

Rio de Janeiro, RJ / Novembro, 2025

Análise multivariada de dados espectroscópicos obtidos por ressonância magnética nuclear para a caracterização de biomassas utilizadas na produção de biocombustíveis de segunda geração



***Empresa Brasileira de Pesquisa Agropecuária
Embrapa Solos
Ministério da Agricultura e Pecuária***

ISSN 1517-2627 / e-ISSN 2966-2443

Documentos 251

Novembro, 2025

Análise multivariada de dados
espectroscópicos obtidos por ressonância
magnética nuclear para a caracterização
de biomassas utilizadas na produção de
biocombustíveis de segunda geração

Etelvino Henrique Novotny

Bianca Braz Mattos

Rodrigo Henrique dos Santos Garcia

Eduardo Ribeiro de Azevedo

Michele Fabri de Resende

Embrapa Solos
Rio de Janeiro, RJ
2025

Embrapa Solos

Rua Jardim Botânico, 1024.
Jardim Botânico.
Rio de Janeiro, RJ
22460-000
www.embrapa.br/solos
www.embrapa.br/fale-conosco

Comitê Local de Publicações

Presidente

Cláudia Pozzi Jantalia

Secretário-executivo

Marcos Antônio Nakayama

Membros

*Bernadete da Conceição Carvalho
Gomes Pedreira, David Vilas Boas
de Campos, Evaldo de Paiva Lima,
Helga Restum Hissa, José Francisco
Lumbreras, Joyce Maria Guimarães
Monteiro, Lucia Raquel Queiroz Pereira
da Luz, Maurício Rizzato Coelho e
Wenceslau Geraldes Teixeira*

Edição executiva

Marcos Antônio Nakayama

Revisão de texto

Marcos Antônio Nakayama

Normalização bibliográfica

Luciana Sampaio de Araujo

Projeto gráfico

Leandro Sousa Fazio

Diagramação

Alexandre Abrantes Cotta de Mello

Foto da capa

Neide Makiko Furukawa

Publicação digital: PDF

Todos os direitos reservados

A reprodução não autorizada desta publicação, no todo ou em parte,
constitui violação dos direitos autorais (Lei nº 9.610).

Dados Internacionais de Catalogação na Publicação (CIP)

Embrapa Solos

Análise multivariada de dados espectroscópicos obtidos por ressonância magnética nuclear para a caracterização de biomassas utilizadas na produção de biocombustíveis de segunda geração / Etelvino Henrique Novotny ... [et al.]. — Rio de Janeiro : Embrapa Solos, 2025.

PDF (28 p.) : il. color. — (Documentos / Embrapa Solos, e-ISSN 2966-2443 ; 251).

1. Espectrometria. 2. Método Estatístico. 3. Biocombustível. I. Novotny, Etelvino Henrique. II. Mattos, Bianca Braz. III. Garcia, Rodrigo Henrique dos Santos. IV. Azevedo, Eduardo Ribeiro de. V. Resende, Michele Fabri de. VI. Embrapa Solos. VII. Série.

CDD (23. ed.) 543.66

Luciana Sampaio de Araujo (CRB-7/5165)

© 2025 Embrapa

Autores

Etelvino Henrique Novotny

Agrônomo, doutor em Química, pesquisador da Embrapa Solos, Rio de Janeiro, RJ

Bianca Braz Mattos

Bióloga, mestra em Microbiologia, analista da Embrapa Solos, Rio de Janeiro, RJ

Rodrigo Henrique dos Santos Garcia

Químico, doutor em Ciência, bolsista (pós-doutorado) do Instituto de Física de São Carlos, SP

Eduardo Ribeiro de Azevedo

Físico, doutor em Física, professor do Instituto de Física de São Carlos, SP

Michele Fabri de Resende

Química, doutora em Química, bolsista (pós-doutorado — PronaSolos/Finep) da Embrapa Solos, Rio de Janeiro, RJ

Apresentação

A Resolução Multivariada de Curvas (MCR) é uma técnica analítica multivariada frequentemente empregada para a análise de dados espectroscópicos, fornecendo o número de componentes presentes em um sistema químico, os espectros puros desses componentes e seus perfis de concentração. Essa abordagem usualmente depende da presença de variabilidade sistemática significativa entre os espectros de várias amostras, o que é atribuído a diversas causas de variação relacionadas a diferenças na origem, composição e tratamentos físico-químicos das amostras. Na Ressonância Magnética Nuclear (RMN) no estado sólido, a MCR tem sido empregada como uma técnica de pós-processamento para eliminar o ruído ou para edição espectral com base em certas propriedades da RMN. Nesse uso, induz-se a variabilidade ao incrementar um parâmetro específico na sequência de pulsos, que codifica a propriedade de interesse para a separação dos espectros.

Neste manuscrito, da série Documentos, propõe-se a utilização de uma sequência de pulso específica para gerar variabilidade controlada nos espectros de RMN de ^{13}C no estado sólido, empregando posteriormente a MCR para isolar os espectros puros dos componentes com base nas características ligadas à variabilidade criada. A primeira prova do conceito envolveu a análise de bagaço de cana-de-açúcar, submetido a diferentes pré-tratamentos químicos visando aumentar a eficiência enzimática para produção de etanol de segunda geração, em que uma série de espectros de RMN de ^{13}C no estado sólido, obtidos utilizando a sequência de pulso Torchia- T_1 com períodos de relaxação spin-rede (T_1) crescentes, teve sua matriz de dados resultante analisada por MCR, estimando-se os espectros individuais dos distintos componentes moleculares (celulose, xilano e lignina) dessa complexa biomassa. Utilizando a mesma sequência de pulsos, apresentamos exemplos adicionais de aplicação para ilustrar o potencial, otimizar parâmetros e/ou delinear as limitações da técnica. Como uma segunda prova do conceito, implementamos a

metodologia em polietileno de média densidade semicristalino (PE) e poli(éter-éter-cetona) (PEEK), obtendo estimativas dos espectros individuais de RMN de ^{13}C das cadeias poliméricas nas regiões amorfa (T_1 curto) e cristalina (T_1 longo). A análise também forneceu as intensidades relativas de cada espectro puro previsto, que correspondem aos decaimentos típicos de T_1 das frações dos domínios amorfo e cristalino. Além disso, aplicamos o estudo ao poli(1-butenio) isotático (iPB-I) como um exemplo em que a variabilidade induzida de T_1 surge da disparidade de mobilidade entre a espinha dorsal do polímero e sua cadeia lateral. Uma abordagem de jack-knifing e um teste t de Student permitiram determinar o número mínimo de espectros e a faixa de tempos de relaxação necessários para uma estimativa precisa dos espectros puros individuais e suas intensidades relativas.

