



UNIVERSIDADE FEDERAL DE SÃO CARLOS  
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA QUÍMICA

ALEX RIBEIRO DE PAULA

**Monitoramento da Hidrólise Enzimática de Bagaço de  
Cana-de-Açúcar por Métodos Espectrofotométricos**

Orientador: Professor Marcelo Perencin de Arruda Ribeiro Ph.D.

São Carlos – SP

2020

Alex Ribeiro de Paula

# **Monitoramento da Hidrólise Enzimática de Bagaço de Cana-de-Açúcar por Métodos Espectrofotométricos**

Exame de dissertação de mestrado no  
Programa de Pós-Graduação em  
Engenharia Química, Universidade  
Federal de São Carlos.

Área de Pesquisa: Simulação e Controle de Processos Químicos  
Orientador: Professor Marcelo Perencin de Arruda Ribeiro Ph.D.

São Carlos – SP

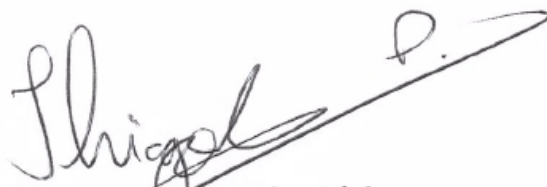
2020

MEMBROS DA BANCA EXAMINADORA DA DEFESA DE DISSERTAÇÃO DE ALEX RIBEIRO DE PAULA APRESENTADA AO PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA QUÍMICA DA UNIVERSIDADE FEDERAL DE SÃO CARLOS, EM 23 DE MARÇO DE 2020.

BANCA EXAMINADORA:



Marcelo Perencin de Arruda Ribeiro  
**Orientador, UFSCar**



Thiago Faggion de Pádua  
**UFSCar**



Marcel Otávio Cerri  
**UNESP**

## DEDICATÓRIA

*Dedico o presente trabalho para os meus pais, em retribuição a todo amor, carinho e apoio que me deram. Sem eles nada disso seria possível. A Deus por me dar forças, coragem e sabedoria para enfrentar todas as dificuldades da vida, questionar realidades, e assim propor sempre um novo mundo de possibilidades.*



## **AGRADECIMENTOS**

*A meu orientador Dr. Marcelo Perencin de Arruda Ribeiro pelo imenso suporte e orientação. Pela paciência e por ser um exemplo pessoal e profissional.*

*Ao patrocínio das agências CAPES (Código de financiamento: 001), CNPq (Número do processo: 130253/2019-7) e FAPESP (Número do processo: 2016 / 10636-8).*

*A todos amigos que me deram suporte presencialmente e virtualmente durante toda essa jornada.*

*E a todos aqueles que participaram diretamente ou indiretamente da realização deste trabalho.*

## RESUMO

O etanol de 2ª geração (2G), produzido a partir do bagaço de cana-de-açúcar, é uma maneira promissora de aumentar a produção de etanol sem aumentar a área de cultivo. No entanto, o processo vem sendo estudado para se tornar economicamente viável. Esse fato decorre de vários fatores, entre eles: extração de compostos fermentativos a partir do bagaço da cana, separação de inibidores enzimáticos e grande demanda de energia para a agitação da mistura. Nesse processo, o bagaço de cana é utilizado como substrato e deve ser submetido ao processo de pré-tratamento para expor as fibras e facilitar a hidrólise da celulose em glicose. Embora existam várias maneiras de realizar a hidrólise do bagaço pré-tratado, a rota enzimática reduz a quantidade de inibidores produzidos pelas rotas químicas tradicionais, que afetam adversamente o processo de fermentação. Por outro lado, o complexo enzimático é caro, logo aumenta o custo do processo. Assim, o projeto do sistema de reação e a definição das condições do processo devem ser analisados para maximizar a produtividade específica da enzima. Um sistema de reatores não convencional foi construído no Laboratório de Automação e Desenvolvimento de Bioprocessos (LADaBio – UFSCar) para contornar os problemas discutidos anteriormente. O sistema reacional é complexo e a tentativa de usar modelos fenomenológicos para descrevê-lo tornaria o problema intratável. Portanto, modelos cinéticos simplificados foram desenvolvidos para este sistema. Devido ao desvio esperado entre modelo e reação, um sistema de monitoramento utilizando a espectroscopia na região do infravermelho próximo e do ultravioleta visível, aliado ao método dos mínimos quadrados parciais e validação cruzada, foi estudado para monitorar a reação. Dados de quatro ensaios realizados foram utilizados para ajustar modelos matemáticos de predição das concentrações de celobiose, glicose e xilose. Por se tratar de um sistema complexo, concluiu-se que os modelos gerados foram capazes de prever as concentrações dos carboidratos com uma ótima precisão.

Palavras chaves: Etanol 2G, Modelagem de processos, Monitoramento de hidrólise enzimática, Bagaço de cana-de-açúcar, Reator Semicontínuo.

## ABSTRACT

The production process of 2nd generation ethanol (2G), using sugarcane bagasse, is a promising way to increase ethanol production without increasing the cultivation area. However, the process has been studied to become economically viable. This fact stems from several factors, among them: extraction of fermentative compounds from sugarcane bagasse, separation of enzymatic inhibitors and the great energy demand for the agitation of the mixture. In this process, sugarcane bagasse is used as a substrate, and must undergo the pretreatment process to expose the fibers and facilitate the hydrolysis of the cellulose into glucose. Although there are several ways of performing the hydrolysis of the pretreated bagasse, the enzymatic route reduces the amount of inhibitors produced by the traditional chemical routes, which adversely affect the fermentation process. On the other hand, the enzyme complex is expensive, thus increasing the cost of the process. Thus, the design of the reaction system and the definition of the process conditions must be analyzed to maximize the specific productivity of the enzyme. A non-conventional reactors system has been built in the Bioprocess Automation and Development Laboratory (LADaBio – UFSCar) to work around the issues discussed earlier. The reactionary system is complex and the attempt to use phenomenological models to describe it would make the problem intractable. Therefore, simplified kinetic models have been developed for this system. Due to the expected deviation between model and reaction, a monitoring system using spectroscopy in the near infrared and visible ultraviolet region, combined with the method of partial least squares and cross validation, was studied to monitor the reaction. Data from four assays performed were used to adjust mathematical models for predicting the concentrations of cellobiosis, glucose and xylose. As it is a complex system, it was concluded that the models generated were able to predict carbohydrate concentrations with great precision.

Keywords: 2G Ethanol, Process Modeling, Enzymatic Hydrolysis Monitoring, Sugarcane Bagasse, Semicontinuous Reactor.

## SUMÁRIO

<b>1. INTRODUÇÃO E JUSTIFICATIVA.....</b>	<b>11</b>
<b>2. OBJETIVO.....</b>	<b>14</b>
2.1. OBJETIVO GERAL.....	14
2.2. OBJETIVOS ESPECÍFICOS.....	14
2.2.1. Amostragem.....	14
2.2.2. Pré-tratamento dos dados.....	14
2.2.3. Modelagem.....	14
<b>3. REVISÃO BIBLIOGRÁFICA.....</b>	<b>15</b>
3.1. MATERIAIS LIGNOCELULÓSICOS.....	15
3.2. PRÉ-TRATAMENTO.....	15
3.3. HIDRÓLISE ENZIMÁTICA.....	17
3.4. ESPECTROSCOPIA DE ABSORÇÃO NA REGIÃO DO INFRAVERMELHO PRÓXIMO (NIR).....	19
3.5. ESPECTROSCOPIA DE ABSORÇÃO NA REGIÃO DO ULTRAVIOLETA E VISÍVEL (UV-VIS).....	20
3.6. CALIBRAÇÃO MULTIVARIADA.....	20
3.7. PRÉ-TRATAMENTO DOS DADOS ESPECTROMÉTRICOS.....	23
3.7.1. Derivadas.....	24
3.7.2. Média Móvel.....	24
3.7.3. <i>Locally Estimated Scatterplot Smoothing</i> .....	24
<b>4. MATERIAIS E MÉTODOS.....</b>	<b>26</b>
4.1. MATERIAL LIGNOCELULÓSICO, TRATAMENTO HIDROTÉRMICO E COMPLEXO ENZIMÁTICO.....	26
4.2. DETERMINAÇÃO DAS CONCENTRAÇÕES DE REFERÊNCIA.....	27
4.3. ENSAIOS.....	27
4.4. MONITORAMENTO DO PROCESSO.....	28
4.4.3. Modelagem matemática.....	30
<b>5. RESULTADOS E DISCUSSÃO.....</b>	<b>31</b>
<b>6. CONCLUSÕES.....</b>	<b>51</b>
<b>7. REFERÊNCIAS.....</b>	<b>52</b>

**ANEXO A – GRÁFICOS DE PREDIÇÕES DAS CONCENTRAÇÕES DE  
CELOBIOSE, GLICOSE E XILOSE PELOS MODELOS PLS NIR E UV-VIS.....59**

## LISTA DE ILUSTRAÇÕES

Figura 1: Esquema do aparato experimental utilizado. Adaptado de FURLONG, 2015.....	13
Figura 2: Sinergismo das celulasas na hidrólise da celulose. EG: Endoglicanases, CBH I e II: Celobiohidralases, MOP: Monooxigenases e bG: $\beta$ -glicosidades. Adaptado de FURLONG, 2019.....	18
Figura 3: Espectrofotômetro Frontier NIR - Perkin Elmer®.....	29
Figura 4: Espectrofotômetro Lambda 456 da Perkin Elmer®.....	30
Figura 5: Gráfico dos dados espectrais brutos (a) e selecionados (b) no NIR obtidos de amostras coletadas de todos ensaios realizados.....	32
Figura 6: Gráfico dos dados espectrais brutos (a) e selecionados (b) no UV-Vis obtidos de amostras coletadas de todos ensaios realizados.....	32
Figura 7: Gráficos dos erros associados (RMSECV e RMSEcal) ao número de regressores utilizados no teste da validação cruzada para cada carboidrato (Celobiose, Glicose e Xilose) no ajuste dos dados espectrais UV-Vis. Condições do ajuste descritas na tabela 4 pelo modelo 1.....	35
Figura 8: Gráfico dos espectros selecionados e pré-tratados do NIR, modelos 1 (a) e 2 (b) para as condições de pré-tratamento: janela de alisamento de 11 pontos, sem derivada e faixa de número de onda de 4000 a 10000 $\text{cm}^{-1}$ (modelo 2 - b).....	36
Figura 9: Gráfico dos espectros selecionados e pré-tratados do UV-Vis, modelos 1 (a) e 2 (b) para as condições de pré-tratamento: janela de alisamento de 11 pontos, sem derivada, faixa de comprimento de onda de 450 a 900 nm (modelo 2- b) e absorvância máxima de 1.2 (modelo 2 - b).....	36
Figura 10: Gráfico da predição da absorvância pelo modelo CLS aplicado aos dados utilizados no ajuste dos modelos 1 e 2 do NIR.....	37
Figura 11: Gráfico da predição da absorvância pelo modelo CLS aplicado aos dados utilizados no ajuste dos modelos 1 e 2 do UV-Vis.....	37
Figura 12: Gráficos das concentrações preditas de cada carboidrato, com a barra do desvio padrão de previsão na CV (RMSECV), dos modelos PLS 1 e 2 do NIR, dado a bissetriz de referência. No eixo x tem-se as concentrações aferidas pelo HPLC (referência, g/L) e no eixo y as concentrações preditas pelos modelos (g/L) para a Celobiose (a), Glicose (b) e Xilose (c). .....	39
Figura 13: Gráficos das concentrações preditas de cada carboidrato, com a barra do desvio padrão de previsão na CV (RMSECV), dos modelos PLS 1 e 2 do UV-Vis, dado a bissetriz de	

referência. No eixo x tem-se as concentrações aferidas pelo HPLC (referência, g/L) e no eixo y as concentrações preditas pelos modelos (g/L) para a Celobiose (a), Glicose (b) e Xilose (c). .....	39
Figura 14: Curva de calibração da celobiose, glicose e xilose utilizando as técnicas NIR e UV-Vis.....	40
Figura 15: Gráfico dos espectros selecionados e pré-tratados do NIR, modelo 3, para as condições de pré-tratamento: LOESS (alpha de 0,2, order 1), sem derivada e faixa de número de onda de 1100 a 10000 $\text{cm}^{-1}$ .....	41
Figura 16: Gráfico dos espectros selecionados e pré-tratados do NIR, modelo 4, para as condições de pré-tratamento: LOESS (alpha de 0,2, order 2), sem derivada e faixa de número de onda de 3500 a 8000 $\text{cm}^{-1}$ .....	42
Figura 17: Gráfico dos espectros selecionados e pré-tratados do UV-Vis, modelo 3, para as condições de pré-tratamento: LOESS (alpha de 0,2, order 1), sem derivada e faixa de comprimento de onda de 190 a 1100 nm.....	42
Figura 18: Gráfico dos espectros selecionados e pré-tratados do UV-Vis, modelo 4, para as condições de pré-tratamento: LOESS (alpha de 0,2, order 1), primeira derivada e faixa de comprimento de onda de 190 a 500 nm.....	42
Figura 19: Regiões de concentrações quantificadas pelo HPLC (g/L) de cada carboidrato (celobiose, glicose e xilose). Os dados ajustados estão representados por C.vc, G.vc e X.vc, e os dados não ajustados, representados por C.ve, G.ve e X.ve. Na primeira coluna tem-se os dados ajustados e não ajustados do NIR (a, b, c) e na segunda coluna do UV-Vis (d, e, f).....	46
Figura 20: Predição das concentrações de celobiose, glicose e xilose pelos modelos PLS NIR 1 e 2 (a, b, c), 3 e 4 (d, e, f) - Batelada alimentada com bagaço 48h (dados ajustados).....	47
Figura 21: Predição das concentrações de celobiose, glicose e xilose pelos modelos PLS NIR 1 e 2 (a, b, c), 3 e 4 (d, e, f) - Batelada alimentada com bagaço 48h (validação externa).....	48
Figura 22: Predição das concentrações de celobiose, glicose e xilose pelos modelos PLS UV-Vis 1 e 2 (a, b, c), 3 e 4 (d, e, f) - Batelada alimentada com bagaço 48h (dados ajustados).....	48
Figura 23: Predição das concentrações de celobiose, glicose e xilose pelos modelos PLS UV-Vis 1 e 2 (a, b, c), 3 e 4 (d, e, f) - Batelada alimentada com bagaço 48h (validação externa)..	49
Figura 24: Predição das concentrações de celobiose, glicose e xilose pelos modelos PLS NIR 1 e 2 (a, b, c), 3 e 4 (d, e, f) - Batelada 48h (dados ajustados).....	60
Figura 25: Predição das concentrações de celobiose, glicose e xilose pelos modelos PLS NIR 1 e 2 (a, b, c), 3 e 4 (d, e, f) - Batelada 48h (validação externa).....	60

Figura 26: Predição das concentrações de celobiose, glicose e xilose pelos modelos PLS NIR 1 e 2 (a, b, c), 3 e 4 (d, e, f) - Batelada alimentada com enzimas 48h (dados ajustados).....	61
Figura 27: Predição das concentrações de celobiose, glicose e xilose pelos modelos PLS NIR 1 e 2 (a, b, c), 3 e 4 (d, e, f) - Batelada alimentada com enzimas 48h (validação externa).....	61
Figura 28: Predição das concentrações de celobiose, glicose e xilose pelos modelos PLS UV-Vis 1 e 2 (a, b, c), 3 e 4 (d, e, f) - Batelada 12h (dados ajustados).....	62
Figura 29: Predição das concentrações de celobiose, glicose e xilose pelos modelos PLS UV-Vis 1 e 2 (a, b, c), 3 e 4 (d, e, f) - Batelada 12h (validação externa).....	62
Figura 30: Predição das concentrações de celobiose, glicose e xilose pelos modelos PLS UV-Vis 1 e 2 (a, b, c), 3 e 4 (d, e, f) - Batelada 48h (dados ajustados).....	63
Figura 31: Predição das concentrações de celobiose, glicose e xilose pelos modelos PLS UV-Vis 1 e 2 (a, b, c), 3 e 4 (d, e, f) - Batelada 48h (validação externa).....	63
Figura 32: Predição das concentrações de celobiose, glicose e xilose pelos modelos PLS UV-Vis 1 e 2 (a, b, c), 3 e 4 (d, e, f) - Batelada alimentada com enzimas 48h (dados ajustados)..	64
Figura 33: Predição das concentrações de celobiose, glicose e xilose pelos modelos PLS UV-Vis 1 e 2 (a, b, c), 3 e 4 (d, e, f) - Batelada alimentada com enzimas 48h (validação externa). .....	64



## LISTA DE TABELAS

Tabela 1: Principais tipos de pré-tratamento e seus efeitos.....	16
Tabela 2: Ensaio da hidrólise enzimática.....	27
Tabela 3: Modelos ajustados NIR.....	33
Tabela 4: Modelos ajustados UV-Vis.....	34
Tabela 5: Todos os modelos ajustados pelos dados obtidos no NIR.....	44
Tabela 6: Todos modelos ajustados pelos dados obtidos no UV-Vis.....	45

## 1. INTRODUÇÃO E JUSTIFICATIVA

Novas alternativas de se obter uma fonte de energia renovável são um dos meios para contornar o problema do aquecimento global e esgotamento das fontes de energia convencionais. No Brasil, o etanol obtido pelo processo de primeira geração (1G) da cana-de-açúcar foi uma das alternativas renováveis utilizadas para comutar o uso de combustíveis fósseis. Nesse processo, o caldo extraído da cana-de-açúcar é fermentado e destilado, para obtenção do etanol. Durante o processo 1G, a biomassa, composta pelo bagaço da cana-de-açúcar e outros materiais lignocelulósicos, é queimada para produzir calor e gerar energia para os processos da planta industrial. Contudo, esses materiais lignocelulósicos também podem ser utilizados como matéria-prima para a produção de mais etanol, pelo processo de segunda geração (processo 2G) (DIAS et al., 2011).

No Brasil, a utilização de um processo 2G pode ser vantajoso para indústria, pois pode ser integrado ao processo 1G aproveitando ainda mais a matéria-prima já disponível, gerando mais produto e reduzindo o custo do processo. Além disso, o processo 2G não compete com a produção de alimentos pelo fato de utilizar como matéria-prima os materiais lignocelulósicos, como o bagaço de cana-de-açúcar, que possuem um baixo custo e é abundante (DIAS et al., 2012).

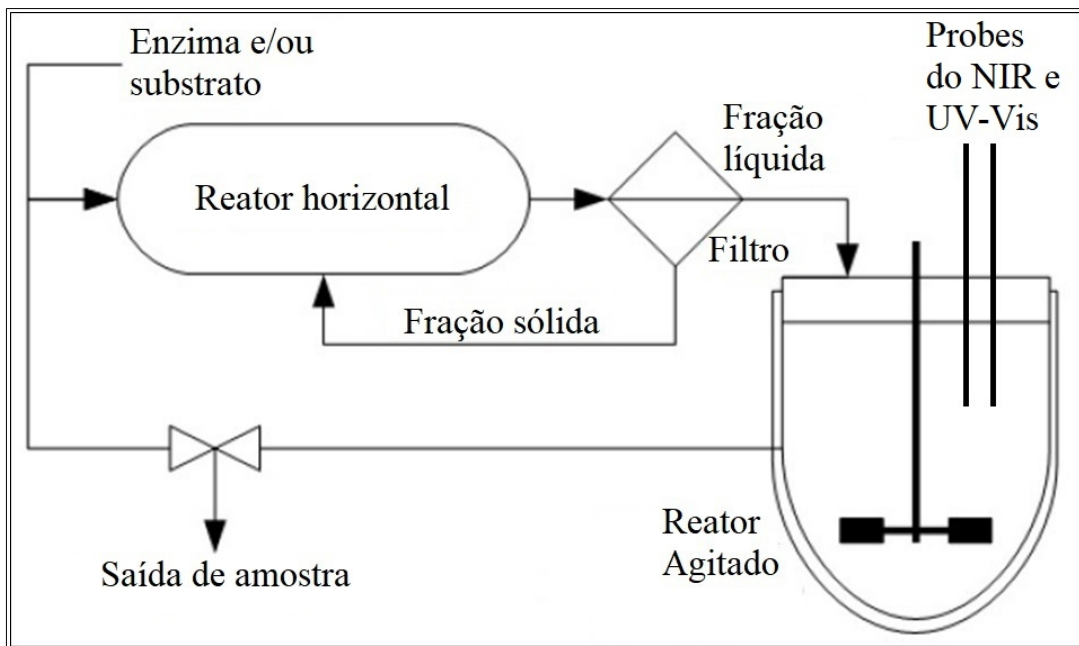
O processo de produção de etanol de segunda geração é basicamente dividido nas etapas de pré-tratamento do material lignocelulósico, hidrólise dos compostos lignocelulósicos e fermentação. O pré-tratamento desorganiza a estrutura das fibras, separando os compostos para que sejam hidrolisados posteriormente. Dentre os diversos tipos de pré-tratamento, o tratamento hidrotérmico é abordado neste trabalho, pois é uma técnica fácil e barata de operar. Por utilizar água no estado líquido entre as temperaturas de 180 °C e 240 °C, em um reator pressurizado sob um tempo determinado, as fibras da biomassa sofrem modificações, aumentando a digestibilidade da celulose e liberando uma baixa concentração de compostos inibitórios da etapa de fermentação (KIM; MOSIER; LADISCH, 2009; NEGRO et al., 2003). Além disso, como não é necessária a utilização de produtos químicos adicionais, o efluente gerado nesse tratamento é facilmente tratado quando comparado às outras alternativas de pré-tratamento.

Há dois principais tipos de hidrólise empregados na biomassa, por solução ácida ou enzimas. Assim como no pré-tratamento, visando a futura aplicação industrial, vem sendo

abordado nesse trabalho a hidrólise enzimática, pois nesse processo há uma baixa produção de furfural e hidroximetilfurfural, e também por ter um efluente facilmente tratado. Esses compostos são provindos da degradação de pentoses e hexoses a temperaturas altas (acima de 200 °C) e são inibidores do processo de hidrólise e fermentação. Além disso, a hidrólise enzimática ocorre em condições amenas de temperatura, pressão e pH, diferentemente da hidrólise ácida, que utiliza soluções concentradas de ácidos (0,5 m/m%). A maior desvantagem da hidrólise enzimática é o custo do complexo enzimático. Entretanto, ao se comparar com a hidrólise química, na hidrólise enzimática, há uma menor formação de inibidores e não utiliza ácido concentrado no processo, tornando o processo sustentável (LIMAYEM; RICKE, 2012).

