetilfurfural (HMF), furfural e compostos fenólicos derivados da lignina. A presença desses inibidores, especialmente furfural e HMF nos licores hidrolisados, é resultado da degradação de açúcares como xilose e glicose em meio ácido, como ilustrado na imagem abaixo. Esses dois compostos são formados por meio da desidratação catalítica da xilose e da glicose, respectivamente. Eles causam atrasos no processo fermentativo no início da fermentação, inibem o crescimento da levedura e a síntese de proteínas (DUARTE, 2005).

Os compostos fenólicos, provenientes da lignina, são os principais inibidores e causam danos à integridade das membranas dos microrganismos. Com concentrações superiores a 0,1 g/L, tanto o crescimento celular quanto a produção de xilitol são inibidos (MUSSATTO, 2004; SÁ, 2015).

O ácido acético, subproduto da degradação de grupos acetil na hemicelulose durante o tratamento ácido, difere dos demais inibidores ao interagir com a célula do microrganismo, provocando sua morte instantânea devido a uma alteração no pH interno da célula.

Figura 2 – Principais produtos formados a partir do material lignocelulósico



Fonte: ALBUQUERQUE (2014).

## 2.4 Produção biotecnológica de xilitol

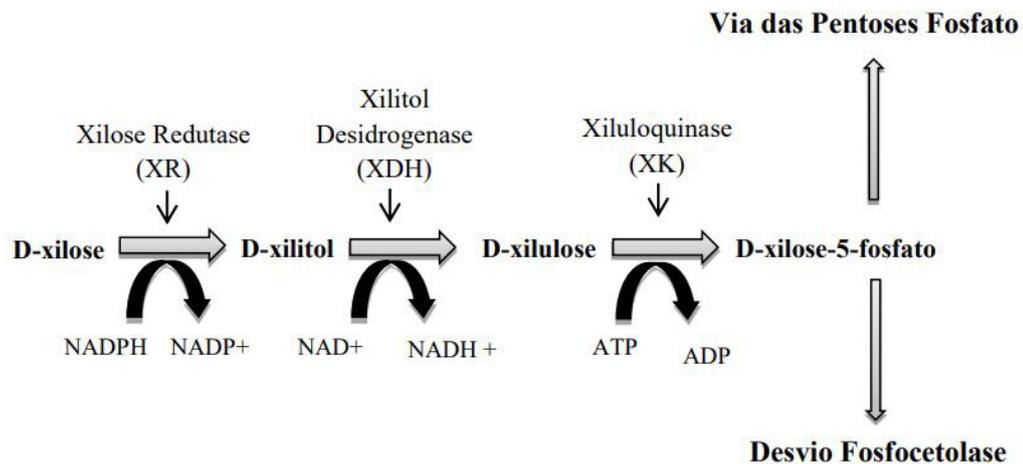
A obtenção de xilitol via biotecnológica, utilizando hidrolisados hemicelulósicos como meio de cultivo, surge como um processo mais vantajoso em comparação com a conversão química da xilose, método tradicional que apresenta rendimentos baixos (entre 50-60%) e custos elevados (PEREIRA et al., 2011). A abordagem biotecnológica oferece benefícios, como a utilização de condições mais suaves de pressão e temperatura (CUNHA et al., 2005). Além disso, ela elimina a necessidade de xilose purificada, uma vez que micro-

organismos específicos conseguem converter xilose em xilitol diretamente a partir do hidrolisado hemicelulósico, atingindo rendimentos de até 78% (SARROUH & SILVA, 2008; MUSSATO & ROBERTO, 2004).

Acredita-se que a adoção dessa abordagem biotecnológica resultará em custos de produção mais baixos em comparação com o método químico convencional, potencialmente ampliando o uso do xilitol e seus benefícios para a saúde humana. Uma vantagem adicional é a geração de efluentes facilmente tratáveis, contribuindo significativamente para a redução da poluição ambiental.

As leveduras produzem xilitol como um intermediário durante o metabolismo da D-xilose, conforme ilustrado na figura abaixo. A xilose redutase (XR) é geralmente uma enzima dependente de NADPH, enquanto a xilitol desidrogenase (XDH) requer NADP<sup>+</sup>. Dessa forma, a conversão do D-xilitol ocorre em duas etapas, uma de redução seguida por outra de oxidação. D-xilulose é inicialmente reduzida a D-xilitol por NADPH e, posteriormente, esse metabólito é oxidado a D-xilulose por NADP<sup>+</sup>. Essas duas reações são consideradas limitantes na fermentação de D-xilose e na produção de xilitol.

Figura 3 - Esquema geral da fermentação de D-xilose



Fonte: Adaptado de Chen et al. (2010).

Em condições microaeróbicas, é comum que as leveduras tenham a capacidade de converter resíduos contendo D-xilose em etanol ou xilitol. A produção desses compostos varia conforme o micro-organismo em questão (YABLOCHKOVA; BOLOTNIKOVA; MIKHAILOVA, 2003a).

### 3 MODELO MATEMÁTICO DE PROCESSOS FERMENTATIVOS

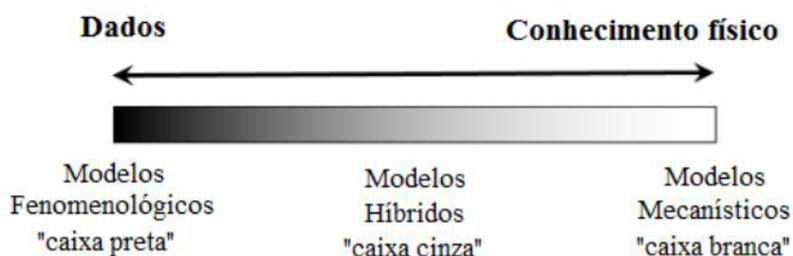
A modelagem matemática de processos fermentativos consiste na tentativa de representar, de forma matemática, os balanços de massa para cada componente da reação, relacionando-os com as complexas transformações bioquímicas que ocorrem no processo e as velocidades com que essas transformações acontecem (BONOMI; SCHMIDELL, 2001). Os processos fermentativos reais apresentam um alto grau de complexidade devido às leis físico-químicas, bioquímicas e genéticas envolvidas. Uma alternativa para lidar com essa complexidade é o uso de modelos baseados na idealidade do sistema, os quais descrevem apenas algumas propriedades do processo (VOLESKY; VOTRUBA, 1992). Segundo Sinclair e Kristiansen (1987), um modelo fermentativo deve contemplar apenas os aspectos relevantes que despertam interesse, permitindo a quantificação da taxa de crescimento celular, do consumo de substrato e da formação de produtos. Dessa forma, o modelo se configura como um conjunto de relações entre as variáveis de interesse em um sistema em estudo, com o propósito, como destacado por Tosetto (2002), de antecipar o comportamento do processo. Isso se torna essencial, uma vez que é impraticável realizar testes experimentais para todas as condições operacionais e escalas possíveis do processo em análise.

Além disso, o modelo busca identificar as condições operacionais economicamente ótimas do processo e estabelecer limites nas condições operacionais. Essa abordagem permite analisar alterações no processo visando à sua otimização. Nesse contexto, é fundamental definir a estratégia de controle e avaliar a sensibilidade do processo diante de perturbações.

De maneira geral, os modelos podem ser classificados em duas categorias: modelos mecanísticos e modelos fenomenológicos (VELTEN; FRIEDE, 2009; PERRY & GREEN, 2008; TOWLER & SINNOTT, 2012; SMITH, 2005). Os modelos mecanísticos, também conhecidos como caixa branca, fundamentam-se no formalismo matemático, gerando, geralmente, equações diferenciais conduzidas pelos princípios fundamentais, como conservação de massa, energia ou quantidade de momento. Eles também incorporam equações de velocidade, que são expressões cinéticas descrevendo a geração ou consumo de espécies dentro do sistema, além de equações termodinâmicas.

Por outro lado, os modelos fenomenológicos, ou caixa-preta, são constituídos apenas com base em dados experimentais, estabelecendo relações com observações do comportamento de dados medidos do sistema a ser modelado. Um exemplo desse modelo fenomenológico é o uso de redes neurais artificiais (RNA) neste trabalho. Existem modelos híbridos (caixa cinza): abordagens que integram as descrições dos modelos anteriores.

Figura 4 – Classes de modelos



Fonte: PEREIRA (2016).

### 3.1 Redes neurais artificiais

As redes neurais artificiais (RNAs) são modelos computacionais fundamentados na estrutura dos neurônios biológicos. Esses modelos, também conhecidos como caixas pretas, são particularmente eficazes para a modelagem de processos devido à sua capacidade de desenvolver rapidamente modelos matemáticos precisos, mesmo em situações em que o conhecimento detalhado do processo é limitado.

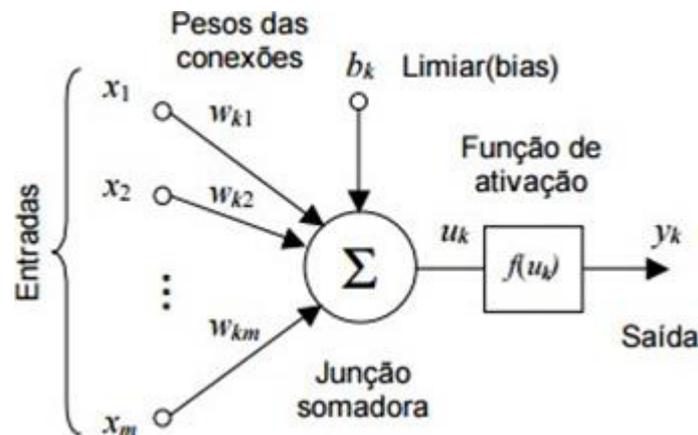
No contexto das RNAs, o tipo mais prevalente de modelo de caixa preta é a própria rede neural artificial. Essas redes são compostas por unidades interconectadas chamadas neurônios artificiais ou nós. Cada neurônio tem uma função específica de entrada/saída, determinada por sua função de transferência, e está interligado a outros neurônios por conexões ou sinapses. Tais redes neurais operam com base em princípios fundamentais, que são:

- **Processamento de Informação:** Neurônios artificiais, também conhecidos como nós, são as unidades básicas responsáveis pelo processamento da informação;
- **Interconexão:** Os neurônios estão interconectados, formando redes neurais que constituem o modelo de caixa preta;
- **Transmissão de Informação:** A transmissão de informações entre neurônios ocorre por meio de conexões ou sinapses;
- **Eficiência Sináptica:** A eficácia de uma sinapse, representada por um peso associado, reflete a informação armazenada pelo neurônio e, por extensão, pela rede neural;
- **Aprendizagem:** O conhecimento é adquirido do ambiente por meio de um processo de aprendizagem. Esse processo ajusta os pesos das conexões em resposta aos estímulos recebidos do ambiente.

Essa abordagem de caixa preta oferece rapidez na obtenção de modelos matemáticos, embora apresente a limitação de não permitir extrapolações. Isso significa que os experimentos devem abranger todo o domínio de aplicação do modelo para garantir sua eficácia. As RNAs, notadamente as redes neurais artificiais, são, portanto, uma ferramenta valiosa na modelagem de processos complexos.

Segue-se abaixo o funcionamento de um neurônio artificial, que pode ser resumido nos seguintes passos: (a) os sinais são recebidos do neurônio anterior; (b) os sinais são multiplicados por pesos correspondentes a cada ligação; (c) a junção somadora soma todos os sinais de entrada ponderados pelos pesos das conexões; (d) a função de ativação é geralmente utilizada com o propósito de limitar a saída do neurônio e introduzir não-linearidade no modelo; (e) o limiar  $b_k$  tem o papel de aumentar ou diminuir a influência do valor da entrada líquida para a ativação do neurônio  $k$ ; (f) os sinais ativos seguem para o neurônio subsequente (PELLICCI, 2001).