Este trabalho atende ao Objetivo de Desenvolvimento Sustentável (ODS) estipulado pelas Nações Unidas (ONU) ODS 7: Energia Limpa e Acessível, mais especificamente à meta 7.2, que pretende, “até 2030, aumentar substancialmente a participação de energias renováveis na matriz energética global”.

Marisa Teixeira Mattioli

Chefe-Geral em exercício da Embrapa Solos

Sumário

Introdução	9
Desenvolvimento da metodologia	13
Avaliação do efeito de pré-tratamentos na estrutura e morfologia do bagaço de cana	13
Aplicações em polímeros semicristalinos	14
Processamento de dados por Resolução de Curvas Multivariadas (MCR)	15
Caracterização e quantificação dos constituintes do bagaço de cana submetidos a pré-tratamentos visando aumentar a eficiência enzimática para produção de etanol de segunda geração	16
Otimização do método proposto utilizando polímeros sintéticos	19
Conclusão	25
Agradecimentos	26
Referências	26

Introdução

A busca por alternativas aos combustíveis fósseis é tema prioritário nas agendas políticas ao redor do mundo e, nesse contexto, biocombustíveis de segunda geração já provaram seu potencial (Espírito Santo et al., 2022). Porém, qualquer programa de produção desses requer a correta caracterização da composição, configuração, conformação e morfologia das potenciais matérias-primas, usualmente resíduos lignocelulósicos agroindustriais. Além disso, a avaliação da eficiência de pré-tratamentos da biomassa, visando otimizar ou até mesmo viabilizar a obtenção desses biocombustíveis, requer métodos analíticos apropriados.

A biomassa vegetal seca é composta principalmente de celulose (35 a 50%), seguida por hemicelulose (20 a 35%) e lignina (10 a 25%). A celulose é um homopolímero linear de unidades de D-glicose ligadas por ligações 1,4- β -D-glicosídicas, conhecida por ser semicristalina. As hemiceluloses são compostas principalmente por pentoses (principalmente xilose e arabinose), com contribuições menores de hexoses (manose, glicose e galactose) e heteropolímeros de ácidos urônicos (tais como ácidos glucurônico e galacturônico). A lignina é um polímero fenólico que forma uma rede estrutural tridimensional ao redor das frações de celulose e hemicelulose, sendo um dos fatores mais importantes na determinação da recalcitrância da parede celular à hidrólise (Paul, 2007; Wall et al., 2008; Mielenz, 2009; Ding et al., 2012).

Visando à produção de biocombustíveis a partir de resíduos renováveis (como o bagaço de cana-de-açúcar, por exemplo), é essencial converter os carboidratos da biomassa em açúcares simples. O desenvolvimento de estratégias específicas de despolimerização da biomassa requer uma compreensão global das características físicas, químicas e morfológicas das paredes celulares das plantas, bem como dos efeitos dos pré-tratamentos e da ação enzimática nessas propriedades da biomassa. Como o rendimento da hidrólise enzimática está intimamente relacionado à composição, conformação e configuração desses biopolímeros, assim como à anatomia e

morfologia das paredes celulares das plantas, sua análise e compreensão requerem o uso de métodos especializados, com capacidade de elucidar cada uma dessas características. Existem várias metodologias que podem ser utilizadas para esse fim (Rezende et al., 2011). Entretanto, neste texto, dá-se ênfase à Ressonância Magnética Nuclear (RMN) no estado sólido, que, além da composição química, fornece importantes informações sobre a cristalinidade, configuração e conformação estrutural da amostra (Bernardinelli et al., 2016; Alessi et al., 2018; Espírito Santo et al., 2019, 2022).

A espectroscopia de RMN de ^1H , ^{13}C , ^{15}N e ^{31}P tem sido um dos métodos analíticos mais importantes para analisar biomassas e seus produtos de transformação. Ela pode ser utilizada para analisar carboidratos, proteínas, lipídios e compostos policondensados em biomassa bruta e processada, de plantas, animais, fungos, algas e outras origens. A RMN tem sido utilizada em análises qualitativas e quantitativas de biomassas; determinação da composição química; estrutura e dinâmica de monômeros, oligômeros e polímeros naturais (Bernardinelli et al., 2016). Para citar alguns exemplos de estudos em biomassas, a RMN foi amplamente utilizada para caracterização de transformações em biomassas induzidas por tratamentos químicos e enzimáticos utilizados na produção de etanol de segunda geração. Utilizando tais técnicas, foi possível determinar as mudanças na composição química das biomassas devido a pré-tratamentos químicos (Bernardinelli et al., 2016; Alessi et al., 2018; Espírito Santo et al., 2019, 2022); o consumo de carboidratos por enzimas durante a hidrólise enzimática e a variação no grau de cristalinidade de carboidratos (Rezende et al., 2011; Bernardinelli et al., 2016; Alessi et al., 2018; Espírito Santo et al., 2019, 2022); e a alteração na porosidade e acessibilidade a fluidos de biomassas (Besser et al., 2018; Espírito Santo et al., 2022). Tais estudos contribuíram efetivamente para entender os principais fatores que levam à recalcitrância de biomassa lignocelulósica a enzimas, o que ajudou a melhorar os procedimentos para a produção de etanol de segunda e terceira gerações a partir dessas biomassas (Lima et al., 2014; Bernardinelli et al., 2016; Alessi et al., 2018; Espírito Santo et al., 2019, 2022).

O uso de métodos multivariados para auxiliar no processamento e interpretação de dados de RMN de ^{13}C no estado sólido (RMN ^{13}C) é um procedimento relativamente comum. No procedimento mais comum, os espectros de RMN ^{13}C de um conjunto de diferentes amostras que apresentam variabilidade (por exemplo, amostras de diferentes espécies vegetais ou submetidas a diferentes tratamentos, como localização, idade, cultivares de programas de melhoramento ou submetidos a tratamentos físicos ou químicos específicos) são utilizados para criar uma matriz de dados com a intensidade espectral, em que uma das dimensões corresponde aos deslocamentos químicos e a outra às diferentes amostras. Com a análise multivariada, como a Análise por Componentes Principais (PCA), uma ferramenta matemática usual, é possível resumir a informação original em um número reduzido de novas variáveis ortogonais (vetores, eixos), avaliando a variabilidade principal na matriz de dados e detectando a estrutura subjacente dos dados de acordo com a forma como as variáveis (deslocamentos químicos, i.e. estrutura da biomassa) estão relacionadas (Rocha et al., 2016).

Além disso, a PCA facilita o reconhecimento de padrões na distribuição das amostras de acordo com sua semelhança (agrupamento) e particularidades, bem como tendências (variabilidade contínua entre as amostras em vez de grupos bem definidos). As diferenças espectrais que levam à distribuição das amostras no novo eixo ortogonal podem ser identificadas nos carregamentos das PCs e, depois, associadas a variações específicas na composição e estrutura das amostras. Outras abordagens, como a Resolução Multivariada de Curvas (MCR), também são comumente usadas para determinar o número de componentes coexistentes em um sistema químico, permitindo a estimativa dos espectros puros dos componentes (análise qualitativa), bem como os perfis de concentração de cada componente (análise quantitativa) (Novotny et al., 2012).