Na etapa de fermentação, geralmente se alimenta açúcares a uma concentração de aproximadamente 180 g.L<sup>-1</sup>. Essa concentração é comum no processo 1G, logo a sacarose obtida no processo 2G pode ser misturada com a do 1G, sem afetar o rendimento das leveduras, o que também diminui a demanda de energia e custo do processo (FURLAN et al., 2013). Para se atingir essa concentração de glicose, o reator deve ser alimentado com uma grande concentração de substrato. Porém, com uma alta concentração de substrato, a enzima pode sofrer inibição por substrato, o que minimiza o rendimento do processo (KRISTENSEN; FELBY; JØRGENSEN, 2009). Além disso, a alta carga de sólidos causa uma alteração na densidade aparente e na viscosidade do fluido dentro do reator, o que demanda ainda mais energia para a agitação do meio, quando comparado ao mesmo reator sendo operado com uma carga de sólidos baixa (FURLONG, 2019).

A utilização de um reator horizontal operando em batelada alimentada com alta fração de sólidos, seguido em série por reator do tipo tanque agitado para a fração líquida foi uma proposta desenvolvida durante o trabalho de doutorado de Vitor Badiale Furlong (PPGEQ-UFSCar) para contornar os problemas discutidos anteriormente (FURLONG, 2015). O reator do tipo tambor rotativo, por ser rotacionado como um todo, não demanda alta potência de agitação e isso atrelado ao regime de batelada alimentada aumenta o rendimento global do processo (DU et al., 2014). Já o reator agitado concentra os compostos solúveis e separa a fração sólida da líquida, o que contribui para o aumento da produtividade e conversão do substrato em produto. Para aumentar mais a conversão, esse esquema é operado em reciclo, o que também resolve a possível lixiviação das enzimas da parte sólida (Figura 1).



**Figura 1:** Esquema do aparato experimental utilizado. Adaptado de FURLONG, 2015.

Por ser um sistema complexo, com inúmeras reações e variáveis, sua descrição utilizando modelos puramente fenomenológicos tornaria o problema intratável. Modelos matemáticos simplificados do processo podem ser utilizados para controle e otimização do processo. Porém, um desvio entre o modelo e o processo seria esperado. Para contornar esse problema, um sistema de monitoramento em tempo real pode ser construído para monitorar o processo, averiguar grandes desvios e controlar a reação. A utilização da espectroscopia na região do infravermelho próximo (*Near Infrared*, NIR) e do ultravioleta visível (*Ultraviolet-visible*, UV-Vis) podem ser de grande utilidade para a caracterização quantitativa e qualitativa dos compostos orgânicos (GOUVEIA et al., 2009; KLINE et al., 2010; SLUITER et al., 2008). O monitoramento através de sondas de fibra ótica permitiria uma análise rápida, sem a destruição ou a necessidade de preparação das amostras. Para a interpretação dos dados medidos, procedimento de calibração multivariada deve ser realizado, utilizando dados de referência quantificados por técnicas analíticas como a cromatografia líquida de alta eficiência (*High Performance Liquid Chromatography*, HPLC).

A utilização da calibração multivariada dos dados espectroscópicos NIR e UV-Vis para criação de modelos preditivos e monitoramento de bioprocessos é bem documentado na literatura (BLANCO et al., 2000; NASCIMENTO, 2016; PINTO; RIBEIRO; FARINAS, 2018). No entanto, não há trabalhos com foco na hidrólise enzimática do bagaço de cana-de-açúcar.

## **2. OBJETIVO**

### **2.1. OBJETIVO GERAL**

O objetivo do presente trabalho foi de utilizar técnicas espectroscópicas (UV-Vis e NIR), associadas com o método dos mínimos quadrados parciais (*Partial Least Squares* – PLS) e validação cruzada, para monitorar os principais componentes (Celobiose, Glicose e Xilose) obtidos da hidrólise enzimática do bagaço de cana-de-açúcar pré-tratado hidrotermicamente.

### **2.2. OBJETIVOS ESPECÍFICOS**

#### **2.2.1. Amostragem**

Obter amostras e dados espectroscópicos (UV-Vis e NIR) da fração aquosa do meio reacional durante todo o intervalo da hidrólise enzimática.

#### **2.2.2. Pré-tratamento dos dados**

Analisar os dados obtidos pelas técnicas espectroscópicas e quantificadas pela cromatografia líquida, de modo identificar e remover possíveis erros sistemáticos, erros no software e ruídos espectrais.

#### **2.2.3. Modelagem**

Desenvolver modelos PLS capazes de prever as concentrações de celobiose, glicose e xilose, durante a hidrólise enzimática do bagaço de cana-de-açúcar pré-tratado hidrotermicamente. Validar os modelos obtidos utilizando validação externa (utilização de dados não utilizados no ajuste dos modelos PLS).

### 3. REVISÃO BIBLIOGRÁFICA

#### 3.1. MATERIAIS LIGNOCELULÓSICOS

As biomassas lignocelulósicas têm uma composição variada dependendo do vegetal de origem, porém as frações principais de composição são celulose (25-50%), hemicelulose (20-30%) e lignina (10-25%), que podem ser direcionadas a produção de biocombustíveis, produtos químicos e energia (KUMAR et al., 2009).

A celulose é um polímero rígido, de conformação linear, com fórmula empírica  $(C_6H_{10}O_5)_n$ , onde resíduos de glicose se ligam gerando interações moleculares que dificultam a quebra desse polissacarídeo (HAMELINCK; VAN HOOIJDONK; FAAIJ, 2005). Dependendo de como essas interações moleculares ocorrem, algumas regiões da molécula de celulose podem ser cristalinas (resistente a hidrólise) ou amorfa (suscetível a sacarificação) (FURLONG, 2019).

Já a estrutura da hemicelulose é composta por hexoses,  $C_6$ , (galactose, glicose e manose), pentoses,  $C_5$ , (arabinose e xilose) e ácidos hexurônicos ( $\beta$ -D-glucurônico, D-4-O-metilglucurônico,  $\beta$ -D-galacturânico e deoxiexoses). Diferente da celulose, a hemicelulose não possui regiões cristalinas, sendo assim facilmente retirada da fonte natural (EDLUND; ALBERTSSON, 2008; YOSHIDA et al., 2008).

Por sua vez, a lignina não possui moléculas simples de açúcar e é formada por unidades de álcool p-cumaril, álcool coniferil e álcool sinapil. Ela é responsável pela proteção da biomassa, formando uma barreira contra agentes externos (MAULIDIYAH et al., 2017).

#### 3.2. PRÉ-TRATAMENTO

O pré-tratamento da biomassa lignocelulósica aumenta sua porosidade e desconstrói os biopolímeros, liberando celulose, hemicelulose e lignina, assim como diminui a cristalinidade da celulose. Este processo deve aumentar o rendimento de formação de açúcares e evitar a formação de inibidores que afetam a etapa de hidrólise. Na literatura existem diversas metodologias para o pré-tratamento da biomassa, como pode ser observado na Tabela 1 (GANDLA; MARTÍN; JÖNSSON, 2018; MOSIER et al., 2005; RASTOGI; SHRIVASTAVA, 2017). No presente trabalho, vem sendo abordado o tratamento hidrotérmico.

**Tabela 1:** Principais tipos de pré-tratamento e seus efeitos.

Pré-tratamento	Ação	Vantagens	Desvantagens	Remoção de hemicelulose	Remoção de lignina	Referência
Mecânico	Moagem	Aumento da área superficial	Alto consumo de energia	-	-	SINGH et al, 2014; MAURYA et al., 2015.
Água líquida quente / Hidrotérmico	Reator pressurizado com água líquida para uma ampla faixa de tempos de retenção	Baixa produção de inibidores e baixo impacto ambiental	Alto consumo de água e energia	Alto	Baixo	RUIZ et al., 2013
Explosão de vapor	Exposição ao vapor por curto tempo de retenção	Aumento da área superficial, baixa exigência de água e impacto ambiental	Alta produção de inibidores e alto custo de instalação	Alto	Baixo	MARTÍN et al., 2002; MOSIER et al., 2005; RASTOGI, M., SHRIVASTAVA, S., 2017.
Ácido diluído	Reatores pressurizados e não pressurizados com soluções ácidas diluídas	Alta despolimerização e hidrólise da hemicelulose	Alto consumo de água e alto custo de tratamento de efluente	Alto	Baixo	VON SIVERS, M., ZACCHI, G., 1995; MOSIER et al., 2005; YANG, B., WYMAN, C. E., 2008; MAURYA et al., 2015.
Alcalino	Reatores pressurizados e não pressurizados com soluções básicas diluídas	Alta despolimerização e facilita a recuperação de lignina	Alto consumo de água e alto custo de tratamento de efluente	Baixo	Alto	MCINTOSH, S., VANCOV, T., 2010; BHATIA et al., 2012.
Organosolvente	Pressurizado e não pressurizado reatores com compostos orgânicos e polares	Alta despolimerização da hemicelulose e lignina	Alto custo de recuperação de solvente, risco de incêndio e explosões	Baixo	Alto	DA SILVA et al., 2018; GANDLA et al., 2018.

No pré-tratamento hidrotérmico a biomassa é colocada em um reator, com pressão elevada para manter a água no estado líquido, por um período de tempo que varia de 1-10 min a alta temperatura, entre 180 e 240 °C. Com isso, ocorre a liberação de ácidos da estrutura da hemicelulose, pois as ligações o-acetil se rompem e assim tais ácidos potencializam a hidrólise da própria hemicelulose em monossacarídeos. Além disso, os açúcares monoméricos liberados podem sofrer degradação térmica formando aldeídos, como furfural e 5-hidroxiacetilfurfural, que apresentam ação inibitória na etapa de fermentação (PALMQVIST; HAHN-HÄGERDAL, 2000). Contudo, a liberação de compostos inibitórios é mínima, quando trabalhado em condições propícias (tempo de 5 a 30 min e temperatura de 190 a 220 °C) (BIOCHEMISTRY, 1996; HONGDAN; SHAOHUA; SHUBIN, 2013; YU et al., 2013), onde se aumenta o rendimento da sacarificação da celulose e hemicelulose (YU et al., 2010).

### 3.3. HIDRÓLISE ENZIMÁTICA

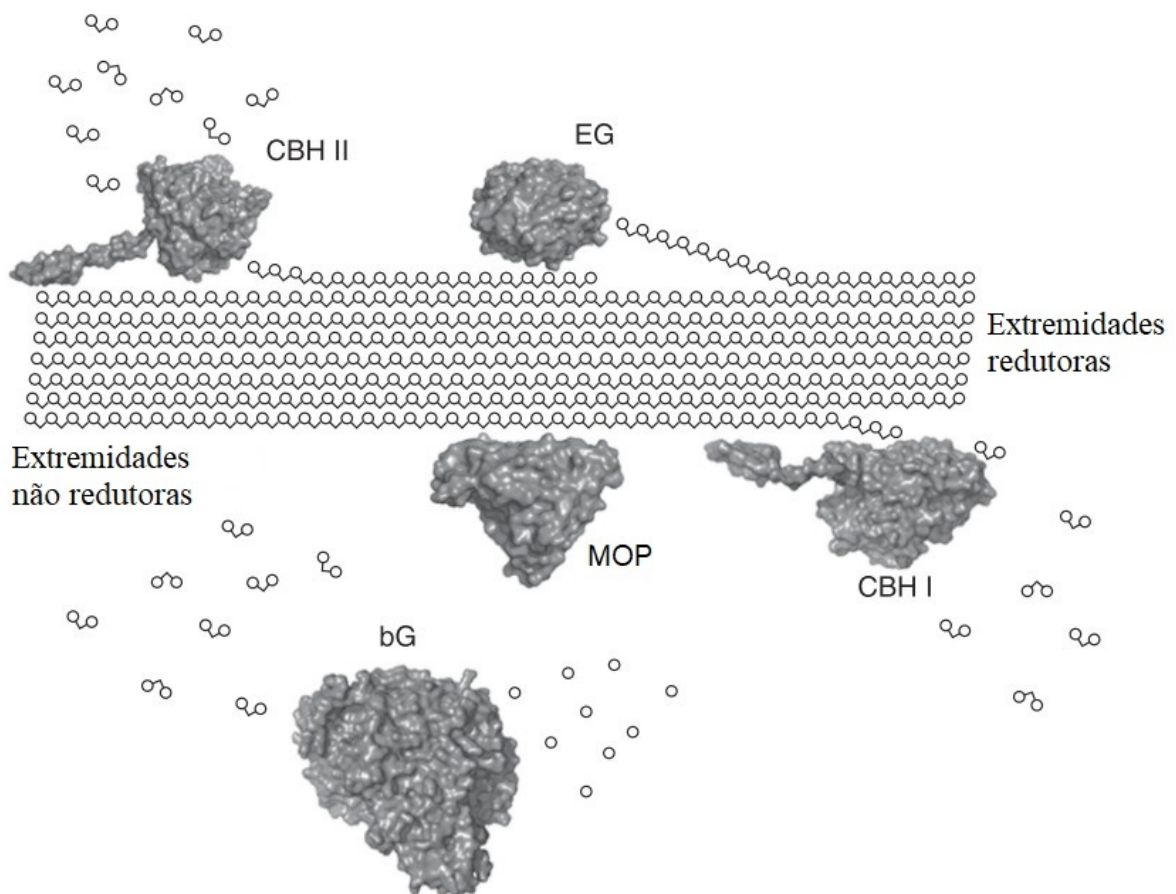
Após o pré-tratamento do bagaço, o material lignocelulósico se torna mais suscetível à hidrólise. Como esse material possui uma alta complexidade, um coquetel de enzimas é utilizado na hidrólise enzimática, que possui diversas enzimas, cada uma interagindo com uma parte do substrato lignocelulósico.

As celulasas são as principais enzimas que interagem com a celulose e hidrolisam os produtos intermediários da mesma. Elas são divididas em quatro grupos, que compõem uma solução típica de enzimas utilizadas na etapa de hidrólise enzimática (Figura 2) (FURLONG, 2019; SÁNCHEZ, 2009):

- Endoglicanases (1,4- $\beta$ -glicanohidrolase, abreviação EG): agem no interior da cadeia da celulose, onde a cristalinidade é baixa, liberando oligômeros;
- Celobiohidrolases ou Exoglicanases (exo-1,4- $\beta$ -glicanase, abreviação CBH): responsáveis pela liberação de  $\beta$ -celobiose (dissacarídeos) por meio da degradação de oligômeros. São subdivididas em celobiohidrolases I e II (CBH I e CBH II), responsáveis pela hidrólise das extremidades redutoras e não redutoras, respectivamente, da celulose (PRATTO, 2015);
- $\beta$ -glicosidades (bG): quebram os dissacarídeos formando assim duas moléculas de glicose;



- Monooxigenases de polissacarídeos (abreviação MOP): ainda há debates sobre o mecanismo de reação, porém sabe-se que os MOP inserem oxigênio molecular nas ligações carbono-hidrogênio, desestabilizando-as e levando à separação;
- Xilanases (XN): responsáveis pela hidrólise das xilanas (frações hemicelulósicas) e liberação de xilooligossacarídeos (XOS). São subdividas em endoxilanases (endo-1,4-xilanase), que quebram as ligações glicosídicas da cadeia principal da xilana, tanto em regiões próximas a ramificação, quanto extremidades não redutoras, exoxilanases, que atua em extremidade redutora da xilana, e as  $\beta$ -xilosidades que degradam pequenos XOS (extremidades não redutoras) em monômeros de xilose (BALDEZ, 2018).



**Figura 2:** Sinergismo das celulases na hidrólise da celulose. EG: Endoglicanases, CBH I e II: Celobiohidralases, MOP: Monooxigenases e bG:  $\beta$ -glicosidades. Adaptado de FURLONG, 2019.

Para se alcançar as condições ótimas em uma hidrólise enzimática é necessário controlar as condições operacionais, como temperatura, pH, concentração inicial de substrato e enzima. Em geral, a temperatura de 50 °C e pH de 4,5 a 5,5 são considerados parâmetros ótimos de acordo com a literatura (COTANA et al., 2014; GODOY; MACHADO; COSTA, 2019; KOSTYLEV; WILSON, 2013; LIANG et al., 2019). Além disso, deve-se considerar o tipo de pré-tratamento e sua eficiência, a atividade enzimática e também o tipo de substrato, pois são parâmetros que influenciam diretamente o rendimento da hidrólise enzimática.

### **3.4. ESPECTROSCOPIA DE ABSORÇÃO NA REGIÃO DO INFRAVERMELHO PRÓXIMO (NIR)**

A região do infravermelho é dividida em três grupos principais, próximo (NIR), médio e distante, onde cada região é definida por um intervalo de comprimentos de ondas e possui características próprias. Nesse estudo foi abordada a região do infravermelho próximo, 13000 – 4000 cm<sup>-1</sup> (769 – 2500 nm). Associada a técnicas de quimiometria, como a calibração multivariada, a espectroscopia além de ser rápida e não destrutiva, possibilita determinar diversos componentes em meio reacional (PASQUINI, 2003).

Basicamente, a espectroscopia no infravermelho próximo atua na vibração molecular quantificando a razão entre a radiação incidente e a radiação absorvida em uma amostra. A energia da radiação eletromagnética, transmitida pelo equipamento, é relacionada com a frequência de vibração dos átomos da amostra, que pode ser calculada pela equação de Bohr:

$$E=h\cdot\nu$$

onde  $h$  é a constante de Planck ( $h = 6.626 \times 10^{-34}$  J.s) e  $\nu$  é a frequência de vibração.

Durante a vibração dos átomos, pode ocorrer a transição de elétrons entre os níveis de energia, que podem ser interpretados por modelos de oscilação harmônico (mudança entre níveis consecutivos) ou anarmônico (mudança entre níveis não adjacentes). Assim, surgem os sobretons, que são as bandas de absorção originadas nessa transição eletrônica (PASQUINI, 2003). Também podem surgir as bandas de combinação que são originadas em moléculas que possuem dois ou mais tipos de vibração, conhecidas como poliatômicas (BLANCO; VILLARROYA, 2002).

Por conseguinte, os sobretons e as bandas de combinação, contidos no NIR, podem ser usados para quantificar a concentração de compostos orgânicos, pois é possível identificar

ligações como C-H, N-H, O-H e S-H (CEN; HE, 2007), assim como grupos funcionais O-H, C-H, N-H e C=O. Portanto, a espectroscopia NIR tem o potencial de ser utilizada na caracterização da biomassa lignocelulósica.

### **3.5. ESPECTROSCOPIA DE ABSORÇÃO NA REGIÃO DO ULTRAVIOLETA E VISÍVEL (UV-VIS)**

A região do espectro eletromagnético correspondente a radiação UV-Vis, compreende no intervalo de comprimento de onda de 200 – 800 nm ( $50000 - 12500 \text{ cm}^{-1}$ ). A energia dos fótons dessa região é maior que na região do NIR e com a absorção dessa energia pode ocorrer a transição de elétrons entre os orbitais eletrônicos de uma molécula. Essa transição de elétrons é detectada pelo espectrofotômetro, gerando assim o espectro do UV-Vis (PERKAMPUS, 1992; PINTO, 2017).

A espectroscopia UV-Vis pode ser utilizada para detecção de grupos funcionais mais reativos, como, por exemplo, compostos com ligações múltiplas e aromáticos, pois esses grupos geralmente estão localizados nos níveis eletrônicos de baixa energia. Logo é possível identificar diversos compostos presentes no material lignocelulósico. Além disso, essa técnica tem o potencial de aplicação no monitoramento dos compostos de um meio reacional, quando associada a calibração multivariada (GOUVEIA et al., 2009; KLINE et al., 2010; SLUITER et al., 2008).

### **3.6. CALIBRAÇÃO MULTIVARIADA**

Os métodos de calibração multivariada podem ser aplicados para quantificação simultânea de vários analitos de uma mesma amostra. Na hidrólise de bagaço, os espectros de absorvância da biomassa compõem uma gama de compostos que são representados por diversas bandas de absorção. Logo é possível utilizar a calibração multivariada, relacionando os dados de absorção das amostras com um ou mais conjuntos de dados, na quantificação dos compostos contidos em cada amostra (ANDRADE et al., 2003).

Grande parte dos métodos de multicalibração tem como base a lei de *Lambert-Beer*, que diz que em um determinado comprimento de onda há uma dependência linear entre a

concentração de um composto e sua absorvância. Com base nisso, define-se a transmitância ( $T$ ) pela equação 2:

$$T = \frac{I}{I_0} \quad \text{Equação 2}$$

onde  $I_0$  é a intensidade de luz incidente e  $I$  a transmitida, e a lei de *Lambert-Beer* pela equação 3:

$$A = -\log(T) = \varepsilon \cdot L \cdot c \quad \text{Equação 3}$$

onde  $A$  é a absorvância,  $c$  é a concentração do analito na solução da amostra em análise,  $\varepsilon$  é a absorvância do analito e  $L$  o caminho óptico (RIBEIRO, 2007).

Dentre os métodos de multicalibração, o método dos mínimos quadrados clássicos (*classical least squares*, CLS), inversa dos mínimos quadrados (*inverse least-squares*, ILS) e mínimos quadrados parciais (*partial least squares*, PLS) foram empregados neste trabalho para interpretar os dados gerados pelos espectrofotômetros.

O modelo CLS deriva diretamente da lei de *Lambert-Beer*, representado de forma matricial pela equação 4:

$$A = CK + E_A \quad \text{Equação 4}$$

onde  $A$  é a matriz composta com  $m$  amostras e  $n$  comprimentos de onda ( $m \times n$ ) da resposta do instrumento (espectro de calibração),  $C$  é a matriz de concentrações dos analitos composta com  $m$  amostras e  $l$  componentes ( $m \times l$ ),  $K$  é a matriz de parâmetros composta por  $l$  componentes e  $n$  comprimentos de onda ( $l \times n$ ) que representam o produto entre a absorvância específica do componente puro pelo comprimento do percurso óptico e  $E_A$  é a matriz dos erros espectrais não ajustados pelo modelo ( $m \times n$ ). Os parâmetros da matriz  $K$  são ajustáveis e pode-se obter o melhor ajuste ( $K_{opt}$ ) minimizando o erro quadrado médio através da clássica solução dos mínimos quadrados, descrita na equação 5 (NELLES, 2001):

$$K_{opt} = (C^T C)^{-1} C^T A \quad \text{Equação 5}$$

Após o cálculo do  $K_{opt}$  é possível prever as concentrações ( $c_{pred}$ ) também pelos mínimos quadrados, equação 6:

$$c_{pred} = a K_{opt}^T (K_{opt} K_{opt}^T)^{-1} \quad \text{Equação 6}$$

onde  $c_{pred}$  ( $1 \times l$ ) e  $a$  ( $1 \times n$ ) são vetores linha. Para prever a concentração pela equação 5, deve-se conhecer todos os componentes químicos contidos na amostra, assim como a calibração destes na parte do espectro de interesse, o que não torna vantajoso (HAALAND; THOMAS, 1988).

