

A era digital e suas implicações sociais: Desafios e contribuições

QUANTIFICAÇÃO DE PROPRIEDADES FÍSICO-QUÍMICAS DE TANINOS VEGETAIS POR ATR-FTIR E CALIBRAÇÃO MULTIVARIADA

Ellisson Henrique de Paulo^{1,2}(PG), Fábio Henrique Weiler¹(PG), Fábio Santos Grasel¹(PQ), Paulo Roberto Filgueiras²(PQ), Marco Flôres Ferrão¹(PQ).

¹Universidade Federal do Rio Grande do Sul, Avenida Bento Gonçalves, 9500, Agronomia – 90650-001 – Porto Alegre-RS, Brasil, ellisson.hp@gmail.com, fabio.weiler@poa.ifrs.edu.br, fsgrasel@gmail.com, mfferrao@gmail.com.

²Universidade Federal do Espírito Santo, Avenida Fernando Ferrari, 514, Goiabeiras - 29075-910 - Vitória-ES, Brasil, paulo.filgueiras@ufes.br.

Resumo

Os taninos vegetais, presentes em plantas como mimosa (*Acacia mearnsii*) e quebracho (*Schinopsis sp.*), são compostos polifenólicos com uma série de aplicações na indústria. A caracterização deles envolve análises de enxofre, nitrogênio e cinzas, revelando informações cruciais sobre sua composição. No entanto, essas análises são experimentalmente laboriosas. A espectroscopia no infravermelho (ATR-FTIR) aliada a quimiometria é uma abordagem não destrutiva e eficiente, que identifica grupos funcionais e permite a quantificação de elementos como enxofre e nitrogênio. Neste estudo a seleção de variáveis através da seleção de preditores ordenados (OPS) aliada à regressão por mínimos quadrados parciais (PLS) e regressão por vetores de suporte (SVR) foram usadas para estimar enxofre, nitrogênio e cinzas em taninos vegetais a partir de seus espectros de ATR-FTIR. Os modelos OPS-PLS apresentaram melhor desempenho para as três propriedades com erros de predição (RMSEP) igual a 0,249 (%m/m), 0,063 (%m/m) e 0,595 (%m/m) para enxofre, nitrogênio e cinzas, respectivamente.

Palavras-chave: Taninos vegetais. ATR-FTIR. OPS. PLS. SVR.

Área do Conhecimento: Ciências Exatas e da Terra.

Introdução

Os taninos vegetais, compostos químicos polifenólicos encontrados nas plantas como a mimosa (*Acacia mearnsii*) e quebracho (*Schinopsis sp.*), são categorizados como taninos hidrolisáveis e condensados. Sua função primordial nas plantas inclui defesa contra herbívoros e estresses ambientais, além de exibirem propriedades antioxidantes. Com presença em alimentos como chá, café e vinho, os taninos afetam tanto o sabor quanto a cor desses produtos. Eles também têm aplicação na medicina tradicional devido às suas propriedades antimicrobianas e adstringentes, além de desempenharem um papel na indústria do couro, onde são usados para o curtimento de peles. Resumidamente, os taninos possuem relevância em campos como botânica, alimentação, farmacologia e indústria (FALCÃO, 2014).

A caracterização de taninos vegetais envolve análises de enxofre, nitrogênio e cinzas, que fornecem informações vitais sobre sua composição química e propriedades. O enxofre está associado a aminoácidos, indicando interações com proteínas; o nitrogênio é indicativo de compostos nitrogenados que afetam a solubilidade e a reatividade; e as cinzas revelam a presença de minerais e sais, impactando a estabilidade e o sabor. No entanto, essa análise é frequentemente complementada por outras técnicas para uma compreensão completa dos taninos (FALCÃO, 2013; TONDI, 2015). As análises tradicionais visando a caracterização de taninos vegetais têm limitações consideráveis. A complexidade da composição dos taninos, repleta de diversos compostos e interações, dificulta a interpretação isolada dos resultados. Além disso, a presença de outros componentes nas amostras, como proteínas e carboidratos, pode gerar interferências nos resultados. Algumas técnicas tradicionais possuem sensibilidade e especificidade limitadas. Processos de preparação e análise, assim como variações na matriz devido a fatores sazonais e de processamento, também afetam os resultados. Geralmente, técnicas avançadas como cromatografia e espectroscopia são usadas em conjunto para

A era digital e suas implicações sociais: Desafios e contribuições

obter informações mais confiáveis, embora essas abordagens complexas possam ser mais demoradas e dispendiosas, limitando seu uso em situações com prazos restritos ou recursos limitados (FALCÃO, 2013; TONDI, 2015; WEILER, 2019).