Figura 5 – Processamento do sinal em um neurônio de uma RNA



Fonte: PEREIRA (2016)

A Equação (1) expressa matematicamente a saída do neurônio  $k$ , onde  $x_0$  é um sinal de entrada de valor 1 e peso associado  $w_{k0} = b_k$ .

$$y_k = f(u_k) = f\left(\sum_{j=0}^m w_{kj} \cdot x_j\right) = f\left(\sum_{j=1}^m w_{kj} \cdot x_j + b_k\right) \quad (1)$$

O desenvolvimento de um projeto para uma rede neural artificial envolve a definição de três características principais: a estrutura, o algoritmo de aprendizagem e a função de ativação (HAYKIN, 1994).

A estrutura de uma RNA refere-se aos padrões de conectividade entre os neurônios da rede. Esses padrões podem incluir apenas conexões diretas, sempre no sentido da entrada para a saída da rede, ou também conexões diretas e recorrentes, responsáveis pela realimentação de informação.

O algoritmo de aprendizagem é o método utilizado para determinar os parâmetros livres da rede. Isso inclui ajustes nos pesos das conexões, escolha da função de ativação, determinação do número de camadas e do número de neurônios por camada na rede. Existem basicamente três paradigmas de aprendizagem: não-supervisionada, supervisionada e por reforço.

A função de ativação desempenha um papel crucial, representando o efeito da entrada e do estado atual de um neurônio na definição do seu próximo estado. Ela determina como a saída do neurônio é calculada com base nas entradas recebidas e no estado atual do neurônio.

Essas três características são detalhadas nos tópicos seguintes e quando bem definidas, constituem a base para o desenvolvimento eficaz de uma rede neural artificial, permitindo que ela aprenda e se adapte aos padrões presentes nos dados de entrada.

### ***3.1.1 Estrutura da RNA***

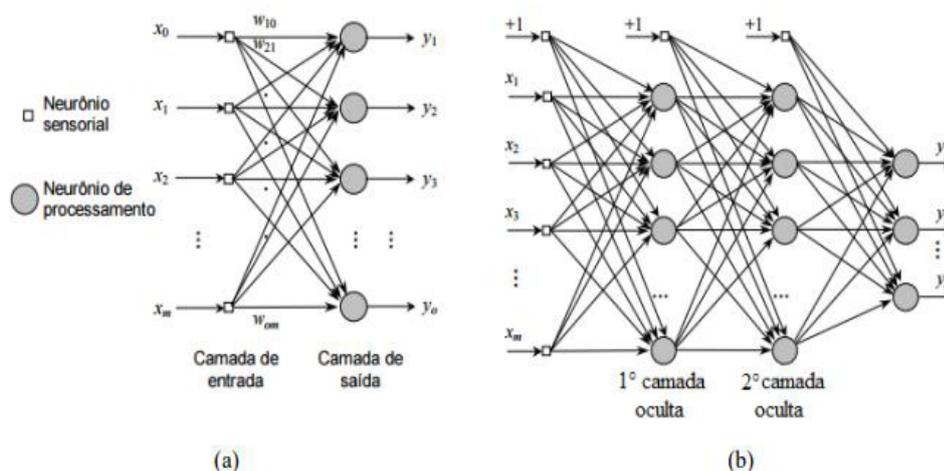
Os métodos de propagação da informação para a topologia dos neurônios são diferenciados entre redes de propagação para frente (feedforward) e redes realimentadas (feedback).

Na modelagem e simulação de processos químicos e bioquímicos, a arquitetura de RNA mais amplamente adotada é a feedforward de múltiplas camadas com treinamento supervisionado por correção do erro, também conhecido como retropropagação - que é capaz de aproximar, com maior ou menor precisão, dependendo do número de neurônios da RNA, qualquer função não-linear (BARRETO, 2002; TONIN, 2005).

Essa configuração é comumente empregada em problemas de regressão de dados, incluindo a previsão de séries temporais.

Nas redes de propagação para frente (feedforward), o fluxo de informação é unidirecional. Essas redes são geralmente, são representadas em camadas.

Figura 6 – Redes feedforward: (a) única camada e (b) múltiplas camadas



Fonte: PEREIRA (2016).

As RNAs podem ser divididas em três tipos de camadas: camada de entrada, camadas intermediárias ou ocultas (podendo não possuir) e camada de saída. A forma pela qual os neurônios estão interconectados está intimamente relacionada ao algoritmo a ser utilizado no seu treinamento.

As principais desvantagens dessa estrutura são: a incerteza se a rede irá parar em um ponto de equilíbrio e o tempo longo necessário ao processamento da rede, podendo levar minutos ou horas.

### 3.1.2 Algoritmo de aprendizagem

Após a determinação do tipo e da topologia da rede, passamos para a fase de treinamento, cujo objetivo é estabelecer os pesos de cada interconexão (ALVES, 2003). O treinamento pode seguir um algoritmo supervisionado, no qual os pesos são ajustados com base na discrepância entre a saída desejada e a saída obtida. Alternativamente, pode adotar um método não supervisionado, dispensando a necessidade de uma saída desejada conhecida, permitindo que a própria rede crie automaticamente um mapa com os dados de entrada para prever o conjunto de saída - como aprendizado auto supervisionado, nesse contexto, a rede se ajusta às regularidades estatísticas presentes nos dados de entrada, criando representações internas que codificam as características dos dados.

Como no presente trabalho a intenção é prever através de RNA os dados de velocidade cinética dos compostos, não podemos extrapolar os dados, assim o treinamento deve

seguir o algoritmo supervisionado que verifica a saída da rede e, se não estiver de acordo com a saída desejada, realiza ajustes nos pesos das conexões visando minimizar essa diferença.

Alguns desafios de engenharia podem ser formulados como problemas de aprendizado supervisionado, abrangendo áreas como classificação e reconhecimento de padrões, predição de séries temporais, identificação de sistemas, controle de processos etc.

### ***3.1.3 Função de ativação***

A função de ativação é uma característica crucial em redes neurais artificiais com múltiplas camadas, desempenhando papéis fundamentais como limitar a saída do neurônio e introduzir não-linearidade ao modelo. A escolha de funções de ativação não lineares e diferenciáveis para as camadas ocultas é essencial para resolver problemas não linearmente separáveis, orientando o ajuste de pesos (HAYKIN, 2001).

### ***3.1.4 Arquitetura da RNA***

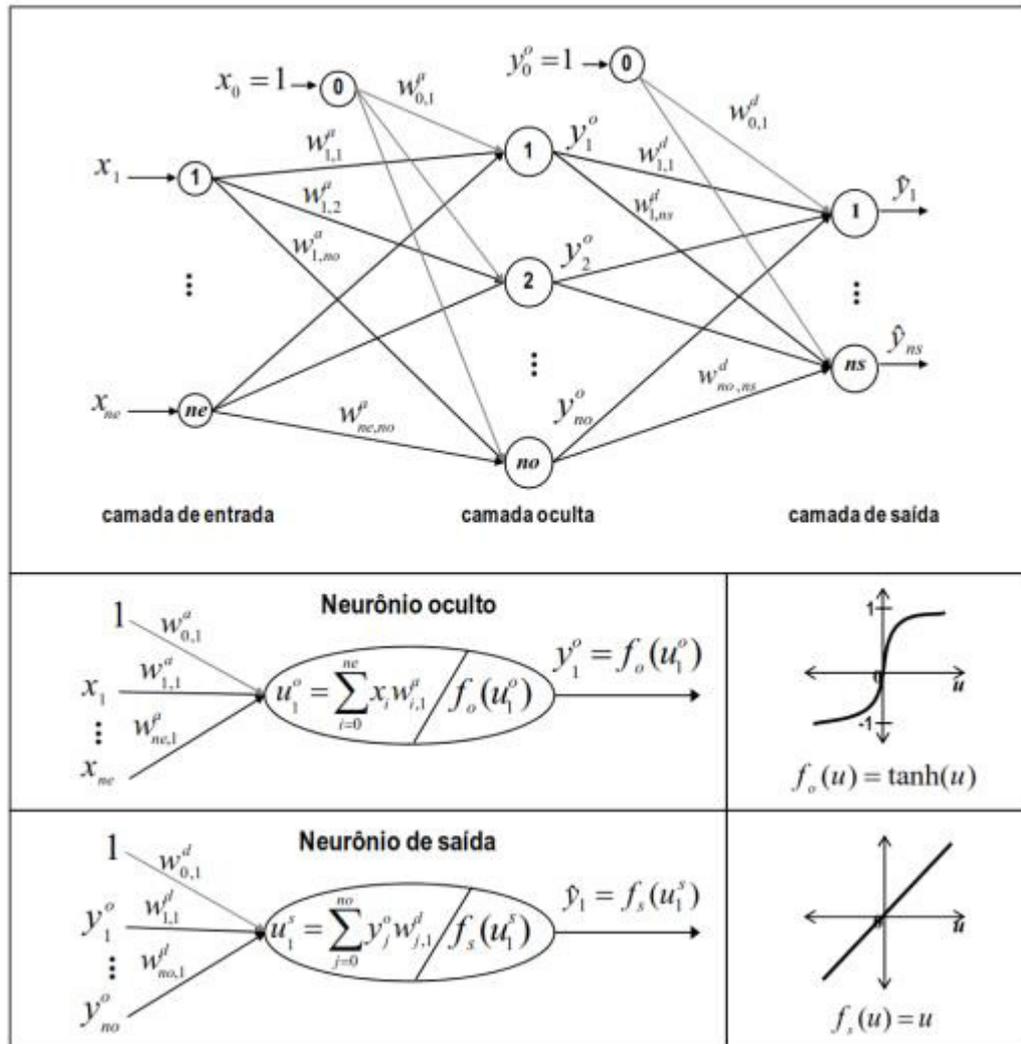
A arquitetura de uma RNA com algoritmo de retropropagação está apresentado na imagem abaixo, que após combinação linear das funções de ativação internas aplicadas nos neurônios das camadas oculta (tipo sigmoideal) e de saída (tipo linear), são formadas as saídas finais de cada neurônio (PUMA-VILLANUEVA, 2011).

Uma consideração importante refere-se ao valor das taxas de aprendizagem: o algoritmo pode adotar uma abordagem de aprendizagem fixa, o que pode tornar a convergência das superfícies de erro lenta, ou uma abordagem adaptativa, ajustando-se a diferentes superfícies de erro, como observado no algoritmo de retropropagação denominado "trainrp" (PEREIRA, 2016).

A dinâmica do algoritmo trainrp pode ser resumida em duas etapas principais: a propagação do efeito de um vetor de entrada pela rede e a retropropagação do sinal de erro gerado na saída da rede, seguindo o sentido oposto das conexões sinápticas.

Quando os padrões são apresentados às entradas e o efeito se propaga pela rede, os pesos sinápticos são atualizados de maneira suave, mesmo quando estão distantes dos valores ótimos (RIEDMILLER e BRAUN, 1993).

Figura 7 – Arquitetura de RNA do tipo retropropagação



Fonte: VILLANUEVA (2011).

### 3.2 Cinética do processo

Determinar parâmetros confiáveis e adequados para modelos de processos biotecnológicos é desafiador devido à complexidade do metabolismo microbiano e à não linearidade de sua cinética. A modelagem que descreve as cinéticas desconhecidas do processo, unindo equações de balanço material com redes neurais artificiais, que funcionam como estimadores das velocidades específicas de formação de biomassa e produto e consumo de substrato, podem ser descritas a seguir.

