

## Notas Científicas

### Predicting rapeseed oil content with near-infrared spectroscopy

Roberta Rossato<sup>(1)</sup>, Cássio Egídio Cavenaghi Prete<sup>(2)</sup>, César de Castro<sup>(3)</sup>, Gilberto Omar Tomm<sup>(4)</sup>,  
Rodrigo Santos Leite<sup>(3)</sup>, José Marcos Gontijo Mandarino<sup>(3)</sup>, Pedro Mário de Araújo<sup>(5)</sup>  
and Claudio Guilherme Portela de Carvalho<sup>(3)</sup>

<sup>(1)</sup>Universidade Estadual de Londrina (UEL), Departamento de Química, Rodovia Celso Garcia Cid, Km 380, Campus Universitário, CEP 86051-990 Londrina, PR, Brazil. E-mail: beta.rossato@hotmail.com <sup>(2)</sup>UEL, Departamento de Agronomia. E-mail: cassio@uel.br <sup>(3)</sup>Embrapa Soja, Caixa Postal 231, CEP 86001-970 Londrina, PR, Brazil. E-mail: cesar.castro@embrapa.br, rodrigo.leite@embrapa.br, josemarcos.gontijo@embrapa.br, portela.carvalho@embrapa.br <sup>(4)</sup>Embrapa Trigo, Rodovia BR 285, Km 294, CEP 99001-970 Passo Fundo, RS, Brazil. E-mail: gilberto.tomm@embrapa.br <sup>(5)</sup>Instituto Agronômico do Paraná, Rodovia Celso Garcia Cid, Km 375, CEP 86047-902 Londrina, PR, Brazil. E-mail: pmaraujo@iapar.br

**Abstract** – The objective of this work was to establish a calibration equation and to estimate the efficiency of near-infrared reflectance (NIR) spectroscopy for evaluating rapeseed oil content in Southern Brazil. Spectral data from 124 half-sib families were correlated with oil contents determined by the chemical method. The accuracy of the equation was verified by coefficient of determination ( $R^2$ ) of 0.92, error of calibration (SEC) of 0.78, and error of performance (SEP) of 1.22. The oil content of ten genotypes, which were not included in the calibration with NIR, was similar to the one obtained by the standard chemical method. NIR spectroscopy is adequate to differentiate oil content of rapeseed genotypes.

**Index terms:** *Brassica napus*, NIR, oilseed, soxhlet, spectral analysis.

### Predição do teor de óleo em sementes de canola por espectroscopia de infravermelho próximo

**Resumo** – O objetivo deste trabalho foi estabelecer uma equação de calibração e estimar a eficiência da espectroscopia de reflectância no infravermelho próximo (NIR), para avaliar o teor de óleo em sementes de canola, no Sul do Brasil. Informações espectrais de 124 famílias de meio-irmãos foram correlacionadas com os teores de óleo determinados pelo método químico. A precisão da equação foi verificada por coeficiente de determinação ( $R^2$ ) de 0,92, erro de calibração (SEC) de 0,78 e erro de desempenho (SEP) de 1,22. O teor de óleo de dez genótipos, não incluídos na calibração por NIR, foi similar ao obtido pelo método químico padrão. A espectroscopia NIR é adequada para diferenciar teores de óleo de genótipos de canola.

**Termos para indexação:** *Brassica napus*, NIR, oleaginosa, soxhlet, análise espectral.

The oil content of rapeseed (*Brassica napus* L.) is usually determined with the Soxhlet chemical extraction method (Zenebon et al., 2008). However, the use of these solvents is harmful for the environment and human health. Furthermore, this method is destructive, expensive, and time consuming. Consequently, the Soxhlet method is not suitable for analyzing a large number of samples or for preserving samples for subsequent studies.

Alternately, oil content may be predicted with near-infrared reflectance (NIR) spectroscopy. The use of NIR is fast, inexpensive, nonpolluting, effective, nondestructive, and uses little manpower, allows for the analysis of intact seeds, and does not require hygroscopic stability (Sato, 2002).

NIR spectroscopy is an indirect method that requires equipment calibration. For calibration, NIR spectral data are correlated with data obtained with the reference method, such as the chemical method. However, environmental features, including soil type, temperature, and humidity, can influence spectra absorbance at seed production sites and the resulting calibration equations (Siemens & Daun, 2005). In addition, other factors, such as seed genetic variability, morphology, and coat color, besides the mathematical treatment and equipment model, may influence this estimate (Pandorf & Man, 1990).

The NIR method has been used for various plant species. In rapeseed, the efficient use of this method was demonstrated by Sato et al. (1998) and Siemens

& Daun (2005). However, the calibration curves that were predicted in these studies were generated in other countries that have different edaphoclimatic conditions than Brazil. The present study focused on the edaphoclimatic conditions of Southern Brazil, because rapeseed production mainly occurred in two states of the region, Rio Grande do Sul (61.7%) and Paraná (31.8%) (Companhia Nacional de Abastecimento, 2012).