A espectroscopia no infravermelho por transformada de Fourier e reflexão total atenuada (ATR-FTIR), uma técnica analítica não destrutiva e robusta, que pode desempenhar um papel crucial na caracterização de enxofre, nitrogênio e cinzas em taninos vegetais. Baseada na absorção de radiação infravermelha por ligações químicas nas amostras, essa técnica identifica e quantifica grupos funcionais específicos. Para a caracterização de taninos especificamente, a espectroscopia ATR-FTIR oferece vantagens significativas. Ela facilita a identificação de grupos funcionais, como ligações C-S para enxofre, grupos amida para nitrogênio e diversos grupos inorgânicos para cinzas, contribuindo para a determinação da composição química dos taninos. A técnica também permite a quantificação dos elementos com base na intensidade das bandas de absorção, desde que seja devidamente calibrada. Além disso, por ser não destrutiva, a espectroscopia ATR-FTIR preserva a integridade das amostras, sendo especialmente valiosa em estudos que demandam a manutenção das características originais. Com rapidez na análise, eficiência e ausência de necessidade de reagentes químicos, essa técnica se destaca como uma escolha principal na caracterização de taninos. Adicionalmente, sua capacidade de lidar com diferentes formas de amostras e o uso de análise de dados multivariados ampliam sua aplicabilidade em estudos de composição química de taninos (WEILER, 2019).

A quimiometria pode ser utilizada em combinação com o ATR-FTIR para a estimativa dos teores de enxofre, nitrogênio e cinzas em taninos vegetais. A abordagem quimiométrica que combina métodos de seleção de variáveis como a seleção de preditores ordenados (OPS) com técnicas de regressão por mínimos quadrados parciais (PLS) e a regressão de vetores de suporte (SVR) é altamente eficaz. Esse procedimento envolve etapas cruciais, começando pela coleta dos espectros, seguida do pré-processamento para aprimorar a qualidade dos dados. Em seguida, os métodos de seleção de variáveis são aplicados para identificar as mais relevantes entre as variáveis espectrais, otimizando a construção dos modelos PLS e SVR. Com as variáveis mais representativas selecionadas, os modelos são treinados e validados utilizando dados de teste independentes, aprimorando sua precisão e eficácia na estimativa dos componentes desejados. Essa estratégia visa maximizar a acurácia das estimativas, melhorando a eficiência do processo de análise e garantindo que apenas informações relevantes dos espectros sejam utilizadas na construção dos modelos. Como resultado, essa abordagem proporciona modelos mais confiáveis e robustos para a determinação dos componentes em taninos vegetais, utilizando espectros de infravermelho como base. Neste estudo, o algoritmo OPS foi associado as regressões PLS e SVR para quantificar os teores de enxofre, nitrogênio e cinzas em amostras de taninos vegetais a partir dos espectros de ATR-FTIR.

Metodologia

Este estudo envolveu a avaliação de amostras de mimosa e quebracho obtidas da Tanac S/A (Brasil) e da Silvateam (Argentina), respectivamente. Um total de 64 amostras foram avaliadas, sendo 31 amostras de quebracho e 33 amostras de mimosa. Esses extratos foram submetidos à secagem por *spray dryer* para criar pós com tamanhos de partícula inferiores a 210 µm e um teor de umidade variando de 4,0% a 6,0%. Para a determinação do nitrogênio total e do enxofre, as amostras passaram por processos específicos. As amostras foram primeiro secas em um forno a 60 °C por 24 horas para atingir um peso constante, prevenindo a degradação da matéria orgânica. Em seguida, as amostras foram maceradas e peneiradas para obter uma granulometria de 250 µm. Cerca de 5 g de amostra foi colocado em um tubo de falcon de 50 mL e misturado com água Milli-Q para atingir um volume de 45 mL. A mistura foi agitada, centrifugada e submetida a um processo de lavagem para remover halogênios potenciais. Após o processo de lavagem, os tubos foram secos em um forno, macerados e enviados para análise de nitrogênio total e enxofre total usando um analisador elementar. A análise envolveu a queima das amostras em analisadores elementares específicos sob atmosferas controladas de oxigênio e temperaturas. A determinação do teor de cinzas foi realizada de acordo com a norma ABNT NBR 11042:2013 (ABNT, 2013). Essa análise foi realizada em duplicata, e os resultados foram relatados como a média aritmética em base seca.