Observa-se que na equação 4, as concentrações de cada analito são as variáveis independentes (regressores). Para aumentar a precisão dessa predição de concentrações, inverte-se o problema tornando as concentrações como variáveis dependentes (resposta) e as absorvâncias como variáveis independentes (regressores). Esse método é conhecido como ILS, equação 7:

$$C = AP + E_C \quad \text{Equação 7}$$

onde  $P$  é a matriz de coeficientes que deve ser ajustada ( $n \times l$ ) e  $E_C$  é a matriz dos erros de calibração não ajustados pelo modelo ( $m \times l$ ), pois a matriz  $P$  ótima minimiza o erro das concentrações  $\|e_C^T e_C\|$  no passo de calibração (RIBEIRO, 2007). Logo,  $P_{opt}$  é descrito pela equação 8:

$$P_{opt} = (A^T A)^{-1} A^T C \quad \text{Equação 8}$$

e a concentração predita passa a ser dada pela equação 9:

$$C_{pred} = A P_{opt} \quad \text{Equação 9}$$

Destaca-se que no método ILS a inversa da matriz  $A^T A$ , equação 8, tem dimensão  $n \times n$  mas o *rank* excede o número de amostras  $m$ , que geralmente é a menor dimensão da matriz  $A$ . Ademais, problemas de multicolinearidade podem surgir em  $A^T A$  (mesmo que  $m \gg n$ ) quando utilizado um grande número de  $n$ . Contudo, há dois grupos de abordagens para contornar esse problema (RIBEIRO, 2007):

aquelas que reduzem a singularidade de  $A^T A$  ao selecionar comprimentos de onda (colunas de  $A$ ) específicos em vez de utilizar todo espectro (ARAÚJO et al., 2001); e aqueles que tratam todo o espectro em um novo sistema de coordenadas (PCR e PLS) e então reduzem o número de entradas (geralmente chamadas variáveis latentes) do modelo.

A vantagem em se utilizar o PLS está na seleção de um conjunto de eixos ortogonais derivados da decomposição das matrizes  $C$  e  $A$ , onde a covariância é maximizada. Assim é

possível escolher um número de eixos que representam bem os dados e não ultrapassem o número de amostras, contornando assim o problema do *rank* e multicolinearidade de  $\mathbf{A}^T\mathbf{A}$  (RIBEIRO, 2007).

Os eixos ortogonais são calculados de forma que a covariância entre os dados da matriz  $\mathbf{A}$  e da matriz  $\mathbf{C}$  seja maximizada ao longo destes eixos, onde o primeiro eixo representa a maior parte da covariância entre o conjunto de dados e o último praticamente não contém informações úteis sobre os dados, representando informações de ruído. O conjunto de eixos ortogonais também são conhecidos como regressores do modelo.

Para que modelos obtidos tenham um bom desempenho frente a dados não utilizados no ajuste, faz-se o uso da validação cruzada. Basicamente, na validação cruzada uma parte dos dados de ajuste do modelo PLS é separada do conjunto total para ser testada em cada passo da calibração. Nesse teste, o modelo estima os dados não usados no ajuste e calcula o erro associado da predição (NELLES, 2001).

Logo, com a utilização conjunta do PLS com uma validação cruzada é possível determinar o número de regressores que melhor representa os dados de  $\mathbf{A}$  e  $\mathbf{C}$ , e que possui um menor erro associado a essa validação cruzada (*root-mean-square error of cross-validation*, RMSECV), equação 10:

$$RMSECV^{(i)} = \sqrt{\frac{1}{m} \sum_{k=1}^m (c_k^{(i)} - \hat{c}_k^{(i)})^2} \quad \text{Equação 10}$$

onde  $c_k^{(i)} - \hat{c}_k^{(i)}$  são os valores de referência e preditos das concentrações do analito  $i$  na  $k$ -ésima amostra de validação e  $m$  é o número de amostras usadas na validação cruzada (RIBEIRO, 2007).

### 3.7. PRÉ-TRATAMENTO DOS DADOS ESPECTROMÉTRICOS

Os métodos de pré-tratamento de dados são usados para remover erros aleatórios ou sistemáticos dos dados obtidos nas análises quimiométricas. Assim, é possível melhorar a inferência de modelos ajustados a partir desses dados (Brereton, 2007; Rinnan *et al.*, 2009). Dentre os métodos existentes, destaca-se que neste trabalho foi utilizado a derivada de primeira ordem, o alisamento pela média móvel e LOESS (*Locally Estimated Scatterplot Smoothing*).

### 3.7.1. Derivadas

As derivadas são usadas para corrigir a linha de base de dados espectrométricos, como do NIR e UV-Vis, removendo qualquer inclinação ou deslocamento. O efeito *offset* é um deslocamento da linha de base de uma quantidade constante de todo espectro. Espectros que possuem um *offset* afetam as tarefas de calibração e, conseqüentemente, aumentam o erro da predição dos modelos ajustados. Um conjunto de dados que possui um *offset* pode ser corrigido tomando-se a primeira derivada (Brereton, 2007).

A segunda derivada e as seguintes ordens, também podem ser usadas para a correção da linha de base. No entanto, podem aumentar os erros de cálculo nos dados tratados (Ferreira *et al.*, 1999).

### 3.7.2. Média Móvel

O alisamento dos dados espectrais pela média móvel centraliza as absorbâncias na média de um determinado intervalo, conservando-se a variação dessas bandas de absorção. O único parâmetro desse alisamento é o tamanho da janela móvel  $m$ , que consiste em um ponto central e  $k$  pontos anteriores e posteriores,  $m = 2.k + 1$  (Zakamulin, 2016). Assim nesse método, o valor médio da absorbância ( $\bar{a}$ ) no tempo  $n$ ,  $a_n$  é calculado pela equação 11:

$$\bar{a} = \frac{x_{n-k} + \dots + x_n + \dots + x_{n+k}}{m} = \frac{1}{m} \sum_{i=-k}^k x_{n+i} \quad \text{Equação 11}$$

### 3.7.3. *Locally Estimated Scatterplot Smoothing*

O alisamento LOESS possui um grande número de etapas distintas envolvidas no ajuste de qualquer conjunto de dados.

Para um melhor entendimento, suponha duas variáveis, X (eixo horizontal) e Y (eixo vertical), em um conjunto de dados com  $n$  observações. Os pontos plotados são pares ordenados  $(x_i, y_i)$ , onde  $i$  varia de 1 a  $n$ . Então, seleciona-se uma série de  $m$  pontos de avaliação,  $v_j$ , com  $j$  variando de 1 a  $m$ , que serão espaçados igualmente no intervalo de X. Após, é realizado uma série de análises de regressão ponderada “locais” para cada  $v_j$ , ou seja, cada série possui um subconjunto de dados que estão próximos desse ponto avaliado ao longo do eixo X. Para cada subconjunto de dados, utiliza-se o parâmetro *alpha* para especificar a

proporção de dados totais que serão incluídas. Esse parâmetro determina a largura da janela deslizante e possui um valor entre 0 e 1, onde valores próximos a 1 produzem uma curva mais reta e que não é tão sensível a variações de ruídos espectrais. Também, tem-se o parâmetro *order* para especificar o grau do polinômio que será usado (equações lineares, *order* 1, ou quadráticas, *order* 2). Esses parâmetros asseguram que os pontos mais próximos de  $v_j$  terão mais influência na resposta da regressão local (JACOBY, 2000).

A cada regressão local são gerados coeficientes,  $\hat{g}(v_j)$ , para estimar um valor previsto ou ajustado do ponto que está sendo avaliado. Após a finalização de todas regressões locais, os  $m$  pares ordenados diferentes  $(v_j, \hat{g}(v_j))$  são plotados sobrepostos aos  $n$  pontos de dados que já estão no gráfico. Por fim, os pontos ajustados adjacentes são ligados por segmentos de linha  $(v_j, \hat{g}(v_j);$  para  $v_j$ s sucessivos), e por estarem próximos um do outro, os segmentos de linha formam uma curva suave passando por todo conjunto de dados (JACOBY, 2000).

### 3.8. APLICAÇÃO DO NIR E UV-VIS NO MONITORAMENTO DE BIOPROCESSOS

A utilização da calibração multivariada associada aos dados espectrométricos obtidos no NIR e/ou UV-Vis, possibilita a criação de modelos preditivos que podem ser usados no monitoramento de bioprocessos. Ariane *et al.* (2018) construíram um modelo de predição das concentrações de compostos fenólicos e furaldeídos no licor gerado durante o pré-tratamento com água líquida quente do bagaço de cana, utilizando a regressão PLS dos dados obtidos pela espectroscopia UV-Vis. Para a obtenção dos espectros de UV-Vis, as amostras foram primeiramente filtradas através de uma membrana de 0,45  $\mu\text{m}$  para eliminar a interferência de qualquer material sólido. O RMSECV foi por volta de 5,0 % e, de modo geral, os resultados demonstraram o potencial poder de predição do modelo, em tempo real, para aplicação no monitoramento dos compostos fenólicos e furaldeídos.

Nascimento, R. J. A. D. (2016) estudou a espectroscopia NIR combinada a calibração PLS para o monitoramento das concentrações de açúcares durante a hidrólise enzimática da fibra da casca do coco verde. Os espectros NIR foram obtidos por meio de uma sonda de fibra ótica que foi inserida diretamente no meio reacional. Os melhores modelos de calibração para a predição da concentração dos açúcares redutores totais e glicose apresentaram um RMSECV de 0,3303 e 0,0527  $\text{g.L}^{-1}$ , respectivamente. O método se mostrou eficiente e que pode ser utilizado no monitoramento das concentrações em tempo real.



Blanco *et al.* (2000) utilizaram o NIR e a regressão PLS para monitorar as concentrações equivalentes de dextrose, glicose e polissacarídeos, durante a hidrólise enzimática do amido e dextrinas. Os autores utilizaram uma célula de fluxo com caminho de luz de 1 mm, onde circulava a solução do meio reacional, para ler o espectro NIR. Os modelos de calibração para a dextrose equivalente, dextrose e polissacarídeos tiveram um RMSECV de 1,7, 4,4 e 5,4%, respectivamente. Os autores concluíram que essa abordagem foi uma alternativa eficaz no monitoramento da hidrólise do amido e que tem o potencial de ser aplicada no controle de qualidade de xaropes industriais.

Na literatura há diversos exemplos de monitoramento de bioprocessos, utilizando a espectroscopia NIR e/ou UV-Vis associada a calibração multivariada, porém não há trabalhos com foco na hidrólise enzimática do bagaço de cana-de-açúcar.

## 4. MATERIAIS E MÉTODOS

### 4.1. MATERIAL LIGNOCELULÓSICO, TRATAMENTO HIDROTÉRMICO E COMPLEXO ENZIMÁTICO.

O bagaço de cana-de-açúcar utilizado nesse trabalho foi doado pela Usina Ipiranga S/A (Descalvado, SP), derivado do processo 1G da empresa.

O tratamento hidrotérmico do bagaço foi feito no reator de 5 L da *Parr Instruments* (Modelo 4580), pressurizado (pressão máxima de 200 psi) a 200 RPM. A concentração empregada no reator de bagaço seco por mililitro de reator foi de 240 gramas e 2,4 litros de água (10 % m/v), onde foi mantido por 10 min a uma temperatura de 195 °C. Então, o bagaço pré-tratado foi colocado em um filtro e lavado com água destilada por 2 h. Por fim, o bagaço pré-tratado foi seco na estufa a uma temperatura de 60 °C por 24 h, onde foi removido aproximadamente 95% da água (SLUITER *et al.*, 2008).

O complexo enzimático utilizado na hidrólise enzimática de todos ensaios foi a CELIC CTEC 2, doada pela Novozymes Latin America (Araucária, Paraná, Brasil). A concentração utilizada desse complexo foi de  $12,5 \text{ mg}_{\text{proteína}} \cdot \text{g}_{\text{sólidos}}^{-1}$  (aproximadamente  $10 \text{ FPU} \cdot \text{g}_{\text{substrato}}^{-1}$ ) que foi calculada a partir da massa inicial de substrato, onde o valor de um FPU (*Filter Paper Units*) é a quantidade de enzimas necessária para gerar 2 mg de glicose a partir

de 50 mg do papel filtro *Whatman*<sup>®</sup> *Number 1* (4% conversão) em 60 min a 50 °C (ADNEY; BAKER, 2008; FURLONG, 2019; GHOSE, 1987).

#### 4.2. DETERMINAÇÃO DAS CONCENTRAÇÕES DE REFERÊNCIA

A concentração dos principais carboidratos liberados (celobiose, glicose e xilose) nas amostras da fase líquida, durante os ensaios de hidrólise, foi determinada usando o HPLC Schimadzu SCL-10A<sup>®</sup> com detector de índice de refração Schimadzu RID10-A<sup>®</sup>, coluna Aminex HPX-87H<sup>®</sup> Bio-rad e ácido sulfúrico 5 mM a 0,6 mL.min<sup>-1</sup> como fase móvel (SLUITER et al., 2008).

Para construção dos padrões foi utilizada uma faixa de concentração (g/L) para cada carboidrato utilizando 5 pontos. O padrão de glicose variou de 0,5 a 15,0 g/L, o de celobiose de 0,2 a 2,0 g/L e o de xilose de 0,5 a 8 g/L. Todas as amostras foram suspensas no tampão utilizado na hidrólise enzimática.

#### 4.3. ENSAIOS

**Tabela 2:** Ensaios da hidrólise enzimática

Modo de operação	Duração do ensaio (h)	Quantidade de substrato no início do ensaio (g)	Volume inicial de tampão no reator agitado (mL)	Retirada de amostras líquidas
Batelada	12	20	400	2 – 2h
	48	20	500	2 – 2h nas primeiras 16 h e 4 – 4 h no restante do tempo
Batelada alimentada de bagaço	48	5	400	2 – 2h
Batelada alimentada de enzima	48	20	500	2 – 2h nas primeiras 16 h e 4 – 4 h no restante do tempo

Os ensaios de hidrólise enzimática foram realizados no sistema de reatores apresentado anteriormente utilizando o bagaço seco pré-tratado suspenso em uma solução tampão de citrato de 50 mM, com pH de 4,8 e temperatura 50 °C. Todos ensaios foram

realizados em duplicata, onde um ensaio foi usado para o ajuste PLS e outro para a validação externa.

Os detalhes dos ensaios estão apresentados resumidamente na Tabela 2. Em todos ensaios foram utilizados 20 g<sub>base-seca</sub> do bagaço pré-tratado com 85 % de umidade, adicionando a quantidade necessária do tampão de citrato para atingir essa umidade. O coquetel de enzimas foi adicionado ao bagaço pré-tratado com 85 % de umidade, a concentração de 12,5 mg<sub>proteína.g<sub>sólidos</sub><sup>-1</sup></sub>. Nos ensaios de batelada (12 h) e batelada alimentada com bagaço foram adicionados 400 mL da solução tampão no reator agitado para iniciar o processo, porém esse volume foi alterado para 500 mL nos outros ensaios, pois foi observado uma diminuição no nível do líquido no reator agitado, devido à evaporação da água contida no tampão, o que deixava parte da instrumentação fora do meio líquido. Também, foram adicionados a cada 4 h, 50 mL de água ultrapura no reator agitado para contornar essa evaporação, até o tempo de 36 h do início do ensaio.

No ensaio de hidrólise operado em batelada alimentada de bagaço, 5 g de bagaço pré-tratado hidrotermicamente foram adicionadas no início da reação e o restante foi alimentado constantemente, de modo que as 15 g fossem adicionadas no reator horizontal durante 36 h (vazão mássica aproximadamente de 7 mg.min<sup>-1</sup>). Já no ensaio de batelada alimentada de enzimas, foi adicionada a metade da carga enzimática no início da reação e o restante também foi alimentado constantemente durante 36 h (vazão volumétrica de aproximadamente 0,21 mL.min<sup>-1</sup>). Ainda nesse ensaio, não foram adicionados 50 mL de água destilada a cada 4 h, pois a carga do complexo enzimático foi diluída no volume total de água ultrapura, 450 mL, que seria adicionado (concentração de enzima no meio, 10 FPU.g<sub>substrato</sub><sup>-1</sup>).

A retirada de amostras líquidas para as operações em batelada de 48 h e batelada alimentada de enzimas foram alterados de 2 em 2 h para 4 em 4 h, após 16 h, pois a composição não se altera tão rapidamente como nas primeiras 16 h (FURLONG, 2019).

#### 4.4. MONITORAMENTO DO PROCESSO

Para a construção dos modelos de monitoramento foram coletados dados de NIR e UV-Vis pelas sondas de fibra ótica, que foram adicionadas dentro do reator agitado. A combinação desses dados com os dados quantificados pelo HPLC, pelo método dos mínimos

quadrados parciais (PLS), gerou modelos matemáticos capazes de prever as concentrações dos carboidratos durante a hidrólise enzimática.

#### 4.4.1. Espectro do infravermelho próximo (NIR)



**Figura 3:** Espectrofotômetro Frontier NIR - Perkin Elmer®.

Uma sonda de refletância Falcata 6 foi usada no reator agitado e estava acoplada ao espectrofotômetro Frontier NIR da Perkin Elmer® (Figura 3). Realizaram-se varreduras a cada 6 min na faixa de número de onda de 10000 a 1200  $\text{cm}^{-1}$  (1000 a 8333.33 nm), com resolução de 4  $\text{cm}^{-1}$ .

#### 4.4.2. Espectro do ultravioleta e visível (UV-Vis)



**Figura 4:** Espectrofotômetro Lambda 456 da Perkin Elmer®.

Uma sonda de refletância também foi usada no reator agitado e estava acoplada ao espectrofotômetro Lambda 456 da Perkin Elmer® (Figura 4). Realizaram-se varreduras a cada 6 min, na região do ultravioleta e visível (UV-Vis) que varia de 190 a 1100 nm, com resolução de 1 nm.

#### 4.4.3. Modelagem matemática

Os dados obtidos a partir das análises cromatográficas, dos carboidratos nas amostras da fase líquida, foram utilizados como a concentração de referência no PLS, com os dados de cada ensaio das varreduras do espectro NIR e UV-Vis. A duplicada de cada ensaio foi usada como validação externa, para avaliar a capacidade de predição da concentração dos modelos ajustados.

No entanto, os dados espectrais foram pré-tratados com os softwares Microsoft Excel e Scilab, para diminuir a interferência dos ruídos na obtenção dos modelos PLS, onde foi avaliado o RMSECV para a escolha do número ótimo de parâmetros dos modelos. Nesse pré-tratamento:

- Foram avaliados, em cada ensaio, os espectros obtidos em cada hora, referente as amostras retiradas, para identificar espectros fora do padrão, onde tais espectros foram

desconsiderados no ajuste do modelo. O padrão considerado foi baseado na comparação do perfil do espectro de absorção com uma janela de alisamento de 11 pontos;

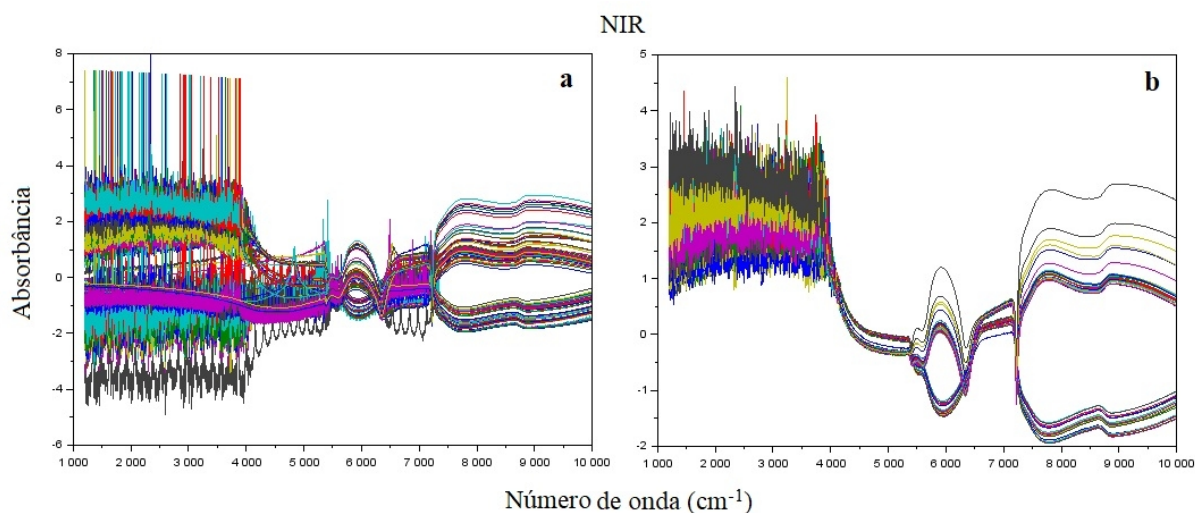
- Avaliou-se a utilização da primeira e segunda derivada;
- Avaliou-se a restrição de algumas regiões do espectro e também a máxima absorbância;
- Avaliou-se a fração de dados aleatórios para testes na validação cruzada;
- Avaliou-se o alisamento dos espectros pela média móvel, com variação da janela de alisamento;
- Avaliou-se o alisamento LOESS dos espectros, com a variação os parâmetros *alpha* e *order*.

Além disso, foi aplicado o método CLS para avaliar os espectros selecionados, onde é avaliado o perfil de predição das absorbâncias. Dentre todos os testes, dois modelos foram selecionados para a discussão, por apresentar um menor RMSECV e uma melhor predição das concentrações quantificadas no HPLC.

