

PROPOSTA DE METODOLOGIA PARA O ESTABELECIMENTO DE MODELO DE CALIBRAÇÃO MULTIVARIADO PARA A PREVISÃO DO TEOR DE TRIGONELINA EM CAFÉ CRU BASEADO EM ESPECTROSCOPIA NO INFRAVERMELHO PRÓXIMO¹

Juliano Souza Ribeiro² e Terezinha de Jesus Garcia Salva³

¹ Trabalho financiado pela FAPESP e Consórcio Brasileiro de Pesquisa e Desenvolvimento do café – Consórcio Pesquisa Café

² Professor, Dr., Instituto Federal do Espírito Santo - ES, julianoribeiro@ifes.edu.br

³ Pesquisador, Dr., Instituto Agronômico de Campinas - SP, tsalva@iac.sp.gov.br

RESUMO: A análise espectroscópica na região do infravermelho próximo (NIRS) de amostras de cafés crus e torrados é uma técnica analítica proposta para a identificação de espécies, na definição do ponto de torra, na qualidade final da bebida e na quantificação de alguns compostos químicos do grão. A exploração dos resultados deste tipo de análise, na maioria das situações, no entanto, pressupõe o confronto prévio entre dados espectrais e resultados de análises das mesmas amostras obtidos pelo emprego de metodologia analítica de referência, gerando os modelos ou curvas de calibração. A construção de um modelo de calibração robusto requer que as duas metodologias sejam empregadas, simultaneamente, para grande número de amostras, o que consiste em fator limitante para o amplo emprego da espectroscopia. Neste trabalho, uma curva de calibração multivariada para a quantificação de trigonelina foi construída e validada mediante análises espectroscópicas na região do infravermelho próximo da substância pura e de amostras de café, empregando-se reduzido número de análises pela metodologia de referência (Cromatografia líquida de alta eficiência). Para isso, foram utilizadas amostras de café do banco de germoplasma do Centro de Café Alcides Carvalho, no Instituto Agronômico de Campinas. A regressão pelo método de quadrados mínimos parciais (PLS) foi utilizada para a construção do modelo. Através da seleção de variáveis realizada, foram identificadas 16 regiões espectrais relacionadas com a trigonelina. Para o modelo calculado com 8 variáveis latentes, o erro de validação cruzada (RMSECV) foi de $0,048 \pm 0,02$, o coeficiente de correlação (r_{vc}) foi de $0,98 \pm 0,02$ e o erro de previsão (RMSEP) foi de $0,056$.

Palavras-chave: Infravermelho próximo por reflectância difusa, trigonelina, quimiometria.

A PROPOSAL OF METHODOLOGY FOR THE ESTABLISHMENT OF THE MULTIVARIATE CALIBRATION MODEL FOR PREDICTION OF TRIGONELLINE CONTENT IN RAW COFFEE BEANS BASED ON NEAR INFRARED SPECTROSCOPY

ABSTRACT: The near infrared spectroscopic (NIRS) applied to raw and roasted coffee samples is an analytical technique proposed to the identification of species, in defining the roasting degree and the final quality of the beverage and, also, in the quantification of some chemical compounds of the seeds. The exploitation of the results of this analysis, in most situations, however, requires the confrontation between spectral data and the results of analysis by the reference method of analysis applied to the same samples, through the calibration models or curves. Despite the advantages of the NIRS analysis, the construction of a robust calibration model requires that a large number of samples is analysed, which consists a limiting factor for the widespread use of spectroscopy. In this study, a multivariate calibration curve for quantification of trigonelline was built and validated by near infrared spectroscopic analysis of the pure compound and of the coffee samples, using low number of analyses by the reference methodology (High pressure liquid chromatography). For this, coffee samples from germplasm bank of the Coffee Center 'Alcides Carvalho' of the Agronomic Institute of Campinas were used. Partial least squares regression method (PLS) was used to build the model. Through the variable selection performed, 16 spectral regions related to trigonelline were identified. The cross validation error (RMSECV) was 0.048 ± 0.02 , the correlation coefficient (r_{vc}) was 0.98 ± 0.02 and the prediction error (RMSEP) was 0.056 , for the model calculated with 8 latent variables

Key words: Diffuse reflectance near infrared, trigonelline and chemometrics.