A espectroscopia ATR-FTIR foi conduzida usando um espectrofotômetro Perkin Elmer Spectrum 400 com um detector DTGS e um acessório de reflectância total atenuada universal. A faixa espectral

A era digital e suas implicações sociais: Desafios e contribuições

de 4000 a 700 cm^{-1} foi utilizada com uma resolução de 4 cm^{-1} e 32 varreduras. As amostras foram analisadas sem tratamento prévio, e as análises foram realizadas em duplicata, sendo os espectros médios utilizados para os modelos de regressão multivariada.

A análise de dados foi realizada no software do MATLAB® R2013a (The Mathworks, Natick, EUA). Os espectros ATR-FTIR foram organizados na matriz X com todas as amostras e alisados usando o algoritmo airPLS (ZHANG, 2009). As amostras foram divididas em 70% para o conjunto de calibração e 30% para a predição pelo algoritmo de Kennard-Stone (KENNARD & STONE, 1969). Os resultados de enxofre, nitrogênio e cinzas foram organizados em vetores y . Uma análise exploratória com a matriz de correlação foi conduzida nos resultados de enxofre, nitrogênio e cinzas. A seleção de variáveis com o OPS foi testada para cada propriedade usando o vetor produto, janela de 10 variáveis com 5 de incremento e 100% do conjunto original. Os dados espectrais foram pré-processados por normalização, centrados na média, variação normal padrão (SNV, do inglês *standard normal variate*), correção de espalhamento multiplicativo (MSC, do inglês *multiplicative scatter correction*) ou pela primeira derivada antes da modelagem com PLS e SVR. Os modelos PLS foram otimizados pelo método k -fold no procedimento de validação cruzada na qual a melhor variável latente (LV) é selecionada e os modelos SVR foram otimizados usando o algoritmo genético (KATOCH, 2021). Para avaliar qual modelo PLS e SVR apresentou o ajuste e a capacidade de previsão mais eficazes, foram calculados os valores do RMSE de calibração (RMSEC) e do RMSE de previsão (RMSEP). Esses parâmetros foram obtidos seguindo as Equações 01 e 02, em que y_i representa o valor de y o valor de referência da propriedade, $\hat{y}_i$ é o valor da propriedade previsto pelo modelo, $\bar{y}_i$ é o valor médio de referência, $ncal$ é o número de amostras utilizadas para a calibração e $npred$ é o número de amostras usadas para a previsão. Além disso, os coeficientes de determinação (R^2) foram calculados conforme a Equação 03, tanto para a calibração (R^2_c) quanto para a previsão (R^2_p) (OLIVIERI, 2014).

$$RMSEC = \sqrt{\sum_{i=1}^{ncal} \frac{(y_i - \hat{y}_i)^2}{ncal}} \quad (01)$$

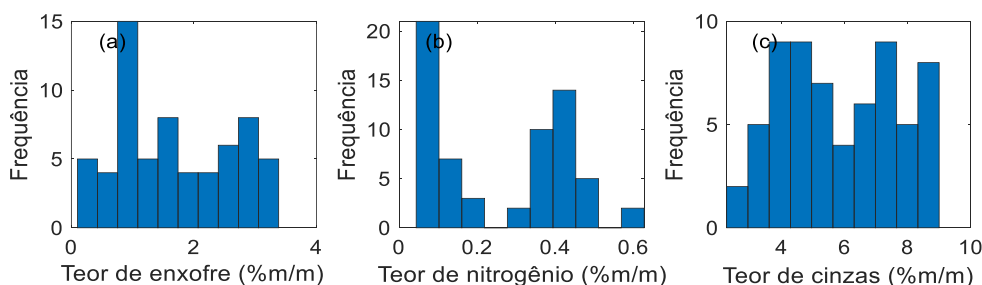
$$RMSEP = \sqrt{\sum_{i=1}^{npred} \frac{(y_i - \hat{y}_i)^2}{npred}} \quad (02)$$