Um pré-requisito óbvio desses métodos é a existência de variabilidade sistemática (estrutura subjacente dos dados) entre as amostras, e essa variabilidade é causada por diferentes fontes de variação presentes na matriz de dados.

Na abordagem multivariada, a variabilidade total e a covariância em todo o espectro de RMN são sistematicamente levadas em conta, e a distribuição das amostras é baseada na variabilidade sistemática encontrada no conjunto completo de amostras, pois o objetivo é geralmente distribuí-las de acordo com suas características físico-químicas reveladas pelos seus espectros de RMN. Na maioria das vezes, a variabilidade capturada pela análise multivariada resulta de diferenças na preparação das amostras, localização geográfica, contaminação ou tratamento físico ou químico, como demonstrado em Espírito Santo et al. (2019).

Outra abordagem foi recentemente proposta (Espírito Santo et al., 2022; Novotny et al., 2023), em que, em vez de utilizar um conjunto de amostras com variabilidade aleatória ou natural, essa variabilidade é induzida em cada amostra individual de forma controlada, utilizando sequências de pulsos de RMN especiais. A proposta baseia-se na sensibilidade conhecida da RMN às mudanças nas interações entre os spins nucleares, o que se traduz em alta seletividade para caracterizar as propriedades estruturais e dinâmicas dos materiais. Mais especificamente, a sequência de pulsos Torchia- T_1 (Torchia, 1978) foi utilizada para modular as intensidades espectrais (variabilidade induzida) dos espectros de RMN ^{13}C resultantes por um fator que depende do tempo de relaxação longitudinal (T_1) dos núcleos de ^{13}C . Assim, variando um parâmetro ajustável na sequência de pulsos Torchia- T_1 , é possível obter um conjunto de espectros de RMN ^{13}C com intensidades moduladas pelo T_1 de cada grupo químico de cada amostra individualmente. Em outras palavras, usando a sequência de pulsos Torchia- T_1 , uma variabilidade dependente de T_1 foi induzida nos espectros de RMN ^{13}C . Em seguida, o procedimento MCR foi realizado a esse conjunto de espectros para separar os diferentes componentes do espectro de acordo com seus tempos de relaxação T_1 do ^{13}C .

Assim, o objetivo deste trabalho foi desenvolver um método que utiliza a variabilidade induzida e controlada pela sequência de pulsos Torchia- T_1 , seguida da análise multivariada de dados por MCR, para a caracterização do bagaço de cana e a otimização desse método utilizando polímeros sintéticos.

Desenvolvimento da metodologia

Avaliação do efeito de pré-tratamentos na estrutura e morfologia do bagaço de cana

Amostras

O bagaço de cana utilizado neste estudo foi coletado na usina Raizen (Costa Pinto, Piracicaba, SP, Brasil) logo após a moagem da cana. O material foi seco a 40 °C e então armazenado em recipientes plásticos à temperatura ambiente. O bagaço de cana foi submetido a quatro pré-tratamentos: alcalino; ácido; alcalino seguido do ácido e ácido seguido do alcalino, sendo essas quatro amostras as utilizadas neste estudo. Esses pré-tratamentos foram realizados em autoclave a 120 °C e pressão de 1,05 bar por 40 min. Foi utilizada a razão de 1 g de bagaço para 10 mL de soluções diluídas (1% m/v) de hidróxido de sódio (pré-tratamento alcalino) ou de ácido sulfúrico (ácido). Os demais pré-tratamentos consistiram na combinação sequencial desses, após resfriamento, lavagem até pH 7 e secagem subsequente ao primeiro passo.

RMN ^{13}C no estado sólido

Os espectros de RMN ^{13}C no estado sólido foram adquiridos no espectrômetro Varian Unit Inova, operando em frequências de 100,5 MHz e 400,0 MHz para ^{13}C e ^1H , respectivamente. A sequência de pulsos utilizadas foi a Torchia- T_1 com τ de 0,1; 0,5; 1; 1,5; 2; 2,5; 3; 4; 5; e 8 s. Para excitação, utilizou-se a polarização cruzada com amplitude variável, com tempo de contato de 1 ms e rotação em torno do ângulo mágico (MAS) a uma taxa de 10 kHz. O desacoplamento foi do tipo dipolar de alta potência (70 kHz), com tempo de reciclagem de 2 s (cinco vezes o T_1^{H} mais longo determinado experimentalmente).

Aplicações em polímeros semicristalinos

Amostras

- a) Polietileno de média densidade: $[\text{CH}_2\text{CH}_2]_n$, densidade (d) de $0,94 \text{ g/cm}^3$ a 25°C , transição vítrea (T_g) a -100°C e temperatura de fusão (T_f) de $\sim 130^\circ\text{C}$. É um polímero semicristalino com estrutura alifática, obtido da Aldrich e utilizado como recebido (PE);
- b) Poli(éter-éter-cetona): $[-\text{C}_6\text{H}_4-\text{O}-\text{C}_6\text{H}_4-\text{O}-\text{C}_6\text{H}_4-\text{CO}-]_n$, d: $1,32 \text{ kg/cm}^3$ a 25°C , $T_g \sim 150^\circ\text{C}$, $T_f \sim 350^\circ\text{C}$. É um polímero semicristalino com estrutura aromática, obtido da Aldrich e utilizado como recebido (PEEK);
- c) Poli(1-butenos) isotático: $[-\text{CH}_2\text{CRH}-]_n$ com $\text{R} = [\text{CH}_2\text{CH}_3]$, d: $0,95 \text{ g/cm}^3$ a 25°C , $T_g \sim -20^\circ\text{C}$, $T_f \sim 135^\circ\text{C}$. É um polímero semicristalino poliolefinico, com 2 carbonos sp^2 (rígidos) no esqueleto e outros 2 hibridizados sp^3 (móveis) na cadeia lateral por unidade repetitiva. Também foi obtido da Aldrich. Os pellets de polímero, assim recebidos, foram preparados para os estudos de RMN por fusão a aproximadamente 150°C , seguidos de resfriamento lento e armazenamento à temperatura ambiente por várias semanas para garantir a conversão completa para a forma cristalina I (iPB-I). O procedimento detalhado pode ser encontrado em Azevedo et al. (2000).