INTRODUÇÃO

Análises químicas de café têm justificado trabalhos que visam ao conhecimento da (i) herdabilidade em híbridos interespecíficos (Priolli et al, 2008), (ii) à possível relação dos diferentes compostos com a qualidade da bebida e, mais recentemente, no que diz respeito a compostos, que não a cafeína, (iii) a prospectar sobre o possível emprego do grão na área da cosmetologia e de alimentos funcionais (Wagemaker et al., 2011).

Métodos analíticos para a quantificação dos principais compostos presentes no grão de café têm sido empregados em rotinas laboratoriais. Utilizando cromatografia líquida de alta eficiência (CLAE) têm sido determinadas, por exemplo, as concentrações de cafeína (Mazzafera et al., 1997; De Maria, et al. (1995), trigonelina (Casal et al., 2000; Perrone et al., 2008a), ácidos clorogênicos (Perrone et al., 2008b), carboidratos e sacarose (Campa et al., 2004), lipídios (González et al., 2001) e proteínas e aminoácidos (Mazzafera e Gonçalves, 1999). Os mesmos compostos têm sido quantificados, também, por ressonância magnética nuclear de prótons - RMN H^1 (Tavares e Ferreira, 2006), espectroscopia no ultravioleta (UV) e no infravermelho (IV), gravimetria (Davrieux et al., 2004) e, ainda, por análises enzimáticas e eletroquímicas (Rambla et al., 1997; Giannoccaro et al., 2008).

A maioria destas técnicas tem em comum a necessidade de tratamentos complexos das amostras (*clean up*), muitas vezes demorados, caros e destrutivos. Para superar essa limitação, tem sido dada ênfase ao desenvolvimento de técnicas espectroscópicas de análise.

A difusão das metodologias baseadas em informações de espectros dos produtos obtidos em diferentes comprimentos de onda se deve em grande parte ao desenvolvimento da quimiometria. Por meio desta ferramenta tem sido possível interpretar melhor os resultados analíticos (Pasquini, 2003) e empregar a espectroscopia em análises qualitativas e quantitativas de diversos materiais (Siebielec et al., 2004; Pedro e Ferreira, 2005; Li et al., 2006; Chen et al., 2006; Cozzolino et al., 2006; Karoui et al., 2006; Cen e He, 2007; Pizarro et al., 2007a).

Entre as análises espectroscópicas em uso, na análise de alimentos destacam-se as realizadas na região do infravermelho próximo (NIR). A espectrometria na região de infravermelho próximo (NIRS) apresenta a vantagem de ser aplicável a análises tanto de amostras líquidas, como azeite de oliva, sucos e vinhos (Downey et al., 2003; Cozzolino et al., 2003; Shao e He, 2007), quanto de amostras sólidas, neste caso utilizando as técnicas de reflectância (Esteban-Díez et al., 2004; Karoui et al., 2006; Alessandrini et al., 2008).

Especificamente com café, vários estudos vêm sendo realizados com NIRS. Dentre eles, encontram-se a discriminação de espécies (*Coffea arabica* e *Coffea canephora*) em amostras puras e em *blends* (Pizarro et al., 2007b; Esteban-Díez et al., 2007), a quantificação de cafeína (Pizarro et al., 2007a), trigonelina, ácidos clorogênicos (Bertrand et al., 2005), açúcares totais (Morgano et al., 2007) e lipídios (Davrieux et al., 2004). a definição do grau de torra e da umidade (Alessandrini et al., 2008) Mais recentemente, vem se estudando também a aplicação do NIRS como alternativa para avaliação da qualidade sensorial da bebida (Ribeiro et al., 2011).

Contudo, para se determinar o teor de um componente químico presente no grão de café, ou em outras matrizes, por meio dos espectros de infravermelho próximo, é necessário primeiramente estabelecer a relação entre as informações fornecidas pelos espectros no infravermelho e os valores das medidas fornecidos por métodos de referência, tais como polarimetria (Da Costa Filho, 2009) e CLAE (Xie et al., 2009), por exemplo. Essa relação consiste no modelo multivariado de regressão linear, ou modelo quimiométrico, que pode ser construído mediante a aplicação dos métodos MLR (regressão linear múltipla) ou PLS (regressão linear por quadrados mínimos parciais).

