

Espectroscopia NIR na Avaliação da Composição da Biomassa de Capim-Elefante

Maria Lúcia Ferreira Simeone¹, Patrícia Abrão de Oliveira Molinari², Raquel Bombarda Campanha², Juarez Campolina Machado³

¹Embrapa Milho e Sorgo, Sete Lagoas, MG, Brasil, marialucia.simeone@embrapa.br

²Embrapa Agroenergia, Brasília, DF, Brasil

³Embrapa Gado de Leite, Juiz de Fora, MG, Brasil

Resumo: Este estudo propõe utilizar a espectroscopia no infravermelho próximo (NIR) associada à regressão de mínimos quadrados parciais (PLS) como uma alternativa rápida, eficaz e não destrutiva para análise da composição da biomassa de capim-elefante. Modelos de calibração NIR-PLS foram desenvolvidos para avaliar os teores de glucana, xilana, arabinana, extrativos, lignina e cinzas e se mostraram eficientes e rápidos, facilitando a seleção de genótipos de capim-elefante para produção de biocombustíveis.

Palavras-chave: Conversão de biomassa; Composição química, Biocombustíveis; Análise multivariada.

INTRODUÇÃO

O capim-elefante (*Pennisetum purpureum* Schum.) se destaca como uma forrageira robusta e altamente produtiva, adaptável a diversos climas e solos, tornando-o crucial para a pecuária em regiões tropicais e subtropicais. Além de seu valor nutricional para a pecuária, pesquisas recentes apontam para o promissor potencial energético do capim-elefante, abrindo novas perspectivas para sua utilização na produção de biocombustíveis e bioenergia (Rocha et al., 2019).

Carboidratos estruturais, como celulose e hemicelulose, constituem a maior parte da matéria seca do capim-elefante e podem ser decompostos em açúcares fermentáveis, servindo de matéria-prima para a produção de biocombustíveis, como o biogás e o bioetanol de segunda geração. A lignina, por sua vez, é um polímero complexo que confere rigidez à planta, mas também pode ser utilizada na produção de biocombustíveis e bioprodutos.

A caracterização detalhada dos componentes da biomassa do capim-elefante é essencial para otimizar seu uso na produção de bioenergia. A espectroscopia NIR associada a métodos quimiométricos oferece uma metodologia de forma rápida, não destrutiva e de baixo custo para a análise de diversos componentes da biomassa, permitindo a obtenção de informações importantes para a caracterização do potencial energético do capim-elefante (Abreu et al., 2022).

Ao analisar os diferentes constituintes da biomassa vegetal em conjunto, utilizando a espectroscopia NIR, é possível obter de forma eficiente e rápida uma compreensão abrangente do potencial energético do capim-elefante. Essa informação é fundamental para identificar oportunidades para otimizar seu uso na produção de

bioenergia, contribuindo para um futuro mais verde e sustentável.

O objetivo desse trabalho foi desenvolver uma metodologia analítica robusta e eficiente utilizando dados de espectroscopia NIR e calibração multivariada para avaliar os teores de glucana, xilana, arabinana, extrativos em água, extrativos em etanol, lignina solúvel, lignina insolúvel e cinzas na biomassa de capim-elefante, facilitando a caracterização do seu potencial energético.

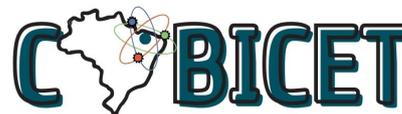
MATERIAL E MÉTODOS

2.1 Preparo das amostras

As amostras de 106 genótipos de capim-elefante foram obtidas de experimentos do Programa de Melhoramento Genético de Capim-Elefante da Embrapa Gado de Leite, Juiz de Fora, MG. As amostras recém-colhidas foram colocadas em sacos de papel e submetidas a um processo de secagem por 72 horas, a uma temperatura de 55 °C. Posteriormente, foram moídas em moinho de facas (Thomas Scientific, Swedesboro, NJ, EUA) na granulometria de 1 mm, em preparação para a análise química laboratorial.

