



Quantificação de gordura, de proteína e de matéria seca do leite bovino utilizando espectroscopia Raman: uma possibilidade de aplicação em campo¹

Giovanni Montesano Schettino², Landulfo Silveira Jr.³, Livia Helena Moreira⁴, Hugo Viveiros Schettino⁵, Marcos Vinícius Gualberto Barbosa da Silva⁶

¹Pesquisa realizada durante período de estágio do primeiro autor na Embrapa Gado de Leite, Juiz de Fora, MG.

²Mestrando do Programa de Pós-Graduação em Bioengenharia – UNICASTELO, Estr. Dr. Altino Bondesan, 500, São José dos Campos, SP. e-mail: giovanni.schettino@ipasoft.com.br

³Pesquisador do Instituto de Engenharia Biomédica – UNICASTELO, Estr. Dr. Altino Bondesan, 500, São José dos Campos, SP. e-mail: landulfo.silveira@gmail.com

⁴Pesquisadora do Instituto de Engenharia Biomédica – UNICASTELO, Estr. Dr. Altino Bondesan, 500, São José dos Campos, SP. e-mail: lhsil@uol.com.br

⁵Mestrando do Programa de Pós-Graduação em Engenharia Elétrica – UFJF, Rua José Lourenço Kelmer S/N, Juiz de Fora, MG. Bolsista do CNPq/Capes. e-mail: hugoschettino@gmail.com

⁶Pesquisador da Embrapa Gado de Leite, Rua Eugênio do Nascimento 610, Juiz de Fora, MG. Bolsista CNPq. e-mail: marcos.vb.silva@embrapa.br

Resumo: As espectroscopias Raman e no infravermelho podem fornecer informações sobre a composição química de amostras biológicas, sendo as contribuições das moléculas no espectro proporcionais às suas concentrações relativas, capacitando ao diagnóstico quantitativo. Este estudo utilizou espectroscopia Raman associada aos métodos quimiométricos e de pré-processamento de sinais para determinar a análise quantitativa de amostras de leite bovino cru. Os resultados revelaram que o modelo espectral, baseado na aplicação da regressão pelos mínimos quadrados parciais sobre os espectros Raman, pode oferecer um método não destrutivo, robusto e custo-efetivo para mensuração de componentes do leite, apresentando baixo erro de predição, podendo ser implementado em espectrômetros Raman portáteis visando à medição de amostras no campo em tempo real.

Palavras-chave: infravermelho, diagnóstico, métodos quimiométricos, Savitzky-Golay

Quantification of fat, protein and dry matter in bovine milk using Raman spectroscopy: a possibility of field application

Abstract: Raman and Infrared spectroscopy can provide information about the chemical composition of biological samples. Molecules contributions to the spectrum are proportional to their relative concentrations, providing quantitative diagnostic ability. This study aimed Raman spectroscopy associate to signal preprocessing and chemometrics methods to determine quantitative analysis of raw milk samples. The results suggest that the spectral model, based on partial least squares regression applied to Raman spectra may provide a nondestructive, robust and cost-effective method for bovine milk compounds measurement with low prediction error, and can be implemented in portable Raman spectrometers for field measure of sample with real time results.

Keywords: infrared, diagnostic, chemometrics methods, Savitzky-Golay

Introduction

Optical spectroscopy can provide information about the composition of matter on molecular level. Infrared (IR) and Raman spectroscopy can give the most detailed information about the chemical composition of biological samples for biomedical applications. Raman spectroscopy, like IR, is a non-destructive analytical technique which shows a unique spectral fingerprint and do not require sample preparation. When the light interacts with a molecule, it can be scattered or absorbed. While IR spectroscopy is based on absorption, the Raman spectroscopy is based on inelastic scattering (Raman effect) by the molecule. Raman scattered photon quantitatively reflects the sample chemical composition and can be collected by a spectrometer and displayed as a spectrum. The spectrum consists of a fingerprint were the molecular composition can be determined. The molecules contributions in the Raman and IR spectrum are proportional to their relative concentrations, providing to Raman spectroscopy the quantitative diagnostic ability (Halon et al., 2000; Ellis et al., 2011).

Several analytical methods have been developed to determine the milk chemical composition. The reference methods used for milk analysis (i.e. Röse-Gottlieb for fat determination, and Kjeldahl for proteins) were replaced by techniques that reduce significantly the cost and time-consuming. Considering the increasing of the global production of milk and milk-derived products, the analytical procedures rationalization would be extremely



advantageous. The IR spectroscopy was adopted for use in the commercialized milk analyzers and became a standard method for the quantification of milk components (Mazurek et al., 2015).