Para a construção de um modelo quimiométrico robusto e confiável, o número de amostras que devem ser analisadas no NIR e pelo método de referência é, geralmente, da ordem de centenas, e o esforço demandado é bastante grande. Outra característica da metodologia de NIRS, da maneira como é usualmente empregada, é não definir completamente os comprimentos de onda (λ) importantes para a quantificação de cada composto, sendo extraídas apenas poucas informações sobre as regiões espectrais. Essa limitação quanto à definição dos comprimentos de onda importantes para a quantificação dos compostos torna os modelos estabelecidos de aplicação apenas local, ou seja, eles são restritos aos conjuntos de dados estudados, sendo impraticável a sua utilização por outros laboratórios.

Apenas recentemente foram publicados alguns trabalhos em que foram definidos os comprimentos de onda específicos dos componentes químicos encontrados nas matrizes estudadas mediante a determinação dos espectros das substâncias puras. É o caso, por exemplo, dos trabalhos que procuram determinar as regiões espectrais referentes à cafeína e à sacarose em chás e chocolates, respectivamente (Sinija e Mishra, 2009; Da Costa Filho, 2009; Ribeiro et al., 2011).

A inovação do método que se apresenta neste resumo consiste em estabelecer os espectros de matrizes originais e artificiais de café. Essa inovação, associada ao espectro de trigonelina puro e à seleção de variáveis empregadas no desenvolvimento do modelo, admite a redução do número de análises pelo método de referência e o desenvolvimento de um modelo quimiométrico universal para a quantificação do composto.

MATERIAL E MÉTODOS

Na extração de trigonelina da matriz de café e nas análises por CLAE foram empregadas trigonelina 99 % (Sigma-Aldrich -Munich, Alemanha), solventes grau espectroscópico (J.T.Baker) e solventes de grau P.A (Merck).

Para a construção da curva de calibração multivariada foram utilizadas 9 amostras de grão de café cru de variedades de *C. arabica* e *C. canephora* do banco de germoplasma do Centro de Café do IAC (Tabela 1), selecionadas por diferirem no conteúdo de trigonelina. A amostra de café *C. canephora* var. Conillon coleção 66-3 (número 4 na Tabela 1) foi utilizada também para a geração de uma matriz de café com baixo teor de cafeína, trigonelina e ácidos clorogênicos (amostra 10) após tratamento com metanol:água (70:30, v:v).

Frutos da safra 2009/2010 das plantas de interesse foram colhidos no estágio de maturação cereja, despolpados em despolpador piloto e desmucilados por 24 horas. Após a desmucilagem, as sementes foram lavadas e colocadas ao

sol para secar. O processo de secagem de cada amostra foi finalizado quando ela apresentou umidade de aproximadamente 11 %.

As "matrizes artificiais" foram desenvolvidas a partir da mistura de diferentes proporções das 9 amostras originais com a amostra tratada com metanol: água. Foram analisadas também todas as matrizes originais.

Tabela 1. Amostras utilizadas para a construção dos modelos de regressão multivariados.

Amostras	Espécie	Variedade
1	<i>Coffea arabica</i>	Bourbon amarelo
2	<i>Coffea arabica</i>	Material com baixo teor de cafeína em seleção
3	<i>Coffea arabica</i>	IAC 3996 - Ep 416 - Cova 1152 - Planta B
4	<i>Coffea canephora</i>	Conillon coleção 66 - cova 3
5	<i>Coffea canephora</i>	Apoatã - Coleção 3599 - Planta 9
6	<i>Coffea canephora</i>	Apoatã - Coleção 3597 - Planta 1
7	<i>Coffea canephora</i>	Bukobensis - Coleção 3
8	<i>Coffea canephora</i>	Ep 133 - 350 - 1653 - Cova 4
9	<i>Coffea canephora</i>	Ep 189 - Cova 128
10	<i>Coffea canephora</i>	Matriz extraída com metanol / água

Os espectros de infravermelho próximo foram adquiridos em espectrômetro modelo Nirsystem 6500 (Foss NIRSystems, Raamsdonksveer, Holanda) de reflectância difusa com módulo de transporte de amostra e controlado pelo software Vision 2.22 (Foss NIRSystems, Raamsdonksveer, Holanda).