Os balanços materiais na etapa de fermentação com crescimento microbiano, consumo de xilose e glicose e formação de etanol e xilitol, em reator batelada, estão apresentados nas Equações (2), (3) e (4) respectivamente:

$$\frac{dX}{dt} = r_x \quad (2)$$

$$\frac{dS}{dt} = -r_s \quad (3)$$

$$\frac{dP}{dt} = r_p \quad (4)$$

Os modelos não estruturados e não segregados, são os mais indicados na literatura para descrever fermentações (MOULING; BONE; GALZY, 1980), baseiam-se na determinação da velocidade específica de crescimento do micro-organismo ( $\mu$ ), definido como a taxa de crescimento microbiano dividido pela concentração de biomassa contida no meio - dado que a taxa de fermentação muda continuamente conforme o substrato é consumido e o produto é formado.

A Equação (5) relaciona a taxa de crescimento microbiano com a velocidade específica de crescimento celular:

$$\mu = \frac{1}{X} * \frac{dX}{dt} = \frac{r_x}{X} \quad (5)$$

$$r_x = \mu * X$$

A equação cinética mais simples e comum para descrever a velocidade específica de crescimento é a equação de Monod, aplicável quando um único substrato é considerado como limitante para o crescimento microbiano (HAN; LEVENSPIEL, 1988), observada na Equação (6) abaixo:

$$\mu = \frac{\mu_{m\acute{a}x} * S}{K_s + S} \quad (6)$$

### 3.3 Discriminação entre modelos

A validação de um modelo é comumente conduzida por meio da inspeção visual do comportamento das curvas do modelo em comparação com conjuntos de dados experimentais não utilizados em sua elaboração ou na etapa de estimativa de parâmetros, acompanhada por uma análise estatística específica, considerando os erros experimentais inerentes aos processos biológicos (BONOMI e SCHMIDELL, 2001). Além disso, é essencial realizar uma análise do desvio padrão residual.

Dessa forma, o modelo pode ser considerado estatisticamente adequado se o erro experimental médio dos dados for menor que o erro experimental aparente ( $\mathcal{E}_{exp} < \mathcal{E}_{aparente}$ ). É conhecido que erros experimentais típicos para variáveis medidas em processos fermentativos são de aproximadamente 10%, devido à falta de reprodutibilidade e flutuações naturais do processo (PAIVA et al., 1996). Erros percentuais têm a vantagem de serem livres de unidades e, portanto, são frequentemente usados para comparar desempenhos de previsão entre conjuntos de dados.

O MAPE (Erro Percentual Absoluto Médio, na sigla em inglês) é uma medida estatística que avalia a precisão de um modelo em prever valores em termos percentuais. Ele é calculado como a média da diferença absoluta entre os valores observados e os valores previstos, normalizada pelos valores observados. A fórmula do MAPE é a da Equação (7) a seguir:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right| \times 100 \quad (7)$$

O MSE (Erro Quadrático Médio) é uma medida que avalia a média dos quadrados das diferenças entre os valores observados e os valores previstos. A fórmula do MSE é dada pela Equação (8):

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (8)$$

O MAE (Erro Absoluto Médio) é uma medida que avalia a média das diferenças absolutas entre os valores observados e os valores previstos. A fórmula do MAE é dada pela Equação (9) abaixo:

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i| \quad (9)$$

Assim, o MAPE fornece uma medida percentual da precisão, o MSE fornece uma medida do erro quadrático médio e o MAE fornece uma medida do erro absoluto médio entre as previsões e os valores reais. Essas métricas são muito usadas na avaliação de modelos estatísticos e de previsão, e são usadas no código para avaliar os modelos gerados deste trabalho. Destaco os valores obtido para MAE, já que basta uma análise estatística certificante. Apesar de esforços para os colaboradores no âmbito educacional, não foi detectada atividades voltadas para os sistemas infantis, sendo um ponto de especial atenção por parte da corporação a ser analisado.

## 4 MATERIAIS E MÉTODOS

O conjunto de dados experimentais utilizados para elaboração e validação das predições propostas aqui, foram obtidos por Tiago Lima de Albuquerque: Produção biotecnológica de xilitol a partir de hidrolisado de bagaço de caju (2014).

Posteriormente, fez-se um tratamento dos dados experimentais através de regressão não linear, interpolação e normalização deles, de forma a amenizar os ruídos das medidas experimentais e aumentar o conjunto de dados, além de padronizar escalas para evitar que uma dimensão sobreponha outras.

### 4.1 Metodologia para obtenção dos dados experimentais

Os conjuntos de dados referem-se aos hidrolisados hemicelulósicos do bagaço de caju (HBC) e suas respectivas composições.

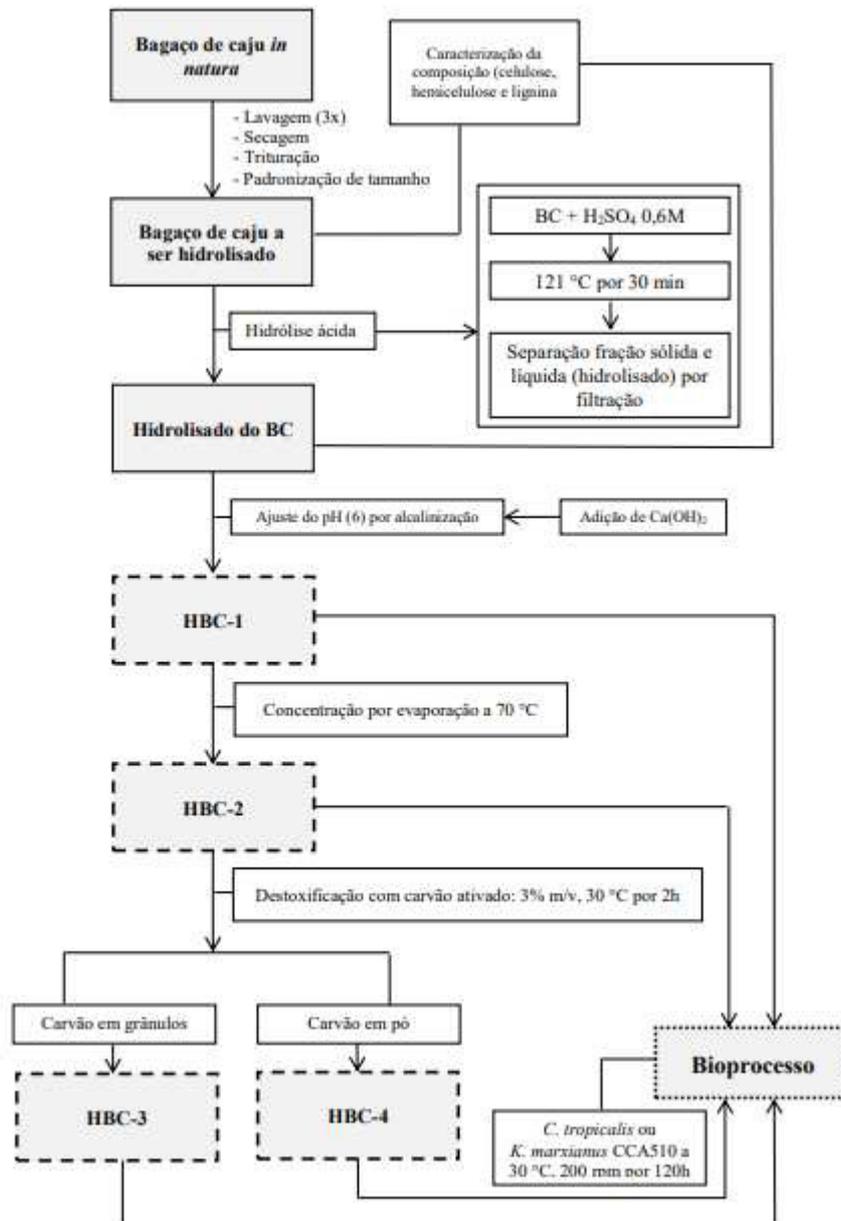
O bagaço de caju foi submetido, após prévia lavagem, secagem e padronização de tamanho, a hidrólise ácida com  $\text{H}_2\text{SO}_4$   $0,6 \text{ mol.L}^{-1}$ , alcançando concentração de glicose e xilose de  $12,09 \text{ g.L}^{-1}$  e  $19,02 \text{ g.L}^{-1}$ , respectivamente.

Depois, o bagaço teve pH ajustado para 6,0 com adição de  $\text{Ca(OH)}_2$ , gerando o primeiro conjunto de hidrolisado chamado HBC-1.

Este conjunto, por sua vez, foi concentrado por evaporação a  $70^\circ\text{C}$ , gerando o segundo conjunto de hidrolisado: HBC-2.

O HBC-2 passou por destoxificação com carvão ativado a  $30^\circ\text{C}$  por 2h - um tipo de carvão em grânulos e outro em pó, formando os últimos dois conjuntos estudados nesse trabalho: HBC-3 e HBC-4 respectivamente, o de maior e menor granulometria de carvão.

Figura 8 – Fluxograma dos experimentos realizados visando avaliar a influência de destoxificação do HBC na produção de xilitol



Fonte: ALBUQUERQUE (2014).

As composições de cada hidrolisado hemicelulósico estão mapeadas na Tabela 1, com a geração de açúcares fermentescíveis e os inibidores detectados no meio. Cada hidrolisado hemicelulósico de bagaço de caju sob esses 4 diferentes tratamentos: HBC-1, HBC-2, HBC-3, HBC-4, foram meio para fermentação a 30°C por 120h a 200rpm, usando a cepa de levedura *Kluyveromyces marxianus* CCA510, a fim de obter uma produção de xilitol.

O acompanhamento dos dados e medidas foram feitos, obtendo-se os seguintes gráficos da Figura 9 para cada meio, e observando-se o comportamento da biomassa, xilose, glicose, xilitol e etanol (ALBUQUERQUE, 2014).