RMN ^{13}C no estado sólido com supressão das bandas laterais rotacionais

Os experimentos de RMN ^{13}C foram realizados utilizando um espectrômetro Bruker Avance III, operando nas frequências de 100,5 MHz e 400,0 MHz para ^{13}C e ^1H , respectivamente. A sequência de pulsos utilizada foi a Torchia- T_1 com τ de 0,1; 0,5; 1;

1,5; 2; 2,5; 3; 4; 5; e 8 s. Para excitação, utilizou-se a polarização cruzada com amplitude variável. Exceto quando especificado, o tempo de polarização cruzada foi escolhido para fornecer a maior intensidade total, sendo de 1 ms para amostras de PE e PEEK e de 2 ms para iPB-I. Devido à limitação da máxima frequência de MAS da sonda disponível, utilizou-se a supressão total das bandas laterais rotacionais (TOSS) combinada com MAS de 4,5 kHz. A técnica de desacoplamento foi a dipolar de alta potência (70 kHz), com tempo de reciclagem de 2 a 5 s (cinco vezes o T_1^H mais longo determinado experimentalmente), de acordo com a amostra.

Processamento de dados por Resolução de Curvas Multivariadas (MCR)

Para a separação dos componentes espectrais com diferentes T_1 nos dados obtidos por RMN ^{13}C no estado sólido, foi utilizada uma técnica de processamento de dados chamada MCR, utilizando o software The Unscrambler® v10.4.1 (CAMO Software AS). A MCR é uma ferramenta matemática que permite determinar o número de componentes coexistentes em um sistema químico, permitindo a extração dos espectros puros dos componentes (análise qualitativa), bem como os perfis de concentração de cada componente (análise quantitativa), desde que tenham variabilidade diferente em função das fontes de variação presentes na matriz de dados. Isso é obtido aplicando uma PCA padrão para estimar o número provável de componentes na mistura e, em seguida, calculando uma rotação dos PCs sem as restrições de ortonormalidade, mas implementando novas restrições: concentrações não negativas e espectros não negativos.

Caracterização e quantificação dos constituintes do bagaço de cana submetidos a pré-tratamentos visando aumentar a eficiência enzimática para produção de etanol de segunda geração

A ideia geral da abordagem proposta é induzir uma variabilidade controlada nos espectros de RMN ^{13}C , utilizando uma sequência de pulsos específica, e aplicar a MCR para estimar os espectros dos diferentes componentes presentes nas amostras, de acordo com a propriedade do sistema associada à variabilidade induzida. Isso foi realizado modulando o espectro de RMN ^{13}C de acordo com os tempos de relaxação T_1 dos diferentes carbonos presentes na amostra. Para alcançar isso, foi utilizada a sequência de pulso Torchia- T_1 (ou CPT_1), ou seja, um período (tempo de contato) de polarização cruzada (CP) ^1H - ^{13}C seguido por um filtro z, definido experimentalmente, de duração τ e ciclagem de fase projetada para produzir uma modulação na magnetização do carbono. A aquisição do sinal é feita sob MAS e desacoplamento de alta potência dos ^1H . Essa sequência de pulsos modula as intensidades do espectro de RMN ^{13}C por um decaimento exponencial do tipo $e^{-\tau/T_1}$. Assim, para o mesmo τ , sinais com um tempo de relaxação T_1 dos ^{13}C mais curto são mais fortemente atenuados. Como, neste caso particular, os decaimentos das intensidades eram as variações relevantes, as normalizações usuais dos espectros (por exemplo, normalização pela área ou normalização pelo vetor unitário de cada espectro) não seriam úteis. Em vez disso, o ideal é utilizar os espectros com suas intensidades absolutas preservadas, ou utilizar o decaimento médio de cada amostra (decaimento da área total dos espectros em função de τ), ajustado por uma função monoexponencial e utilizar o I_0 (intensidade em $\tau = 0$) para normalizar cada espectro individualmente.

Nessa abordagem, em vez de utilizar um conjunto de amostras com variabilidade aleatória ou natural, essa variabilidade foi induzida em cada amostra individual, de forma controlada, utilizando uma sequência de pulsos de RMN especial. Em outras palavras, a sequência de pulsos Torchia- T_1 foi utilizada para modular as intensidades dos diversos sinais de ^{13}C (variabilidade induzida) por um fator que depende do tempo de relaxação longitudinal (T_1) dos núcleos desses ^{13}C . Assim, variando um parâmetro ajustável na sequência de pulsos Torchia- T_1 , é possível obter um conjunto de espectros de RMN ^{13}C com intensidades dependentes do T_1 de cada grupo químico, de cada amostra, individualmente. Com a sequência de pulsos Torchia- T_1 , uma variabilidade dependente de T_1 foi induzida nos espectros de RMN ^{13}C . Em seguida, empregou-se a MCR a esse conjunto de espectros para separar os diferentes componentes do espectro de acordo com seus tempos de relaxação T_1 dos ^{13}C .

Para o caso do bagaço de cana, as diferenças em T_1 podem estar diretamente associadas às diferenças na mobilidade molecular e no empacotamento da celulose, hemicelulose e lignina. Assim, a utilização desse método, em amostras de bagaço de cana-de-açúcar quimicamente pré-tratadas, permitiu decompor os espectros de RMN ^{13}C em dois subespectros, um correspondente à hemicelulose e outro à celulose, com traços de lignina (Figura 1).

O procedimento proposto também permite obter a intensidade dos sinais de cada componente em função de τ . Assim, é possível, por meio de um ajuste exponencial ordinário, estimar a concentração de cada componente em $\tau = 0$ (a intensidade total), bem como os correspondentes tempos de relaxação T_1 (Tabela 1) e a proporção de xilano em forte interação com a celulose, formando uma estrutura em fita dupla envolvendo as microfibrilas de celulose. O xilano é o principal constituinte da hemicelulose de biomassa vegetal.

Esses resultados, inéditos, são fundamentais para a avaliação dos diferentes pré-tratamentos, pois a eficiência dos tratamentos enzimáticos subsequentes, para a obtenção da glicose livre, que será

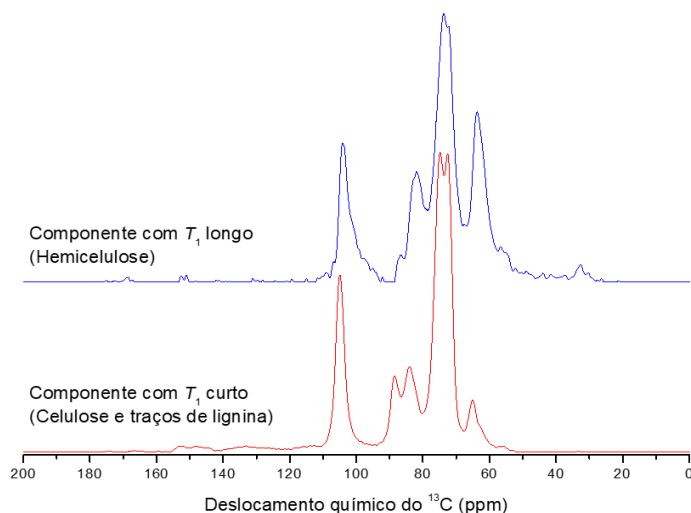


Figura 1. Espectros estimados por MCR após indução de variabilidade utilizando a sequência Torchia- T_1 . Com este procedimento, é possível, com apenas uma amostra, determinar a composição precisa e acurada da biomassa.