Os espectros foram obtidos diretamente sobre as amostras moídas de café cru usando 256 varreduras na faixa de 1100 - 2500 nm, com 4 cm⁻¹ de resolução. Um padrão interno de poliestireno (4 picos) foi utilizado como branco. Três diferentes porções de cada amostra eram analisadas (triplicatas).

A quantificação da trigonelina nas matrizes originais e isentas de trigonelina foi realizada por cromatografia líquida de alta eficiência (CLAE), de acordo com adaptação feita em método de Casal et al. (2000). Empregou-se cromatógrafo Shimadzu, equipado com detector de UV-Vis e coluna C-18 Shim-pack CLC-ODS (M) de 4,6 x 250 mm e 5 µm de diâmetro de partícula, com pré-coluna de 4 x 1 mm. A eluição foi isocrática com fase móvel composta por metanol : ácido acético: água (30:0,5:69,5; v:v:v) com vazão de 1 ml/minuto a 30 °C. O comprimento de onda de leitura das amostras foi 272 nm. A concentração da trigonelina foi determinada pela relação entre as áreas dos picos de trigonelina da amostra e a do padrão de concentração conhecida. As amostras foram analisadas em triplicata.

Os dados obtidos na análise espectroscópica foram transformados em matrizes **X** (*I* x *J*), na qual cada replicata foi utilizada como uma amostra. O processamento dos dados foi realizado utilizando o software Matlab 6.5 (The MathWorks, Co., Natick, MA, USA) e o pacote computacional PLS_Toolbox (Eigenvector Research, Inc. – PLS_Toolbox version 3.02.) (Wise et al., 2004).

A regressão por quadrados mínimos parciais (PLS) foi utilizada como método de regressão (Ferreira et al., 1999). Para reduzir as variações aleatórias (ruído experimental) foi realizado um alisamento, segundo Savitzky-Golay (1964), com janela de tamanho 11. As variações sistemáticas foram reduzidas utilizando a 2ª derivada e os dados foram centrados na média.

RESULTADOS E DISCUSSÃO

Na construção do modelo de regressão foram utilizados os valores médios das concentrações (%) de trigonelina medidas por Cromatografia líquida de alta eficiência e 345 espectros referentes à 115 matrizes artificiais de café. Para a etapa de validação externa foram utilizadas 49 amostras analisadas por CLAE e NIRS.

As variáveis selecionadas para a construção do modelo PLS de previsão da concentração de trigonelina foram obtidas a partir da análise visual dos espectros sobrepostos de algumas amostras, e de trigonelina pura. Assim, 16 regiões espectrais (109 variáveis) que absorveram a trigonelina foram selecionadas. Essas regiões estão indicadas em faixas verticais na Figura 1.

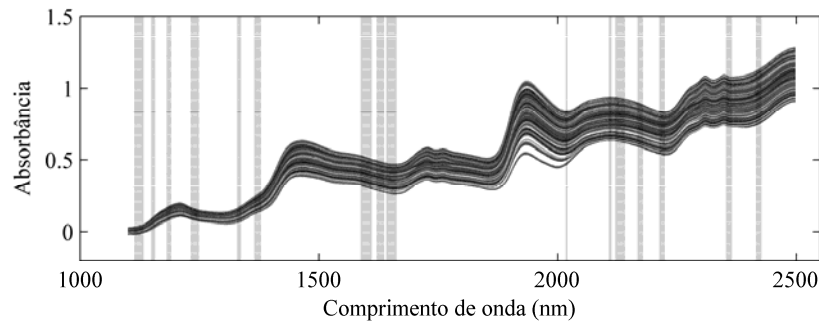


Figura 1- Espectros sobrepostos das amostras de café cru moído e regiões do infravermelho próximo selecionadas para a construção de modelo de calibração para a trigonelina.