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y}_i)^2} \quad (03)$$

Resultados

Na Figura 1 estão os histogramas com a distribuição dos valores de enxofre (Figura 1a) nitrogênio (Figura 1b) e cinzas (Figura 1c) nas amostras. Os teores de enxofre, nitrogênio e cinzas variaram na faixa de 0,098–3,383%, 0,044–0,628%, 2,410–8,753%, respectivamente.

Figura 1 - Histogramas de enxofre, nitrogênio e cinzas.



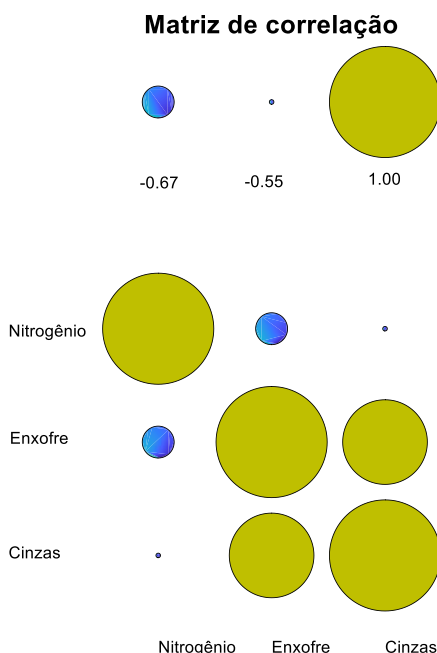
Fonte: Os autores.

A Figura 2 mostra os resultados da matriz de correlação pelos *multiblocks* para enxofre, nitrogênio e cinzas. Quanto maior o círculo, maior a correlação entre as propriedades. A correlação pode ser direta (amarelo) ou inversa (azul). A Figura 3(a) apresenta os espectros ATR-FTIR brutos das 64 amostras e a Figura 3(b) são os espectros após a correção da linha de base com o algoritmo airPLS. Após o pré-processamento, a seleção de variáveis pelo OPS foi performada e os resultados estão apresentados para enxofre na Figura 4(a), para nitrogênio na Figura 4(b) e para cinzas na Figura 4(c). Na qual, os números de onda selecionados estão representados pelas colunas em verde. A partir das variáveis

A era digital e suas implicações sociais: Desafios e contribuições

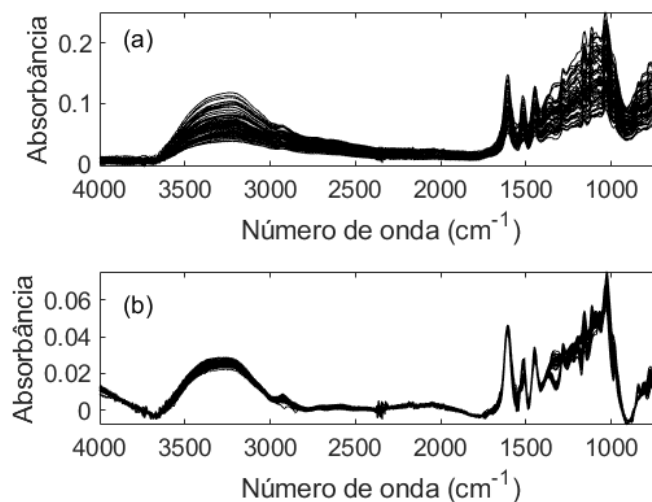
selecionadas, os modelos de regressão foram construídos. Também foram criados modelos com toda a informação espectral para a comparação. Os resultados de todos os modelos estão na Tabela 1 e os melhores modelos de regressão estão apresentados na Figura 5, na qual os círculos pretos representam as 45 amostras usadas para a calibração e os losangos verdes são as 19 amostras do conjunto de predição. Os melhores modelos para enxofre, nitrogênio e cinza foram os modelos OPS-PLS.

Figura 2 - Matriz de correlação entre enxofre, nitrogênio e cinzas.



Fonte: Os autores.