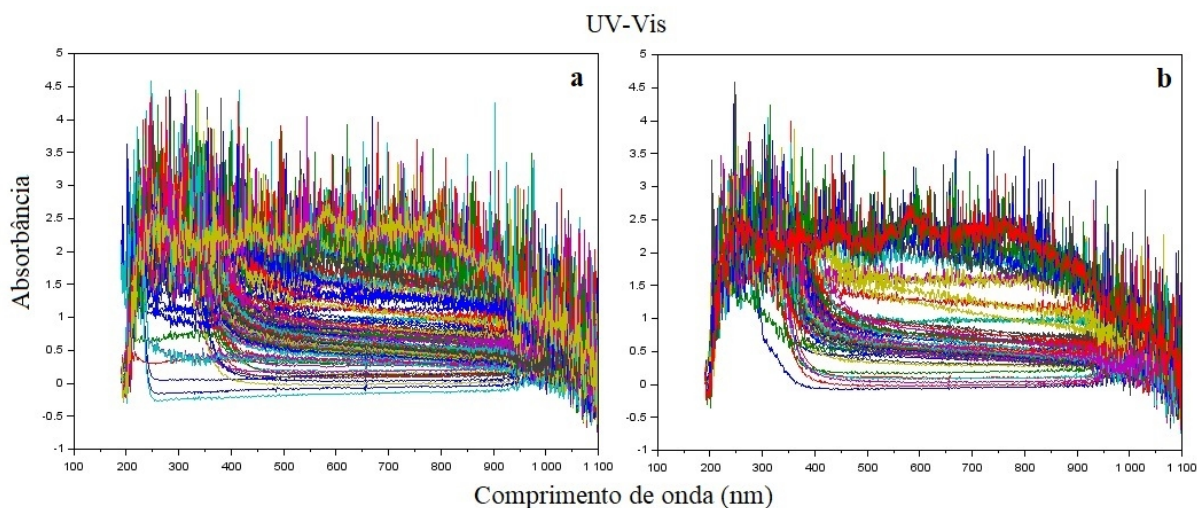
## 5. RESULTADOS E DISCUSSÃO

A primeira etapa do ajuste dos modelos foi a análise dos espectros e seleção dos dados. Em cada ensaio os espectros foram gerados de 6 em 6 minutos e as amostras retiradas em períodos diferentes, como mostrado na tabela 2 (Materiais e Métodos, Seção 4.3). Então, foi analisado o espectro no tempo exato que foi retirado a amostra, assim como cinco espectros anteriores e posteriores a esse ponto, de modo a identificar aqueles fora do padrão do restante. Assim, foram selecionados os dados que melhor representava a amostra, em cada tempo, para serem usados no ajuste PLS.

As figuras 5 e 6 apresentam os espectros NIR e UV-Vis com os dados brutos e selecionados das amostras coletadas nos ensaios de hidrólise enzimática operado em batelada, batelada alimentada com bagaço de cana-de-açúcar e batelada alimentada com enzimas.



**Figura 5:** Gráfico dos dados espectrais brutos (a) e selecionados (b) no NIR obtidos de amostras coletadas de todos ensaios realizados.



**Figura 6:** Gráfico dos dados espectrais brutos (a) e selecionados (b) no UV-Vis obtidos de amostras coletadas de todos ensaios realizados.

No espectro do NIR, figura 5 b, a região de 1100 a 4000  $\text{cm}^{-1}$  é ruidosa devido ao próprio equipamento, por ser uma faixa fora da especificação do instrumento, porém ela foi considerada nesse trabalho de modo a avaliar a influência dessa região no ajuste dos modelos e predição das concentrações dos analitos. Há também um diferente perfil de absorção nas regiões de 5500 a 6500  $\text{cm}^{-1}$  e 7200 a 10000  $\text{cm}^{-1}$ . Analisando o espectro de cada ensaio, em pontos de 2 em 2 h, foi possível observar que nessa região ocorre variações no espectro de absorção. Como os números de ondas 4411 e 5980  $\text{cm}^{-1}$  correspondem a absorção da lignina,

essas variações podem estar ligadas à lignina e outros compostos não identificados (SCHWANNINGER; RODRIGUES; FACKLER, 2011; ZHOU et al., 2015).

Já no espectro do UV-Vis, figura 6 b, a lignina é absorvida na região de 200 a 400 nm (JABLONSKÝ et al., 2015). Em algumas amostras, na faixa de 400 a 900 nm, ocorre um aumento na absorção da radiação. Analisando também o espectro de cada ensaio, pontos de 2 em 2 h, observou-se que esse aumento ocorreu nos períodos em que houve formação de bolhas nas sondas de fibra ótica e também após 30 h do início dos ensaios. Nesse período de 30 h do início do ensaio, observou-se também um aumento da turbidez do líquido contido no reator agitado, que está relacionada a concentração dos compostos que foram hidrolisados durante o ensaio (heterogeneidade do meio).

Para diminuir a interferência desses ruídos na obtenção dos modelos PLS, os dados foram pré-tratados, como descrito no item 4.4.3. As Tabelas 3 e 4 abaixo apresentam os parâmetros e os erros associados da calibração, de cada modelo para cada técnica espectroscópica.

**Tabela 3:** Modelos ajustados NIR.

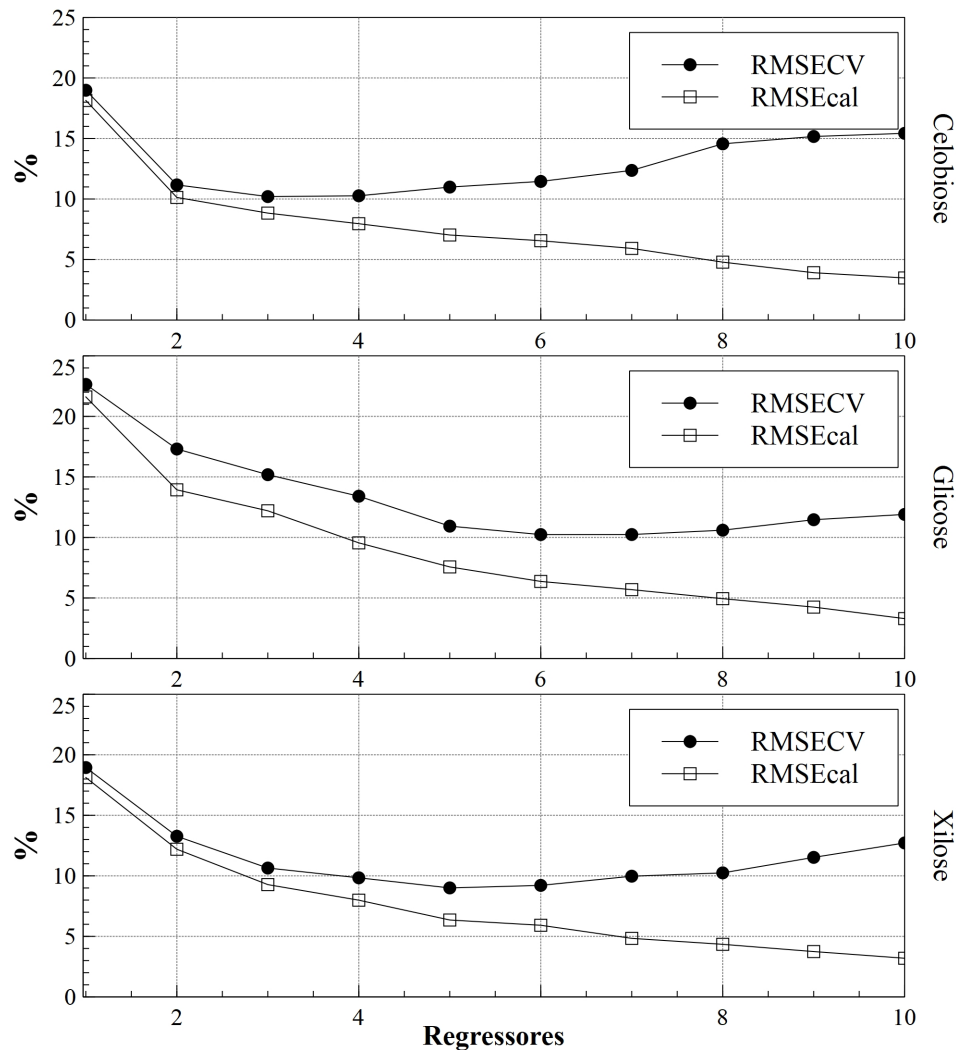
	<b>NIR</b>					
	<b>MODELO 1</b>			<b>MODELO 2</b>		
	<b>Celobiose</b>	<b>Glicose</b>	<b>Xilose</b>	<b>Celobiose</b>	<b>Glicose</b>	<b>Xilose</b>
<b>Número de Regressores</b>	3	4	4	3	3	3
<b>RMSECV</b>	13,6%	16,2%	13,8%	17,1%	18,9%	16,4%
<b>RMSECV (g/L)</b>	0,079	1,575	0,215	0,099	1,842	0,256
<b>RMSEP (g/L)</b>	0,292	0,218	0,229	0,056	0,064	0,059
<b>Fração de teste</b>	12,0%			12,0%		
<b>Janela de alisamento</b>	11			11		
<b>Faixa de número de onda</b>	1100 – 10000 cm <sup>-1</sup>			4000 – 10000 cm <sup>-1</sup>		



Tabela 4: Modelos ajustados UV-Vis.

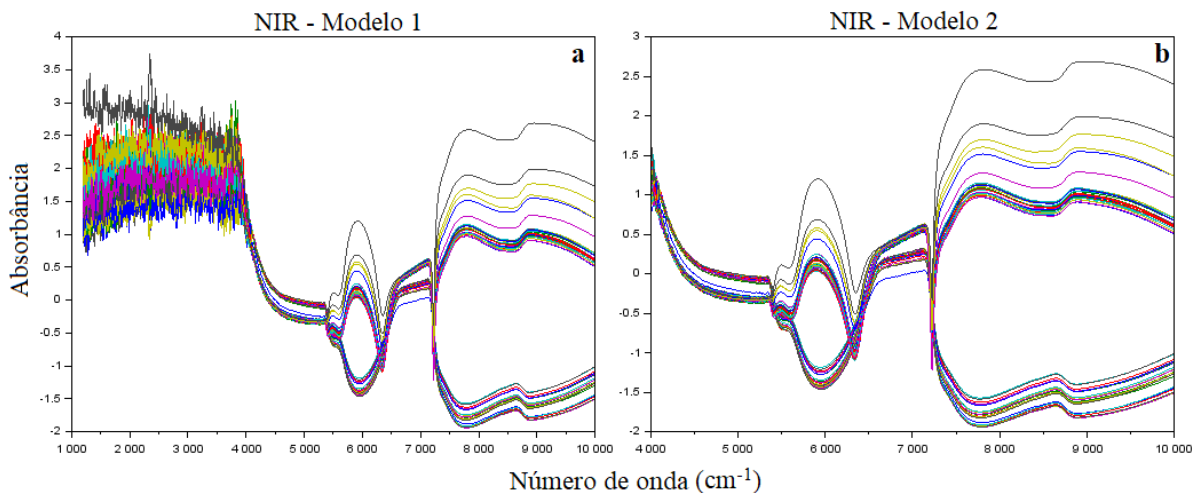
	UV-Vis					
	MODELO 1			MODELO 2		
	Celobiose	Glicose	Xilose	Celobiose	Glicose	Xilose
<b>Número de Regressores</b>	2	5	5	2	3	3
<b>RMSECV</b>	11,3%	11,3%	9,1%	15,0%	12,5%	10,1%
<b>RMSECV (g/L)</b>	0,065	1,216	0,142	0,082	1,348	0,133
<b>RMSEP (g/L)</b>	0,072	0,122	0,094	0,286	0,159	0,175
<b>Fração de teste</b>	12,0%			12,0%		
<b>Janela de alisamento</b>	11			11		
<b>Faixa de comprimento de onda</b>	190 – 1100 nm			450 – 900 nm		
<b>Absorbância máxima</b>	8			1,2		
<b>Derivada</b>	-			-		

Para determinar o número de regressores em cada ajuste, foi utilizado um procedimento de validação cruzada, onde 88 % dos dados foram usados para o treinamento/validação e 12 % foram usados como dados de teste do modelo. A Figura 7 apresenta o gráfico dos erros associados à validação cruzada em função do número de regressores utilizados para cada carboidrato, quando aplicado o ajuste PLS nos dados do UV-Vis, com as condições de fração de teste, janela de alisamento, faixa de comprimento de onda e absorbância máxima descritas pelo modelo 1 na Tabela 4. O RMSEcal representa uma aproximação do desvio padrão do erro do ajuste e o RMSECV, o desvio padrão do erro de validação cruzada. Para um melhor entendimento, suponha um ajuste de 100 dados. Na validação cruzada, retira-se 1 dado a cada passo do ajuste. O RMSECV representa o erro desse dado que foi retirado e o RMSEcal representa o erro dos 99 dados restantes. Esses dois erros possuem valores semelhantes quando não há um sobre-ajuste (*overfitting*) do modelo, ou seja, quando o modelo segue a tendência dos dados de ajuste, porém não o suficiente para seguir a tendência do ruído desses dados. Logo, o número de regressores ótimo para cada carboidrato foi determinado a partir do menor RMSECV, ou seja, um número de regressores maiores que o ponto ótimo pode aumentar o RMSECV e causar o *overfitting*.



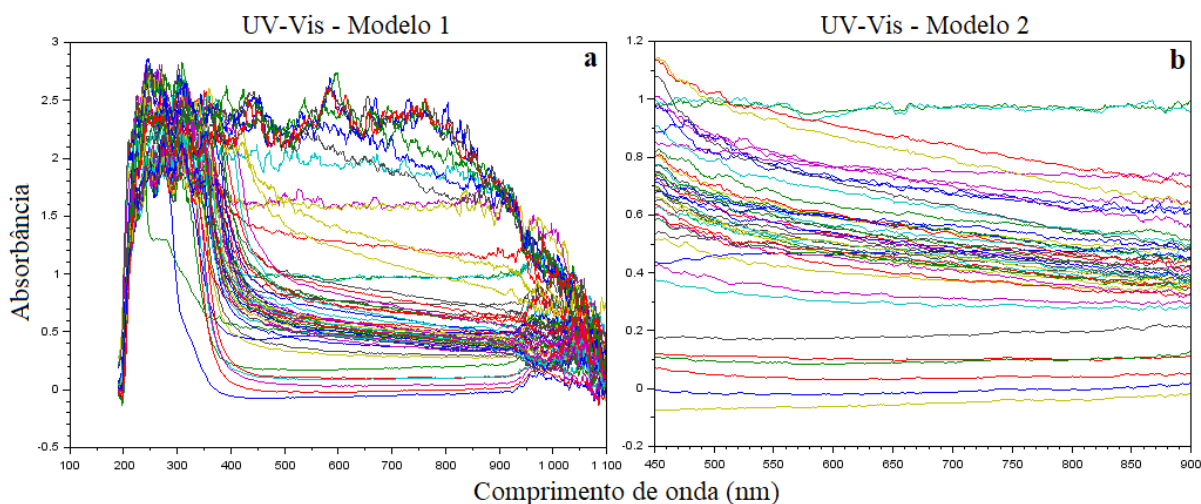
**Figura 7:** Gráficos dos erros associados (RMSECV e RMSEcal) ao número de regressores utilizados no teste da validação cruzada para cada carboidrato (Celobiose, Glicose e Xilose) no ajuste dos dados espectrais UV-Vis. Condições do ajuste descritas na tabela 4 pelo modelo 1.

Os modelos 1 e 2 utilizados no espectro do NIR diferem apenas na faixa de comprimento de onda. Há um aumento do RMSECV para o modelo 2 e isso pode estar associado à remoção da região ruidosa,  $1100 - 4000 \text{ cm}^{-1}$ , o que diminui a sensibilidade do modelo em relação aos dados que foram ajustados. A figura 8 apresenta os espectros selecionados e pré-tratados dos modelos 1 e 2, respectivamente, do NIR.



**Figura 8:** Gráfico dos espectros selecionados e pré-tratados do NIR, modelos 1 (a) e 2 (b) para as condições de pré-tratamento: janela de alisamento de 11 pontos, sem derivada e faixa de número de onda de 4000 a 10000 cm<sup>-1</sup> (modelo 2 - b).

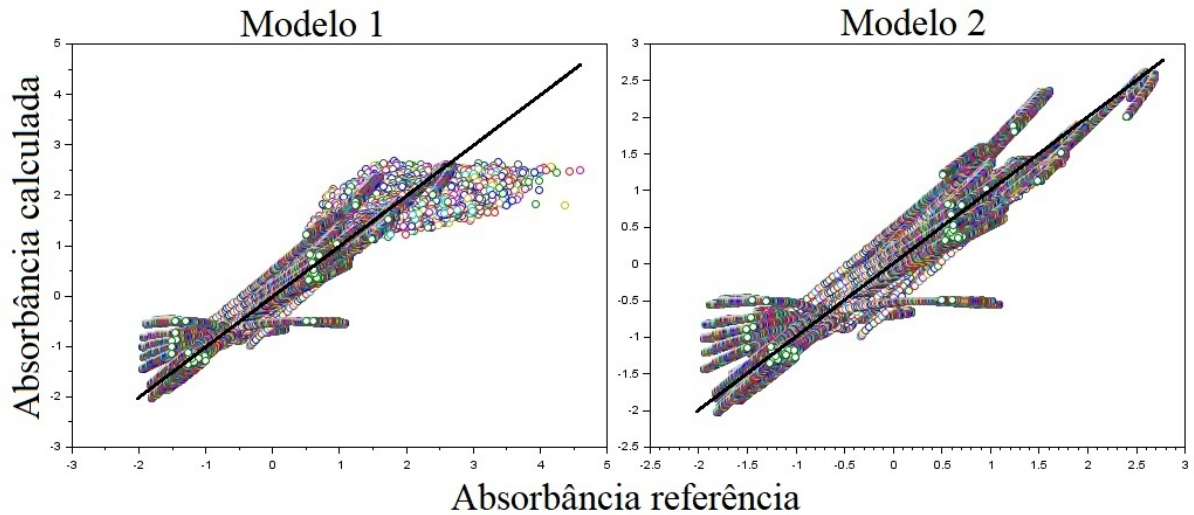
Já a diferença entre os modelos do UV-Vis está na faixa de comprimento de onda e absorbância máxima. É evidente um aumento do RMSECV para o modelo 2, provavelmente pela mesma razão dos modelos do NIR (restrição no espectro). A figura 9 apresenta os espectros selecionados e pré-tratados dos modelos 1 e 2, respectivamente, do UV-Vis.



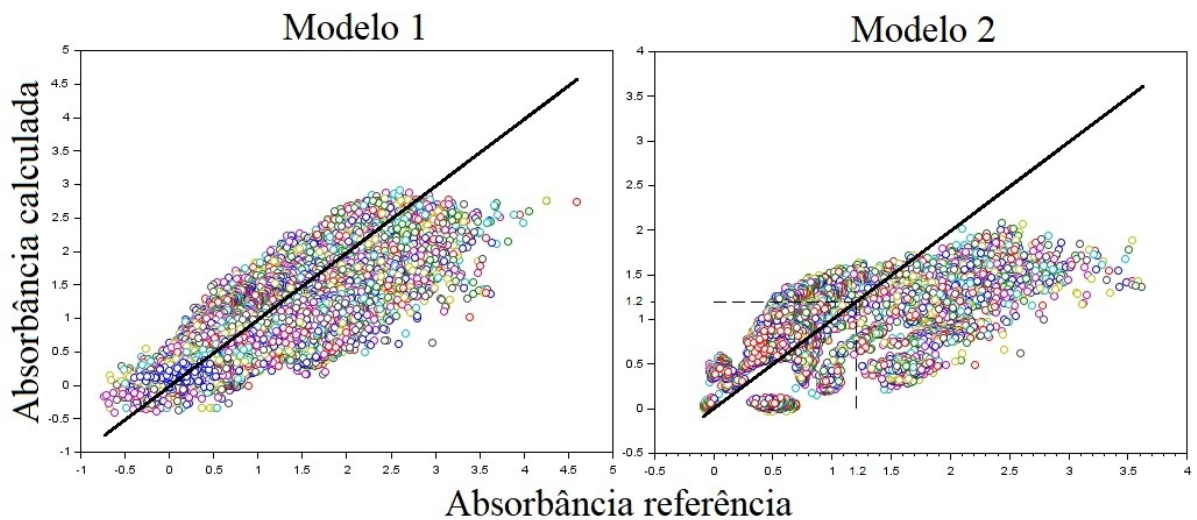
**Figura 9:** Gráfico dos espectros selecionados e pré-tratados do UV-Vis, modelos 1 (a) e 2 (b) para as condições de pré-tratamento: janela de alisamento de 11 pontos, sem derivada, faixa de comprimento de onda de 450 a 900 nm (modelo 2 - b) e absorbância máxima de 1.2 (modelo 2 - b).

Para avaliar os espectros selecionados no ajuste dos modelos, aplicou-se o método CLS. Nesse método, cria-se um modelo de predição das absorbâncias a partir das

concentrações. As figuras 10 e 11 apresentam o resultado do método CLS aplicado aos dados utilizados no ajuste dos modelos 1 e 2 do NIR e UV-Vis, respectivamente.



**Figura 10:** Gráfico da predição da absorvância pelo modelo CLS aplicado aos dados utilizados no ajuste dos modelos 1 e 2 do NIR.



**Figura 11:** Gráfico da predição da absorvância pelo modelo CLS aplicado aos dados utilizados no ajuste dos modelos 1 e 2 do UV-Vis.