2.2 Análises químicas

Os constituintes químicos da biomassa foram avaliados nos laboratórios da Embrapa Agroenergia (Brasília, DF) e seguindo procedimentos analíticos laboratoriais (LAPs) desenvolvidos pelo NREL (Sluiter et al., 2010). Os constituintes químicos analisados foram glucana, xilana, lignina solúvel, lignina insolúvel, extrativos em água, extrativos em etanol e cinzas. Todas as análises foram realizadas em triplicata, e os valores médios estão apresentados



na seção de resultados e discussão. 2.3 Obtenção dos espectros NIR

Para a coleta dos espectros NIR, 15 g de amostra de capim-elefante moída foram colocados em placas de Petri de borossilicato (9 cm de diâmetro) para obtenção dos espectros, em triplicata. Os espectros foram obtidos com o equipamento NIRFlex 500 (Buchi Labortechnik, Flawil, Suíça) na região de 1.000 a 2.500 nm (4.000 a 10.000 cm^{-1}) com uma resolução de 32 scans por espectro.

2.4 Desenvolvimento e validação dos modelos NIR

Foram utilizadas 106 amostras de capim-elefante, sendo que o conjunto de dados foi dividido em conjuntos de calibração e validação para cada componente químico avaliado (ou seja, glucana, xilana, lignina insolúvel, lignina solúvel, extrativos aquosos e etanólico, cinzas). Primeiro, o conjunto de dados foi classificado em ordem crescente a fim de garantir uma cobertura abrangente do componente-alvo, seguido de agrupamentos de três dados consecutivos. Em seguida, a primeira e a terceira amostras de cada grupo sucessivo foram alocadas ao conjunto de calibração, e a segunda amostra foi para o conjunto de validação. O desempenho do modelo foi avaliado no conjunto de dados de validação. Utilizaram-se 71 amostras no conjunto de calibração e 35 amostras de capim-elefante que não foram utilizadas na construção do modelo para prever cada componente químico avaliado.

Os dados espectrais foram pré-processados utilizando os métodos transformação Padrão Normal de Variação (SNV), Derivada Primeira (DP) com filtro Savitzky-Golay ou Derivada Segunda (DS), variando-se o tamanho das janelas de 5, 9 e 11 pontos. Essas etapas de pré-processamento foram aplicadas para remover variações como desvios de linha de base e espalhamento de luz, melhorando assim a qualidade da informação espectral. Os dados espectrais pré-processados foram centralizados na média antes da construção dos modelos pelo algoritmo de mínimos quadrados parciais (PLS – Partial Least Squares) com todas as variáveis espectrais e com a seleção de variáveis baseada nos coeficientes de regressão significativos, Martens & Westad, (2000). Os modelos obtidos para todos os constituintes químicos analisados foram validados internamente (validação cruzada) e com as amostras que não foram utilizadas na construção do modelo de calibração.

A detecção de amostras anômalas foi realizada utilizando os gráficos dos resíduos versus o *leverage*, na matriz X (espectros) e em y

(propriedade de interesse). As amostras que apresentaram resíduos e *leverage* altos foram excluídas do conjunto de dados. Em seguida, novos modelos foram construídos, e os parâmetros, como coeficiente de correlação (R^2) e raiz quadrada do erro médio de validação cruzada, RMSECV (do inglês, *Root Mean Squared Error of Cross Validation*), foram reavaliados. O número ideal de variáveis latentes de cada modelo PLS foi obtido a partir da análise do gráfico da variância explicada versus RMSECV, gráfico dos *loadings* e do resíduo dos modelos. O software Unscrambler® 10.5 (Camo Software, Oslo, Norway) foi empregado no pré-processamento dos dados e na obtenção dos modelos PLS.

A capacidade preditiva dos modelos PLS foi avaliada pelos erros de previsão RMSEP (do inglês, *Root Mean Squared Error of Prediction*), para o conjunto de validação e pelos seguintes parâmetros estatísticos: raiz quadrada do erro médio de calibração, RMSEC (do inglês, *Root Mean Squared Error of Calibration*); raiz quadrada do erro médio de calibração cruzada, RMSECV (do inglês, *Root Mean Squared Error of Cross Validation*); R^2 para o conjunto de calibração e validação, PRL (do inglês, *Prediction Error Relative to Laboratory*); razão de intervalo de erro, RER (do inglês, *Range Error Ratio*) e relação de desempenho do desvio, RPD (do inglês, *Residual Prediction Deviation*) (Yang et al., 2017).