Para o conjunto de dados considerado neste trabalho, o modelo de regressão PLS construído com 8 variáveis latentes apresentou algumas replicatas com valores elevados de *leverage* e de resíduo de *Student*. Neste caso, as replicatas foram retiradas e o modelo recalculado. A Figura 2 apresenta o gráfico *leverage vs* resíduo de *Student* do modelo recalculado (limite de confiança do valor de *leverage* de 95 %).

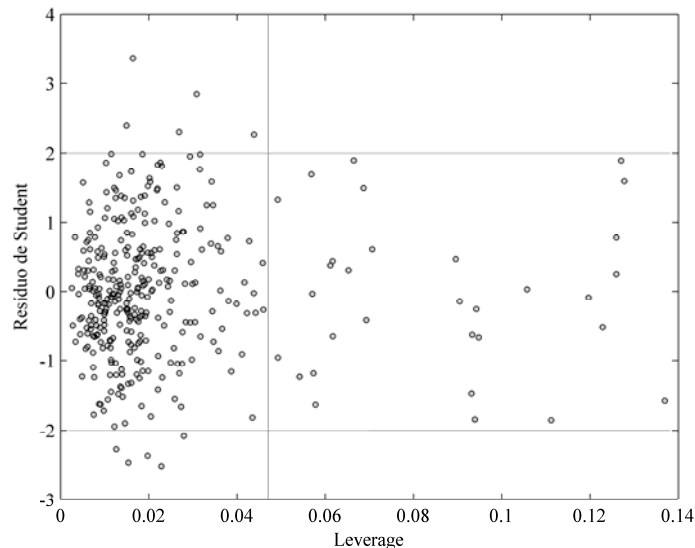


Figura 2 - Gráfico de *leverage vs* resíduo de *Student* para visualização de possíveis amostras atípicas.

Esse modelo apresentou erro de validação cruzada (RMSECV, *root mean square error of cross validation*) igual a $0,048 \pm 0,002$ e coeficiente de correlação da validação cruzada (r_{vc}) igual a $0,98 \pm 0,02$. Com as 8 variáveis latentes utilizadas foi possível descrever 97,11 % e 97,55 % da variância dos blocos **Y** e **X** utilizados nas calibrações.

O modelo foi validado por validação externa, utilizando o conjunto de 49 amostras, relativas às 147 replicatas. A Figura 3 representa os valores determinados por CLAE *versus* os valores preditos pelo modelo para cada amostra do conjunto de validação externa distribuídos entre os valores do conjunto de validação cruzada. O erro de previsão das amostras externas (RMSEP, *root mean square error of prediction*) foi de 0,056.

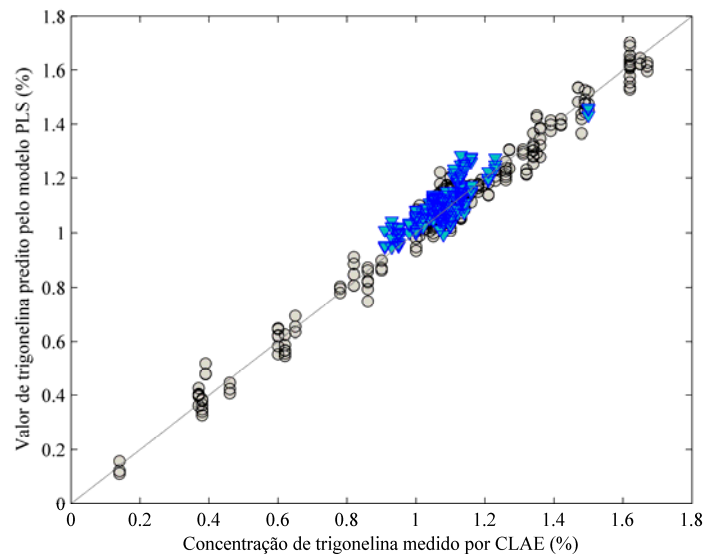


Figura 3 - Representação gráfica dos valores de trigonelina obtidos por CLAE vs os valores previstos pelo modelo PLS. Amostras de validação cruzada (○) e validação externa (△).