Tabela 1 – Composição dos meios formulados a partir de hidrolisado do bagaço

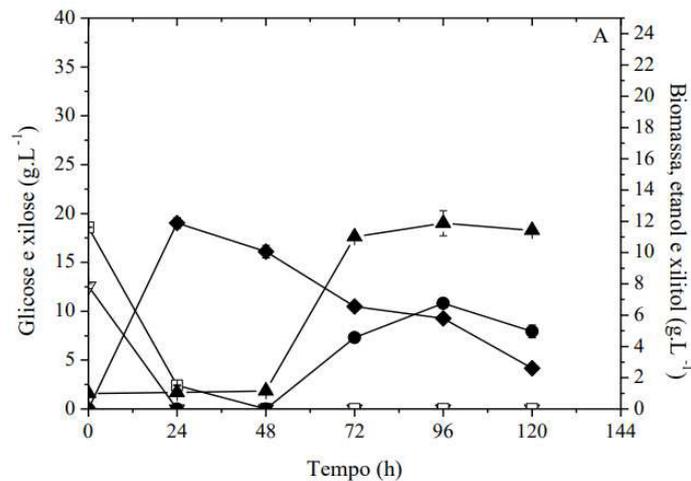
Composição (g.L <sup>-1</sup> )	Concentração			
	(HBC-1)	(HBC-2)	(HBC-3)	(HBC-4)
Glicose	12,57 ± 0,25	30,64 ± 0,02	29,87 ± 0,14	29,71 ± 0,25
Xilose	18,59 ± 0,04	36,44 ± 0,16	36,43 ± 0,56	35,86 ± 0,22
Arabinose	10,85 ± 0,19	32,23 ± 0,09	32,42 ± 0,93	30,96 ± 0,12
Ácido acético	1,89 ± 0,08	4,50 ± 0,09	4,64 ± 0,06	4,06 ± 0,03
Ácido fórmico	0,08 ± 0,00	0,38 ± 0,01	0,36 ± 0,02	0,00 ± 0,00
Furfural	n.d.	n.d.	n.d.	n.d.
5-HMF	n.d.	n.d.	n.d.	n.d.
Fenóis	3,30 ± 0,07	3,34 ± 0,08	2,39 ± 0,01	1,23 ± 0,11

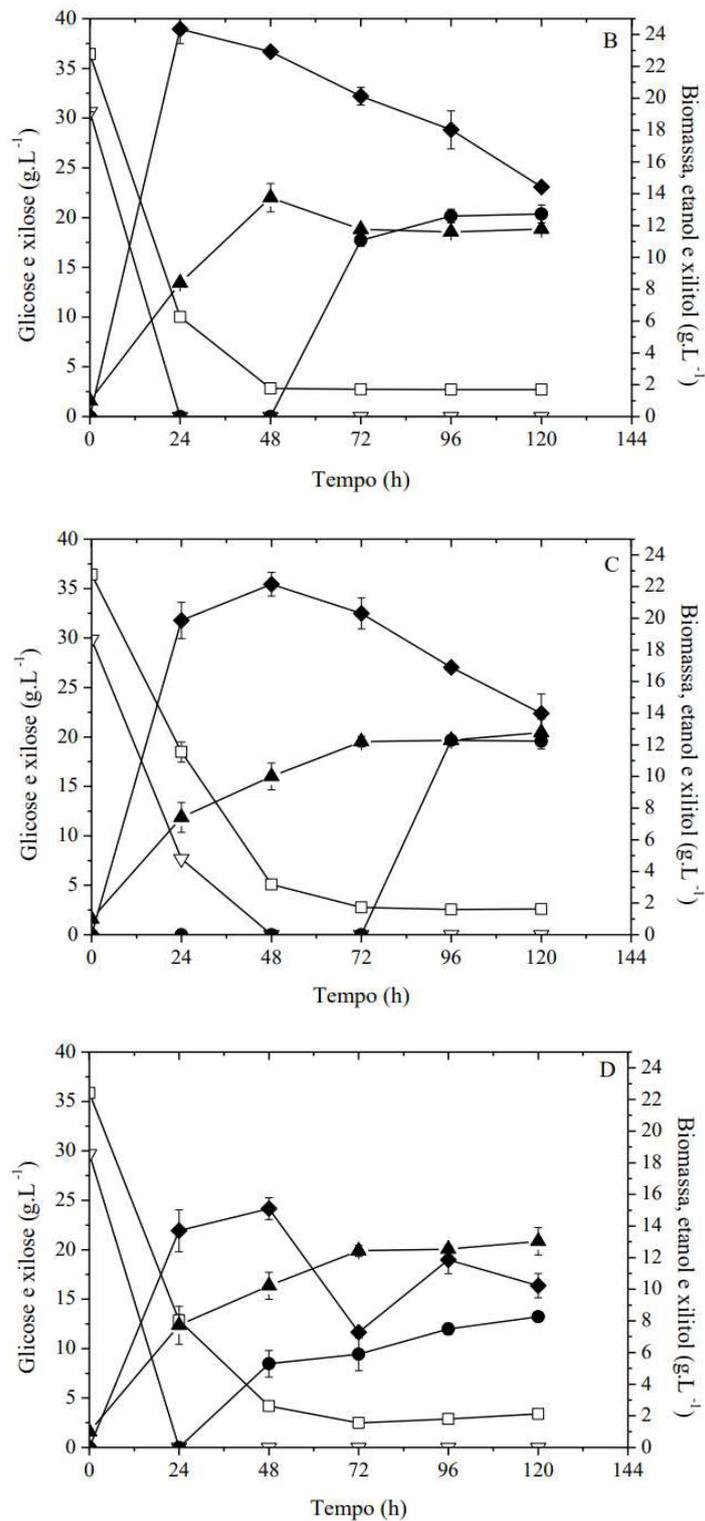
\*n.d.: não detectado

\*\*HBC-1:hidrolisado não concentrado após ajuste do pH, HBC-2: hidrolisado concentrado sem tratamento, HBC-3: hidrolisado concentrado tratado com carvão ativado em grânulos e HBC-4:hidrolisado concentrado tratado com carvão ativado em pó.

Fonte: ALBUQUERQUE (2014).

Figura 9 – Fermentação, por *Kluyveromyces marxianus* CCA 510, de meios formados por HBC com diferentes tratamentos: (A) meio não concentrado, (B) meio concentrado e sem tratamento, (C) meio concentrado com carvão (grânulos) e (D) meio concentrado com carvão (pó). Biomassa (▲), etanol (◆), xilitol (●), glicose (▽) e xilose (□): 200 rpm a 30 °C por 120h





Fonte: ALBUQUERQUE (2014).

#### 4.1.1 Base de dados

Os pontos fermentativos dos quatro gráficos da Figura 9 acima, foram lidos através da ferramenta Plotdigitizer, que é uma ferramenta online de extração de dados, permitindo aos

usuários extrair dados de imagens em formato numérico. Ou seja, ela faz a engenharia reversa de gráficos visuais em números.

Foram 4 meios analisados: HBC-1, HBC-2, HBC-3 e HBC-4, cada um passando pelo processo de fermentação a 30°C, 200rpm e gerando: biomassa, açúcares - xilose e glicose, produtos - xilitol e etanol, durante o intervalo de tempo de 120h, onde houve coleta experimental de dados a cada 24h. Totalizando 120 pontos obtidos experimentalmente. 24 desses pontos, estão exemplificados na Tabela 2, que serviram como base de dados para as previsões para biomassa. As matrizes para os outros componentes estudados: xilose, glicose, xilitol e etanol estão catalogadas no Apêndice A.

Tabela 2 – Matriz das condições operacionais iniciais para Biomassa

<i>Biomassa</i>		
<b>Método</b>	<b>Tempo (h)</b>	<b>Concentração Inicial (g/L)</b>
<b>HBC-1</b>	0	1,00719
<b>HBC-1</b>	24	1,07914
<b>HBC-1</b>	48	1,11511
<b>HBC-1</b>	72	11,07914
<b>HBC-1</b>	96	12,05036
<b>HBC-1</b>	120	11,51079
<b>HBC-2</b>	0	1,13636
<b>HBC-2</b>	24	8,51911
<b>HBC-2</b>	48	13,8535
<b>HBC-2</b>	72	11,9697
<b>HBC-2</b>	96	11,7803
<b>HBC-2</b>	120	11,89394
<b>HBC-3</b>	0	1,06461
<b>HBC-3</b>	24	7,41557
<b>HBC-3</b>	48	9,98532
<b>HBC-3</b>	72	12,15125
<b>HBC-3</b>	96	12,3348
<b>HBC-3</b>	120	12,81204
<b>HBC-4</b>	0	0,93795
<b>HBC-4</b>	24	7,82828
<b>HBC-4</b>	48	10,20924
<b>HBC-4</b>	72	12,44589
<b>HBC-4</b>	96	12,59098
<b>HBC-4</b>	120	12,95488

Fonte: Autoria própria.

## 4.2 Tratamento dos dados experimentais

### 4.2.1 Regressão não linear e Interpolação

Entende-se por tratamento dos dados experimentais, a correção ou transformação dos mesmos buscando adequá-los à análise desejada (BORZANI et al, 2001). É recomendável que os dados destinados à calibração de modelos matemáticos e treinamento de redes neurais sejam ajustados de modo a reduzir os erros de medidas experimentais, bem como para aumentar a quantidade de dados disponíveis por predição de pontos intermediários (HORTA, 2008). A regressão é uma técnica de análise que utiliza a relação entre variáveis quantitativas para determinar um modelo matemático de forma que o efeito de uma possa ser previsto através da outra variável. Além da regressão, a interpolação também foi necessária, pois os algoritmos de ajustes do tipo RNA requerem uma quantidade relativamente grande de dados.

Para o alisamento dos dados experimentais, utilizou-se um modelo de regressão não linear, aplicado ao método de mínimos quadrados para determinar os parâmetros da função para cada condição operacional. Em suma, o algoritmo implementado para o tratamento dos dados requer como entrada o conjunto de dados experimentais de todos os ensaios, a função para regressão e o intervalo de tempo no qual os dados serão interpolados. Como saída, obtêm-se todos os gráficos de regressão e de velocidades específicas, assim como os dados interpolados para ambos, chamados de dados pseudo-experimentais.

Nesta etapa, utilizou-se do *software* Origin, pois o programa importa dados, processa-os estatisticamente, usando por exemplo regressão não-linear - ajusta os dados experimentais de modo a evidenciar as características do fenômeno observado (integral, filtro de ruídos, linha de base) e procura a melhor curva que representa os dados coletados. Retornando os pontos de concentração já interpolados, montagem da curva de concentração instantânea e posterior derivação da curva para encontrar as velocidades específicas. O tratamento feito no software Origin, de regressão não-linear, interpolação, e derivada dos pontos, são mostrados para o comportamento das concentrações e das velocidades específicas são detalhados a seguir.

Os modelos citados no Quadro I, foram os que melhor se adequaram aos conjuntos de dados e variações de biomassa, etanol, xilitol, xilose e glicose em cada um dos meios HBC-1, HBC-2, HBC-3 e HBC-4. Os modelos de regressão não-linear foram gerados na ferramenta Origin para o ajuste da curva de concentrações e para ampliar o conjunto de dados, gerando

maior volume de dados pseudo-experimentais servindo ao treinamento da máquina com eles já interpolados. Os gráficos desses ajustes encontram-se no Apêndice D.

Quadro 1 – Ajustes de regressão não linear aplicados em Origin

	<b>Biomassa</b>	<b>Xilose</b>	<b>Glicose</b>	<b>Xilitol</b>	<b>Etanol</b>
<b>HBC-1</b>	Boltzmann	Ajuste Assintótico	Ajuste Assintótico	Boltzmann	Polinomial de 4º grau
<b>HBC-2</b>	Polinomial de 4º grau	Ajuste Assintótico	Ajuste Assintótico	Ajuste Logístico (sigmoidal)	Polinomial de 4º grau
<b>HBC-3</b>	Polinomial de 4º grau	Boltzmann	Boltzmann	Boltzmann	Polinomial de 4º grau
<b>HBC-4</b>	Boltzmann	Boltzmann	Ajuste Assintótico	Gompertz	Polinomial de 4º grau

Fonte: Autoria própria.

As equações referentes aos ajustes são as Equações (10), (11), (12), (13) e (14):

i. Boltzmann

$$y = \frac{A_1 - A_2}{1 + e^{\frac{x-x_0}{dx}}} - A_2 \quad (10)$$

ii. Polinomial de 4º grau

$$y = A_4x^4 + A_3x^3 + A_2x^2 + A_1x + A_0 \quad (11)$$

iii. Ajuste Logístico (sigmoidal)

$$y = \frac{a}{1 + e^{-b(x-c)}} \quad (12)$$

iv. Gompertz

$$y = a * e^{-\exp(-b(x-c))} \quad (13)$$

v. Ajuste Assintótico

$$y = a - b * c^x \quad (14)$$

Nesse cenário, obteve-se a ampliação de 5 para 20 pontos em cada substância, totalizando assim 2280 pontos a mais, que poderão ser interpolados – mas não podem ser extrapolados. Tornou-se conveniente estimar as concentrações e as velocidades específicas em

intervalos menores de tempo - ao invés de manter de 24h em 24h, passou-se para pontos de 6h em 6h, ampliando a quantidade de dados destinados ao posterior treinamento das RNAs.