Tabela 1. Conteúdo dos componentes com T_1 curto (principalmente celulose) e T_1 longo (hemicelulose), seus correspondentes tempos de relaxação spin-rede (T_1) e o conteúdo de xilano na conformação de fita dupla, fortemente associado à celulose.

Tratamento	T_1 curto (%) – Celulose	T_1 longo (%) – Hemi- celulose	T_1 curto (s) – Celulose	Xilano em fita dupla (%)
H ₂ SO ₄ e NaOH	25 ± 1	75 ± 1	0,8 ± 0,1	7 ± 1
NaOH e H ₂ SO ₄	33 ± 1	67 ± 1	1,6 ± 0,2	16 ± 1

fermentada para a produção de etanol de segunda geração, depende fortemente não só do conteúdo de hemicelulose, que reduz a eficiência enzimática, mas também, e principalmente, da sua conformação e do grau de associação com a celulose. Além do mais, o T_1 mais curto

da celulose no tratamento ácido primeiro (H_2SO_4 e NaOH), indica que ela tem uma dinâmica molecular mais rápida, ou seja, está numa forma menos cristalina, o que também favorece sua hidrólise enzimática. Esses pressupostos foram confirmados experimentalmente, com maior eficiência enzimática após o pré-tratamento ácido primeiro (Espírito Santo et al., 2022).

É interessante mencionar que o tratamento ácido primeiro forneceu menos celulose (25%), porém ela tinha uma dinâmica mais rápida, sendo mais amorfa (T_1 da celulose mais curto, 0,8 s) e com menos xilano envolvendo suas microfibrilas (7%), resultando em uma maior eficiência enzimática (Espírito Santo et al., 2022).

Esses resultados confirmam as potencialidades do método proposto para a caracterização de biomassas e pré-tratamentos para otimizar a obtenção de biocombustíveis. Além do mais, esse método é aplicável a outras amostras complexas, como a matéria orgânica do solo.

Otimização do método proposto utilizando polímeros sintéticos

Visando demonstrar sua generalidade e estabelecer suas limitações, bem como otimizar o número mínimo de experimentos (espectros) necessários para a obtenção dos resultados, foram utilizados três polímeros sintéticos mais simples (Figuras 2 a 4).

Como primeira prova do princípio, a abordagem proposta foi aplicada ao polímero PE, cujo espectro de RMN ^{13}C é composto por apenas duas linhas espectrais, a ~ 33 e ~ 31 ppm (Figura 2a), correspondentes aos sinais nos domínios cristalinos e amorfos, respectivamente. Devido às conhecidas diferenças no empacotamento molecular e na mobilidade nos domínios cristalinos e amorfos, espera-se que o valor de T_1 do ^{13}C dos segmentos nos domínios amorfos seja muito mais curto do que nas regiões cristalinas. Isso é claramente observado na Figura 2a, com o sinal da componente amorfa (~ 31 ppm) decaindo muito mais rapidamente que sua contraparte cristalina (~ 33 ppm).

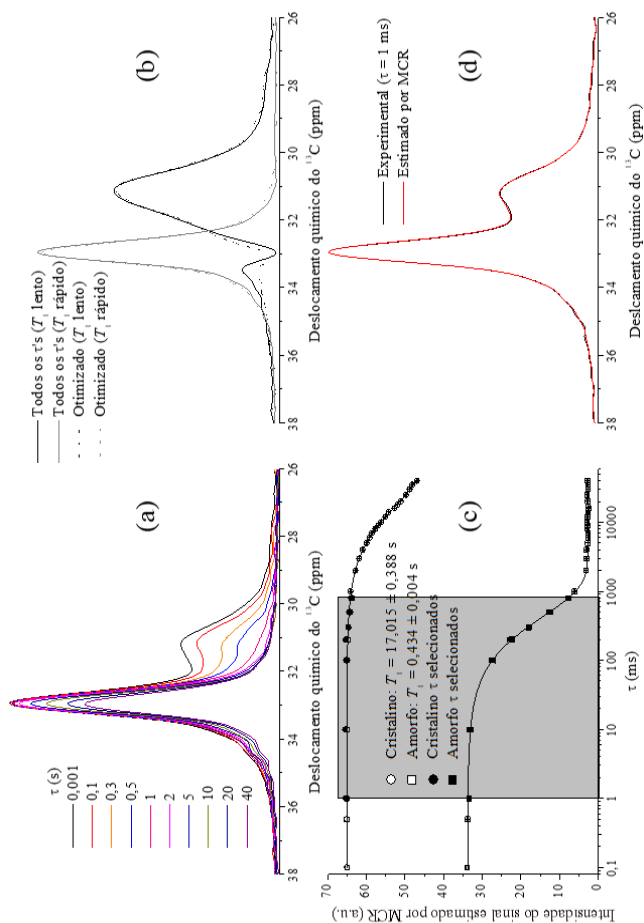


Figura 2. (a) Espectros ^{13}C Torchia- T_1 de polietileno semicristalino de média densidade com diferentes períodos de relaxação τ . (b) Espectros estimados por MCR utilizando o conjunto completo de dados (28 espectros, i.e., 28 valores de τ , tempo experimental total de 7 h) e utilizando o conjunto otimizado, i.e., apenas 7 espectros com valores de τ de 1; 10; 100; 200; 300; 500 e 800 ms (tempo experimental total de 30 min), mostrados pelos símbolos cheios em (c), marcados com o retângulo cinza, que são as intensidades estimadas por MCR como uma função do indutor da variabilidade, o período de relaxação τ juntamente com o ajuste exponencial e respectivos valores de T_1 . (d) Comparação entre o espectro experimental e aquele estimado por MCR utilizando apenas o conjunto otimizado, obtido pela soma dos espectros estimados individuais, multiplicados pela correspondente concentração de cada domínio.

A análise MCR desse conjunto de dados (28 espectros com diferentes tempos τ na sequência Torchia- T_1) confirmou a coexistência de dois componentes com diferentes T_1 , cujos espectros isolados foram adequadamente estimados (Figura 2b). Ademais, as intensidades relativas de cada um desses componentes em função de τ também são estimadas e podem ser utilizadas para o ajuste do fator de modulação de Torchia: $c_i(0)e^{-\tau/T_{1i}}$, obtendo-se, assim, os T_1 de cada domínio (Figura 2c), bem como suas intensidades iniciais, que são proporcionais à concentração de cada domínio, resultando em $66 \pm 2\%$ da fração amorfa e $33 \pm 2\%$ da fração cristalina. Entretanto, devido às limitações técnicas da sonda utilizada (baixa taxa de MAS), foi necessário utilizar TOSS, que torna os espectros de RMN ^{13}C semiquantitativos. Mas como o objetivo deste manuscrito é demonstrar a aplicabilidade do método, com a clara separação dos domínios e a obtenção de importantes parâmetros físico-químicos e da dinâmica molecular, a quantificação absoluta não se faz necessária nesta demonstração do princípio. Além disso, com este polímero simples e com os sinais dos diferentes domínios bem resolvidos, os mesmos resultados poderiam ser obtidos por deconvolução ordinária dos picos. No entanto, reitera-se que o objetivo era apresentar uma prova do princípio e, posteriormente, demonstrar um exemplo em que, mesmo com total sobreposição dos sinais, é possível obter os espectros de cada componente, suas concentrações e seus tempos de relaxação, algo impossível por deconvolução.