CONCLUSÕES

O método proposto para construção da curva de calibração tomando por base os espectros de infravermelho e as concentrações de trigonelina de matrizes de café cru, tanto originais quanto artificiais, bem como o espectro de infravermelho próximo de trigonelina pura, se mostrou bastante eficiente. Ele possibilitou a construção de modelo de regressão multivariada para a determinação do teor de trigonelina em café cru com erro relativo médio entre as análises por infravermelho próximo e CLAE para as 49 amostras de previsão igual 5,6 %, em concentrações que variaram entre 0,16% e 1,67%. De acordo com as médias e desvios-padrão das replicatas analisadas, tanto no conjunto de calibração quanto de validação, o modelo se apresentou bem reproduzível, com média de 3% de variação entre elas..

A confiabilidade nas faixas espectrais selecionadas para este modelo PLS, o credencia para ser utilizado pelo Centro de café do Instituto Agrônomo de Campinas e por outros grupos de pesquisa que utilizem outras amostras (espécies e variedades), equipamentos de diferentes marcas, graus de moagem da amostra e diferentes outras variáveis.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALESSANDRINI, L.; ROMANI, S.; PINNAVAIA, G.; ROSA, M. D.; Near infrared spectroscopy: an analytical tool to predict coffee roasting degree. **Anal. Chim. Acta**, 2008, 625, 95-102.
- BERTRAND, B.; ETIENNE, H.; LASHERMES, P.; GUYOT, B.; DAVRIEUX F. Can near-infrared reflectance of green coffee be used to detect introgression in *Coffea arabica* cultivars ? **J. Sci. Food Agric.**, 2005, 85, 955-962.
- CAMPA, C.; BALLESTER, J. F.; DOULBEAU, S.; DUSSERT, S.; HAMON, S.; NOIROT, M. Trigonelline and sucrose diversity in wild *Coffea species*. **Food Chem.**, 2004, 88, 39-43.
- CASAL, S.; OLIVEIRA, B.; FERREIRA, M. A. HPLC/diode-array applied to the thermal degradation of trigonelline, nicotinic acid and caffeine in coffee. **Food Chem.**, 2000, 68, 481-485.
- CEN, H.; HE, Y. Theory and application of near infrared reflectance spectroscopy in determination of food quality. **Trends Food Sci. Technol.**, 2007, 18, 72-83.
- CHEN, Q.; ZHAO, J.; ZHANG, H.; WANG, X. Feasibility study on qualitative and quantitative analysis in tea by near infrared spectroscopy with multivariate calibration. **Anal. Chim. Acta**, 2006, 572, 77-84.
- COZZOLINO, D.; SMYTH, H. E.; GISHEN, M. Feasibility study on the use of visible and near-infrared spectroscopy together with chemometrics to discriminate between commercial white wines of different varietal origins. **J. Agric. Food Chem.**, 2003, 51, 7703-7708.
- COZZOLINO, D.; SMYTH, H. E.; LATTEY, K. A.; CYNKAR, W.; JANIK, L.; DAMBERGS, R. G.; FRANCIS, I. L.; GISHEN, M. Combining mass spectrometry based electronic nose, visible-near infrared spectroscopy and chemometrics to assess the sensory properties of Australian Riesling wines. **Anal. Chim. Acta**, 2006, 563, 319-324.
- DA COSTA FILHO, P.A. Rapid determination of sucrose in chocolate mass using near infrared spectroscopy. **Anal. Chim. Acta**, 2009, 631, 206-211.
- DAVRIEUX, F.; MANEZ, J. C.; DURAND, N.; GUYOT, B. Determination of the content of six major biochemical compounds of green coffee using NIR. NIR Publications. Near Infrared Spectroscopy: Proceedings of the 11th International Conference, 2004.