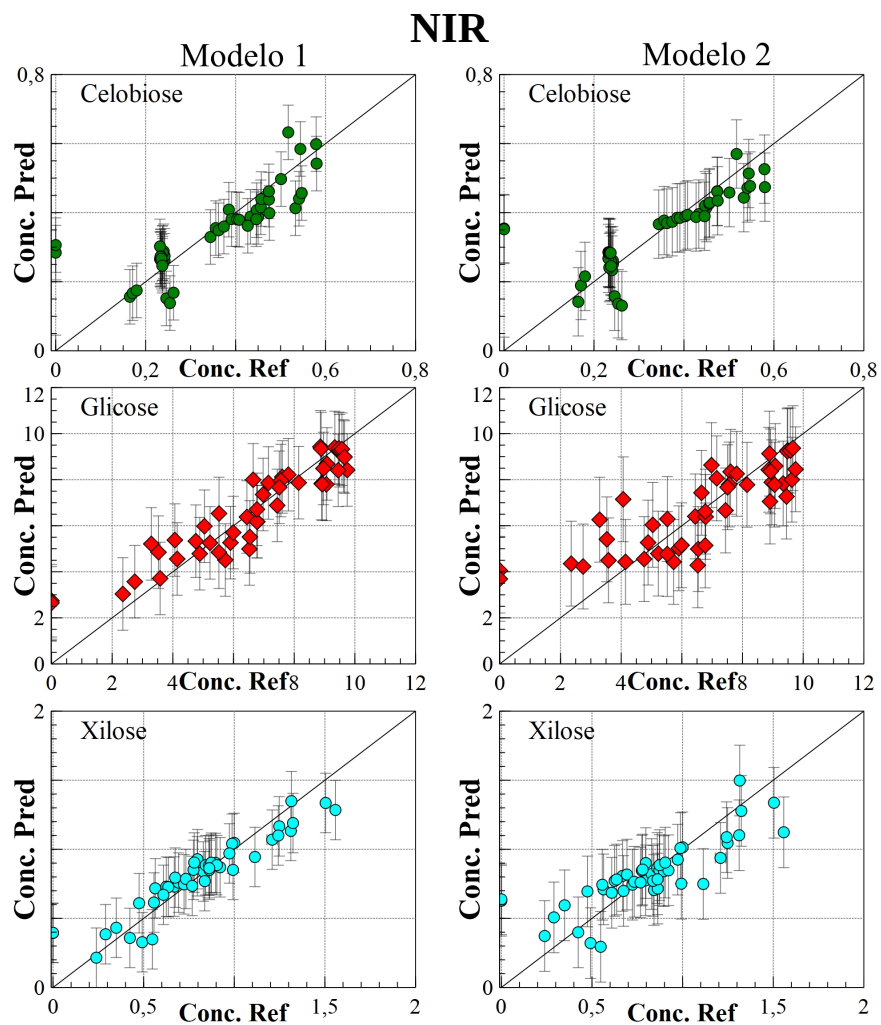
É possível observar em ambas figuras, 10 e 11, que o sistema é linear, porém com alguns desvios devido ao espectro ruidoso e absorvâncias negativas. Na figura 10, o modelo 1 possui desvios da linearidade em absorvâncias acima de 1,5 unidades de absorvância (UA).

Logo, para esse modelo, seria indicado filtrar o espectro limitando a uma absorvância máxima de 1,5 UA. No entanto, em todos ensaios, os espectros obtidos pelo NIR possuem absorvâncias acima de 1,5 UA, o que inviabiliza essa abordagem. Já no modelo 2, o sistema é linear quando analisadas as absorvâncias positivas, ou seja, acima de 0 UA.

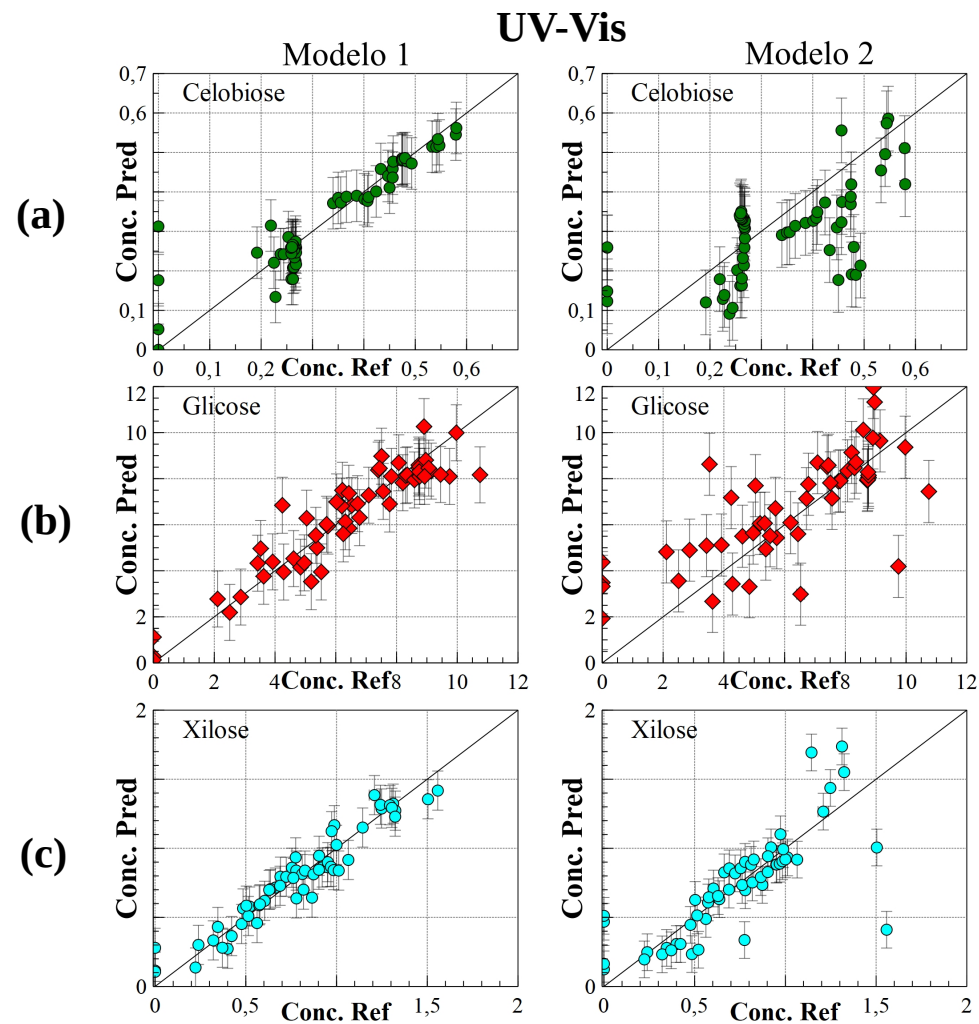
Na figura 11, em ambos modelos o espectro ruidoso afeta na dispersão (uniforme). No modelo 2, há um desvio sistemático do modelo para absorvâncias acima de 1,2 UA. Assim, selecionaram-se espectros abaixo de 1,2 UA, além da região de 450 a 900 nm.

O uso do CLS mostra que em ambas técnicas, NIR e UV-Vis, os dados selecionados para o ajuste dos modelos possuem uma linearidade, mesmo com os desvios provindos dos ruidos, o que torna aplicável o PLS.

Associado ao ajuste PLS de cada modelo, foram gerados os gráficos das concentrações preditas x concentrações de referência, que estão apresentados nas figuras 12 e 13. Em geral as predições do modelo 1 no NIR e UV-Vis foram melhores que do modelo 2, pois as concentrações estão menos dispersas em torno da bissetriz do que no modelo 2. Observa-se que ambos os modelos não predizem bem as concentrações no tempo 0 h, provavelmente pelo alto ruído nos espectros. A princípio, a remoção de *outliers* melhora o ajuste do modelo e conseqüentemente teria que melhorar a predição das concentrações. No entanto, quando o ponto 0 h é removido do conjunto de dados de concentração de referência, durante o ajuste do modelo, o RMSECV de fato diminui (remoção do *outlier*), porém a predição das concentrações perde o perfil de aumento de concentração ao longo do tempo.



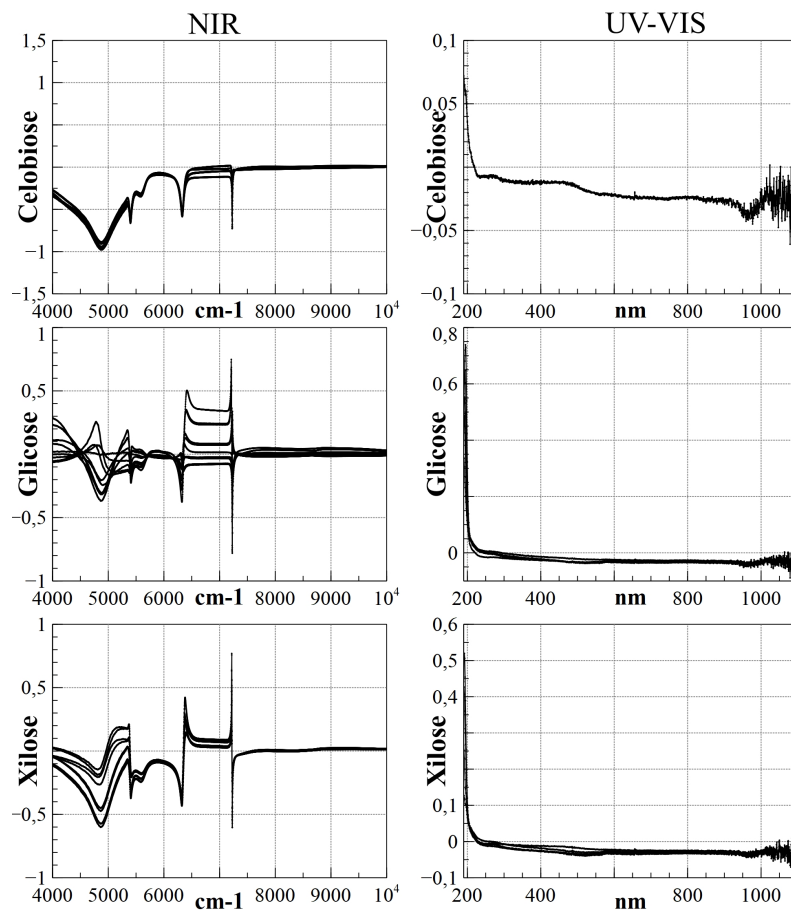
**Figura 12:** Gráficos das concentrações previstas de cada carboidrato, com a barra do desvio padrão de previsão na CV (RMSECV), dos modelos PLS 1 e 2 do NIR, dado a bissetriz de referência. No eixo x tem-se as concentrações aferidas pelo HPLC (referência, g/L) e no eixo y as concentrações previstas pelos modelos (g/L) para a Celobiose (a), Glicose (b) e Xilose (c).



**Figura 13:** Gráficos das concentrações previstas de cada carboidrato, com a barra do desvio padrão de previsão na CV (RMSECV), dos modelos PLS 1 e 2 do UV-Vis, dado a bissetriz de referência. No eixo x tem-se as concentrações aferidas pelo HPLC (referência, g/L) e no eixo y as concentrações previstas pelos modelos (g/L) para a Celobiose (a), Glicose (b) e Xilose (c).

De modo a melhorar a predição das concentrações pelos modelos, foi abordado um estudo do espectro de modo a criar uma curva de calibração que indique as regiões de absorção de cada carboidrato. Como o sistema reacional deste trabalho é complexo, pois contém diversos componentes na mistura, além de outros vários que surgem conforme a hidrólise ocorre, é inviável preparar uma mistura sintética desse meio reacional. Logo, nesse estudo das regiões do espectro NIR e UV-Vis foram utilizados apenas os carboidratos puros.

As amostras utilizadas nesse estudo tiveram concentrações variadas de modo a criar uma curva de calibração com 5 pontos, celobiose 0,2 a 10 g/L, glicose 3 a 50 g/L e xilose 0,2 a 10 g/L. A figura 14 apresenta os espectros obtidos da celobiose, glicose e xilose pelo NIR e UV-Vis.



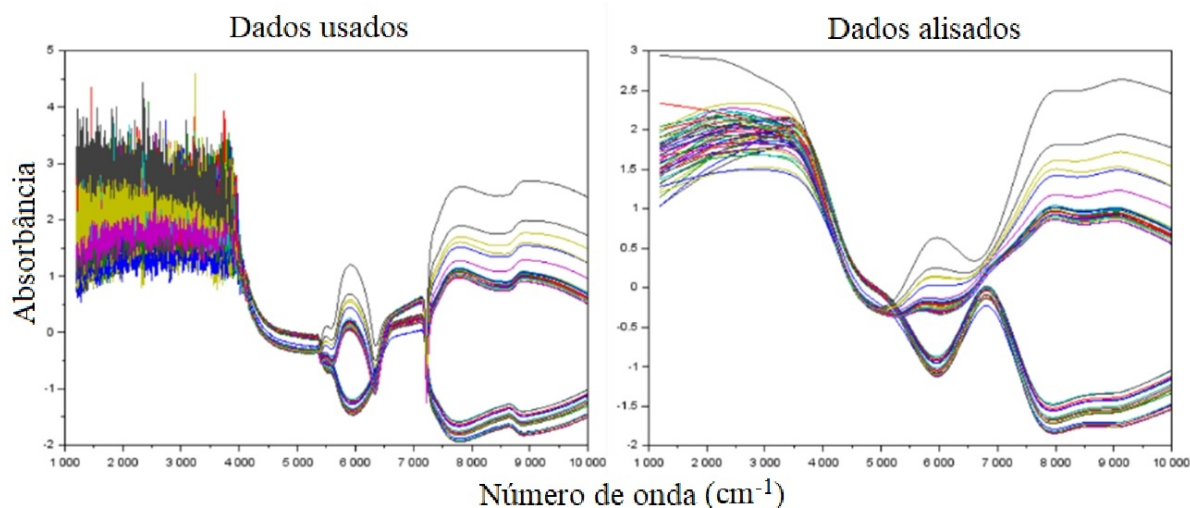
**Figura 14:** Curva de calibração da celobiose, glicose e xilose utilizando as técnicas NIR e UV-Vis.

É possível observar que a região entre  $4000$  e  $8000\text{ cm}^{-1}$  do NIR e  $190$  a  $300\text{ nm}$  do UV-Vis, possui diferentes picos que podem estar relacionados a absorção de radiação dos



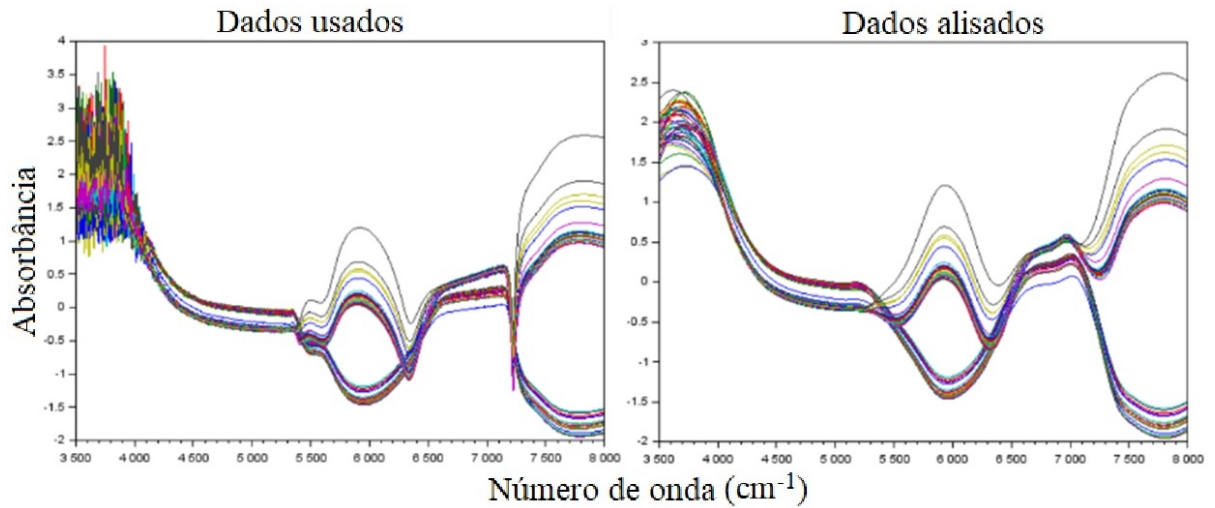
carboidratos. De modo geral, os picos característicos dos carboidratos, no UV-Vis, estão entre os comprimentos de onda de 260 a 290 nm (glicose), 245 a 255 nm (xilose) e 200 a 300 nm (celobiose desidrogenase – celobiose) (FANG; HUANG; GAO, 1999; KAIJANEN et al., 2015; PACSU; HILLER, 1948).

Logo, foram selecionadas as regiões de 3500 a 8000  $\text{cm}^{-1}$ , NIR, e 190 a 500 nm, UV-Vis, para gerar outros modelos PLS. Como o alisamento pela média móvel não se mostrou eficaz na redução do ruído espectral, foi também utilizado o método de alisamento *Locally Estimated Scatterplot Smoothing*, LOESS, para um maior alisamento do espectro gerado no NIR e UV-Vis. As Figuras 15 e 16 apresentam os dados usados no ajuste do modelo 3 e 4, respectivamente, e esses mesmos dados alisados pelo método LOESS. Já as Figuras 17 e 18, do UV-Vis.

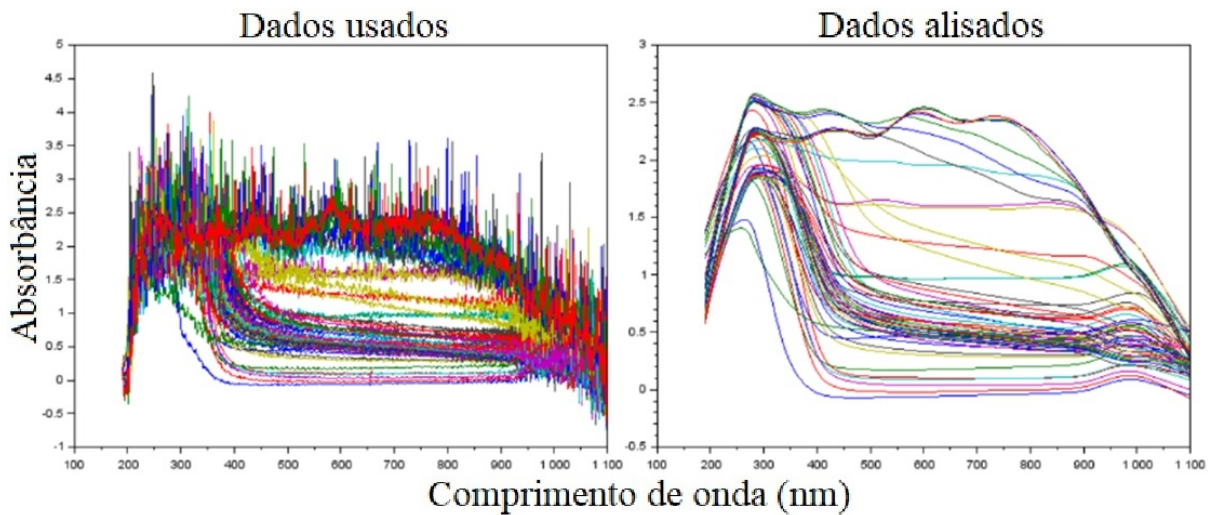


**Figura 15:** Gráfico dos espectros selecionados e pré-tratados do NIR, modelo 3, para as condições de pré-tratamento: LOESS (alpha de 0,2, order 1), sem derivada e faixa de número de onda de 1100 a 10000  $\text{cm}^{-1}$ .

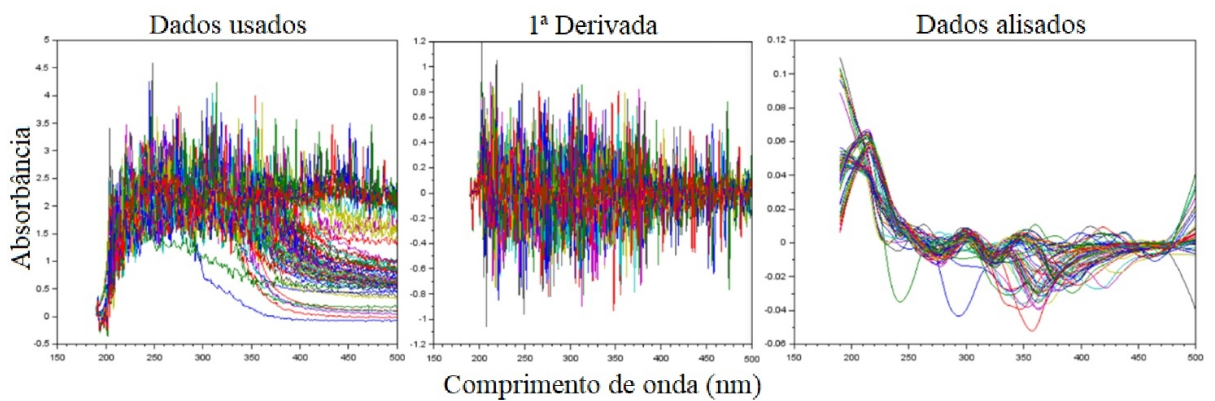




**Figura 16:** Gráfico dos espectros selecionados e pré-tratados do NIR, modelo 4, para as condições de pré-tratamento: LOESS (alpha de 0,2, order 2), sem derivada e faixa de número de onda de 3500 a 8000  $\text{cm}^{-1}$ .



**Figura 17:** Gráfico dos espectros selecionados e pré-tratados do UV-Vis, modelo 3, para as condições de pré-tratamento: LOESS (alpha de 0,2, order 1), sem derivada e faixa de comprimento de onda de 190 a 1100 nm.



**Figura 18:** Gráfico dos espectros selecionados e pré-tratados do UV-Vis, modelo 4, para as condições de pré-tratamento: LOESS (alpha de 0,2, order 1), primeira derivada e faixa de comprimento de onda de 190 a 500 nm.

O método LOESS foi efetivo na suavização das curvas, porém deformou as curvas características de cada amostra. Modelos ajustados com essas curvas perdem, de certa forma, seu caráter fenomenológico. Isso pode reduzir expressivamente o desempenho de modelos lineares. De qualquer forma, como o pré-tratamento é sempre realizado da mesma forma, seus resultados devem ser consistentes e reprodutivos e por essa razão esse pré-tratamento também foi avaliado neste trabalho.

As Tabelas 5 e 6 abaixo apresentam os parâmetros e os erros associados da calibração, de cada modelo para cada técnica espectroscópica. Entre os modelos do NIR, houve um aumento representativo do RMSECV apenas entre os modelos 1 e 3, de 34% para celobiose, 18% glicose e 28% xilose, ou seja, quando aplicado o LOESS no mesmo conjunto de dados, modelo 1, a redução do ruído não implica em uma redução do RMSECV. Já entre os modelos do UV-Vis, houve uma redução do RMSECV de aproximadamente 13% para celobiose entre os modelos 1 e 3 e de 8% entre os modelos 2 e 4, um aumento de 14% para glicose entre os modelos 1 e 3 e 21,6% entre os modelos 2 e 4. Nesse caso, o uso de LOESS não mudou significativamente o RMSECV entre os modelos.