O ponto chave do procedimento MCR proposto é a variabilidade (mais especificamente, variância-covariância). Assim, formulou-se a hipótese de que o número de experimentos (diferentes valores de τ) poderia ser reduzido sem perda de informação, utilizando-se apenas os τ que forneçam a máxima variabilidade. Visto que, no fenômeno físico utilizado para induzir essa variabilidade (o decaimento exponencial devido à T_1), a maior variação sistemática da intensidade do sinal ocorre no início do decaimento, provavelmente essa seja a região (τ curtos) mais útil para ser amostrada. Além disso, trata-se da região com maior relação sinal/ruído e com menor tempo experimental, pois na sequência Torchia- T_1 , o grande gargalo é utilizar τ longos.

Para otimizar e validar essa seleção, visando otimizar o tempo experimental, reduzindo inclusive o custo das análises, utilizou-se o procedimento Jack-knife. Para isso, espectros adquiridos com certos valores de τ foram excluídos da matriz de dados, e os ajustes MCR foram comparados estatisticamente com as médias e desvios padrão (teste t de Student) dos resultados obtidos por validação cruzada aleatória, com o mesmo número de τ dos experimentos otimizados, procedimento similar ao teste de incerteza de Martens e Martens (2000).

Os melhores resultados foram obtidos utilizando apenas τ no intervalo $T_1/50$ a $2 T_1$, considerando o T_1 do componente com relaxação mais rápida. Esses valores se justificam, pois, a razão (variabilidade) entre a magnetização a $\tau = T_1/50$ e a máxima magnetização teórica (M_0) a $\tau = 0$ é somente 2% ($e^{-T_1/50/T_1}/e^{-0/T_1} = e^{-0,02}/e^{-0} = 0,98$), e a $2 T_1$ a intensidade do sinal é menor que 14% de M_0 (e^{-2T_1/T_1}). Os resultados da MCR utilizando o conjunto de dados otimizados (apenas 7 valores de τ : 1; 10; 100; 200; 300; 500 e 800 ms) são apresentados nas Figuras 2B a D. O tempo experimental otimizado é de 30 min por amostra, enquanto, para adquirir o conjunto completo (28 valores de τ), são necessárias 7 h por amostra.

A abordagem também foi aplicada para estimar os espectros de RMN ^{13}C individuais das componentes amorfas e cristalinas de um polímero quimicamente mais complexo, o PEEK, em que foi possível reproduzir os espectros experimentais a partir da combinação dos dois espectros estimados com muito alta precisão, mesmo com a superposição completa dos sinais das fases amorfas e cristalinas (Figura 3). Nessa amostra de PEEK, $71 \pm 3\%$ é amorfa e $29 \pm 3\%$ cristalina.

Da mesma forma que para o PE, o melhor conjunto de τ selecionados situa-se no intervalo $T_1/50$ a $2 T_1$, indicando que esse intervalo ideal não depende da complexidade da amostra. Para o caso do PEEK o tempo experimental total caiu de 25 h para 6,5 h por amostra.

Um terceiro exemplo foi obtido aplicando-se a abordagem ao Poli(-1-buteneno) isotático; porém, para este polímero, a separação por T_1 ocorreu principalmente pela diferença de mobilidade entre o grupo lateral (mais móvel) e a cadeia principal olefínica (mais rígida) do polímero (Figura 4a e b). Isso ocorreu porque a maior variabilidade foi entre os diferentes segmentos das unidades poliméricas, e não em função de uma eventual ocorrência de domínios com diferentes graus de cristalinidade, o que é pouco provável considerando o processo de preparo da amostra, que visou obter apenas a forma cristalina I (vide seção experimental).

Isso mostra outra utilidade do método proposto: distinguir sinais oriundos de segmentos moleculares com diferentes mobilidades locais, o que é muito útil não somente para o estudo de polímeros e

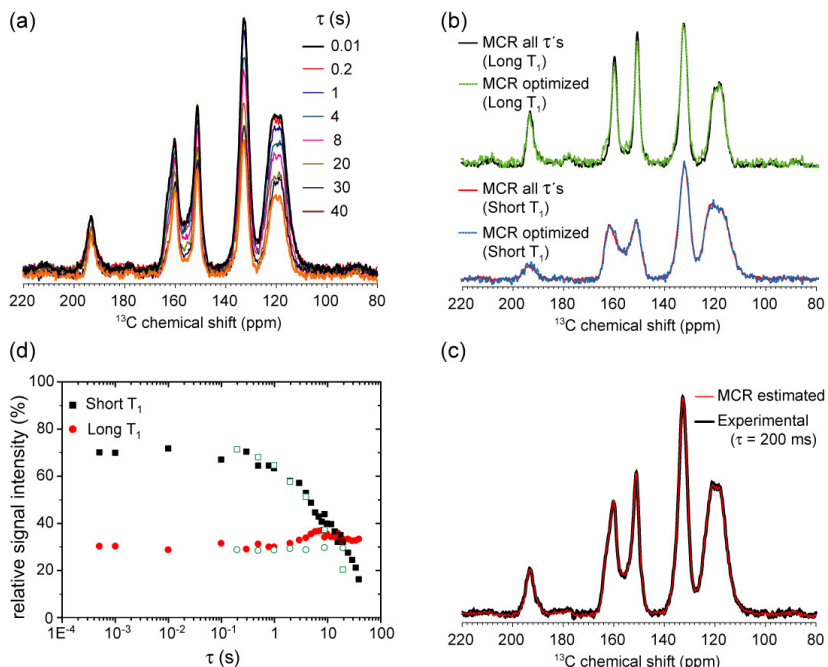


Figura 3. Espectros ^{13}C Torchia- T_1 de Poli(éter-éter-cetona) (PEEK) com diferentes períodos de relaxação τ . (b) Espectros estimados por MCR utilizando o conjunto completo de dados (28 espectros, i.e., 28 valores de τ , tempo experimental total de 25 h) e utilizando o conjunto otimizado, i.e., apenas 7 espectros com valores de τ de 0,2; 0,55; 1; 2; 4; 9 e 20 s (tempo experimental total de 6,5 h), mostrados pelos símbolos abertos em verde em (c) que são as intensidades estimadas por MCR como uma função do indutor da variabilidade, o período de relaxação τ juntamente com o ajuste exponencial e respectivos valores de T_1 . (d) Comparação entre o espectro experimental e aquele estimado por MCR utilizando apenas o conjunto otimizado, obtido pela soma dos espectros estimados individuais, multiplicados pela correspondente concentração de cada domínio.