- De Maria, C. A. B.; Trugo, L.C.; Moreira, R.F.A.; Petracco, M. Simultaneous determination of total chlorogenic acid, trigonelline and caffeine in green coffee samples by high performance gel filtration chromatography. *Food Chemistry*, 1995; 52: 447-449
- DOWNEY, G.; MCINTYRE, P.; DAVIES, A. N.. Geographic classification of extra virgin olive oils from the eastern Mediterranean by chemometric analysis of visible and near-infrared spectroscopy data. *Applied Spectrosc.*, 2003, 57, 158-163.
- ESTEBAN-DÍEZ, I.; GONZÁLEZ-SÁIZ, J. M.; PIZARRO, C. Prediction of sensory properties of espresso from roasted coffee samples by near-infrared spectroscopy. *Anal. Chim. Acta*, 2004, 525, 171-182.
- ESTEBAN-DÍEZ, I.; GONZÁLEZ-SÁIZ, J. M.; SÁENZ-GONZÁLEZ, C.; PIZARRO, C. Coffee varietal differentiation based on near infrared spectroscopy. *Talanta*, 2007, 71, 221-229.
- FERREIRA, M. M. C.; ANTUNES, A. M.; MELGO, M. S.; VOLPE, P. L. O. Chemometrics I – Multivariate calibration, a tutorial. *Quim. Nova*, 1999, 22, 724-731.
- Geromel, C.; Ferreira, L. P.; Guerreiro, S. M. C.; Cavalari, A. A.; Port, D.; Pereira, L. F. P.; Leroy, T.; Vieira, L. G. E.; Mazzafera, P.; Marraccini, P. Biochemical and genomic analysis of sucrose metabolism during coffee (*Coffea arabica*) fruit development. *Journal of experimental Botany*, 2006, 57, 3243-3258.
- GIANNOCCARO, E.; WANG, Y. J.; CHEN, P. Comparison of two HPLC systems and an enzymatic method for quantification of soybean sugars. *Food Chem.*, 2008, 106, 324-330.
- GONZÁLEZ, A. G.; PABLOS, F.; MARTÍN, M. J.; LEÓN-CAMACHO, M.; VALDENEBRO, M. S.. HPLC analysis of tocopherols and triglycerides in coffee and their use as authentication parameters. *Food Chem.*, 2001, 73, 93-101.
- KAROUI, R.; PILLONEL, L.; SCHALLER, E.; BOSSET, J. O.; DE BAERDEMAEKER, J. Prediction of sensory attributes of European Emmental cheese using near-infrared spectroscopy: A feasibility study. *Food Chem.*, 2006, 101, 1121-1129.
- LI, W.; HUANG, D.; LIU, L. The application of near infrared spectroscopy in the quality control analysis of glass/phenolic resin prepreg. *J. Mat. Sci.*, 2006, 41, 7183-7189.
- MALUF, M. P.; SILVA, C. C.; OLIVEIRA, M. P. A.; TAVARES, A. G.; SILVAROLLA, M. B.; GUERREIRO FILHO, O. Altered expression of the caffeine synthase gene in a naturally caffeine-free mutant of *Coffea arabica*. *Gen.Mol. Biol.*, 2009, 32, 802-810.
- MAZZAFERA, P.; SILVAROLLA, M. B.; LIMA, M. M. A.; MEDINA FILHO, H. P. Caffeine content in diploid coffee species. *Ciê. Cult.*, 1997, 49, 216 – 218.
- MAZZAFERA, P.; GONÇALVES, K. V. Nitrogen compounds in the xylem sap of coffee. *Phytochem.*, 1999, 50, 383-386.
- MORGANO, M. A.; FARIA, C. G.; FERRÃO, M. F.; FERREIRA, M. C. Determinação de açúcares total em café cru por espectroscopia no infravermelho próximo e regressão por mínimos quadrados parciais. *Quim. Nova* 2007, 30, 346-350.
- PASQUINI, C. Near Infrared Spectroscopy: fundamentals, practical aspects and analytical Applications. *J. Braz. Chem. Soc.*, 2003, 14, 198-219.
- PEDRO, A. M. K.; FERREIRA, M. M. C. Nondestructive determination of solids and carotenoids in tomato products by near-infrared spectroscopy and multivariate calibration. *Anal. Chem.*, 2005, 77, 2505-2511.
- PERRONE, D.; DONANGELO, C. M.; FARAH, A. Fast simultaneous analysis of caffeine, trigonelline, nicotinic acid and sucrose in coffee by liquid chromatography–mass spectrometry. *Food Chem.*, 2008a, 110, 1030–1035.
- PERRONE, D.; FARAH, A.; DONANGELO, C. M.; PAULIS, T.; MARTIN, P. R. Comprehensive analysis of major and minor chlorogenic acids and lactones in economically relevant Brazilian coffee cultivars. *Food Chem.*, 2008b, 106, 859–867.
- PIZARRO, C.; ESTEBAN-DÍEZ, I.; GONZÁLEZ-SÁIZ, J. M.; FORINA, M. Use of Near-Infrared Spectroscopy and Feature Selection Techniques for Predicting the Caffeine Content and Roasting Color in Roasted Coffees. *J. Agric. Food Chem.*, 2007a, 55, 7477-7488.
- PIZARRO, C.; ESTEBAN-DÍEZ, I.; GONZÁLEZ-SÁIZ, J. M. Mixture resolution according to the percentage of robusta variety in order to detect adulteration in roasted coffee by near infrared spectroscopy. *Anal. Chim. Acta*, 2007b, 585, 266-276.
- PRIOLLI, R. H.G.; MAZZAFERA, P.; SIQUEIRA, W. J.; MÖLLER, M.; ZUCCHI, M. I.; RAMOS, L. C. S.; GALLO, P. B.; COLOMBO, C. A.. Caffeine inheritance in interspecific hybrids of *Coffea arabica* x *Coffea canephora* (Gentianales, Rubiaceae). *Genet. Mol. Biol.* 2008, 31, 498-504.
- RAMBLA, F. J.; GARRIGUES, S.; DE LA GUARDIA, M. PLS-NIR determination of total sugar, glucose, fructose and sucrose in aqueous solutions of fruit juices. *Anal. Chim. Acta*, 1997, 344, 41–53.
- RIBEIRO, J.S.; FERREIRA, M.M.C.; SALVA, T.J.G. Chemometric models for the quantitative descriptive sensory analysis of Arabica coffee beverages using near infrared spectroscopy. *Talanta*, 2011, 83, 1352-1358.
- SAVITZKY, A.; GOLAY, M. J. E. Smoothing + differentiation of data by simplified least squares procedures. *Anal. Chem.*, 1964, 36, 1627-1679.
- SHAO, Y.; HE, Y.. Nondestructive measurement of the internal quality of bayberry juice Vis/NIR spectroscopy. *J. Food Eng.*, 2007, 79, 1015-1019.