Figura 3 - Espectros de ATR-FTIR bruto (a) e alisado por airPLS (b).



Fonte: Os autores.

A era digital e suas implicações sociais: Desafios e contribuições

Figura 4 - Variáveis selecionadas pelo algoritmo OPS para enxofre (a), nitrogênio (b) e cinzas (c).

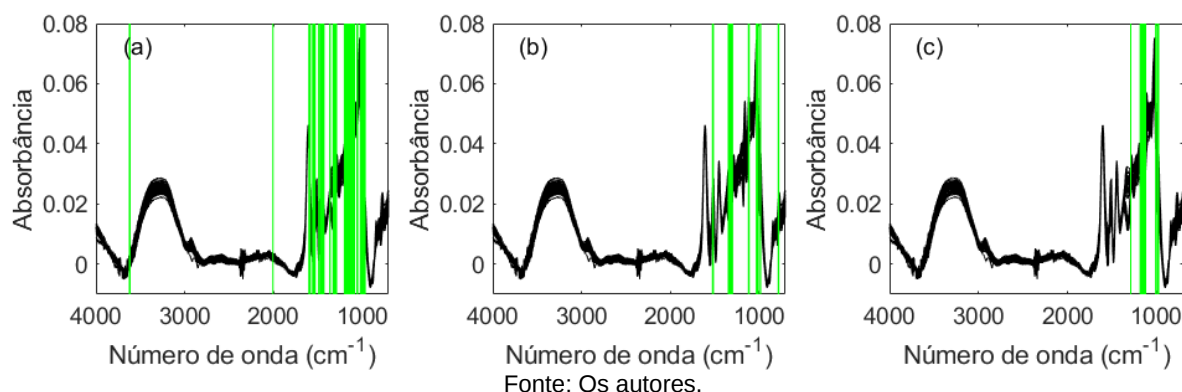


Figura 5. Modelos OPS-PLS para enxofre (a), nitrogênio (b) e cinzas (c).

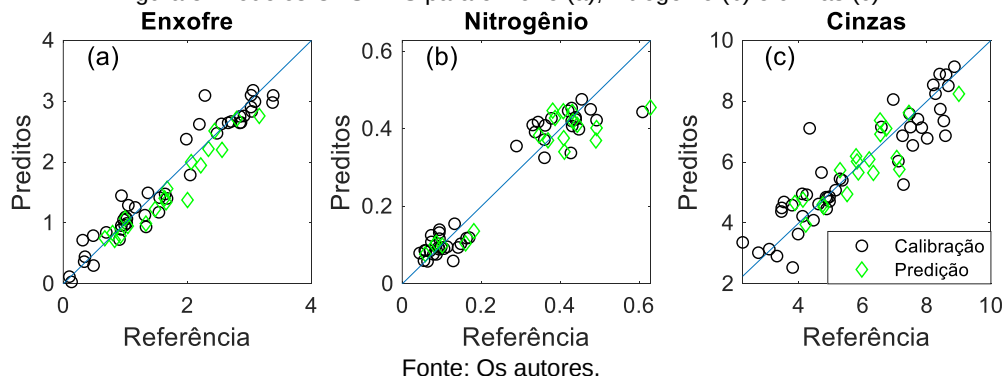


Tabela 1 - Resultados dos modelos de regressão.

Prop	VAR	Modelo	Xcal	Xtest	VL	PRE-TRAT	RMSEC	RMSEP	R ² _C	R ² _P
Enx	3301	PLS	45	19	10	SNV	0,231	0,236	0,960	0,922
Enx	275	OPS-PLS	45	19	9	SNV	0,278	0,249	0,940	0,938
Enx	3301	SVR	45	19	-	SNV	0,220	0,269	0,953	0,881
Enx	275	OPS-SVR	45	19	-	SNV	0,289	0,289	0,917	0,880
Nitro	3301	PLS	45	19	10	SNV	0,035	0,068	0,967	0,841
Nitro	100	OPS-PLS	45	19	6	SNV	0,050	0,063	0,921	0,867
Nitro	3301	SVR	45	19	-	SNV	0,045	0,076	0,936	0,821
Nitro	100	OPS-SVR	45	19	-	SNV	0,049	0,071	0,921	0,851
Cin	3301	PLS	45	19	7	SNV	0,756	0,731	0,881	0,725
Cin	80	OPS-PLS	45	19	7	DER	0,940	0,595	0,816	0,777
Cin	3301	SVR	45	19	-	SNV	0,772	0,696	0,850	0,727
Cin	80	OPS-SVR	45	19	-	SNV	0,561	0,779	0,928	0,712

Prop = propriedade; VAR = variáveis; Xcal = conjunto de calibração; Xtest = conjunto de predição; VL = variável latente; Pre-Trat = pré-tratamento; Enx = enxofre; Nitro = nitrogênio; Cin = cinzas; SNV = variação normal padrão; DER = derivada.