Esses dados encontram-se no Apêndice B, e abaixo um exemplo na Tabela 3 da interpolação feita para os dados experimentais da biomassa, trazendo mais pontos pseudo-experimentais.

Tabela 3 – Interpolação utilizando os dados experimentais da biomassa

<i>Biomassa</i>		
<b>Método</b>	<b>Tempo (h)</b>	<b>Concentração Inicial (g/L)</b>
<b>HBC-1</b>	0	1,04379
<b>HBC-1</b>	6	1,04379
<b>HBC-1</b>	12	1,04379
<b>HBC-1</b>	18	1,04379
<b>HBC-1</b>	24	1,04382
<b>HBC-1</b>	30	1,04401
<b>HBC-1</b>	36	1,0453
<b>HBC-1</b>	42	1,05409
<b>HBC-1</b>	48	1,11381
<b>HBC-1</b>	54	1,5047
<b>HBC-1</b>	60	3,56261
<b>HBC-1</b>	66	8,3109
<b>HBC-1</b>	72	11,07942
<b>HBC-1</b>	78	11,67194
<b>HBC-1</b>	84	11,76457
<b>HBC-1</b>	90	11,77826
<b>HBC-1</b>	96	11,78027
<b>HBC-1</b>	102	11,78056
<b>HBC-1</b>	108	11,78061
<b>HBC-1</b>	114	11,78061
<b>HBC-1</b>	120	11,78061
<b>HBC-2</b>	0	1,03959
<b>HBC-2</b>	6	3,35559
<b>HBC-2</b>	12	5,47314
<b>HBC-2</b>	18	7,3629
<b>HBC-2</b>	24	9,00297
<b>HBC-2</b>	30	10,37877
<b>HBC-2</b>	36	11,48314
<b>HBC-2</b>	42	12,31628
<b>HBC-2</b>	48	12,88579
<b>HBC-2</b>	54	13,20662
<b>HBC-2</b>	60	13,30112
<b>HBC-2</b>	66	13,19901

<b>HBC-2</b>	72	12,93741
<b>HBC-2</b>	78	12,5608
<b>HBC-2</b>	84	12,12104
<b>HBC-2</b>	90	11,67738
<b>HBC-2</b>	96	11,29644
<b>HBC-2</b>	102	11,05223
<b>HBC-2</b>	108	11,02614
<b>HBC-2</b>	114	11,30691
<b>HBC-2</b>	120	11,99071
<b>HBC-3</b>	0	1,09957
<b>HBC-3</b>	6	3,01885
<b>HBC-3</b>	12	4,66126
<b>HBC-3</b>	18	6,05866
<b>HBC-3</b>	24	7,24075
<b>HBC-3</b>	30	8,2351
<b>HBC-3</b>	36	9,0671
<b>HBC-3</b>	42	9,76001
<b>HBC-3</b>	48	10,33494
<b>HBC-3</b>	54	10,81084
<b>HBC-3</b>	60	11,20451
<b>HBC-3</b>	66	11,53061
<b>HBC-3</b>	72	11,80162
<b>HBC-3</b>	78	12,02791
<b>HBC-3</b>	84	12,21766
<b>HBC-3</b>	90	12,37693
<b>HBC-3</b>	96	12,50961
<b>HBC-3</b>	102	12,61745
<b>HBC-3</b>	108	12,70004
<b>HBC-3</b>	114	12,75482
<b>HBC-3</b>	120	12,77708
<b>HBC-4</b>	0	0,9714
<b>HBC-4</b>	6	3,14768
<b>HBC-4</b>	12	4,93782
<b>HBC-4</b>	18	6,41015
<b>HBC-4</b>	24	7,62097
<b>HBC-4</b>	30	8,61665
<b>HBC-4</b>	36	9,43535
<b>HBC-4</b>	42	10,10851
<b>HBC-4</b>	48	10,66196
<b>HBC-4</b>	54	11,11697
<b>HBC-4</b>	60	11,49105
<b>HBC-4</b>	66	11,79858
<b>HBC-4</b>	72	12,05139
<b>HBC-4</b>	78	12,25922
<b>HBC-4</b>	84	12,43007
<b>HBC-4</b>	90	12,57052

<b>HBC-4</b>	96	12,68597
<b>HBC-4</b>	102	12,78087
<b>HBC-4</b>	108	12,85889
<b>HBC-4</b>	114	12,92301
<b>HBC-4</b>	120	12,97573

Fonte: Aatoria própria.

Para evitar extrapolações, padronizou-se os tempos de fermentação total em 120h, como limitado pelo próprio gráfico de fermentação em meios.

O coeficiente de determinação ( $R^2$ ) foi empregado com o objetivo de certificar o grau de explicação de cada ajuste.

Segue na Tabela 4, os limites de interpolação de cada componente:

Tabela 4 – Limites de interpolação

	<b>Concentração Inicial (g/L)</b>	
	Mínimo	Máximo
<b>Biomassa</b>	0,93795	13,8535
<b>Xilose</b>	0,008	36,56051
<b>Glicose</b>	0,013	30,70064
<b>Xilitol</b>	0	12,95828
<b>Etanol</b>	0	24,44268

Fonte: Aatoria própria.

De posse dos perfis de concentração para o substrato, célula e produto de todos os testes, já planilhados no software Origin, calculou-se as velocidades específicas de consumo dos açúcares xilose e glicose, crescimento celular da biomassa e produção de etanol e xilitol. Para isso, aplicou-se a definição demonstrada a partir das Equações (15), (16) e (17), derivando as Equações anteriores (10), (11), (12), (13) e (14) e dividiu-se cada uma pela concentração celular instantânea para todos os testes.

$$\mu_x(t) = \frac{\frac{dX(t)}{dt}}{X(t)} \quad (15)$$

$$\mu_s(t) = \frac{\frac{dS(t)}{dt}}{X(t)} \quad (16)$$

$$\mu_p(t) = \frac{\frac{dP(t)}{dt}}{X(t)} \quad (17)$$

Como o  $R^2$  expressa a quantidade de variabilidade nos dados que é explicado pelo modelo de regressão não-linear, nota-se pela Tabela 5 que os ajustes para os modelos de concentração celular é nível acima de 90%, portanto considerou-se a qualidade das regressões satisfatórias.

Tabela 5 –  $R^2$  para validação dos ajustes de curva de concentração dos compostos

<b><math>R^2</math> dos Ajustes Não-lineares</b>				
	<b>HBC-1</b>	<b>HBC-2</b>	<b>HBC-3</b>	<b>HBC-4</b>
<b>Xilose</b>	91,27%	99,62%	99,58%	99,42%
<b>Biomassa</b>	94,05%	94,90%	97,08%	93,58%
<b>Xilitol</b>	93,56%	99,96%	97,63%	95,34%
<b>Etanol</b>	90,11%	96,78%	92,64%	97,83%
<b>Glicose</b>	95,78%	92,45%	97,56%	90,23%

Fonte: Autoria própria.

#### **4.2.2 Normalização dos dados pseudo-experimentais**

De posse dos dados pseudo-experimentais, oriundos do uso do Origin por regressão não-linear e interpolação, é adequado realizar a normalização dos mesmos para aplicação em métodos de inteligência artificial, pois existe uma melhora significativa na distribuição dos dados. A normalização substitui a necessidade de harmonizar escalas, que numericamente corresponde a minimizar os problemas de adaptação de escalas de grandezas muito diferentes. Nesse trabalho, optou-se pela normalização mín-máx mostrada na Equação (18), com intervalo unitário:

$$z_i^n = \frac{z_i - z_i^{\min}}{z_i^{\max} - z_i^{\min}} \quad (18)$$

As condições operacionais foram normalizadas diretamente no código da RNA em Python, sendo processada como uma das etapas para a implementação do código/ algoritmo da RNA.

### 4.2.3 Metodologia

O código foi desenvolvido no ambiente Google Colab, utilizando a linguagem de programação Python, versão 3.6.9. O Google Colab é um serviço gratuito baseado na nuvem, fornecido pelo Google para Aprendizado de Máquina e IA [COLABORATORY 2020]. Ele oferece aceleradores de GPU gratuitos, bibliotecas pré-instaladas e é construído sobre o Jupyter Notebook. Além disso, permite armazenar os notebooks diretamente no Google Drive de forma online. As principais bibliotecas empregadas no desenvolvimento da Rede Neural Artificial (RNA) foram o TensorFlow 2.0 e o Keras, ambos focados em aprendizado profundo. O Keras é amplamente utilizado devido à sua facilidade de uso e capacidade de prototipagem rápida. Além disso, é uma camada de alto nível para outros frameworks de Deep Learning, como TensorFlow, Theano e CNTK [Keras 2020]. No TensorFlow 2.0, o Keras foi integrado por meio do módulo `tf.keras`, o que significa que toda a API do Keras está presente na nova versão do TensorFlow [TENSORFLOW 2020].

Foram usadas redes neurais do tipo feedforward de múltiplas camadas com treinamento supervisionado por correção do erro (retropropagação com algoritmo `trainrp`). As RNAs funcionam como estimadores das velocidades específicas de crescimento celular (biomassa), de consumo de substrato (xilose e glicose) e formação de produto (xilitol e etanol). Desse modo, foram elaboradas cinco redes respectivamente: RN1 ( $\mu_X$ ), RN2 ( $\mu_{Sx}$ ), RN3 ( $\mu_{Sg}$ ), RN4 ( $\mu_{Px}$ ) e RN5 ( $\mu_{Pe}$ ).

Para o treinamento das redes, a camada de entrada foi alimentada com a concentração instantânea de cada hidrolisado (HBC), para predição das velocidades específicas de biomassa, xilose, glicose, xilitol e etanol. Além das concentrações instantâneas (dados pseudo-experimentais de velocidade específica divididos pelas concentrações encontradas para os pontos pseudo-experimentais no Origin), também houve o input da composição já citada na Tabela 1 – Composições de cada meio de hidrolisado hemicelulósico do bagaço de caju (HBC), sendo elas formadas por: xilose, glicose, arabinose, ácido acético, ácido fórmico e fenóis (como 5-HMF e furfural não houve detecção, então não virou input da rede).

Já a camada de saída, são as velocidades específicas da reação. Todos normalizados via código de RNA em Python. Portanto, as velocidades específicas são funções dos produtos mostrados na Equação (19):

$$\mu^n = f(S_0^n, HBC^n, Inibidores^n) \quad (19)$$

Várias arquiteturas foram testadas e avaliadas com base no coeficiente de determinação ( $R^2$ ) de forma que as redes conseguissem reproduzir os dados de saída com maior

precisão. Os erros foram pontuados através do MAE, que ficou abaixo de 1, portanto validando os modelos de RNA.

Para a predição das velocidades específicas em função do tempo, utiliza-se as concentrações no instante  $t + \Delta t$  para a predição das RNAs. A validação do modelo foi realizada através de análises estatísticas e de inspeção do comportamento frente aos dados não utilizados para o treinamento das RNAs, seguindo a proposta de 20% a 30% dos dados ficarem de fora do treinamento para possibilitar comparações da qualidade de predição dentro do algoritmo.

O modelo cinético foi determinado de acordo com as arquiteturas mostradas a seguir, onde as velocidades específicas de xilose, glicose, biomassa, xilitol e etanol são estimadas como função da concentração instantânea de HBC-1, 2, 3 e 4 e da composição dos meios.