- SIEBIELEC, G.; MCCARTY, G. W.; STUCZYNSKI, T. I.; REEVES, J. B. Near- and Mid-infrared diffuse reflectance spectroscopy for measuring soil metal content. **J. Environ. Qual.**, 2004, 33, 2056-2069.
- SINIJA, V. R.; MISHRA, H. N.. FT-NIR spectroscopy for caffeine estimation in instant green tea powder and granules. **LWT - Food Sci. Technol.**, 2009, 42, 998–1002.
- TAVARES, L. A.; FERREIRA, A. G.. Análise quali- e quantitativa de cafés comerciais via ressonância magnética nuclear. **Quím. Nova**, 2006, 29, 911-915.
- WAGEMAKER, TAIS ALERIANA LUCON ; CARVALHO, CASSIA REGINA LIMONTA ; MAIA, NILSON BORLINA ; GUERREIRO FILHO, OLIVEIRO ; BAGGIO, SUELI REGINA . Sun protection factor, content and composition of lipid fraction of green coffee beans. **Industrial Crops and Products**, 2011, 33, 469-473,
- WISE, B. M.; GALLAGHER, N. B.; BRO, R.; SHAVER, J. M.; WINDIG, W.; KOCH, R. S. PLS_Toolbox 3.5, for Use with Matlab™, Eigenvector Research. 2004.
- XIE, L.; YE, X.; LIU, D.; YING, Y. Quantification of glucose, fructose and sucrose in bayberry juice by NIR and PLS. **Food Chem.**, 2009, 114, 1135–1140.