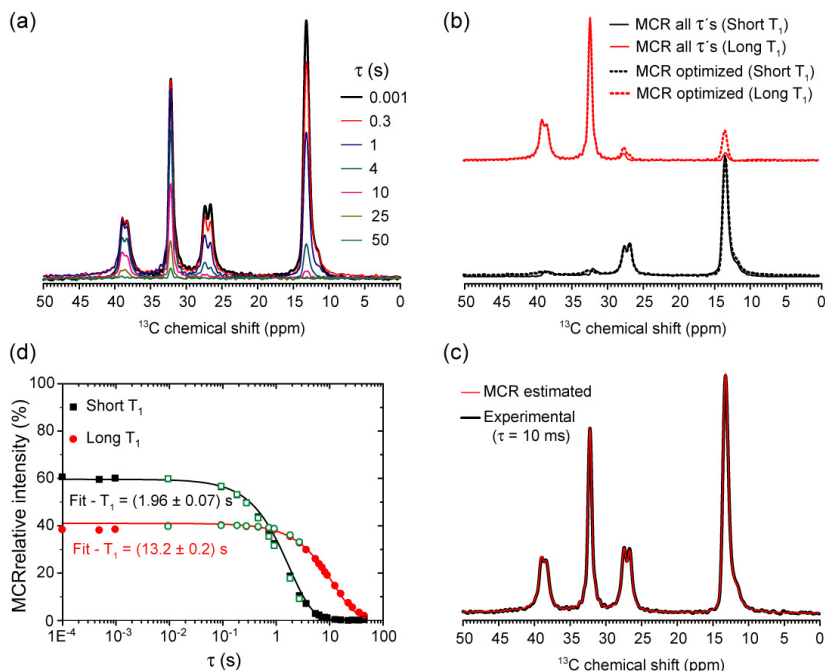


Figura 4. (a) Espectros ^{13}C Torchia- T_1 de polibuteno isotático polimérico semi-cristalino (iPB-I) com diferentes períodos de relaxação τ . (b) Espectros estimados por MCR utilizando o conjunto completo de dados (24 espectros, i.e., 24 valores de τ , tempo experimental total de 11 h) e utilizando o conjunto otimizado, i.e., apenas 9 espectros com valores de τ de 10; 100; 200; 300; 500; 800; 1000; 2000 e 3000 ms (tempo experimental total de 1,5 h), mostrados pelos símbolos abertos em verde em (c) que são as intensidades estimadas por MCR como uma função do indutor da variabilidade, o período de relaxação τ juntamente com o ajuste exponencial e respectivos valores de T_1 . (d) Comparação entre o espectro experimental e aquele estimado por MCR utilizando apenas o conjunto otimizado, obtido pela soma dos espectros estimados individuais, multiplicados pela correspondente concentração de cada domínio.

biomassas, mas também para analisar a matéria orgânica do solo, por exemplo, estimar recalcitrância e afinidade por agrotóxicos.

Como anteriormente, o conjunto otimizado está no intervalo $T_1/50$ a $2 T_1$, reduzindo o tempo total da análise de 11 h para apenas 1,5 h. Apesar da excelente concordância entre o espectro experimental e o

estimado pela combinação dos dois segmentos da unidade polimérica, o segmento molecular com restrição dinâmica (T_1 mais longo) é claramente subestimado (Figura 4c), visto que, estequiometricamente, as proporções deveriam ser idênticas. Isso pode ser devido ao pequeno contraste de T_1 entre os dois segmentos (2 vs. 13 s, menos de uma ordem de magnitude), limitando a correta separação dos segmentos, o que pode ser observado também nos espectros estimados, pois há sinais residuais do segmento móvel (13 e 27 ppm) na componente rígida e, vice-versa, sinais do segmento rígido (32 e 38 ppm) na componente móvel (Figura 4b), indicando uma limitação do método que necessita de mais investigação. Há outras sequências de pulsos que induzem variabilidade em outras escalas de tempo, como T_2 , $T_{1\rho}$ etc. ou mesmo a combinação delas. De qualquer forma, a metodologia está apresentada, o conceito foi provado e basta induzir a variabilidade que forneça a informação desejada e tratar os dados como proposto.

Conclusão

As análises de espectroscopia de RMN no estado sólido de ^{13}C utilizando a sequência de pulsos Torchia- T_1 , seguidas da análise multivariada de dados, mostraram-se eficientes na caracterização de amostras de bagaço de cana. Os dados obtidos permitiram acessar informações inéditas e extremamente relevantes relacionadas à mobilidade molecular e ao empacotamento da celulose, hemicelulose e lignina, diretamente relacionadas com a eficiência dos tratamentos enzimáticos posteriores. A generalidade do método foi confirmada por meio da análise de três polímeros sintéticos, em que também se otimizou o tempo de análise, reduzindo-a de 3,8 até 14 vezes, o que representa uma enorme diminuição nos custos das análises. Portanto, o método proposto é de uso geral, podendo ser empregado para a caracterização de polímeros sintéticos e naturais (biomassa), na avaliação de pré-tratamentos visando otimizar a geração de biocombustíveis de segunda e terceira geração, no estudo da degradação de polímeros, assim como no estudo da matéria orgânica do solo.

Agradecimentos

Os autores agradecem o apoio financeiro do CNPq (processos: 442377/2023-1 e 308760/2022-0); Finep/CT-Agro/FNDCT (Convênio 01.22.0081.00, Ref. 1218/21) e Mapa (TED 396/2020). Etelvino Henrique Novotny agradece ao CNPq a bolsa de produtividade em pesquisa (307694/2023-2) e à Faperj pela bolsa Cientista do Nosso Estado (E-26/200.470/2023).

Referências

ALESSI, A. M.; BIRD, S. M.; OATES, N. C.; LI, Y.; DOWLE, A. A.; NOVOTNY, E. H.; AZEVEDO, E. R. de; BENNETT, J. P.; POLIKARPOV, I.; YOUNG, J. P. W.; MCQUEEN-MASON, S. J.; BRUCE, N. C. Defining functional diversity for lignocellulose degradation in a microbial community using multi-omics studies. **Biotechnology for Biofuels**, v. 11, 166, Jun. 2018. DOI: <https://doi.org/10.1186/s13068-018-1164-2>.