Foram realizados aproximadamente 150 treinamentos com as mais variadas arquiteturas de redes neurais do tipo simétrica, decrescente e crescente. No Quadro 2, as arquiteturas utilizadas nas RNAs que entregaram  $R^2 > 90\%$ , onde sequencialmente estão as redes neurais para: biomassa, xilose, glicose, xilitol e etanol.

Quadro 2 – Arquitetura das RNAs

	<b>Camada</b>	<b>Número de neurônios</b>	<b>Função de ativação</b>
<b>RN1 - Biomassa</b>	1ª camada oculta	32	ReLU (Rectified Linear Unit)
	2ª camada oculta	16	ReLU (Rectified Linear Unit)
	3ª camada oculta	16	ReLU (Rectified Linear Unit)
	Camada de saída	1	Linear
<b>RN2 - Xilose</b>	1ª camada oculta	16	ReLU (Rectified Linear Unit)
	2ª camada oculta	32	ReLU (Rectified Linear Unit)
	3ª camada oculta	32	ReLU (Rectified Linear Unit)
	Camada de saída	1	Linear
<b>RN3 - Glicose</b>	1ª camada oculta	64	ReLU (Rectified Linear Unit)
	2ª camada oculta	24	ReLU (Rectified Linear Unit)
	3ª camada oculta	16	ReLU (Rectified Linear Unit)
	Camada de saída	1	Linear
<b>RN4 - Xilitol</b>	1ª camada oculta	48	ReLU (Rectified Linear Unit)
	2ª camada oculta	32	ReLU (Rectified Linear Unit)
	3ª camada oculta	32	ReLU (Rectified Linear Unit)
	Camada de saída	1	Linear
<b>RN5 - Etanol</b>	1ª camada oculta	64	ReLU (Rectified Linear Unit)
	2ª camada oculta	32	ReLU (Rectified Linear Unit)
	3ª camada oculta	16	ReLU (Rectified Linear Unit)
	Camada de saída	1	Linear

Fonte: Autoria própria.

As simulações das RNAs mostram um espalhamento dos pontos em torno da reta bissetriz, sugerindo que as previsões das redes exibem variações em relação aos dados experimentais. Apesar das inúmeras arquiteturas de redes neurais testadas, os dados experimentais para as concentrações instantâneas possuem comportamento inesperados para as condições operacionais que dificultam a proposição de arquiteturas simples.

Apesar dos inúmeros testes realizados com diferentes arquiteturas de RNAs, observa-se um comportamento imprevisível nas concentrações instantâneas em relação às condições operacionais. Essa imprevisibilidade tem dificultado a formulação de arquiteturas de redes neurais mais simplificadas para este contexto, dado que os dados experimentais apresentam padrões desafiadores de serem modelados.

## 5 RESULTADOS E DISCUSSÃO

### 5.1 Resultados

As arquiteturas utilizadas para estimar as velocidades específicas se mostraram satisfatórias com base no coeficiente de determinação ( $R^2$ ) como exposto na Tabela 6 e nos desvios (máximos e médios) entre os dados teóricos e pseudo-experimentais mostrada pelas bissetrizes na Figura 10.

Para a definição das arquiteturas das RNAs, também foram levadas em consideração a complexidade da estrutura e conseqüentemente o tempo necessário para a simulação, já que variam entre minutos ou horas.

Tabela 6 –  $R^2$  dos modelos preditivos usando RNAs

	<b><math>R^2</math> treino</b>	<b><math>R^2</math> validação</b>
<b>RN1 - Biomassa</b>	99,85%	94,33%
<b>RN2 - Xilose</b>	99,17%	90,12%
<b>RN3 - Glicose</b>	99,11%	93,92%
<b>RN4 - Xilitol</b>	96,05%	90,03%
<b>RN5 - Etanol</b>	98,62%	94,11%

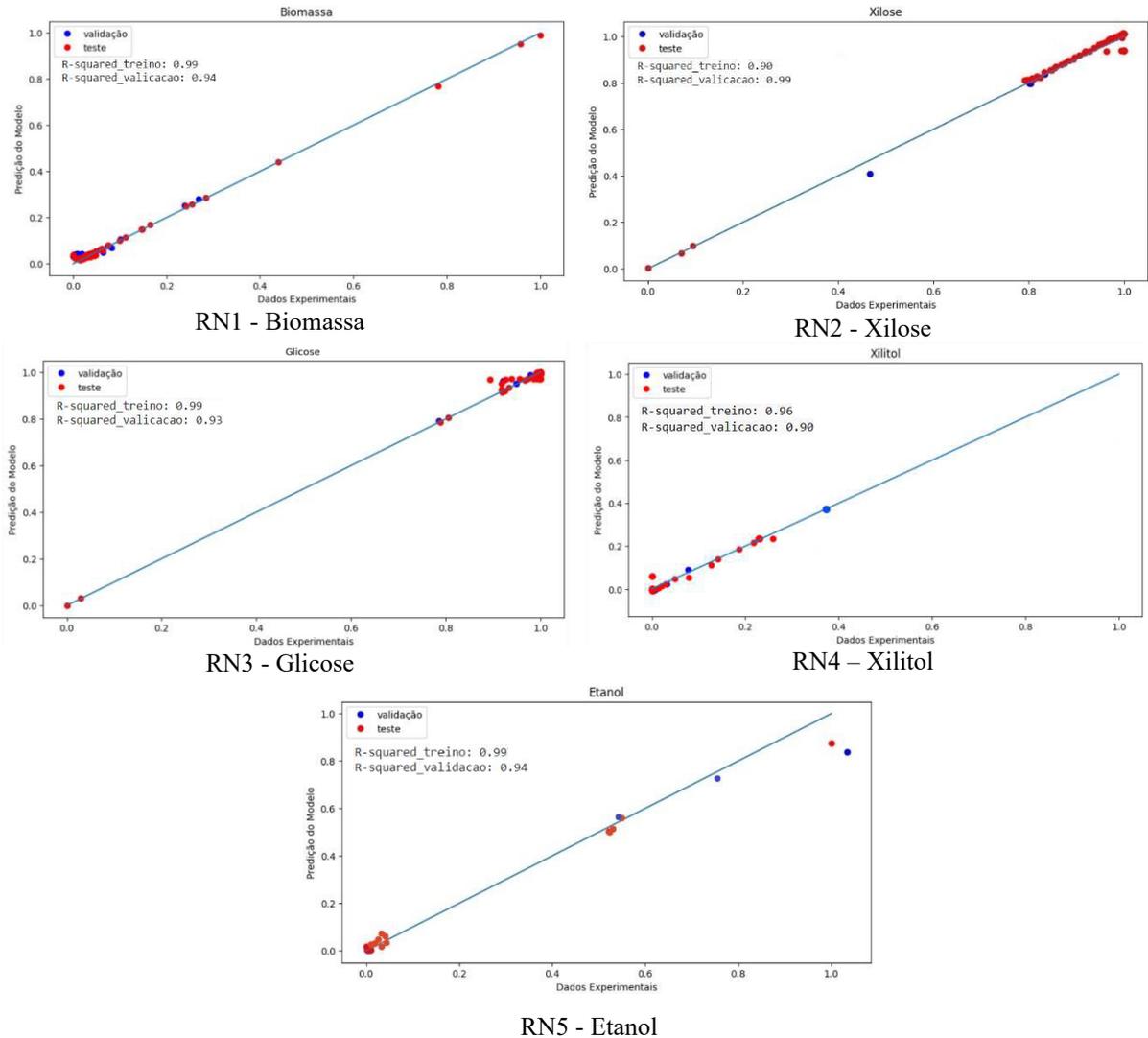
Fonte: Autoria própria.

Tabela 7 – Validação estatística – MAE

	<b>MAE</b>
<b>RN1 - Biomassa</b>	0,005
<b>RN2 - Xilose</b>	0,004
<b>RN3 - Glicose</b>	0,011
<b>RN4 - Xilitol</b>	0,589
<b>RN5 - Etanol</b>	0,260

Fonte: Autoria própria.

Figura 10 – Simulações de RNAs: Comparativo entre as previsões das RNAs e os dados pseudo-experimentais; (—) Reta bissetriz, (●) Dados para validação do modelo, (●) Dados de teste com as concentrações pseudo-experimentais em (A) RN1 – Biomassa; (B) RN2 – Xilose; (C) RN3 – Glicose; (D) RN4 – Xilitol e (E) RN5 – Etanol



Fonte: Autoria própria.

Nota-se que as medidas MAE ficaram muito pequenas, indicando a qualidade de predição das RNAs. O MAE, ou Erro Médio Absoluto, é uma medida estatística frequentemente utilizada para avaliar o desempenho de Redes Neurais Artificiais (RNAs) e outros modelos preditivos. Sua vantagem como avaliador reside na sua interpretação intuitiva e na capacidade de proporcionar uma visão clara da magnitude média dos erros de previsão. O MAE é calculado como a média das diferenças absolutas entre as previsões do modelo e os valores reais.

Motivos para o MAE ser um bom avaliador é que ele representa a média absoluta dos erros de previsão. Isso facilita a interpretação dos resultados, especialmente em contextos em que a compreensão do impacto direto dos erros é crucial. Ao utilizar a média absoluta, o MAE é menos sensível a outliers extremos em comparação com outras métricas, como o Erro Quadrático Médio (MSE).

Quando o MAE é pequeno para um conjunto de dados, indica que as previsões da RNA estão em concordância com os valores reais. Um MAE muito baixo indica que, em média, as previsões do modelo estão muito próximas dos valores observados. Essa situação é altamente desejável, indicando um ajuste preciso do modelo aos dados experimentais. No entanto, é importante considerar outros aspectos, como a possibilidade de overfitting ou a necessidade de validação em conjuntos de dados independentes para garantir a generalização do modelo para novos casos.

## 5.2 Perspectivas para trabalhos futuros

Com o presente trabalho, e a arquitetura obtida para previsão das velocidades específicas dos compostos estudados aqui, há uma sequência de possibilidades para reuso desses dados em outros trabalhos. Sugestões:

- Proposição de modelos matemáticos mecanísticos e híbridos com redes neurais, no intuito de prever o comportamento da fermentação alcoólica: Essa abordagem busca capturar tanto os fundamentos teóricos quanto os padrões empíricos observados no processo. Aqui pode-se empregar as redes neurais artificiais que previram as velocidades específicas de crescimento celular, para prever consumo de substrato e produção biotecnológica de xilitol. A aplicação de técnicas de aprendizado de máquina visa explorar padrões complexos nos dados experimentais, proporcionando uma abordagem alternativa para modelagem e predição;

- Otimização das condições operacionais, com base na maximização da eficiência e da produtividade do processo, pelo método de enxame de partícula ou outros mais adequados a fermentações. Essa abordagem busca identificar as condições ideais que resultam em um desempenho otimizado do sistema.

## 6 CONCLUSÃO

De acordo com toda a discussão proposta pelo presente trabalho, foi possível alcançar o objetivo geral apontado: apresentar uma proposta de utilização de Rede Neural Artificial (RNA) que serve como modelo matemático satisfatório para descrever a predição das velocidades específicas de crescimento celular, consumo de substrato e geração de produtos usando como substâncias: biomassa, xilose, glicose, etanol e xilitol.

Diante do exposto, os resultados das arquiteturas das redes neurais para o  $R^2$  estarem acima de 90% nos testes e nas validações, expressa a adequação do modelo para base de próximos projetos que partam das velocidades específicas e possam construir a predição dos consumos de substratos na produção biotecnológica de xilitol a partir do bagaço de caju. As arquiteturas ficaram com estrutura de 4 camadas, 3 camadas ocultas e 1 camada de saída, alternando o número de neurônios entre elas.