De modo geral, os modelos apresentaram um baixo RMSECV, dada as condições operadas. No item 3.8 foi apresentado trabalhos que utilizaram o ajuste PLS, onde os valores do RMSECV foram abaixo de 10%. Neste trabalho, foi abordado um cenário mais próximo das biorrefinarias, ou seja, um meio reacional mais complexo. Mesmo assim, a média do RMSECV entre os quatro modelos do NIR foi de 16,5% para celobiose, 18,3% glicose e 16,1% xilose, e do UV-Vis, 12,5% celobiose, 13% glicose e 9,5% para xilose.

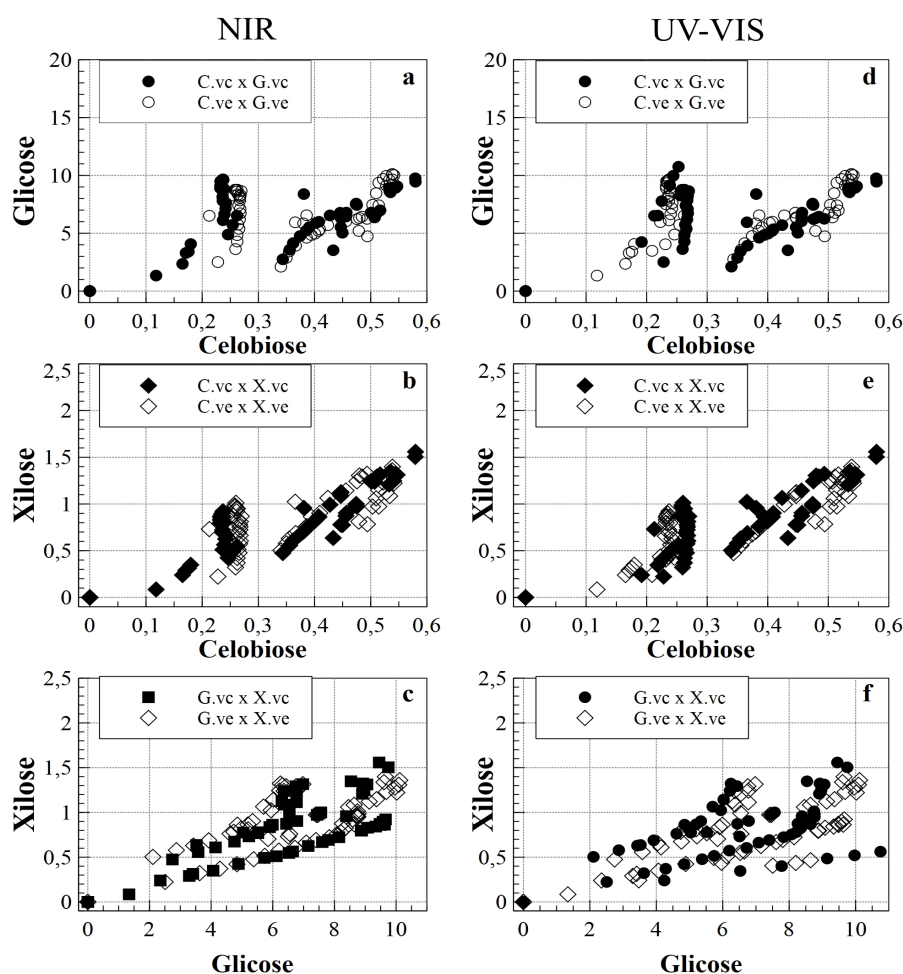
**Tabela 5:** Todos os modelos ajustados pelos dados obtidos no NIR.

NIR												
MODELO 1			MODELO 2			MODELO 3			MODELO 4			
	Celobiose	Glicose	Xilose	Celobiose	Glicose	Xilose	Celobiose	Glicose	Xilose	Celobiose	Glicose	Xilose
<b>Número de Regressores</b>	3	4	4	3	3	3	6	6	6	3	3	3
<b>RMSECV</b>	13,6%	16,2%	13,8%	17,1%	18,9%	16,4%	18,2%	19,2%	17,7%	17,1%	18,9%	16,5%
<b>RMSECV (g/L)</b>	0,079	1,575	0,215	0,099	1,842	0,256	0,106	1,871	0,275	0,099	1,844	0,257
<b>RMSEP (g/L)</b>	0,292	0,218	0,229	0,056	0,064	0,059	0,084	0,098	0,108	0,079	0,046	0,020
<b>Fração de teste</b>	12,0%			12,0%			12,0%			12,0%		
<b>Alisamento</b>	Média móvel: janela = 11			Média móvel: janela = 11			LOESS: alpha = 0,2, order = 1			LOESS: alpha = 0,2, order = 2		
<b>Faixa de número de onda</b>	1100 – 10000 cm <sup>-1</sup>			4000 – 10000 cm <sup>-1</sup>			1100 – 10000 cm <sup>-1</sup>			3500 – 8000 cm <sup>-1</sup>		
<b>Derivada</b>	-			-			-			-		

**Tabela 6:** Todos modelos ajustados pelos dados obtidos no UV-Vis.

	UV-Vis											
	MODELO 1			MODELO 2			MODELO 3			MODELO 4		
	Celobiose	Glicose	Xilose	Celobiose	Glicose	Xilose	Celobiose	Glicose	Xilose	Celobiose	Glicose	Xilose
<b>Número de Regressores</b>	2	5	5	2	3	3	3	6	6	3	3	6
<b>RMSECV</b>	11,3%	11,3%	9,1%	15,0%	12,5%	10,1%	9,8%	12,9%	9,0%	13,8%	15,2%	10,0%
<b>RMSECV (g/L)</b>	0,065	1,216	0,142	0,082	1,348	0,133	0,057	1,384	0,140	0,080	1,630	0,155
<b>RMSEP (g/L)</b>	0,072	0,122	0,094	0,286	0,159	0,175	0,048	0,194	0,084	0,041	0,073	0,057
<b>Fração de teste</b>	12,0%			12,0%			12,0%			12,0%		
<b>Alisamento</b>	Média móvel: janela = 11			Média móvel: janela = 11			LOESS: alpha = 0,2, order = 1			LOESS: alpha = 0,2, order = 1		
<b>Faixa de comprimento de onda</b>	190 – 1100 nm			450 – 900 nm			190 – 1100 nm			190 – 500 nm		
<b>Absorbância máxima</b>	8			1,2			8			8		
<b>Derivada</b>	-			-			-			1 <sup>a</sup>		

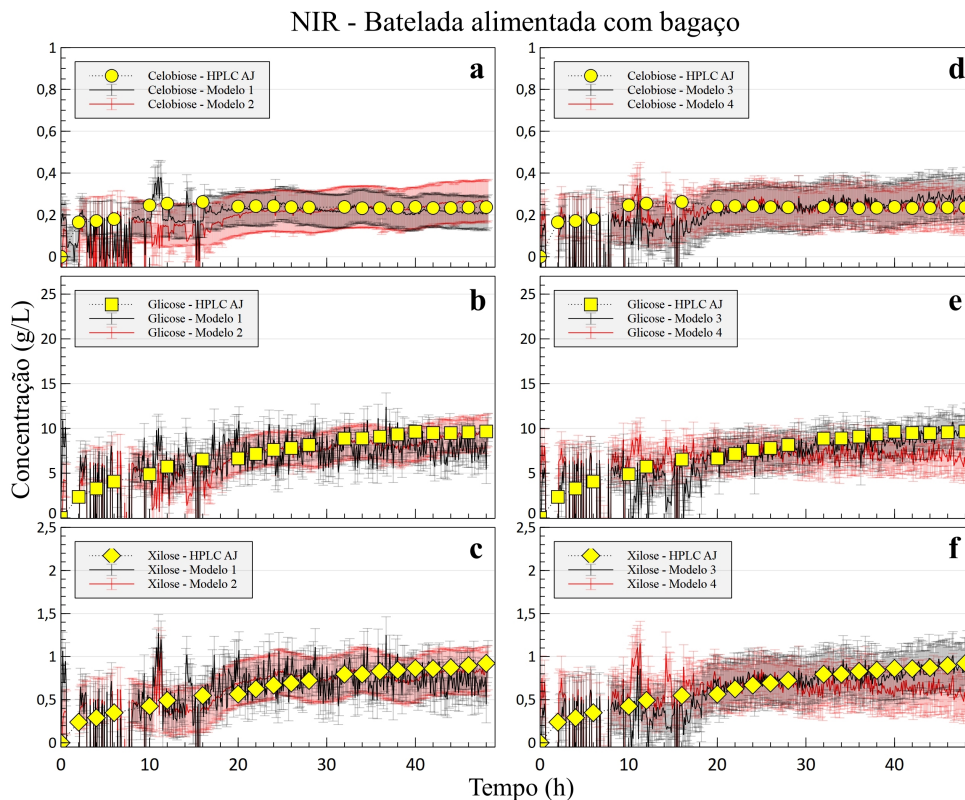
Além da avaliação dos modelos PLS, também foram avaliados as regiões de concentrações de cada carboidrato (celbiose, glicose e xilose) utilizadas no ajuste do modelo PLS, juntamente à validação cruzada, e na validação externa. A figura 14 apresenta o perfil de concentração de cada carboidrato quantificado pelo HPLC para as amostras utilizadas no ajuste PLS e na validação externa. Observa-se que a dispersão das concentrações varia de acordo com cada carboidrato. A maioria dos pontos da celbiose se concentra na região de 0,1 a 0,5 g/L, glicose de 2 a 10 g/L e xilose de 0,3 a 1,5 g/L. Analisando ambas figuras, é possível constatar que as regiões das concentrações utilizadas na validação cruzada (símbolos geométricos cheios) coincidem com a região de concentrações na validação externa (símbolos geométricos vazios). Portanto, a região utilizada na validação externa não é uma extrapolação do modelo ajustado. Logo, é consistente utilizar o modelo obtido nos dados de validação externa.



**Figura 19:** Regiões de concentrações quantificadas pelo HPLC (g/L) de cada carboidrato (celbiose, glicose e xilose). Os dados ajustados estão representados por C.vc, G.vc e X.vc, e os dados não ajustados, representados por C.ve, G.ve e X.ve. Na primeira coluna tem-se os dados ajustados e não ajustados do NIR (a, b, c) e na segunda coluna do UV-Vis (d, e, f).

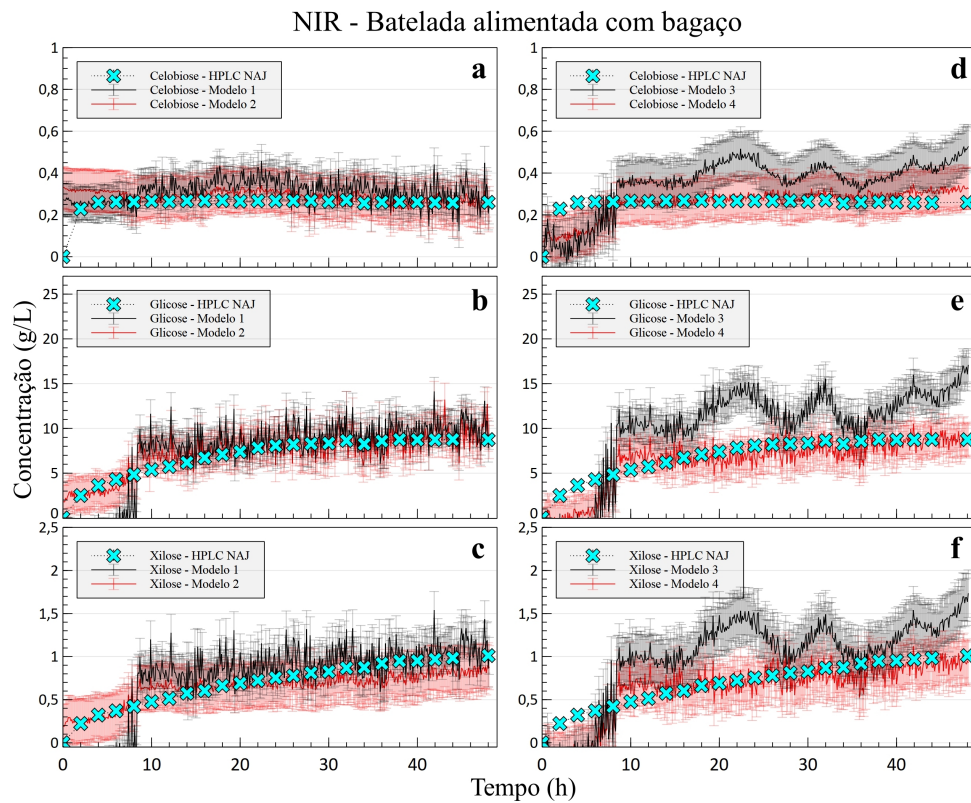
As figuras 20 a 23 ilustram o perfil de predição das concentrações de celobiose, glicose e xilose pelos modelos 1 e 3 na cor preta, 2 e 4 na cor vermelha, de cada uma das técnicas espectroscópicas, NIR (Figuras: 20 e 21) e UV-Vis (Figuras: 22 e 23), com os respectivos RMSECV representados por barras. Os gráficos na primeira coluna (a, b, c), de cada figura, ilustram as predições das concentrações dos carboidratos pelos modelos 1 e 2. Já os gráficos na segunda coluna (d, e, f), ilustram as predições pelos modelos 3 e 4.

Para cada técnica espectroscópica tem-se duas figuras, onde a primeira figura apresenta a predição das concentrações a partir dos dados utilizados no ajuste dos modelos, espectro NIR/UV-Vis e HPLCAJ (pontos ajustados), representados na cor amarela, e a segunda figura, os dados das duplicatas de cada modo de operação, espectro NIR/UV-Vis e HPLCNAJ (pontos não ajustados), representados na cor azul. Alguns pontos quantificados pelo HPLC foram retirados do ajuste por apresentarem valores fora do aumento gradativo da concentração de cada analito, possivelmente causado por algum erro grosseiro durante a retirada de amostras.

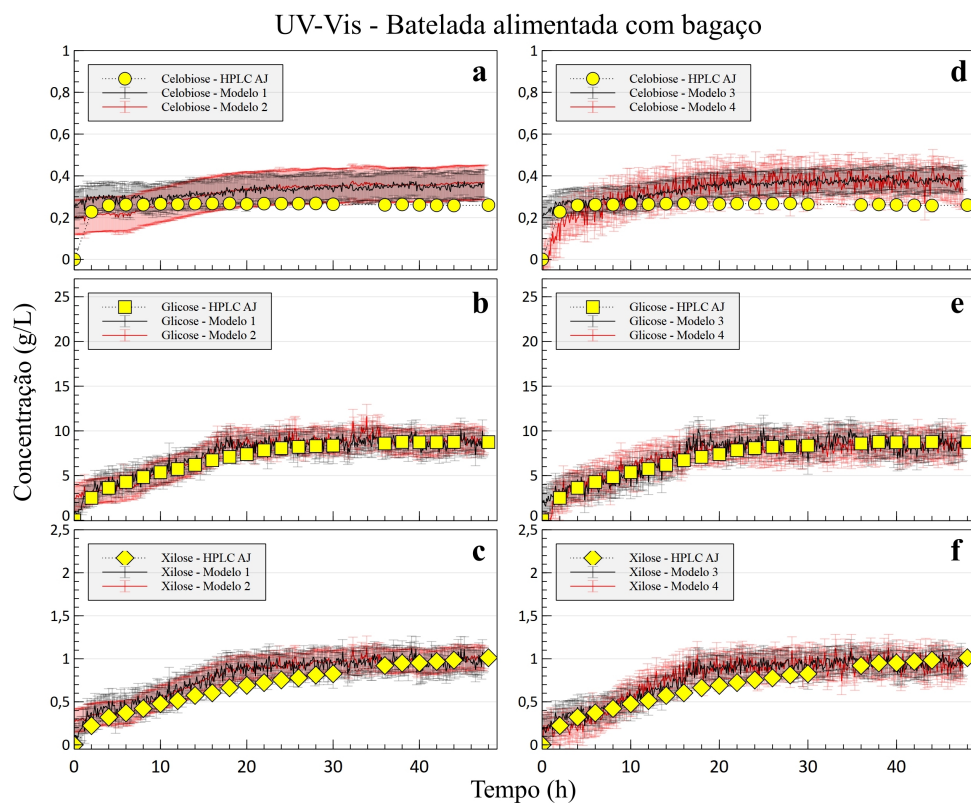


**Figura 20:** Predição das concentrações de celobiose, glicose e xilose pelos modelos PLS NIR 1 e 2 (a, b, c), 3 e 4 (d, e, f) - Batelada alimentada com bagaço 48h (dados ajustados).

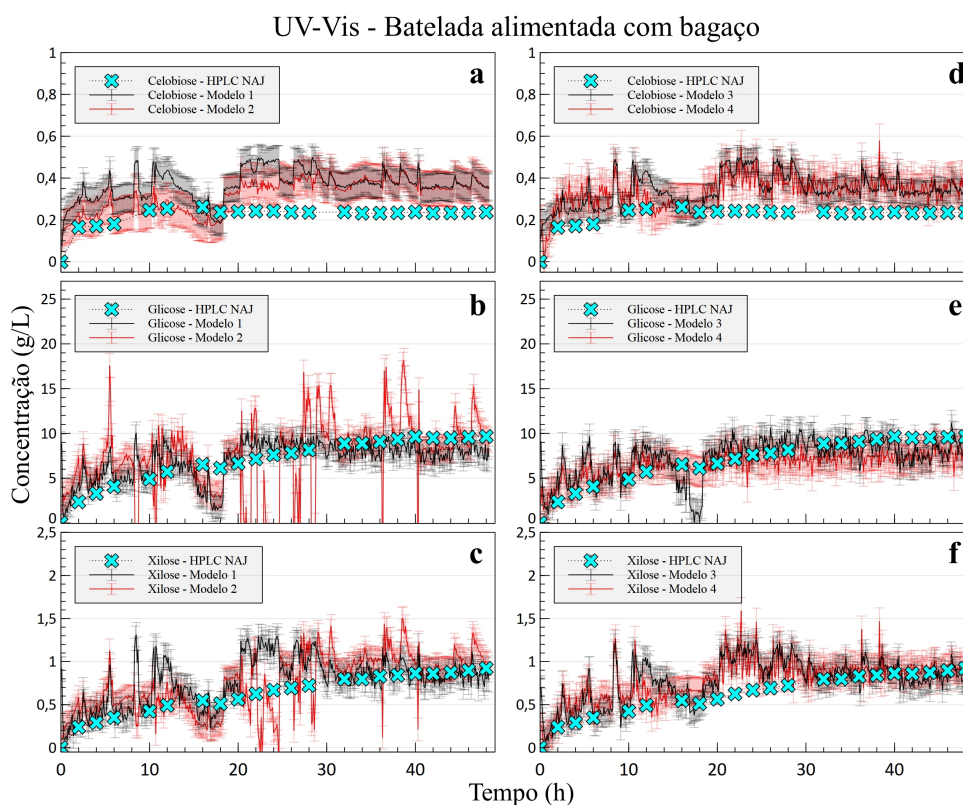




**Figura 21:** Predição das concentrações de celobiose, glicose e xilose pelos modelos PLS NIR 1 e 2 (a, b, c), 3 e 4 (d, e, f) - Batelada alimentada com bagaço 48h (validação externa).



**Figura 22:** Predição das concentrações de celobiose, glicose e xilose pelos modelos PLS UV-Vis 1 e 2 (a, b, c), 3 e 4 (d, e, f) - Batelada alimentada com bagaço 48h (dados ajustados).



**Figura 23:** Predição das concentrações de celobiose, glicose e xilose pelos modelos PLS UV-Vis 1 e 2 (a, b, c), 3 e 4 (d, e, f) - Batelada alimentada com bagaço 48h (validação externa).

Analisando os modelos 1, 2, 3 e 4 do NIR, nota-se que de modo geral a aplicação do alisamento LOESS melhorou a predição em todos os casos. Os modelos 1 e 3 tiveram um desempenho similar, assim como os modelos 2 e 4. No entanto, o modelo 4 teve uma melhor predição dos pontos quantificados, validando a seleção da região de  $3500$  a  $8000\text{ cm}^{-1}$  realizada no estudo dos espectros apresentados na Figura 14.

Entre os modelos 1, 2, 3 e 4 do UV-Vis, os modelos 3 e 4 apresentaram um melhor desempenho nas predições das concentrações dos carboidratos, principalmente o modelo 3. Para o UV-Vis, os modelos 1 e 3 tiveram um desempenho similar, diferente dos modelos 2 e 4. No modelo 4, além de outra seleção da faixa de comprimento de onda, foi utilizado a primeira derivada para uma correção da linha de base e isso melhorou o desempenho do modelo. Os quatro modelos do UV-Vis desempenharam muito bem na figura 22, o que reforça a hipótese de que a qualidade dos dados coletados influenciam diretamente na predição dos modelos PLS, pois esse foi o ensaio menos ruidoso entre todos.



No anexo A contém todos gráficos de predições das concentrações de celobiose, glicose e xilose pelos modelos 1, 2, 3 e 4 do NIR e UV-Vis. De modo geral, para o NIR e UV-Vis, os modelos 2 e 4 (em vermelho) apresentam mais ruídos que os modelos 1 e 3 (em preto). Como nos modelos 1 e 3 não há uma restrição na faixa de comprimento/número de onda, uma maior parte do espectro é ajustada e isso deixa o modelo PLS menos sensível ao ruído espectral. Em outras palavras, o ruído espectral tem um menor impacto no desempenho do modelos 1 e 3.

Entre os modelos do NIR, observa-se que o modelo 4 teve um desempenho melhor nos ensaios de batelada alimentada, principalmente na predição de Glicose e Xilose. Já o modelo 3 teve um melhor desempenho no ensaio de batelada por 48h e apresentou uma melhor predição das concentrações de Celobiose e Xilose. Analisando todos ensaios, o modelo 3 foi o que apresentou a melhor desempenho para todos carboidratos. Isso mostra que para o ajuste dos dados do NIR é melhor considerar toda região dos números de ondas ( $1100 - 10000 \text{ cm}^{-1}$ ).

No UV-Vis, os modelos 1 e 3 apresentaram uma melhor predição das concentrações. Como no NIR, o modelo 3 do UV-Vis apresentou a melhor performance para todos carboidratos, que também considera toda região dos comprimentos de onda (190 – 1100 nm).