O teste *t* pareado foi utilizado para verificar se havia diferença significativa entre os valores de referência e os valores previstos dos constituintes nos conjuntos de validação dos modelos otimizados. Neste teste, calculou-se a probabilidade (*p*-valor) de obter uma diferença de erro absoluto maior que a observada experimentalmente. Um *p*-valor menor que o nível de significância ($\alpha = 0,05$) indica uma diferença significativa nas habilidades preditivas do modelo.

RESULTADOS E DISCUSSÃO

A composição química das amostras de capim-elefante foi obtida pelos métodos químicos de via úmida conforme (Sluiter et al., 2010) e está apresentada na Tabela 1. Os resultados representam os valores médios, valores mínimos (Min), máximo (Max) e desvio padrão (DP), no conjunto de dados de calibração ($n = 71$) e no conjunto de dados de validação ($n = 35$). Os teores de carboidratos estruturais (glucana, xilana, arabinana), lignina (insolúvel e solúvel), extrativos (aquoso e etanólico) e cinzas das amostras de capim-elefante encontravam-se na faixa de 30,42 a 39,14; 12,78 a

19,54; 1,40 a 3,15; 3,49 a 20,91; 3,16 a 19,21; 7,56 a 18,51; 1,80 a 3,07; 3,37 a 8,38%, respectivamente, e mostraram propriedades semelhantes para os conjuntos de dados de calibração e validação de todos os componentes químicos estudados (Tabela 1). Tais variações de composição podem ser devidas à diversidade genética, ao estágio de crescimento, à época de colheita, aos fatores ambientais e ao processamento pós-colheita. Os teores de glucana, xilana, arabinana, lignina solúvel e insolúvel, cinzas, extratos aquosos e etanólico foram muito semelhantes aos valores relatados por outros autores para biomassa herbácea (Ahmed et al., 2024).

Tabela 1. Características dos conjuntos de calibração e validação para os constituintes químicos avaliados na biomassa de capim-elefante.

Constituintes (%)	Média	Max ^a	Min ^b	DP ^c
Calibração (N^d = 71)				
Glucana	34,53	39,14	30,42	1,72
Xilana	15,94	19,54	12,78	1,87
Arabinana	2,31	3,15	1,40	0,49
Lig.insolúvel	11,48	19,21	3,16	5,91
Lig. solúvel	9,54	20,91	3,49	6,59
Ext. aquoso	12,17	18,51	7,56	2,19
Ext. EtOH	2,39	3,07	1,80	0,27
Cinzas	6,02	8,38	3,37	1,08
Validação (N=35)				
Glucana	34,52	38,40	31,20	1,65
Xilana	15,96	19,35	12,83	1,93
Arabinana	2,32	3,10	1,43	0,49
Lig.insolúvel	11,48	18,65	3,44	5,96
Lig. solúvel	9,42	20,18	3,54	6,59
Ext. aquoso	12,06	18,24	8,55	2,12
Ext. EtOH	2,39	3,26	1,76	0,31
Cinzas	6,15	8,40	3,74	1,05

a-valor máximo; b-valor mínimo; c-DP: desvio padrão; d-N: número de amostras.

Lig.: lignina; Ext.: extrativo; Ext. EtOH: extrativo etanólico.

A Figura 1 mostra os espectros médios NIR das 106 amostras de capim-elefante registrados na faixa de 10.000 – 4.000 cm⁻¹. Na região de 1.000 a 2.500 nm, observam-se sobretons e bandas de combinações de vibrações moleculares, como estiramento e deformações das ligações O–H, C–H e N–H. Essas bandas surgem por causa da interação da luz com os modos vibracionais fundamentais das

moléculas. Ao se medir os tons harmônicos e as combinações dessas vibrações, a espectroscopia NIR fornece informações sobre a composição química e a estrutura molecular de várias substâncias. Na Figura 1, podemos observar a primeira banda característica entre 1.409 e 1.530 nm, associada à celulose (1.490 nm), ao estiramento O–H no primeiro sobretom da H₂O (1.450 nm) e à combinação C–H de estruturas aromáticas da lignina (1.446 nm). O segundo pico está relacionado à deformação da ligação O–H e ao estiramento/HOH de uma combinação de celulose a 1.930 nm, e ao segundo harmônico da curva O–H em 1.940 nm (Ahmed et al., 2024). A terceira banda corresponde ao estiramento assimétrico C–O–O (carboxilato) no terceiro sobretom da celulose (2.100 nm). O quarto pico pode ser atribuído à combinação do estiramento C–H com o estiramento C–C em 2.495 nm.