Os modelos de redes neurais artificiais propostos para predição das velocidades específicas de biomassa, xilose, glicose, xilitol e etanol também foram validados estatisticamente através do cálculo do Erro Médio Absoluto (MAE), indicando um ajuste preciso do modelo aos dados experimentais pelos valores menores que 1% mostrados nos resultados.

Ao utilizar os valores deste trabalho, os dados podem ser interpolados, mas não extrapolados, dado que as predições feitas com redes neurais artificiais devem abranger todo o domínio de aplicação do modelo para garantir sua eficácia: o modelo de rede neural artificial não usa os dados estatísticos para gerar novos dados fora do domínio setado (modelo caixa-preta). O modelo de rede neural híbrida, uma das sugestões para trabalhos futuros, pode utilizar os dados trabalhos nessa monografia para expandir até uma otimização do processo estudado e ampliação de escala deste mesmo processo.

## REFERÊNCIAS

- ALBUQUERQUE, T. L. **Produção biotecnológica de xilitol a partir de hidrolisado de bagaço de caju**. 2014. 148 f. Dissertação (Mestrado em Engenharia Química) - Centro de Tecnologia, Universidade Federal do Ceará, Fortaleza, 2014.
- ALVES, R. M. B. **Otimização de um Processo Industrial de Produção de Isopreno via Redes Neurais**. 2003, 283 f. Tese (Doutorado em Engenharia Química) – Escola Politécnica, Universidade de São Paulo, São Paulo, 2003.
- BONOMI, A.; SCHMIDELL, W. **Modelagem matemática e simulação de processos fermentativos**. In: Schmidell, W.; Lima, U. A.; Aquarone, E.; Borzani, W. (coord). *Biotechnology Industrial: Engenharia Bioquímica*. São Paulo: Edgard Blucher, v. 2, p. 123-178, 2001.
- BRIONES, R.; SERRANO, L.; LABIDI, J. Valorization of some lignocellulosic agro-industrial residues to obtain biopolyols. **Journal of Chemical Technology Biotechnology**, v. 87, p. 244-249, 2012.
- CHEN, X.; JIANG, Z.-H.; CHEN, S.; QIN, W. Microbial and bioconversion production of D-xylitol and its detection and application. **International journal of biological sciences**, v. 6, n. 7, p. 834-44, 2010.
- CORREIA, J. A. C.; JUNIOR, J. E. M.; GONÇALVES, L. R. B.; ROCHA, M. V. P. Alkaline hydrogen peroxide pretreatment of cashew apple bagasse for ethanol production: Study of parameters. **Process Biochemistry**, v. 139, p. 249-256, 2013.
- CUNHA, M. A. A.; SILVA, S. S.; CARVALHO, W.; SANTOS, J. C. Uso de células imobilizadas em gel de PVA: uma nova estratégia para a produção biotecnológica de xilitol a partir de bagaço de cana-de-açúcar. **Semina**, v.26, p.59-68, 2005.
- DEMIRBAS, A. Products from lignocellulosic materials via degradation processes. **Energy Sources**, v. 30, p.27-37, 2008.
- DUARTE, L. C.; CARVALHEIRO, F.; NEVES, I.; GIRIO, F.M. Effects of aliphatic acids, furfural, and phenolic compounds on *Debaryomyces hansenii* CCM1 941. **Applied Biochemistry and Biotechnology**, v.121-124, p.413-425, 2005.
- FAKRUDDIN, MD. et al. Analysis of key factors affecting ethanol production by *Saccharomyces cerevisiae* IFST – 072011. **Biotechnology**, v. 11, n. 4, p. 248 - 252, 2012.
- FASOLI, G.; BARRIO, E.; TOFALO, R.; SUZZI, G.; BELLOCH, C. **Multilocus analysis reveals large genetic diversity in *Kluyveromyces marxianus* strains isolated from Parmigiano Reggiano and Pecorino diN Farindola cheeses**. *International Journal of Food Microbiology*, v. 233, p. 1 – 10, september. 2016.
- FERREIRA, P. G.; SILVEIRA, F. A.; SANTOS, R. C. V.; GENIER, H. L. A.; DINIZ, R. H.; JÚNIOR, J. I. R.; FIETTO, L. G.; PASSOS, F. M. L.; SILVEIRA, W. B. **Optimizing ethanol**

**production by thermotolerant *Kluyveromyces marxianus* CCT 7735 in a mixture of sugarcane bagasse and ricotta whey.** Food Science Biotechnology, v. 24, n. 4, p. 1421 – 1427, July. 2015.

FONSECA, B. G.; MOUTTA, R. O.; FERRAZ, F. O.; VIEIRA, E. R.; NOGUEIRA, A. S.; BARATELLA, B. F.; RODRIGUES, L. C.; RUI, Z. H.; SILVA, S. S. Biological detoxification of different hemicellulosic hydrolysates using *Issatchenkia occidentalis* CCTCC M 206097 yeast. **Journal Industrial Microbiology Biotechnology**, v. 38, p. 199-207, 2011.

HAN, K., LEVENSPIEL, O. Extended Monod kinetics for substrate, product, and cell inhibition. **Biotechnology and Bioengineering**, v. 32, p. 430-437, 1988.

HAYKIN, S. **Neural Networks**. A Comprehensive Foundation. Macmillan, New York, NY, 1994.

LANE, M. M.; BURKE, N.; KARREMAN, R.; WOLFE, K.H.; O'BYRNE C.P.; MORRISSEY J.P. **Physiological and metabolic diversity in the yeast *Kluyveromyces marxianus***. Antonie Van Leeuwenhoek, n. 100, p. 507-519, June. 2011.

LANE, M. M.; MORRISSEY, J. P. ***Kluyveromyces marxianus*: a yeast emerging from its sister's shadow**. Fungal Biology Reviews, v. 24, n. 1-2, p. 17 – 26, February. 2010.

LIMA, U. A.; BASSO, L. C.; AMORIM, H. V. Fadiga. In: **Biotecnologia industrial: Processos fermentativos e enzimáticos**. São Paulo-SP. Edgard Blucher, v. 3, p. 1-43, 2001.

MOOD, S. H.; GOLFESHAN, A. H.; TABATABAEI, M. T.; JOUZANI, G. S.; NAJAFI, G. H.; GHOLAMI, M.; ARDJMAND, M. Lignocellulosic biomass to bioethanol, a comprehensive review with a focus on pretreatment. **Renewable and Sustainable Energy Reviews**, v. 27, p. 77–93, 2013.

MOSIER, N.; WYMAN, C.; DALE, B.; ELANDER, R.; LEE, Y. Y.; HOLTZAPPLE, M.; LADISCH, M. Features of promising technologies for pretreatment of lignocellulosic biomass. **Bioresource Technology**, v. 96, p. 673–686, 2005.

MOULING, G.; BOZE, H.; GALZY, P. Inhibition of alcoholic fermentation by substrate and ethanol. **Biotechnology and Bioengineering**, New York, v. 22, p. 235-2381, 1980.

MULAKHUDAIR, ALI & HANOTU, JAMES & ZIMMERMAN, W.B. **Exploiting microbubble-microbe synergy for biomass processing: Application in lignocellulosic biomass pretreatment**. Biomass and Bioenergy. 2016.

MUSSATO, S. I.; ROBERTO, I. C. Optimal experimental condition for hemicellulosic hydrolyzate treatment with activated charcoal for xylitol production. **Biotechnology Progress**, v.20, n.1, p.134-139, 2004.

MUSSATTO, S. I.; ROBERTO, I. C. Xilitol: Edulcorante com efeitos benéficos para a saúde humana. **Brazilian Journal of Pharmaceutical Sciences**. v. 38, n. 4, p. 401-413, 2002.

NACHAIWIENG, W.; LUMYONG, S.; YOSHIOKA, K.; WATANABE, T.; KLANONGNUCH, C. **Bioethanol production from rice husk under elevated temperature simultaneous saccharification and fermentation using *Kluyveromyces marxianus* CK8.** *Biocatalysis and Agricultural Biotechnology*, v. 4, n. 5, p. 543 – 549, october. 2015.

PAIVA, T. C. B. P. et al. Continuous alcoholic fermentation process in a tower reactor with recycling of flocculating yeast. **Appl Biochem Biotechnol**, v. 57, n. 58, p. 535–541, 1996.

PELLICCI, R. L. **Modelagem e Simulação de Um Processo de Extração de Aromáticos Via Redes Neurais.** 2001, 130 f. Dissertação (Mestrado em Engenharia) – Escola Politécnica, Universidade de São Paulo, São Paulo, 2001.

PEREIRA, A. S. **Modelagem e simulação do processo de produção de etanol a partir do suco do pedúnculo de caju, visando a otimização das condições operacionais.** 2016. 167 f. Dissertação (Mestrado em Engenharia Química) – Centro de Tecnologia, Universidade Federal do Ceará, Fortaleza, 2016.

PEREIRA, R. S.; MUSSATTO, S. I.; ROBERTO, I. C. Inhibitory action of toxic compounds present in lignocellulosic hydrolysates on xylose to xylitol bioconversion by *Candida guilliermondii*. **Journal Ind Microbiol Biotechnology**, v.38, p.71–78, 2011.

PÉREZ, J.; MUÑOZ-DORADO, J.; DE-LA-RUBIA, T.; MARTÍNEZ, J. Biodegradation and biological treatments of cellulose, hemicellulose and lignin: an overview. **International Microbiology**, v. 5, p. 53-63. 2002.

PERRY, R. H.; GREEN, D. W. **Perry's chemical engineers' handbook.** 8. ed. New York: McGraw-Hill, 2008.

PRAKASH, G.; VARMA, A. J.; PRABHUNE, A., SHOUCHE, Y.; RAO, M. Microbial production of xylitol from D-xylose and sugarcane bagasse hemicellulose 147 using newly isolated thermotolerant yeast *Debaryomyces hansenii*. **Bioresource Technology**, v. 102, p. 3304-3308, 2011.

PUMA-VILLANUEVA, W. J. **Síntese automática de redes neurais artificiais com conexões à frente arbitrárias.** 2011. 220 f. Tese (Doutorado em Engenharia Elétrica) - Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, Campinas – SP, 2011.

RAFIQUL, I. S. M.; SAKINAH, M. M. Design of process parameters for the production of xylose from wood sawdust. **Chemical Engineering Research and Design**, v.90, n. 9, p. 1307-1312, 2012.

RIEDMILLER, M.; BRAUN, H. **A direct adaptive method for faster backpropagation learning: The RPROP algorithm.** IEEE International Conference on, p. 586-591, 1993.

SÁ, L. R. V.; MOUTTA, R. O.; BON, E. P. S.; CAMMAROTA, M. C.; LEITÃO, V. S. F. Fermentative biohydrogen production using hemicellulose fractions: Analytical validation for C5 and C6-sugars, acids and inhibitors by HPLC. **International journal of hydrogen energy**, v. 40, p. 13888-13900, 2015.

SARKAR, N.; GHOSH, S.K.; BANNERJEE, S.; AIKAT, K. Bioethanol production from agricultural wastes: An overview. **Renewable Energy**, v. 37, p. 19-27, 2012.

SANTOS, R. P.; SANTIAGO, A. A. X.; GADELHA, C. A. A.; CAJAZEIRAS, J. B.; CAVADA, B. S.; MARTINS, J. L.; OLIVEIRA, T. M.; BEZERRA, G. A.; SANTOS, R. P.; FREIRE, V. N. Production and characterization of the cashew (*Anacardium occidentale* L.) peduncle bagasse ashes. **Journal of Food Engineering**, v. 79, p. 1432-1437, 2007.