Raman spectroscopy has an advantage over IR spectroscopy, allowing that the samples be acquired through a variety of transparent materials, providing analysis of food products through plastic or glass packaging. Moreover, Raman spectroscopy is able to detect analytes in solutions with minimal water interference. Another limitation of IR spectroscopy is the quite infrared light penetration depth in most organic samples. The combination of the above properties with the possibility of portable instrumentation makes Raman spectroscopy a promising technique for food science (Ellis et al., 2011).

The use of a tool that allows to quantify the milk constituents in the field, could bring numerous advantages, including reduced costs and time in monitoring animal's performance by producers or researchers, possibility of stipulate a minimum acceptable standards, thus reducing the fraud occurrence in the cooling-tank collection, and possibility of contract customization between the dairy industry and milk producers based on the percentage of solids.

The aim of this study was to develop a robust and cost-effective methodology based on Raman spectroscopy to quantify milk fat, protein and dry matter, which allows to be used in the field with results in real time.

Material and Methods

To collect the spectra we used a dispersive Raman spectrometer (model P-1, Lambda Solutions Inc., MA, USA), which uses a diode laser with a wavelength of 830 nm and adjustable power up to 350 mW. The spectrometer uses CCD camera 1300x100 pixels "back-thinned, deep depleted", cooled to -70 °C (Pixis 100, Princeton Instruments Inc., MA, USA). A Raman probe is used to obtain the Raman spectra. The laser power was adjusted to 250 mW and each spectrum was generated from the average of 10 frames with integration time of 2 s per frame.

We collected 20 samples of raw milk from Girolando cows raised at Embrapa Experimental Station (Coronel Pacheco, MG, Brazil). Measurements were performed using 80 µL of each sample and scanned through the Raman probe, in triplicate, and the average was estimated. The spectra of raw data were preprocessing in this sequence: elimination of ray cosmic artifacts, removal the fluorescence background using modified polynomial fit function and de-noise Raman spectra using the Savitzky-Golay smoothing/derivative method (polynomial order = 2, window size = 7).

To construct the model, we separated 13 samples for calibration and 7 samples for validation. To characterize the predictive abilities of calibration models, the root mean square error of calibration/prediction (RMSEC/RMSEP) and the relative standard error of calibration/prediction (RSEC/RSEP) were calculated.

$$\text{RMSEC} = \text{RMSEP} = \sqrt{\frac{\sum_{i=1}^n (C_{\text{prev}_i} - C_{\text{ref}_i})^2}{n}} \quad (\% \text{ w/v}) \quad (1) \quad \text{RSEC} = \text{RSEP} = \sqrt{\frac{\sum_{i=1}^n (C_{\text{prev}_i} - C_{\text{ref}_i})^2}{\sum_{i=1}^n (C_{\text{ref}_i})^2}} \times 100 \quad (\%) \quad (2)$$

Where C_{prev_i} is the concentration of the sample calculated by the spectral model, C_{ref_i} is the concentration determined by the reference method, and n is the number of samples. The RSEC and RSEP errors were calculated for the calibration and validation datasets, respectively. The reference analyses of the milk chemical composition (fat, protein and dry matter) were performed with the Bentley FTS IR milk analyzer (Bentley Instruments, Chaska, MI, USA), at Embrapa Milk Quality Laboratory (Juiz de Fora, MG, Brazil).

Principal components analysis (PCA), principal components regression (PCR), partial last square regression (PLSR) and all calculations were performed using Matlab (version 7.14.0.739, The MathWorks, Natwick, MA, USA).

Results and Discussion

The calibration model was constructed using PLS over the Raman spectra of 13 samples. Spectra regions were selected from 1000.3 to 1010.8 cm^{-1} for protein, 1403.7 to 1470.1 cm^{-1} for fat and either region for dry matter (DM). PLSR was chosen due to its lower RMSEC value compared to the PCR. In this study, the firsts five PLS components (loading vectors) were considered in the PLSR calculation. The model was constructed as follows:

$$C_{\text{cal}_i} = \text{Spectral_region}_i \times \beta \quad (3)$$

Where C_{cal_i} is the concentration calculated from the sample (used to calculate the calibration and validation values), Spectral_region_i is the spectral region selected from sample, and β is a coefficient vector obtained by the PLSR using the samples spectra and the reference concentration of the calibration samples.

The calibration and validation of fat, protein and DM concentration were calculated using Eq. 3. The results of PLS modeling are summarized in the Table 1.