The objective of this work was to establish a calibration equation and to estimate the efficiency of NIR spectroscopy for evaluating rapeseed oil content in Southern Brazil.

To establish the calibration equation for evaluating rapeseed oil content with NIR spectroscopy, 124 half-sib families grown in the municipalities of Londrina (2008) and Pato Branco (2009), in the state of Paraná, Brazil, by Instituto Agrônômico do Paraná, were used. These families originated from a population of 14 genotype crosses from the rapeseed active germplasm bank of Embrapa. The soil of the Londrina site (23°21'23"S, 51°10'W, at 585 m altitude) is a eutrophic Red Nitrosol (Alfisol), and the one of the Pato Branco site (26°13'43"S, 52°40'10"W, at 760 m altitude) is a dystrophic humic Rhodic Hapludox.

The standard samples (half-sib families) were previously cleaned and homogenized to improve the spectra resolution. The spectra absorbance (Log 1/R) was obtained between wavelengths of 1,100 and 2,500 nm. The spectra were obtained from intact seeds (with coat) and were positioned on the integrating sphere of the NIR analyzer Antaris II (Thermo Fisher Scientific Brasil Instrumentos e Processo Ltda., São Paulo, SP, Brazil). The analyzer was adjusted to obtain 32 scans from each sample with a resolution of 4 cm<sup>-1</sup>.

NIR calibration was based on the regression equation. The regression equation was established by using the partial least squares method to correct for light scattering, which occurred when obtaining spectra, and by using the pre-processing standard normal variation technique. NIR spectral data were correlated with the oil levels obtained in triplicate with the chemical standard reference method, with N-hexane as the extraction solvent in a Soxhlet apparatus (Zenebon et al., 2008). The oil contents were expressed in dry basis. The first derivative and a Savitzky-Golay filter were used to reduce and smooth random signals and noise with an 11 point interval and a polynomial order

of 1. Seven principal components were used to explain the mathematical model.

The accuracy of the equation and the validity of the calibration curve were tested with the TQ Analyst software (Thermo Fisher Scientific Brasil Instrumentos e Processo Ltda., São Paulo, SP, Brazil) to estimate the coefficient of determination (R<sup>2</sup>), the standard error of calibration (SEC), and the standard error of performance (SEP).

The chemically determined oil contents of the ten rapeseed hybrids introduced in Brazil were compared with the oil contents obtained by the NIR calibration curve. The estimated oil contents of the hybrids were obtained by NIR from seeds of two field experiments conducted in the municipalities of Encruzilhada do Sul (2009 harvest season) and Três de Maio (2010 harvest season), in the state of Rio Grande do Sul. These NIR values were not included in the calibration. The soil of the Encruzilhada do Sul site (30°32'28"S, 52°31'19"W, at 432 m altitude) is a dystrophic Red Yellow Ultisol. In contrast, the soil of the Três de Maio site (27°46'24"S, 54°14'24"W, at 343 m altitude) is a eutrophic Haplic Cambisol.

The assays were conducted in a completely randomized block design, with three replicates, in a split-plot arrangement. The plots consisted of hybrids and environments (sites and specific years), and the subplots of two methods for evaluating oil content. Analysis of variance was performed with the Genes software (Cruz, 2006).

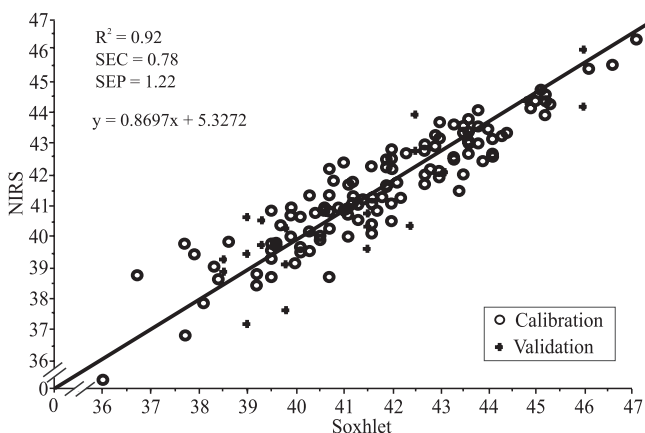
The infrared spectrum region with wavelengths between 1,386 and 2,228 nm was established to determine the calibration equation of the NIR device for evaluating rapeseed oil content. Based on this region, the R<sup>2</sup> of the equation was 0.92, the SEC was 0.78, and the SEP was 1.22 (Figure 1).