No ensaio de batelada alimentada com enzimas, os dados do NIR que não foram ajustados (validação externa) apresentaram muitos ruídos e nenhum dos modelos teve um bom desempenho nas predições das concentrações. Já na validação externa desse mesmo ensaio no UV-Vis, apenas o modelo 2 não performou bem, ou seja, os outros modelos ficaram mais robustos e mantiveram a predição das concentrações mais consistente (contra as perturbações do sistema).

Portanto, deve-se estudar meios de melhorar a aquisição de dados, de modo que seja prático e que reduza as divergências espectrais entre cada ensaio. Entre as possíveis melhorias estão:

- Uma melhor filtragem do meio líquido que fica no reator agitado, onde se encontra as sondas de fibra ótica;
- Controle da temperatura no reator horizontal, para evitar a mudança de temperatura do líquido no reator agitado, que pode diminuir a dissolução de gases e fazer que mais bolhas de gases se formem em torno das sondas e nas paredes do reator;

- Variação da agitação do meio no reator agitado para evitar que bolhas de gás se prendam na sonda de fibra ótica, assim como o acúmulo de bagaço hidrolisado.

## 6. CONCLUSÕES

Neste trabalho foram avaliados diferentes conjuntos de dados espectroscópicos dos ensaios realizados. Os dados foram tratados por calibração multivariada, a fim de obter o melhor modelo que representasse o sistema de reação e que pudesse prever a concentração da celobiose, glicose e xilose durante a hidrólise enzimática. Por mais que o sistema reacional seja complexo e possua diversas perturbações que geram dados ruidosos, os modelos ajustados puderam prever as concentrações com uma ótima precisão, média de 16,96 % (RMSECV) entre os modelos do NIR e de 11,65 % entre os modelos do UV-Vis.

Dentre os quatro modelos ajustados, o trabalho indica que o modelo 3, em ambas técnicas espectroscópicas, teve o melhor desempenho. O modelo 4 também teve uma ótima performance, porém no NIR as predições estavam abaixo da concentração quantificada pelo HPLC e no UV-Vis esse modelo estava mais sensível as perturbações do sistema, e logo não manteve consistente as predições das concentrações dos carboidratos. As predições de celobiose, glicose e xilose pelo modelo 3 foram mais consistentes durante todo ensaio, o que indica que o uso do alisamento LOESS melhorou o poder de predição do modelo.

Assim sendo, foi possível aplicar as técnicas de NIR e UV-Vis para o monitoramento da hidrólise enzimática do bagaço de cana-de-açúcar. Entre os ensaios realizados, o ensaio de batelada alimentada com enzimas foi o que gerou mais ruídos e conseqüentemente dificultou a predição das concentrações dos carboidratos. Isso provavelmente provém de um maior grau de hidrólise do bagaço nesse ensaio e da geração de diversos compostos que absorvem ambas radiações.

Como trabalho futuro, recomenda-se pesquisar possíveis formas de melhorar o ajuste PLS. Como exemplo, estão a diminuição do número de variáveis, utilização de outros métodos de calibração multivariada e uso de redes neurais. Também, a realização de outros ensaios visando a exploração de outras regiões do espectro eletromagnético.

## 7. REFERÊNCIAS

ADNEY, B.; BAKER, J. Measurement of cellulase activities laboratory. **Technical Report NREL/TP-510-42628**, n. January, p. 8, 2008.

ANDRADE, S. R. C. et al. Comparison of multivariate calibration methods to determine simultaneously mebendazole-cambendazole and mebendazole-thiabendazole in pharmaceutical preparations by UV-visible spectrophotometry. **Journal of Pharmaceutical and Biomedical Analysis**, v. 33, n. 4, p. 655–665, 2003.

ARAÚJO, M. C. U. et al. The successive projections algorithm for variable selection in spectroscopic multicomponent analysis. **Chemometrics and Intelligent Laboratory Systems**, v. 57, n. 2, p. 65–73, 2001.

BALDEZ, T. O. **ESTUDO CINÉTICO DA HIDRÓLISE DE XILOOLIGÔMEROS POR XILANASES E BETA-XILOSIDASE NA FORMA LIVRE E IMOBILIZADA**. [s.l.] Universidade Federal de São Carlos, 2018.

BRETERON, R. G. **Applied Chemometrics for Scientists**. . England: John Wiley e Sons Ltd, 2007.

BHATIA, Latika; JOHRI, Sonia; AHMAD, Rumana. An economic and ecological perspective of ethanol production from renewable agro waste: a review. **Amb Express**, v. 2, n. 1, p. 1-19, 2012.

BIOCHEMISTRY, A. **Seventeenth Symposium on Biotechnology for Fuels and Chemicals**. [s.l: s.n.].

BLANCO, M. et al. On-line monitoring of starch enzymatic hydrolysis by near-infrared spectroscopy. **Analyst**, v. 125, n. 4, p. 749–752, 2000.

BLANCO, M.; VILLARROYA, I. NIR spectroscopy: A rapid-response analytical tool. **TrAC - Trends in Analytical Chemistry**, 2002.

CEN, H.; HE, Y. Theory and application of near infrared reflectance spectroscopy in determination of food quality. **Trends in Food Science and Technology**, 2007.

COTANA, F. et al. Production of bioethanol in a second generation prototype from pine wood chips. **Energy Procedia**, v. 45, p. 42–51, 2014.

DA SILVA, André Rodrigues Gurgel; ERRICO, Massimiliano; RONG, Ben-Guang. Techno-economic analysis of organosolv pretreatment process from lignocellulosic biomass. **Clean Technologies and Environmental Policy**, v. 20, n. 7, p. 1401-1412, 2018.

DIAS, M. O. S. et al. Second generation ethanol in Brazil: Can it compete with electricity production? **Bioresource Technology**, v. 102, n. 19, p. 8964–8971, 2011.

DIAS, M. O. S. et al. Integrated versus stand-alone second generation ethanol production from sugarcane bagasse and trash. **Bioresource Technology**, v. 103, n. 1, p. 152–161, 2012.

DU, J. et al. Enzymatic liquefaction and saccharification of pretreated corn stover at high-solids concentrations in a horizontal rotating bioreactor. **Bioprocess and Biosystems Engineering**, 2014.

EDLUND, U.; ALBERTSSON, A. C. A microspheric system: Hemicellulose-based hydrogels. **Journal of Bioactive and Compatible Polymers**, v. 23, n. 2, p. 171–186, 2008.

FANG, J.; HUANG, F.; GAO, P. Optimization of cellobiose dehydrogenase production by *Schizophyllum commune* and effect of the enzyme on kraft pulp bleaching by ligninases. **Process Biochemistry**, v. 34, n. 9, p. 957–961, 1999.

FERREIRA, M. M. C. et al. Quimiometria I: calibração multivariada, um tutorial. **Química Nova**, v. 22, p. 724-731, 1999.

FURLAN, F. F. et al. Bioelectricity versus bioethanol from sugarcane bagasse: Is it worth being flexible? **Biotechnology for Biofuels**, 2013.

FURLONG, Vitor Badiale. **Automation of a reactor for enzymatic hydrolysis of sugar cane bagasse: Computational intelligence-based adaptive control**. [s.l.] Federal University of São Carlos, 2015.

FURLONG, Vitor Badiale. **Design and Automation of a New Reactor For Enzymatic Hydrolysis of Lignocellulosic Biomass at High Solids Loading**. [s.l.] Federal University of São Carlos, 2019.

GANDLA, M. L.; MARTÍN, C.; JÖNSSON, L. J. Analytical Enzymatic Saccharification of Lignocellulosic Biomass for Conversion to Biofuels and Bio-Based Chemicals. **Energies**, v. 11, n. 11, 2018.

GHOSE, T. K. MEASUREMENT OF CELLULASE ACTIVITIES. **Pure and Applied Chemistry**, v. 59, n. 5, p. 257–268, 1987.

GODOY, C. M. DE; MACHADO, D. L.; COSTA, A. C. DA. Batch and fed-batch enzymatic hydrolysis of pretreated sugarcane bagasse – Assays and modeling. **Fuel**, v. 253, n. February, p. 392–399, 2019.

GOUVEIA, E. R. et al. Validação de metodologia para a caracterização química de bagaço de cana-de-açúcar. **Quimica Nova**, 2009.

HAALAND, D. M.; THOMAS, E. V. Partial Least-Squares Methods for Spectral Analyses. 1. Relation to Other Quantitative Calibration Methods and the Extraction of Qualitative Information. **Analytical Chemistry**, v. 60, n. 11, p. 1193–1202, 1988.

HAMELINCK, C. N.; VAN HOOIJDONK, G.; FAAIJ, A. P. C. **Ethanol from lignocellulosic biomass: Techno-economic performance in short-, middle- and long-term Biomass and Bioenergy**, 2005.

HONGDAN, Z.; SHAOHUA, X.; SHUBIN, W. Enhancement of enzymatic saccharification of sugarcane bagasse by liquid hot water pretreatment. **Bioresource Technology**, v. 143, p. 391–396, 2013.

JABLONSKÝ, M. et al. Characterization and comparison by UV spectroscopy of precipitated lignins and commercial lignosulfonates. **Cellulose Chemistry and Technology**, v. 49, n. 3–4, p. 267–274, 2015.

JACOBY, William G. Loess:: a nonparametric, graphical tool for depicting relationships between variables. **Electoral Studies**, v. 19, n. 4, p. 577-613, 2000.

KAIJANEN, L. et al. Ultraviolet detection of monosaccharides: Multiple wavelength strategy to evaluate results after capillary zone electrophoretic separation. **International Journal of Electrochemical Science**, v. 10, n. 4, p. 2950–2961, 2015.

KIM, Y.; MOSIER, N. S.; LADISCH, M. R. Enzymatic digestion of liquid hot water pretreated hybrid poplar. **Biotechnology Progress**, 2009.

KLINE, L. M. et al. Simplified determination of lignin content in hard and soft woods via UV-spectrophotometric analysis of biomass dissolved in ionic liquids. **BioResources**, 2010.

KOSTYLEV, M.; WILSON, D. Two-parameter kinetic model based on a time-dependent activity coefficient accurately describes enzymatic cellulose digestion. **Biochemistry**, 2013.

KRISTENSEN, J. B.; FELBY, C.; JØRGENSEN, H. Yield-determining factors in high-solids enzymatic hydrolysis of lignocellulose. **Biotechnology for Biofuels**, 2009.

KUMAR, P. et al. **Methods for pretreatment of lignocellulosic biomass for efficient hydrolysis and biofuel production** *Industrial and Engineering Chemistry Research*, 2009.

LIANG, C. et al. Development of modified HCH-1 kinetic model for long-term enzymatic cellulose hydrolysis and comparison with literature models. **Biotechnology for Biofuels**, v. 12, n. 1, p. 1–13, 2019.

LIMAYEM, A.; RICKE, S. C. **Lignocellulosic biomass for bioethanol production: Current perspectives, potential issues and future prospects** *Progress in Energy and Combustion Science*, 2012.

MARTÍN, Carlos et al. Comparison of the fermentability of enzymatic hydrolyzates of sugarcane bagasse pretreated by steam explosion using different impregnating agents. In: **Biotechnology for Fuels and Chemicals**. Humana Press, Totowa, NJ, 2002. p. 699-716.

MAULIDIYAH, M. et al. Lignin degradation of oil palm empty fruit bunches using TiO<sub>2</sub> photocatalyst as antifungal of *Fusarium oxysporum*. **Oriental Journal of Chemistry**, v. 33, n. 6, p. 3101–3106, 2017.

MAURYA, Devendra Prasad; SINGLA, Ankit; NEGI, Sangeeta. An overview of key pretreatment processes for biological conversion of lignocellulosic biomass to bioethanol. **3 Biotech**, v. 5, n. 5, p. 597-609, 2015.

MCINTOSH, S.; VANCOV, Tony. Enhanced enzyme saccharification of Sorghum bicolor straw using dilute alkali pretreatment. **Bioresource technology**, v. 101, n. 17, p. 6718-6727, 2010.

MOSIER, N. et al. Features of promising technologies for pretreatment of lignocellulosic biomass. **Bioresource Technology**, v. 96, n. 6, p. 673–686, 2005.

NASCIMENTO, R. J. A. DO. **MONITORAMENTO EM TEMPO REAL DA HIDRÓLISE ENZIMÁTICA DA FIBRA DA CASCA DE COCO VERDE POR**

**ESPECTROSCOPIA NO INFRAVERMELHO PRÓXIMO (NIRS).** [s.l.] Universidade Federal do Rio Grande do Norte, 2016.

NEGRO, M. J. et al. Hydrothermal Pretreatment Conditions to Enhance Ethanol Production from Poplar Biomass. **Applied Biochemistry and Biotechnology**, 2003.

NELLES, O. **Nonlinear system identification : from classical approaches to neural networks and fuzzy models.** [s.l.] Springer Science & Business Media, 2001.

PACSU, E.; HILLER, L. A. Open-Chain Sugars. I. Absorption in Ultraviolet of D-Glucose and L-Arabinose in Acid Solution. **Journal of the American Chemical Society**, v. 70, n. 2, p. 523–526, 1948.

PALMQVIST, E.; HAHN-HÄGERDAL, B. **Fermentation of lignocellulosic hydrolysates. II: Inhibitors and mechanisms of inhibition** **Bioresource Technology**, 2000.

PASQUINI, C. **Near infrared spectroscopy: Fundamentals, practical aspects and analytical applications** **Journal of the Brazilian Chemical Society**, 2003.

PERKAMPUS, H.-H. **UV-VIS Spectroscopy and Its Applications.** [s.l.] Springer Berlin Heidelberg, 1992.

PINTO, Ariane Silveira Sbrice. **MONITORAMENTO POR ESPECTROSCOPIA DOS COMPOSTOS FENÓLICOS E FURALDEÍDOS GERADOS NO PROCESSAMENTO DE BIOMASSA LIGNOCELULÓSICA.** [s.l.] Universidade Federal de São Carlos, 2017.

PINTO, A. S. S.; RIBEIRO, M. P. A.; FARINAS, C. S. Fast spectroscopic monitoring of inhibitors in the 2G ethanol process. **Bioresource Technology**, v. 250, n. November 2017, p. 148–154, 2018.

PRATTO, B. **Hidrólise Enzimática Da Palha De Cana-De-Açúcar : Estudo Cinético E Modelagem Matemática Semi- Mecanística.** [s.l.] Universidade Federal de São Carlos, 2015.

RASTOGI, M.; SHRIVASTAVA, S. Recent advances in second generation bioethanol production: An insight to pretreatment, saccharification and fermentation processes. **Renewable and Sustainable Energy Reviews**, v. 80, n. January, p. 330–340, 2017.

RIBEIRO, M. P. A. **Operação ótima de reator para síntese enzimática de ampicilina com cristalização simultânea dos produtos.** [s.l.] Universidade Federal de São Carlos, 2007.

RINNAN, Åsmund; VAN DEN BERG, Frans; ENGELSEN, Søren Balling. Review of the most common pre-processing techniques for near-infrared spectra. **TrAC Trends in Analytical Chemistry**, v. 28, n. 10, p. 1201-1222, 2009.

RUIZ, Héctor A. et al. Hydrothermal processing, as an alternative for upgrading agriculture residues and marine biomass according to the biorefinery concept: a review. **Renewable and Sustainable Energy Reviews**, v. 21, p. 35-51, 2013.

SÁNCHEZ, C. **Lignocellulosic residues: Biodegradation and bioconversion by fungi** **Biotechnology Advances**, 2009.

SAVITZKY, A.; GOLAY, M. J. E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. **Analytical Chemistry**, v. 36, n. 8, p. 1627-1639, 1964/07/01 1964.

SCHWANNINGER, M.; RODRIGUES, J. C.; FACKLER, K. A review of band assignments in near infrared spectra of wood and wood components. **Journal of Near Infrared Spectroscopy**, v. 19, n. 5, p. 287–308, 2011.

SINGH, Renu et al. A review on delignification of lignocellulosic biomass for enhancement of ethanol production potential. **Renewable and Sustainable Energy Reviews**, v. 32, p. 713-728, 2014.

SLUITER, A. et al. NREL/TP-510-42618 analytical procedure - Determination of structural carbohydrates and lignin in Biomass. **Laboratory Analytical Procedure (LAP)**, 2008.

VON SIVERS, Margareta; ZACCHI, Guido. A techno-economical comparison of three processes for the production of ethanol from pine. **Bioresource Technology**, v. 51, n. 1, p. 43-52, 1995.

YANG, Bin; WYMAN, Charles E. Pretreatment: the key to unlocking low-cost cellulosic ethanol. **Biofuels, Bioproducts and Biorefining: Innovation for a sustainable economy**, v. 2, n. 1, p. 26-40, 2008.



YOSHIDA, M. et al. Effects of Cellulose Crystallinity, Hemicellulose, and Lignin on the Enzymatic Hydrolysis of *Miscanthus sinensis* to Monosaccharides. **Bioscience, Biotechnology, and Biochemistry**, 2008.

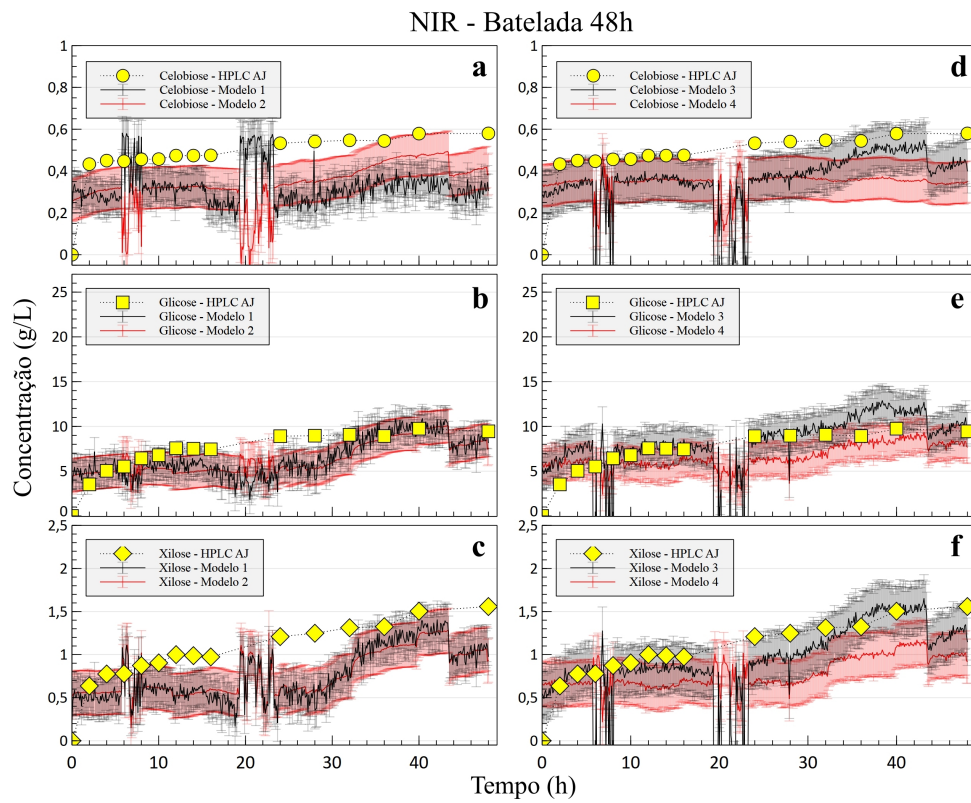
YU, Q. et al. Two-step liquid hot water pretreatment of *Eucalyptus grandis* to enhance sugar recovery and enzymatic digestibility of cellulose. **Bioresource Technology**, 2010.

YU, Q. et al. Liquid hot water pretreatment of sugarcane bagasse and its comparison with chemical pretreatment methods for the sugar recovery and structural changes. **Bioresource Technology**, v. 129, p. 592–598, 2013.

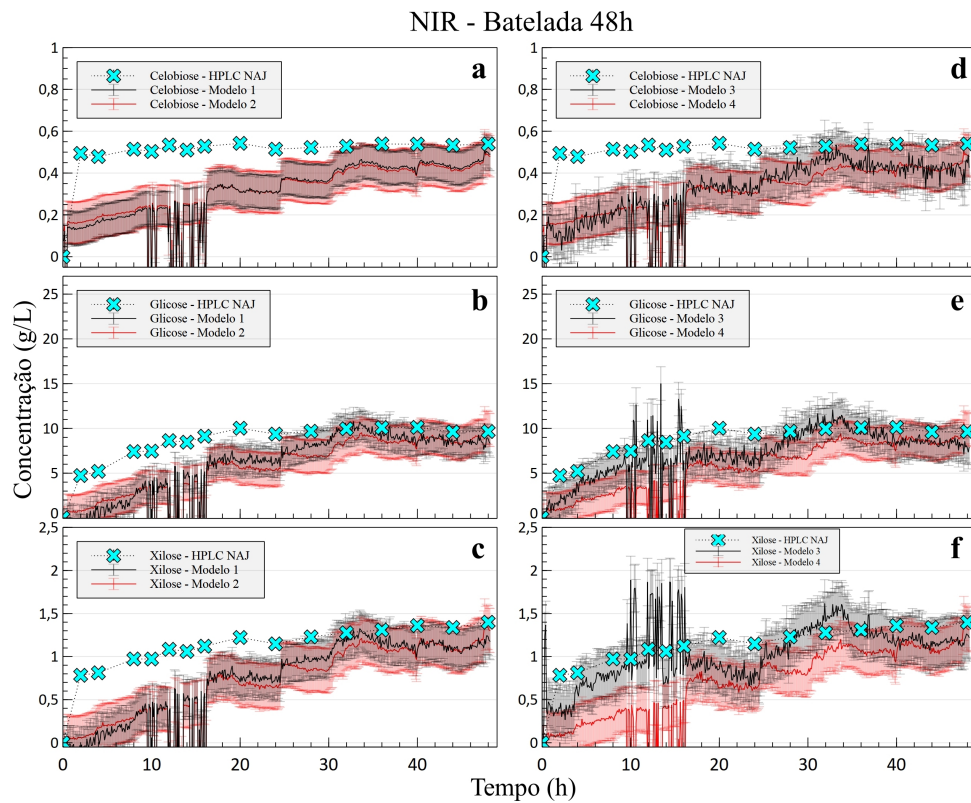
ZAKAMULIN, V. (2016). Moving Averages for Market Timing. *Available at SSRN 2854180*.