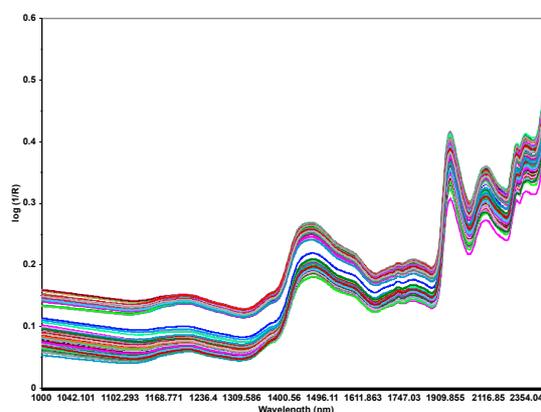
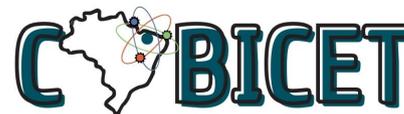


Figura 1. Espectros médios NIR das 106 amostras de capim-elefante. Os dados de reflectância difusa foram convertidos a log(1/R), absorvância.

A biomassa é uma mistura complexa de compostos orgânicos, como celulose, hemicelulose, lignina e vários açúcares. Cada um desses componentes contribui para a absorção da luz na região do NIR, e suas bandas de absorção frequentemente se sobrepõem na faixa espectral NIR (normalmente 850 – 2.500 nm). Observa-se nos espectros uma grande variação sistemática da linha de base ao longo de toda a faixa de trabalho, causada pela não uniformidade das amostras. Isso se deve principalmente às diferenças de granulometria, ao empacotamento e à orientação das partículas de capim-elefante. Portanto, verificou-se a necessidade de usar técnicas de correção antes de se obter os modelos de calibração, a fim de minimizar o efeito do espalhamento de luz presente nos espectros obtidos por reflectância. Os pré-processamentos testados foram: SNV, SNV + primeira derivada com filtro Savitzky-Golay janela de 5, 9 e 11 pontos, e segunda derivada com 11 pontos. O pré-processamento que proporcionou os melhores



resultados foi a segunda derivada com janela de 11 pontos, seguido da seleção de variáveis baseada na importância dos coeficientes de regressão.

A escolha do número de variáveis latentes foi feita por meio da resposta de maior variância explicada e o menor valor de RMSECV. Os resultados obtidos estão apresentados na Tabela 2.

Tabela 2. Resultados das análises estatísticas dos conjuntos de calibração e validação para a biomassa de capim-elefante.

Constituinte	Calibração				
	N ^a	Out ^b	VL ^c	R ^{2d}	RMSE ^e
Glucana	71	-	8	0,89	0,57
Xilana	64	7	3	0,77	0,89
Arabinana	69	2	3	0,90	0,15
Lig. insolúvel	71	-	4	0,97	1,06
Lig. solúvel	70	1	4	0,97	1,21
Ext. aquoso	66	5	5	0,91	0,62
Ext. EtOH	56	15	4	0,87	0,09
Cinzas	69	2	7	0,92	0,30
Validação					
Glucana	32	3	8	0,84	0,64
Xilana	26	9	3	0,75	0,95
Arabinana	34	1	3	0,91	0,14
Lig. insolúvel	34	1	4	0,92	1,61
Lig. solúvel	35	-	4	0,92	1,84
Ext. aquoso	31	4	5	0,72	1,09
Ext. EtOH	31	4	4	0,75	0,15
Cinzas	35	-	7	0,86	0,39

a-N: número de amostras; b-Out.: c-Outlier; d-VL: variável latente; e-R²: coeficiente de determinação; f-RMSE: raiz quadrada do erro médio de calibração ou validação; Lig.: lignina; Ext.: extrativo; Ext. EtOH: extrativo etanólico.

Obteve-se uma boa correlação entre os valores previstos pelo modelo e os valores obtidos pelo método de referência tanto para as amostras do conjunto de calibração (RMSEC) como para as do conjunto de validação (RMSEP), Figura 2. O coeficiente de determinação (R²) obtido para o conjunto de amostras da calibração foi superior a 0,86 para os constituintes analisados, demonstrando uma boa correlação entre os valores previstos pelo modelo multivariado desenvolvido e os valores obtidos pelos métodos de referência, exceto para o teor de xilana que foi de 0,77.