Four regions were established to estimate the equation, according to Hourant et al. (2000), including regions B (1,300–1,600 nm), C (1,600–1,850 nm), D (1,850–2,050 nm), and E (2,050–2,230 nm). Based on the first derivative, the highest absorption peaks occurred in regions C, D, and F (2,230–2,500 nm). Region C contains bands in the spectrum that are characteristic of the first C-H stretching overtones of methylene (CH<sub>2</sub>), methyl (CH<sub>3</sub>), and ethenyl (-CH=CH-) groups. Region D contains bands that are characteristic of a combination of C-H stretching bands from ethenyl and OH (hydroxyl) groups in water (Hourant et al., 2000). Despite the presence of

absorption peaks, regions A (1,100–1,300 nm) and F (2,230–2,500) were not considered further in the present study because they were less similar to the calibration equation than the other regions ( $R^2 = 0.87$ , SEC = 0.98, SEP = 1.31; and  $R^2 = 0.93$ , SEC = 0.71, SEP = 1.20, respectively).

The spectral region between 1,880 and 1,930 nm was associated with water content rather than oil content. This is a result of the stretching vibrations of the functional groups that are more intense in this region, due to atoms with different electronegativity values, such as hydroxyl atoms (Westad et al., 2008). Therefore, because of its association with water, that region may not be adequate for establishing the calibration equation. This result was previously obtained from a study that analyzed the oil contents of sunflowers grown in South-Central Brazil (Grunvald, 2012). However, in the present study, the inclusion of the 1,880 and 1,930 nm region increased the estimation accuracy of the equation that only included regions B, C, and E ( $R^2 = 0.85$ , SEC = 1.07, SEP = 1.22).

To estimate the equation shown in Figure 1, 248 standard samples from a population originated from 14 genotype crosses were used. The oil content of the samples was between 36 and 47%. This variability is typically observed in *Brassica napus*. The accuracy of the equation in Figure 1 (SEC = 0.78 and SEP = 1.22) was similar to the one obtained by Tkachuk (1981) and Hom et al. (2007). However, this accuracy was greater than the one found by Velasco et al. (1999), mainly because these authors used other species from the Brassicaceae family to estimate the equation



**Figure 1.** Calibration correlation of the near-infrared device with the oil content values (%) of intact rapeseed (*Brassica napus*) seeds.

(SEC = 1.85 and SEP = 1.92). Therefore, the greater genetic variability between the different species potentially influenced the final estimates.

When using seed samples with different coat colors, it was possible to estimate an equation with satisfactory accuracy from a low number of samples (248 samples) (Figure 1). Moreover, coat color minimally affected the estimated equations for rapeseed, as observed by Tkachuk (1981), and for soybean (*Glycine max* L.) by Wang et al. (2002). Greater variability in coat color, as found in sunflower, generally decreases the estimation accuracy ( $R^2=0.87$ , SEC=2.79, SEP=2.57) (Grunvald, 2012). Therefore, a greater number of samples was required (901 samples) for the estimation. In addition to the influences of coat color and uniform morphology in rapeseed, a lower number of sampled environments (Londrina and Pato Branco, PR, Brazil, in 2008 and 2009, respectively) and a lower genetic variability in the samples from the population with 14 genotypes resulted in an accurate equation. These results differed from the ones obtained by Grunvald (2012), who used sunflower seed samples with naturally nonuniform morphology that originated from 88 hybrids and 116 lines. These lines were obtained from various sites in South-Central Brazil.

The SEP values were used to assess the accuracy of the equation and to validate the calibration curve, whereas the oil content values obtained by the standard chemical method were compared with the NIR spectroscopy results. The analyses of variance showed that no significant differences occurred between the two methods or between the interaction methods and hybrids (Table 1). This result is an indicator of the similarity between the methods used to compare the mean oil content of the hybrids.

In the present study, the genotypes that were used to assess the accuracy of the calibration equation were different from the ones used to construct the calibration curve. Furthermore, rapeseed samples were obtained from different environments: the seeds from Paraná were used for calibration, and the seeds from Rio Grande do Sul were used for validation. Still, the results of the NIR spectroscopy analysis were similar to the ones of the standard chemical method. This similarity and the  $R^2$ , SEC, and SEP values indicate that the evaluation of seeds with NIR spectroscopy is adequate for determining oil content variability in rapeseed genotypes in Southern Brazil.

**Table 1.** Analysis of variance of the oil contents of the rapeseed (*Brassica napus*) hybrid obtained with near-infrared reflectance spectroscopy (NIR) and the standard chemical method.