AZEVEDO, E. R. de; BONAGAMBA, T. J.; SCHMIDT-ROHR, K. Pure-exchange solid-state NMR. *Journal of Magnetic Resonance*, v. 142, n. 1, p. 86-96, Jan. 2000. DOI: <https://doi.org/10.1006/jmre.1999.1918>.

BERNARDINELLI, O. D.; NOVOTNY, E. H.; AZEVEDO, E. R. de; COLNAGO, L. A. Analyses of biomass products by nuclear magnetic resonance spectroscopy. In: VAZ JUNIOR, S. (ed.). **Analytical techniques and methods for biomass**. Cham: Springer, 2016. cap. 6, p. 143-172. DOI: https://doi.org/10.1007/978-3-319-41414-0_6.

BESSER, K.; MALYON, G. P.; EBORALL, W. S.; CUNHA, G. P. da; FILGUEIRAS, J. G.; DOWLE, A.; GARCIA, L. C.; PAGE, S. J.; DUPREE, R.; KERN, M.; GOMEZ, L. D.; LI, Y.; ELIAS, L.; SABBADIN, F.; MOHAMAD, S. E.; PESANTE, G.; STEELE-KING, C.; AZEVEDO, E. R. de; POLIKARPOV, I.; DUPREE, P.; CRAGG, S. M.; BRUCE, N. C.; MCQUEEN-MASON, S. J. Hemocyanin facilitates lignocellulose digestion by wood-boring marine

crustaceans. **Nature Communications**, v. 9, 5125, Dec. 2018. DOI: <https://doi.org/10.1038/s41467-018-07575-2>.

DING, S.-Y.; LIU, Y.-S.; ZENG, Y.; HIMMEL, M. E.; BAKER, J. O.; BAYER, E. A. How does plant cell wall nanoscale architecture correlate with enzymatic digestibility? **Science**, v. 338, n. 6110, p. 1055-1060, Nov. 2012. DOI: <https://doi.org/10.1126/science.1227491>.

ESPÍRITO SANTO, M. C. do; CARDOSO, E. B.; GUIMARAES, F. E. G.; AZEVEDO, E. R. de; CUNHA, G. P. da; NOVOTNY, E. H.; PELLEGRINI, V. de O. A.; CHANDEL, A. K.; SILVEIRA, M. H. L.; POLIKARPOV, I. Multifaceted characterization of sugarcane bagasse under different steam explosion severity conditions leading to distinct enzymatic hydrolysis yields. **Industrial Crops and Products**, v. 139, 111542, Nov. 2019. DOI: <https://doi.org/10.1016/j.indcrop.2019.111542>.

ESPÍRITO SANTO, M. C. do; THEMA, F. T.; PELLEGRINI, V. de O. A.; KANE, A. O.; GUIMARAES, F. E. G.; FILGUEIRAS, J. G.; NOVOTNY, E. H.; AZEVEDO, E. R. de; POLIKARPOV, I. When the order matters: impacts of lignin removal and xylan conformation on the physical structure and enzymatic hydrolysis of sugarcane bagasse. **Industrial Crops and Products**, v. 180, 114708, Jun. 2022. DOI: <https://doi.org/10.1016/j.indcrop.2022.114708>.

LIMA, M. A.; GOMEZ, L. D.; STEELE-KING, C. G.; SIMISTER, R.; BERNARDINELLI, O. D.; CARVALHO, M. A.; REZENDE, C. A.; LABATE, C. A.; AZEVEDO, E. R. de; MCQUEEN-MASON, S. J.; POLIKARPOV, I. Evaluating the composition and processing potential of novel sources of Brazilian biomass for sustainable biorenewables production. **Biotechnology for Biofuels**, v. 7, 10, Jan. 2014. DOI: <https://doi.org/10.1186/1754-6834-7-10>.

MARTENS, H.; MARTENS, M. Modified Jack-knife estimation of parameter uncertainty in bilinear modelling by partial least squares regression (PLSR). **Food Quality and Preference**, v. 11, n. 1/2, p. 5-6, Jan. 2000. DOI: [https://doi.org/10.1016/S0950-3293\(99\)00039-7](https://doi.org/10.1016/S0950-3293(99)00039-7).

MIELENZ, J. R. (ed.). **Biofuels: methods and protocols**. New York: Humana Press, 2009. DOI: <https://doi.org/10.1007/978-1-60761-214-8>.

NOVOTNY, E. H.; GARCIA, R. H. S.; AZEVEDO, E. R. de. Pulse sequence induced variability combined with multivariate analysis as a potential tool

for ^{13}C solid-state NMR signals separation, quantification, and classification.

Journal of Magnetic Resonance Open, v. 14/15, 100089, Jun. 2023. DOI: <https://doi.org/10.1016/j.jmro.2022.100089>.

NOVOTNY, E. H.; AUCCAISE, R.; VELLOSO, M. H. R.; CORRÊA, J. C.; HIGARASHI, M. M.; ABREU, V. M. N.; ROCHA, J. D.; KWAPINSKI, W.

Characterization of phosphate structures in biochar from swine bones.

Pesquisa Agropecuária Brasileira, v. 47, n. 5, p. 672-676, maio 2012. DOI: <https://doi.org/10.1590/S0100-204X2012000500006>.

PAUL, E. A. **Soil microbiology, ecology, and biochemistry**. 3rd ed. Amsterdam: Academic Press, 2007. DOI: <https://doi.org/10.1016/C2009-0-02816-5>.

REZENDE, C. A.; LIMA, M. A. de; MAZIERO, P.; AZEVEDO, E. R. de; GARCIA, W.; POLIKARPOV, I. Chemical and morphological characterization of sugarcane bagasse submitted to a delignification process for enhanced enzymatic digestibility. **Biotechnology for Biofuels**, v. 4, 54, Nov. 2011. DOI: <https://doi.org/10.1186/1754-6834-4-54>.

ROCHA, D. A. M.; TORRES, J. P. M.; REICHEL, K.; NOVOTNY, E. H.; ESTRELLA, L. F.; MEDEIROS, R. O.; PEREIRA NETTO, A. D. Determination of polychlorinated dibenzo-p-dioxins and dibenzofurans (PCDD/Fs) in Brazilian cow milk. **Science of the Total Environment**, v. 572, p. 177-184, Dec. 2016. DOI: <https://doi.org/10.1016/j.scitotenv.2016.07.179>.

WALL, J. D.; HARWOOD, C. S.; DEMAINE, A. L. (ed.). **Bioenergy**. Washington, DC: ASM Press, 2008. DOI: <https://doi.org/10.1128/9781555815547>.

TORCHIA, D. A. The measurement of proton-enhanced carbon-13 T_1 values by a method which suppresses artifacts. **Journal of Magnetic Resonance**, v. 30, n. 3, p. 613-616, Jun. 1978. DOI: [https://doi.org/10.1016/0022-2364\(78\)90288-3](https://doi.org/10.1016/0022-2364(78)90288-3).