A escolha do melhor modelo de regressão baseou-se na análise do menor erro padrão, associado à calibração e à validação, no coeficiente de determinação (R²) entre os resultados do método de referência e aqueles previstos pelo modelo, e nos índices PRL, RPD e RER (Tabela 3). Segundo Yang et al. (2017), para fins de quantificação, o RMSEP dos modelos de calibração NIR deve variar dentro de duas vezes o valor do SEL. Para essa avaliação utilizamos o índice PRL, o qual variou de 1,0 a 1,9 para todos os modelos e, portanto, dentro da faixa de

2 vezes a precisão dos erros padrão de laboratório (SEL), indicando que os modelos estavam dentro da faixa de erro esperada para uma quantificação analítica.

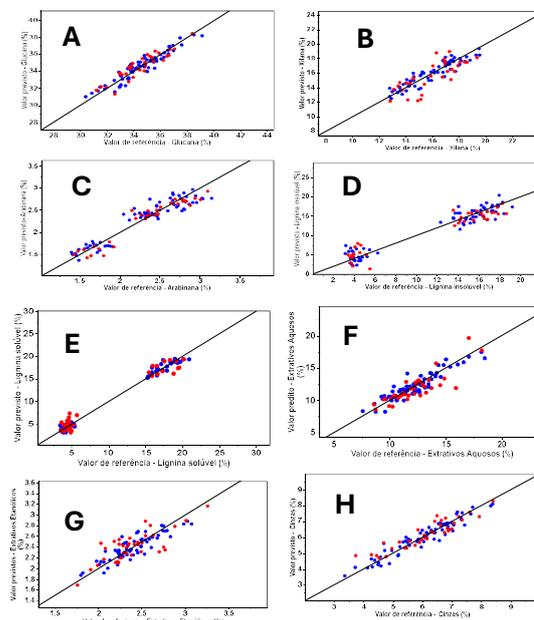
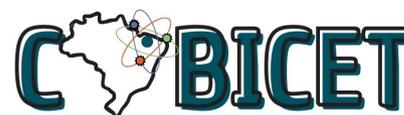


Figura 2. Gráfico dos valores de referência versus valores preditos para os constituintes químicos da biomassa de capim-elefante: A-glucana, B-xilana, C-Arabinana, D-Lignina insolúvel, E-Lignina solúvel, F-Extrato aquoso, G-Extrato etanólico e H-Cinzas. (●: amostras de calibração e ●: validação externa).

O valor de RPD representa a relação entre o desvio padrão dos valores da propriedade medida pelo método convencional baseado na química úmida e o erro padrão das amostras contidas no conjunto de previsão. O RER representa a faixa de valores para o constituinte em questão combinado com um baixo erro de previsão no conjunto de validação. RER >10 aumenta as chances de se obter modelos NIR bem-sucedidos. Segundo a literatura, modelos de calibração utilizando dados de espectroscopia NIR, com boa qualidade preditiva, apresentam valores de RPD acima de 2,5 e RER >10 para determinações quantitativas na maioria das aplicações utilizando a tecnologia NIR em produtos agrícolas. (Zhou et al, 2024).

Todos os modelos desenvolvidos para a qualificação da biomassa de capim-elefante apresentaram RPD acima de 2,5 e RER maior que 10, exceto os modelos para extrativos aquoso e etanólico, que apresentaram RPD de 2,0 e 1,6 e RER de 8,6 e 8,0, respectivamente. Sendo assim, os modelos para extrativos aquosos e etanólico são mais adequados para a realização de *screening* de amostras. Para melhorar a performance desses dois modelos,



será necessária a inclusão de novas amostras no conjunto de calibração e validação.

Tabela 3. Parâmetros estimados para o conjunto de validação dos modelos NIR-PLS para os constituintes do capim-elefante.

Constituinte	Validação				
	SEL	PRL	RPD	RER	<i>p</i> valor
Glucana	0,62	1,0	2,6	11,2	0,33
Xilana	0,63	1,5	2,6	10,1	0,46
Arabinana	0,13	1,1	3,5	11,9	0,07
Lig. insolúvel	0,89	1,8	3,7	10,1	0,75
Lig. solúvel	0,98	1,9	3,6	10,1	0,69
Ext. aquoso	0,62	1,7	2,0	8,6	0,46
Ext. EtOH	0,13	1,2	1,6	7,9	0,42
Cinzas	0,30	1,3	2,7	11,9	0,09

SEL: erro padrão de laboratório; PRL: razão entre erro de predição e erro de laboratório; RPD: relação de desempenho do desvio; RER: razão de intervalo de erro.