SARROUH, B. F.; SILVA, S. S. Evaluation of the performance of a three-phase fluidized bed reactor with immobilized yeast cells for the biotechnological production of xylitol. **International Journal of Chemical Reactor Engineering**, v.6, p.1-15. 2008.

SILVA NETO, R. M. da. **Inspeção em indústria de beneficiamento da castanha de caju visando a implantação das boas práticas de fabricação**. 2000. 128 f. Dissertação (Mestrado em Tecnologia de Alimentos) – Universidade Federal do Ceará, Fortaleza.

SINCLAIR, C. G.; KRISTIANSEN, B. **Fermentation kinetics and modeling**. Grã-Bretanha: Open University Press, 113 f., 1987.

SMITH, E. M. **Advances in thermal design of heat exchangers: a numerical approach: direct-sizing, step-wise rating and transients**. Wiley. 2005.

TOSETTO, G. M. **Influência da matéria-prima no comportamento cinético de levedura na produção de etanol**. Dissertação (Mestrado em Engenharia Química), Universidade Estadual de Campinas, Campinas, Brasil, 2002.

TOWLER, G.; SINNOTT, R. K. **Chemical Engineering Design: Principles, Practice and Economics of Plant and Process Design**. Elsevier. 2012.

VELTEN, K. **Mathematical Modeling and Simulation: Introduction for Scientists and Engineers**. Ilustrada. Wiley, 348 f., 2009.

VOLESKY, B.; VOTRUBA, J. **Modeling and optimization of fermentation processes**. Boca Raton: Elsevier, 1992.

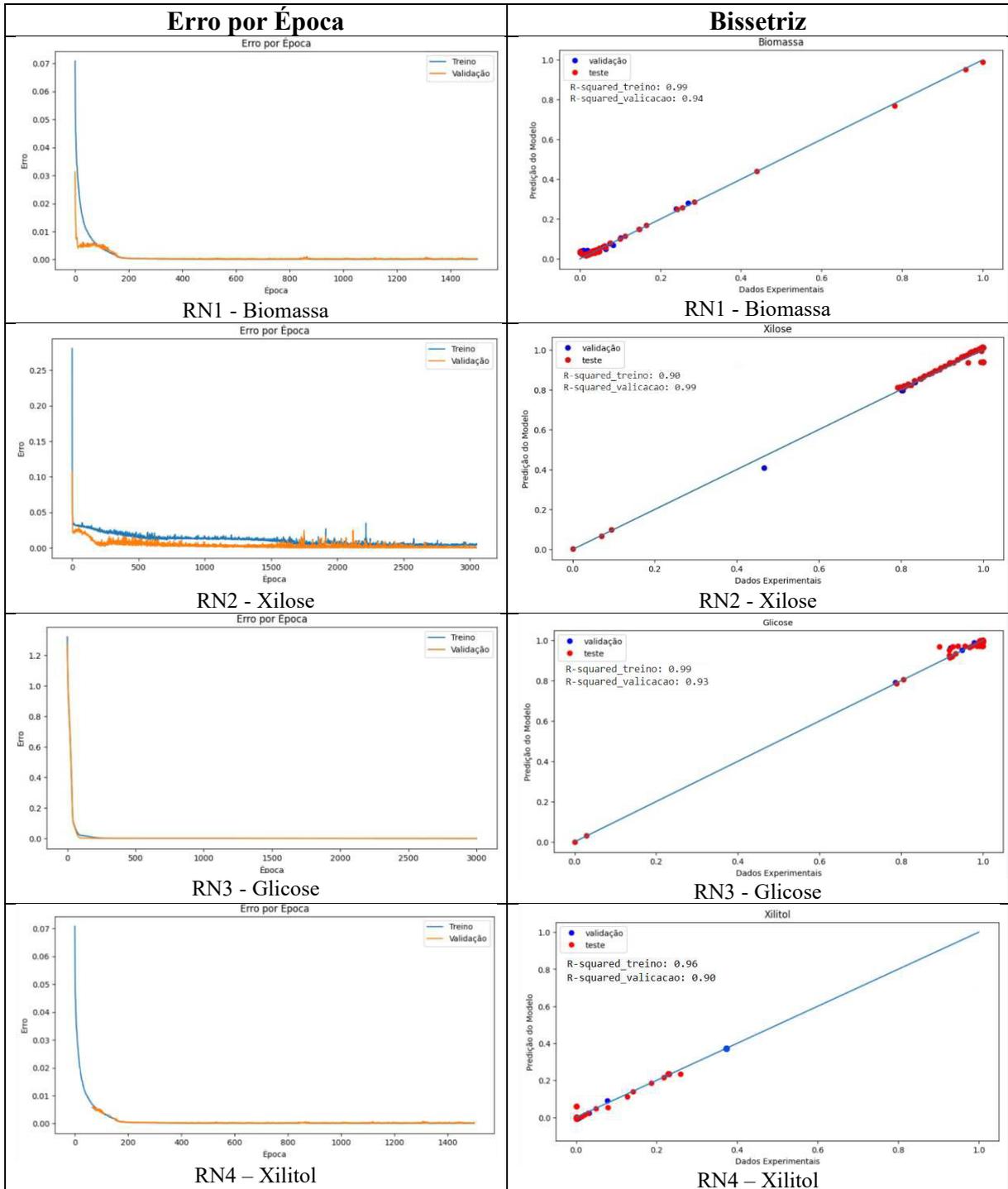
YABLOCHKOVA, E. N.; BOLOTNIKOVA, O. I.; MIKHAILOVA, N. P. The activity of xylose reductase and xylitol dehydrogenase in yeasts. **Microbiology**, v. 72, n. 4, p.414-417, 2003a.

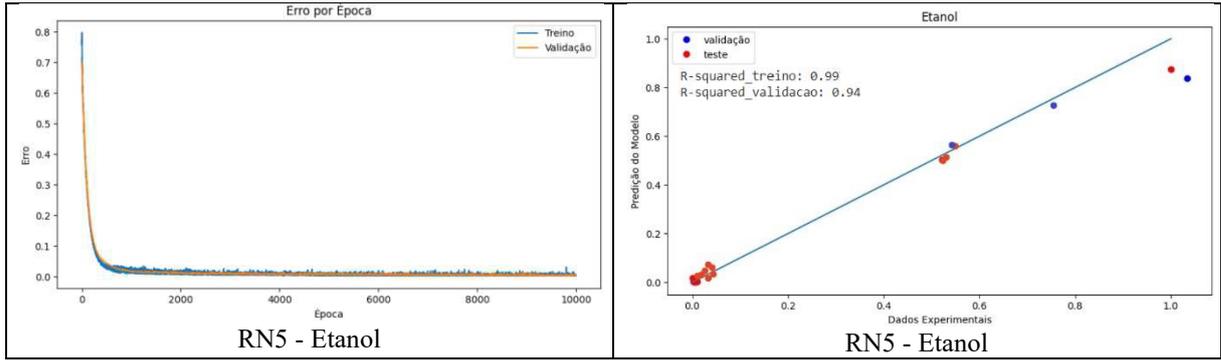
## APÊNDICE A – CONDIÇÕES OPERACIONAIS INICIAIS PARA XILOSE, GLICOSE, XILITOL E ETANOL

Biomassa			Xilose			Glicose			Xilitol			Etanol			
Número do teste	Método	Tempo (h)	Concentração Inicial (g/L)	Número do teste	Método	Tempo (h)	Concentração Inicial (g/L)	Número do teste	Método	Tempo (h)	Concentração Inicial (g/L)	Número do teste	Método	Tempo (h)	Concentração Inicial (g/L)
1	HBC-1	0	1,00719	1	HBC-1	0	18,70503	1	HBC-1	0	12,78571	1	HBC-1	0	0
2	HBC-1	24	1,07914	2	HBC-1	24	2,53237	2	HBC-1	24	0,05	2	HBC-1	24	0,07257
3	HBC-1	48	1,11511	3	HBC-1	48	0,011	3	HBC-1	48	0,05	3	HBC-1	48	0,07357
4	HBC-1	72	11,07914	4	HBC-1	72	0,01	4	HBC-1	72	0,05	4	HBC-1	72	4,56834
5	HBC-1	96	12,05036	5	HBC-1	96	0,009	5	HBC-1	96	0,05	5	HBC-1	96	6,94245
6	HBC-1	120	11,51079	6	HBC-1	120	0,008	6	HBC-1	120	0,05	6	HBC-1	120	5
7	HBC-2	0	1,13636	7	HBC-2	0	36,56051	7	HBC-2	0	30,70064	7	HBC-2	0	0
8	HBC-2	24	8,51911	8	HBC-2	24	10,06369	8	HBC-2	24	0,057	8	HBC-2	24	0,003
9	HBC-2	48	13,8535	9	HBC-2	48	2,92994	9	HBC-2	48	0,057	9	HBC-2	48	0,003
10	HBC-2	72	11,9697	10	HBC-2	72	2,92994	10	HBC-2	72	0,057	10	HBC-2	72	11,1465
11	HBC-2	96	11,7803	11	HBC-2	96	2,92994	11	HBC-2	96	0,057	11	HBC-2	96	12,61943
12	HBC-2	120	11,89394	12	HBC-2	120	2,92994	12	HBC-2	120	0,057	12	HBC-2	120	12,95828
13	HBC-3	0	1,06461	13	HBC-3	0	36,48094	13	HBC-3	0	30,08157	13	HBC-3	0	0
14	HBC-3	24	7,41557	14	HBC-3	24	18,47507	14	HBC-3	24	7,62463	14	HBC-3	24	0,05
15	HBC-3	48	9,98532	15	HBC-3	48	5,10264	15	HBC-3	48	0,32626	15	HBC-3	48	0,06
16	HBC-3	72	12,15125	16	HBC-3	72	2,7566	16	HBC-3	72	0,013	16	HBC-3	72	0,06392
17	HBC-3	96	12,3348	17	HBC-3	96	2,58065	17	HBC-3	96	0,013	17	HBC-3	96	12,37151
18	HBC-3	120	12,81204	18	HBC-3	120	2,52199	18	HBC-3	120	0,013	18	HBC-3	120	12,36138
19	HBC-4	0	0,93795	19	HBC-4	0	35,78644	19	HBC-4	0	30,0452	19	HBC-4	0	0
20	HBC-4	24	7,82828	20	HBC-4	24	12,75613	20	HBC-4	24	0,05	20	HBC-4	24	0,05
21	HBC-4	48	10,20924	21	HBC-4	48	4,09812	21	HBC-4	48	0,05	21	HBC-4	48	5,15873
22	HBC-4	72	12,44589	22	HBC-4	72	2,19336	22	HBC-4	72	0,05	22	HBC-4	72	5,91631
23	HBC-4	96	12,59098	23	HBC-4	96	2,5974	23	HBC-4	96	0,05	23	HBC-4	96	7,46753
24	HBC-4	120	12,95488	24	HBC-4	120	3,11688	24	HBC-4	120	0,05	24	HBC-4	120	8,11111



**APÊNDICE C – ERROS POR ÉPOCA E BISSETRIZES DA PREDIÇÃO COM REDES NEURAIS PARA CADA SUBSTÂNCIA**





RN5 - Etanol

RN5 - Etanol

## APÊNDICE D – GRÁFICOS DE REGRESSÃO NÃO-LINEAR PARA CADA SUBSTÂNCIA