Discussão

A partir da Figura 1 pode-se observar que as três propriedades apresentaram uma longa faixa de distribuição, possibilitando uma maior representatividade dos valores físico-químicos. Na Figura 2, o

A era digital e suas implicações sociais: Desafios e contribuições

teor de nitrogênio está inversamente relacionado com os teores de cinzas e enxofre. Porém, cinzas e enxofre possuem uma correlação elevada. O alisamento dos espectros com airPLS diminuiu a diferença entre as amostras, o que melhorou na modelagem quimiométrica. O OPS selecionou 275, 100 e 80 variáveis para enxofre, nitrogênio e cinzas, respectivamente. Essas variáveis foram as mais significativas para a construção dos modelos de regressão, nos quais os modelos OPS-PLS demonstraram melhor exatidão e linearidade que o PLS com todo o espectro.

Conclusão

A metodologia utilizando ATR-FTIR e técnicas de regressão multivariadas para estimar os teores de enxofre, nitrogênio e cinzas em taninos vegetais se mostrou promissora. O tratamento com OPS selecionou boas variáveis preditoras, o que resultou nas melhores métricas para os modelos de regressão. Os métodos OPS-PLS e OPS-SVR predominaram. A estratégia demonstrou-se rápida, prática para a caracterização das propriedades físico-químicas de taninos vegetais.

Referências

ABNT NBR 11042:2013 Insumos — Corantes, ligantes, pastas pigmentadas, substâncias graxas e tanantes — Determinação do teor de cinza.

FALCÃO, L; ARAÚJO, MEM. **Tannins characterization in historic leathers by complementary analytical techniques ATR-FTIR, UV-Vis and chemical tests.** Journal of Cultural Heritage 14 (2013) 499–508.

FALCÃO, L; ARAÚJO, MEM. **Application of ATR-FTIR spectroscopy to the analysis of tannins in historic leathers: The case study of the upholstery from the 19th century Portuguese Royal Train.** Vibrational Spectroscopy 74 (2014) 98–103.

TONDI, G; PETUTSCHNIGG, P. **Middle infrared (ATR FT-MIR) characterization of industrial tannin extracts.** Industrial Crops and Products 65 (2015) 422–428

WEILER, FH; GRASEL, FA; LOUREGA, RV; RAMOS, AS; FERRÃO, MF. **Simultaneous determination of sulfur, nitrogen and ash for vegetable tannins using ATR-FTIR spectroscopy and multivariate regression.** Microchemical Journal.149 (2019), 10399.

KENNARD, RW; STONE, LA. **Computer Aided Design of Experiments.** Technometric. 1969;11:137–48. <https://doi.org/10.1080/00401706.1969.10490666>

KATOCH, S; CHAUHAN, SS; KUMAR, V. **A review on genetic algorithm: past, present, and future,** Multimed Tools Appl. 80 (2021) 8091–8126. <https://doi.org/10.1007/s11042-020-10139-6>.

OLIVIERI, AC. **Analytical figures of merit: From univariate to multiway calibration,** Chem Rev. 114 (2014) 5358–5378. <https://doi.org/10.1021/cr400455s>.

ROQUE, JV et al. **Comprehensive new approaches for variable selection using ordered predictors selection.** Anal Chim Acta, 2019;1075:57–70. <https://doi.org/10.1016/j.aca.2019.05.039>.

ZHANG, ZM; CHEN, S; LIANG, YZ. **Baseline correction using adaptive iteratively reweighted penalized least squares.** Analyst, 2010, 135, 1138–1146.

Agradecimentos

TANAC S/A, UFRGS, UFES, CAPES, CNPq.