Source of variation	Degrees of freedom	Mean square
Blocks (B) / Environments (E)	4	20.84
Genotypes (G)	9	49.13*
E	1	313.63**
G x E	9	16.69**
Error (a)	36	19.48
Methods (M)	1	1.32 <sup>ns</sup>
M x E	1	1.77*
M x G	9	0.73 <sup>ns</sup>
M x G x E	9	0.73 <sup>ns</sup>
Error (b)	40	0.42
Overall mean	40.00	-
CV (%) <sup>(1)</sup>	11.03	-
CV (%) <sup>(2)</sup>	1.62	-

<sup>(1)</sup>Coefficient of variation referring to the plot. <sup>(2)</sup>Coefficient of variation referring to the subplot. <sup>ns</sup>Nonsignificant. \* and \*\*Significant at 5 and 1% probability, respectively, by the F test.

## References

- COMPANHIA NACIONAL DE ABASTECIMENTO. **Acompanhamento da safra brasileira – grãos – safra 2011/2012 – décimo levantamento – julho/2012.** Brasília: Conab, 2012. 29p. Disponível em: <[http://www.conab.gov.br/OlalaCMS/uploads/arquivos/12\\_07\\_05\\_08\\_41\\_20\\_boletim\\_graos\\_-\\_10julho\\_2012.pdf](http://www.conab.gov.br/OlalaCMS/uploads/arquivos/12_07_05_08_41_20_boletim_graos_-_10julho_2012.pdf)>. Acesso em: 25 jul. 2012.
- CRUZ, C.D. **Programa Genes:** versão Windows: aplicativo computacional em genética e estatística. Viçosa: Ed. da UFV, 2006. 285p.
- GRUNVALD, A.K. **Influência da temperatura na composição dos ácidos graxos e caracterização da qualidade e do teor de óleo de genótipos de girassol por espectroscopia de NIR.** 2012. 74p. Tese (Doutorado) – Universidade Estadual de Maringá, Maringá.
- HOM, N.H.; BECKER, H.C.; MÖLLERS, C. Non-destructive analysis of rapeseed quality by NIRS of small seed samples and single seeds. **Euphytica**, v.153, p.27-34, 2007. DOI: 10.1007/s10681-006-9195-3.
- HOURLANT, P.; BAETEN, V.; MORALES, M.T.; MEURENS, M.; APARICIO, R. Oil and fat classification by selected bands of near-infrared spectroscopy. **Applied Spectroscopy**, v.54, p.1168-1174, 2000. DOI: 10.1366/0003702001950733.
- PANFORD, J.A.; MAN, J.M. Determination of oil content of seeds by NIR: influence of fatty acid composition on wavelength selection. **Journal of the American Oil Chemists' Society**, v.67, p.473-482, 1990.
- SATO, T. New estimation method for fatty acid composition in oil using near infrared spectroscopy. **Bioscience, Biotechnology, and Biochemistry**, v.66, p.2543-2548, 2002. DOI: 10.1271/bbb.66.2543.
- SATO, T.; UEZONO, I.; MORISHITA, T.; TETSUKA, T. Nondestructive estimation of fatty acid composition in seeds of *Brassica napus* L. by near-infrared spectroscopy. **Journal of the American Oil Chemists' Society**, v.75, p.1877-1881, 1998.
- SIEMENS, B.J.; DAUN, J.K. Determination of the fatty acid composition of canola, flax, and solin by near-infrared spectroscopy. **Journal of the American Oil Chemists' Society**, v.82, p.153-157, 2005.
- TKACHUK, R. Oil and protein analysis of whole rapeseed kernels by near infrared reflectance spectroscopy. **Journal of the American Oil Chemists' Society**, v.58, p.819-822, 1981.
- VELASCO, L.; GOFFMAN, F.D.; BECKER, H.C. Development of calibration equations to predict oil content and fatty acid composition in Brassicaceae germplasm by near-infrared reflectance spectroscopy. **Journal of the American Oil Chemists' Society**, v.76, p.25-30, 1999.
- WANG, D.; RAM, M.S.; DOWELL, F.E. Classification of damaged soybean seeds using near-infrared spectroscopy. **American Society of Agricultural Engineers**, v.45, p.1943-1948, 2002.
- WESTAD, F.; SCHMIDT, A.; KERMIT, M. Incorporating chemical band-assignment in near infrared spectroscopy regression models. **Journal of Near Infrared Spectroscopy**, v.16, p.265-273, 2008. DOI: 10.1255/jnirs.786.
- ZENEBO, O.; PASCUET, N.S.; TIGLEA, P. (Coord.). Procedimentos e determinações gerais. In: ZENEBO, O.; PASCUET, N.S.; TIGLEA, P. (Coord.). **Métodos físico-químicos para análise de alimentos.** 4.ed. São Paulo: Instituto Adolfo Lutz, 2008. p.83-158.

Received on October 12, 2012 and accepted on November 26, 2013