ZHOU, C. et al. Prediction of mixed hardwood lignin and carbohydrate content using ATR-FTIR and FT-NIR. **Carbohydrate Polymers**, v. 121, p. 336–341, 2015.

**ANEXO A – GRÁFICOS DE PREDIÇÕES DAS CONCENTRAÇÕES DE  
CELOBIOSE, GLICOSE E XILOSE PELOS MODELOS PLS NIR E UV-VIS.**

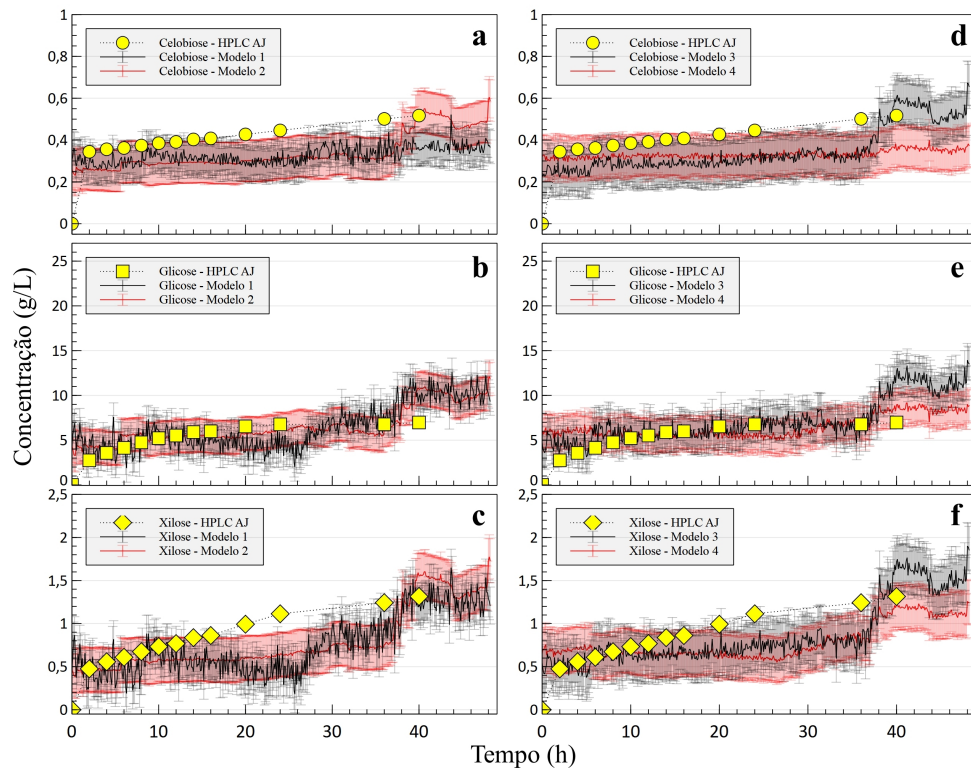


**Figura 24:** Predição das concentrações de celobiose, glicose e xilose pelos modelos PLS NIR 1 e 2 (a, b, c), 3 e 4 (d, e, f) - Batelada 48h (dados ajustados).



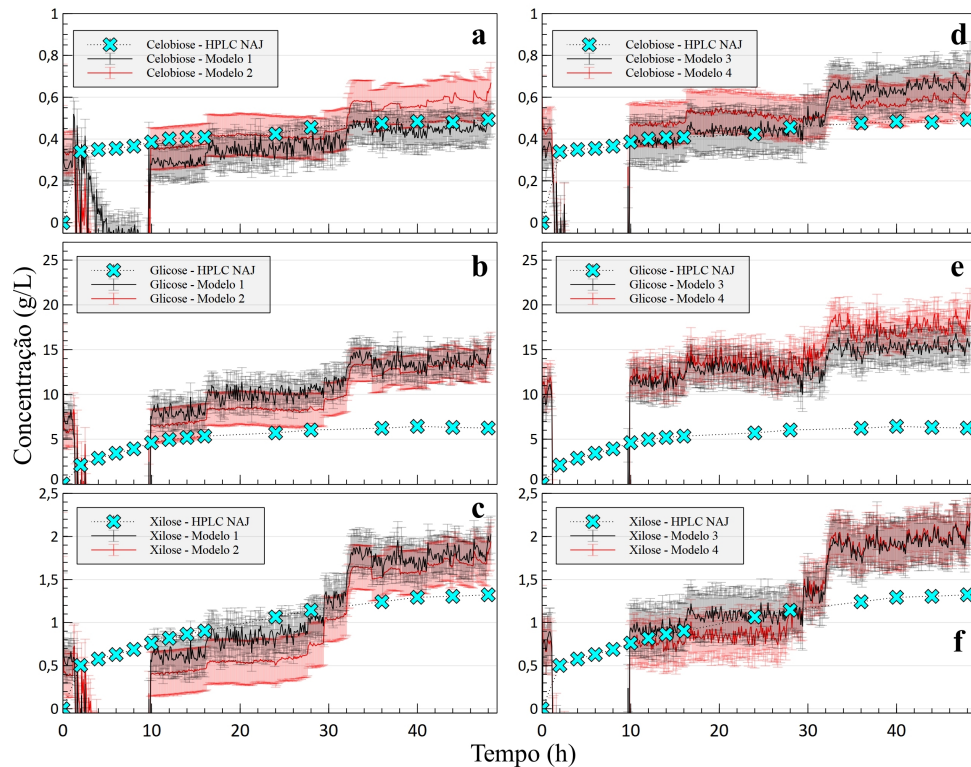
**Figura 25:** Predição das concentrações de celobiose, glicose e xilose pelos modelos PLS NIR 1 e 2 (a, b, c), 3 e 4 (d, e, f) - Batelada 48h (validação externa).

## NIR - Batelada alimentada com enzimas

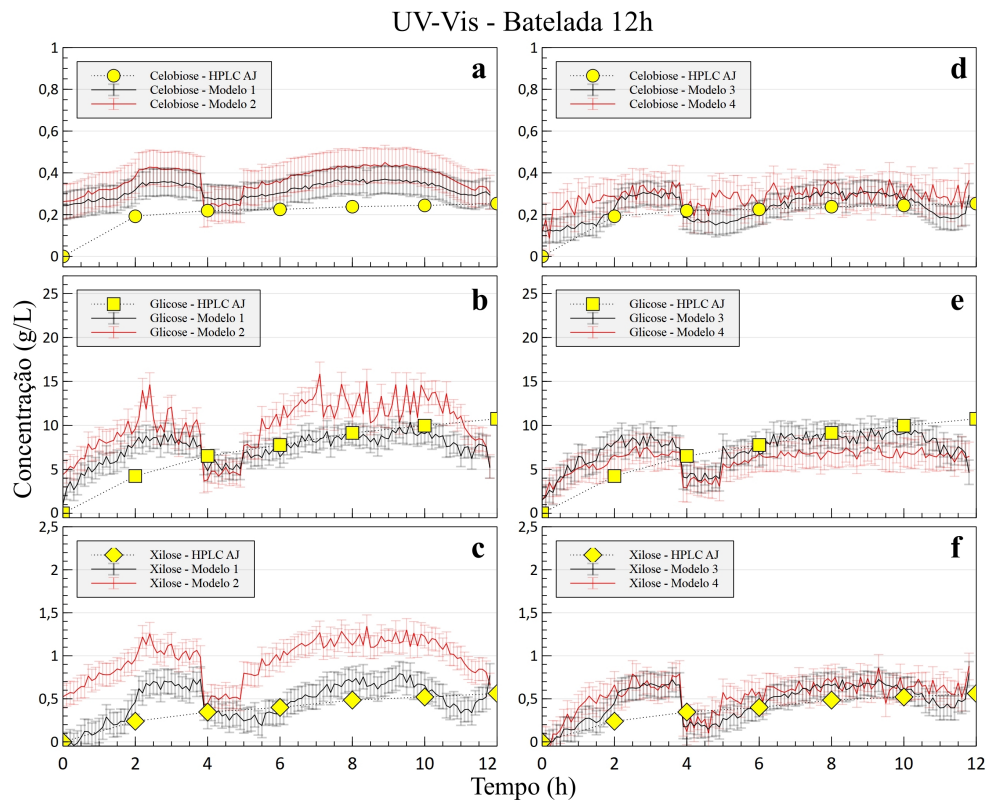


**Figura 26:** Predição das concentrações de celobiose, glicose e xilose pelos modelos PLS NIR 1 e 2 (a, b, c), 3 e 4 (d, e, f) - Batelada alimentada com enzimas 48h (dados ajustados).

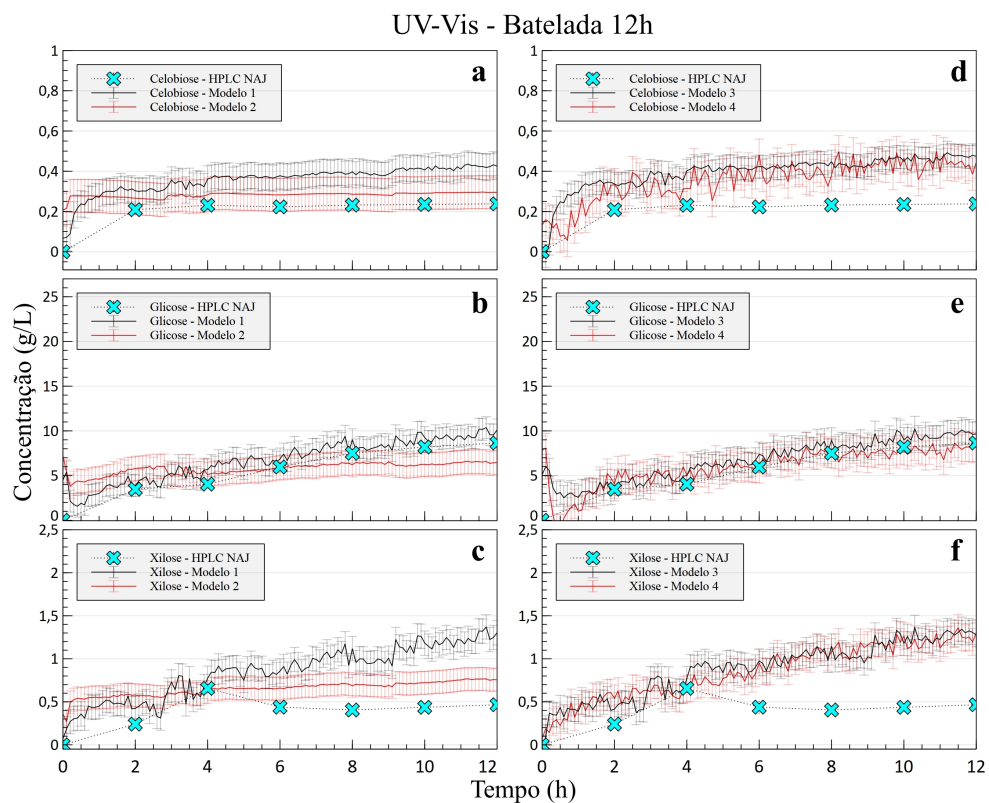
## NIR - Batelada alimentada com enzimas



**Figura 27:** Predição das concentrações de celobiose, glicose e xilose pelos modelos PLS NIR 1 e 2 (a, b, c), 3 e 4 (d, e, f) - Batelada alimentada com enzimas 48h (validação externa).

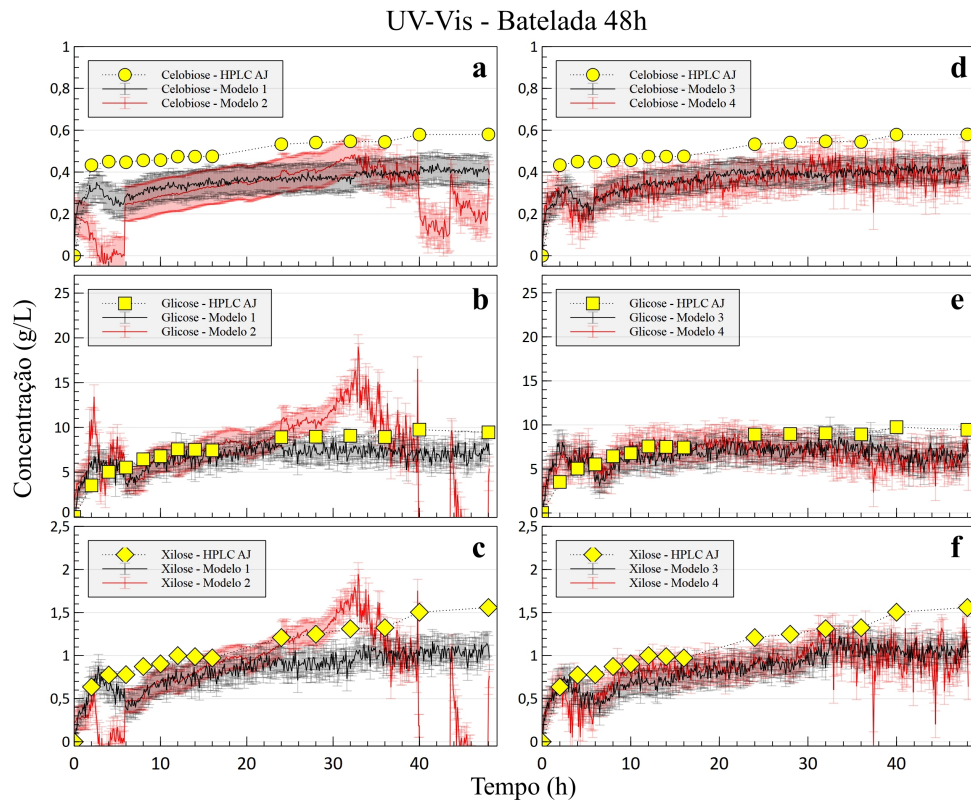


**Figura 28:** Predição das concentrações de celobiose, glicose e xilose pelos modelos PLS UV-Vis 1 e 2 (a, b, c), 3 e 4 (d, e, f) - Batelada 12h (dados ajustados).

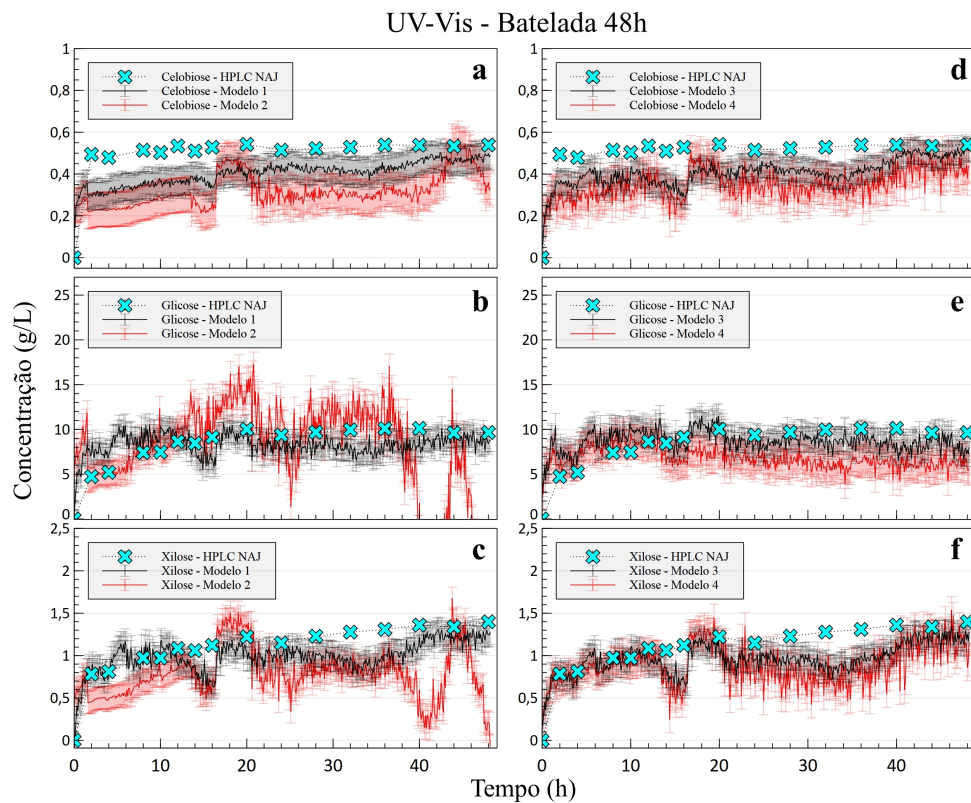


**Figura 29:** Predição das concentrações de celobiose, glicose e xilose pelos modelos PLS UV-Vis 1 e 2 (a, b, c), 3 e 4 (d, e, f) - Batelada 12h (validação externa).

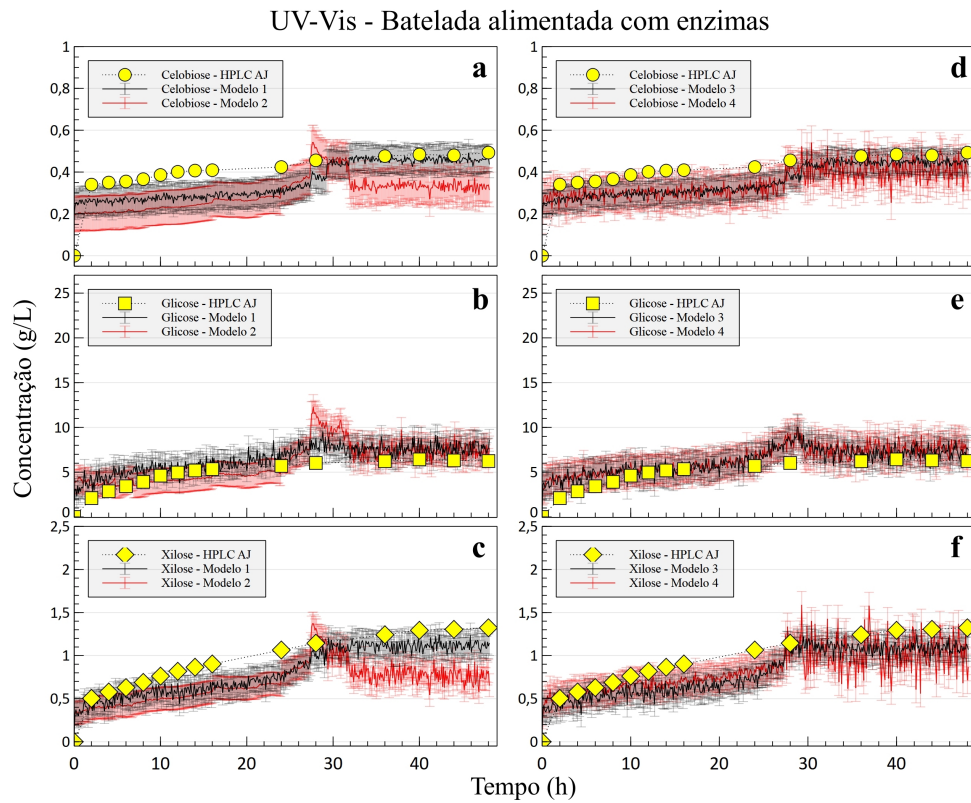




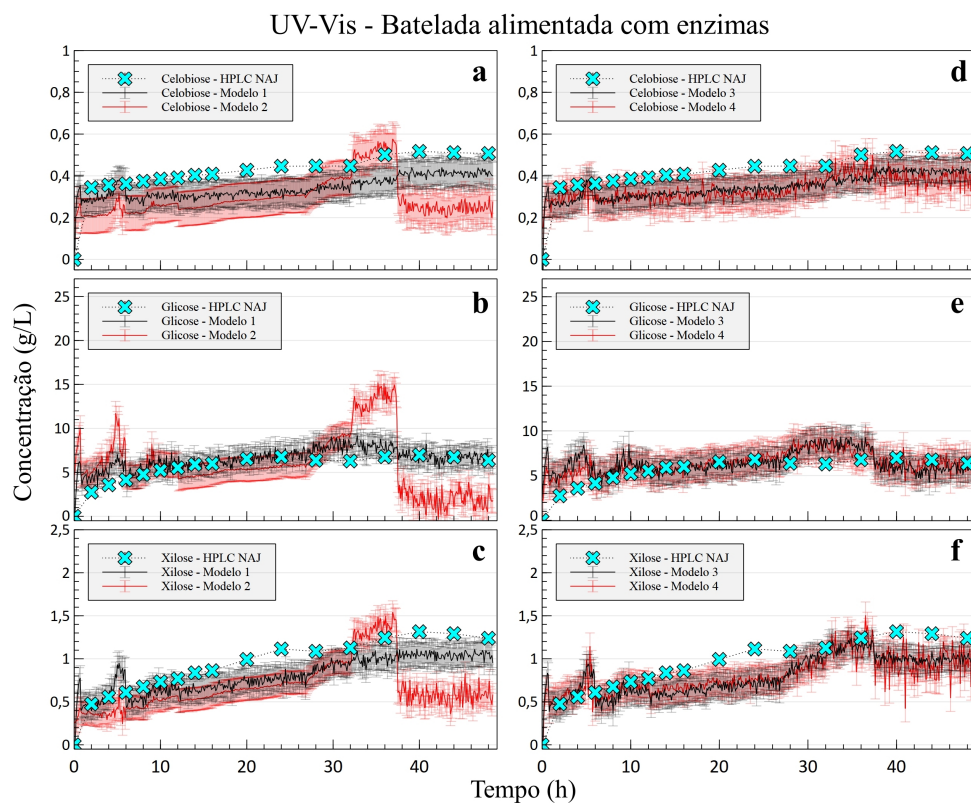
**Figura 30:** Predição das concentrações de celobiose, glicose e xilose pelos modelos PLS UV-Vis 1 e 2 (a, b, c), 3 e 4 (d, e, f) - Batelada 48h (dados ajustados).



**Figura 31:** Predição das concentrações de celobiose, glicose e xilose pelos modelos PLS UV-Vis 1 e 2 (a, b, c), 3 e 4 (d, e, f) - Batelada 48h (validação externa).



**Figura 32:** Predição das concentrações de celobiose, glicose e xilose pelos modelos PLS UV-Vis 1 e 2 (a, b, c), 3 e 4 (d, e, f) - Batelada alimentada com enzimas 48h (dados ajustados).



**Figura 33:** Predição das concentrações de celobiose, glicose e xilose pelos modelos PLS UV-Vis 1 e 2 (a, b, c), 3 e 4 (d, e, f) - Batelada alimentada com enzimas 48h (validação externa).