Table 1 – Results of PLS modeling

	Calibration			Validation		
	R ²	RMSEC (% w/v)	RSEC (%)	R ²	RMSEP (% w/v)	RSEP (%)
Fat	0.996	0.0659	1.34	0.996	0.194	4.08
Protein	0.979	0.0594	1.81	0.932	0.110	3.41
Dry Matter	0.997	0.0953	0.72	0.989	0.196	1.51

R² = coefficient of determination; RMSEC = root mean square error of calibration; RSEC = relative standard error of calibration; RMSEP = root mean square error of prediction; RSEP = relative standard error of prediction

In order to evaluate the determination ability of this method for raw milk fat, protein and DM, the results obtained here were compared with other studies. Wu et al. (2008) used short-wave near-infrared (NIR) spectroscopy at 800–1050 nm region and two regression models for the analysis of main compounds in milk powder: PLS and least-squares support vector machine (LS-SVM). The model that showed best results, LS-SVM, presented calibration R² = 0.983 and 0.996, RMSEC = 0.225% w/v and 0.069% w/v, and validation RMSEP = 0.335% w/v and 0.177% w/v for concentration of fat and protein, respectively. Li et al. (2011) studied near infrared diffuse reflectance spectroscopy using a portable short-wave near-infrared spectrometer and PLS regression for calibrations of raw milk fat, protein and DM. The results showed calibration R² = 0.98, 0.95 and 0.98, RMSEC = 0.187% w/v, 0.105% w/v and 0.217% w/v, and validation RMSEP = 0.187% w/v, 0.120% w/v and 0.296% w/v for the concentrations of fat, protein and DM, respectively. Masurek et al. (2015) used FT-Raman spectroscopy and PLS regression to quantify milk fat, protein, carbohydrates and dry matter. The results showed calibration R² = 0.993, 0.985 and 0.994, RSEC = 4.66%, 3.37%, and 2.13% and validation RSEP = 5.3%, 5.6% and 3.4% for the concentrations variation of fat, protein and DM, respectively. Our study showed that the RMSEC and RMSEP values were lower for fat and protein when they were compared to Wu et al. (2008), despite our protein R² to be smaller. Comparing to Li et al. (2011), all of the variables were slightly lower in our study. Compared to Masurek et al. (2015), our study showed lower RSEC and RSEP values, meaning a smaller variation in the calculated concentrations despite our protein R² to be smaller.

The results presented here indicated that PLSR method applied to bovine raw milk can provide a good estimation of the concentrations of milk fat, protein and DM. The model can be improved increasing the number of calibration samples and applying other preprocessing methods. The additional advantage of the proposed method is its ease implementation of embedded software in a portable Raman spectrometer.

Conclusions

The results showed that the model based on Raman spectroscopy and partial least squares regression may provide a new method for a nondestructive and rapid measurement of milk compounds, and can be implemented in portable Raman spectrometers for field measure, with prediction error as low as 0.194% w/v, 0.110% w/v and 0.196% w/v for fat, protein and dry matter, respectively, using the raw milk.

Acknowledgments

L. Silveira Jr. acknowledges FAPESP (Grant no. 2009/01788-5) and CNPq (Grant no. 483026/2010-7).

References

- Ellis, D.I.; Brewster, V.L.; Dunn, W.B.; Allwood W.J.; Golovanov, A.P. and Goodacre, R. 2001. Fingerprinting food: current technologies for the detection of food adulteration and contamination. *Chem. Soc. Rev.* 41: 5706–5727.
- Hanlon, E.B.; Manoharan, R.; Koo, T.W.; Shafer, K.E.; Motz, J.T.; Fitzmaurice, M.; Kramer, J.R.; Itzkan, I.; Dasari, R.R. and Feld, M.S. 2000. Prospects for *in vivo* Raman spectroscopy. *Phys. Med. Biol.* 45: R1–R59.
- Li, X.Y.; Wang, J.H.; Huang, Y.W. and Han, D.H. 2011. Determination of fat, protein and DM in raw milk by portable short-wave near infrared spectrometer. *Guang Pu Xue Yu Guang Pu Fen Xi.* 1:665–668.
- Mazurek, S.; Szostak, R.; Czaja, T. and Zachwieja, A. 2015. Analysis of milk by FT-Raman spectroscopy. *Talanta* 138:285–289.
- Wu, D.; He, Y. and Feng, S. 2008. Short-wave near-infrared spectroscopy analysis of major compounds in milk powder and wavelength assignment. *Analytica Chimica Acta* 610: 232–242.