Para comparar os resultados obtidos entre o método de referência e o modelo desenvolvido utilizando o algoritmo PLS, foi utilizado o teste-*t* pareado (Tabela 3). O teste-*t* pareado não revelou diferenças estatísticas entre os dois métodos avaliados ao nível de 5% de significância.

Assim, em menos de 1 minuto, é possível qualificar a biomassa de capim-elefante para diferentes constituintes utilizando os modelos de calibração NIR-PLS.

CONCLUSÃO

Este estudo demonstrou o potencial da espectroscopia NIR associada a métodos quimiométricos como uma ferramenta analítica rápida, e de alto rendimento para a caracterização da biomassa de capim-elefante destinada à produção de biocombustíveis e bioprodutos. Modelos PLS-NIR foram desenvolvidos com diferentes genótipos de capim-elefante abrangendo 106 amostras para prever teores de glucana, xilana, lignina insolúvel e solúvel, extrato aquoso e etanólico, cinzas, e demonstraram ser precisos e rápidos. Os resultados da validação não apresentaram diferenças significativas entre os métodos de referência e o método NIR-PLS, sendo que os modelos para glucana, xilana, arabinana, lignina solúvel, lignina insolúvel e cinzas apresentaram PRL < 2, RER > 10 e RPD > 2,5 qualificando-os para quantificação desses constituintes em rotina laboratorial e os modelos para extrato aquoso e etanólico para triagem das amostras de capim-elefante. Os modelos NIR-PLS contribuirão significativamente para possibilitar uma

melhor qualificação da biomassa sobre o potencial energético do capim-elefante, abrindo caminho para o desenvolvimento de novas estratégias sustentáveis para a produção de biocombustíveis e bioenergia a partir dessa importante forrageira.

AGRADECIMENTOS

Os autores agradecem à Embrapa pelo suporte financeiro, projeto SEG n. 20.18.01.016.00.00.

REFERÊNCIAS

- ABREU, J., FERNÁNDEZ-CABANÁS, V., RODRIGUES, J., ABREU, M., FERREIRA, E., & PEIXOTO, W. Estimation of in vitro digestibility and fermentation of elephant grass by near infrared spectroscopy. *Grassland Science*, 69, (1), 33-41, 2022.
- AHMED, M. W.; ESQUERRE, C. A.; EILTS, K.; ALLEN, D. P.; MCCOY, S. M.; VARELA, S.; SINGH, V.; LEAKEY, A. D. B.; KAMRUZZAMAN, M. Rapid and high-throughput determination of sorghum (*Sorghum bicolor*) biomass composition using near-infrared spectroscopy and chemometrics, *Biomass and Bioenergy*, v. 186, 107276, 2024.
- MARTENS, H, WESTAD, F. Variable selection in near-infrared spectroscopy based on significance testing in partial least squares regression. *Journal of Near Infrared Spectroscopy*, 8, (2), 117–124, 2000.
- ROCHA, J., MARÇAL, T., SALVADOR, F., SILVA, A., CARNEIRO, P., RESENDE, M., ... & MACHADO, J. Unraveling candidate genes underlying biomass digestibility in elephant grass (*Cenchrus purpureus*). *BMC Plant Biology*, 19, (1), 2019.
- SLUITER, J. B.; RUIZ, R. O.; SCARLATA, C. J.; SLUITER, A.; W, D.; TEMPLETON, D. W. Compositional Analysis of Lignocellulosic Feedstocks. 1. Review and Description of Methods. *Journal of Agricultural and Food Chemistry*, 58 (16), 9043-9053, 2010.
- YANG, Z., NIE, G., PAN, L., ZHANG, Y., HUANG, L., MA, X., & ZHANG, X. Development and validation of near-infrared spectroscopy for the prediction of forage quality parameters in *Lolium multiflorum*. *PeerJ*, p. 5, e3867, 2017.
- ZHOU, W., LEI, Y., ZHOU, Q., XU, J., XUN, H., & XU, C. A rapid determination of wheat flours components based on near infrared spectroscopy and chemometrics. *Vibrational Spectroscopy*, 130, 103650, 2024.