Journal of Oilseed Brassica, 2012 3(2): 88-98

Determination of oil, protein and moisture content in whole seeds of three oleiferous *Brassica* species using near-infrared reflectance spectroscopy

Deepak Prem[1], Kadambari Gupta[2], Gautam Sarkar[3], Abha Agnihotri[4*]

Bayer Bioscience Pvt. Ltd., Delta Square, Sector-25, MG Road, Gurgaon, Haryana, India
[2] Topworth Group of Companies, A-21, 2nd Floor, Green Park Main, New Delhi- 110016, India
[3] Plant Biotechnology Division, TERI, Habitat Place, Lodhi Road, New Delhi- 110 003, India
[4] Centre for Agricultural Biotechnology, Amity Institute of Microbial Technology, Block E-3, Amity University Uttar Pradesh, Sector 125, Noida – 201303 (NCR), India
**Corresponding author: aagnihotri@amity.edu*

Abstract

The near infrared reflectance spectroscopy (NIRS) was used for non destructive estimation of biochemical components in seed for three *Brassica* species, irrespective of harvest season and seed coat colour heterogeneity of the samples. Self – pollinated seeds of 132 yellow or brown seeded genotypes of *B. juncea*, *B. napus* and *B. rapa* were analysed through wet chemistry and NIRS scanning. The nuclear magnetic resonance (NMR), oven drying, and Kjeldahl nitrogen estimation were used as reference methods for oil, moisture and protein content, respectively. Considerable variations were observed between and within the species for these constituents. The original Log 1/R seed spectrums showed high absorption in the range of 1124- 1300, 1592- 1868, 1920- 2104, and 2208- 2356 nm wavelength. A considerable base line and absorption intensity variations were observed for between and within the species respectively. The 2nd derivative transformation with gap = 2, segment size = 2 nm was established as most appropriate for calibration development. The best combined species calibrations resulted in the best calibration statistics for the estimation of the oil (SECV = 1.30, R^2 = 0.94, SEC = 1.18), moisture (SECV = 0.12, R^2 = 0.87, SEC = 0.39) and protein content (SECV = 12.19, R^2 = 0.91, SEC = 2.18). A high linear correlation ($r^2 > 0.85$) and non – significant difference ($P > 0.05$) was observed between the NIRS and the reference method estimations, proving the robustness of the calibration equations.

Keywords: *De –oiled meal protein, spectral data, oil content, moisture percentage*

Introduction

The seeds of oleiferous *Brassica* crops are valued for their oil content and the de-oiled meal is a rich source of protein valued as animal feed (Anonymous, 1995, 2000). The seed moisture content has a bearing on the storage and commercial value of the seed (Bandel *et al*., 1991, Barnes *et al*, 1989). The primary methods used for the estimation of these three biochemical components, namely, solvent extraction, Kjeldahl nitrogen estimation and oven drying are unsuitable due to seed destructive procedures. Therefore, rapid, cost effective, non-destructive and reliable multi-trait estimation methods are a prerequisite for selection of superior quality plants in crop improvement programs. Near-infrared Spectroscopy (NIRS) is a rapid, reliable and non-destructive alternative procedure for estimation of such quality components (Batten, 1998). It provides a main advantage of simultaneous evaluation of multiple seed components such as oil, protein, glucosinolate content (Cowe, and McNicols, 1985, Dardanne, 1996), fatty acid composition (Daun and Williams, 1995) and seed colour (Daun *et al*., 1994, Downey and Rimmer, 1993) without the need of sample grinding or chemical alterations. However, much work utilizing this versatile tool has been done on *B. napus*, and information on the development of joint calibrations for multi-trait for more than one species is limited (Font *et al*., 1999). The joint species calibrations are valuable to screen/ evaluate commercial

harvests in countries like India, Pakistan, Australia and Canada, where more than one species of *Brassica* is cultivated, and also for screening of inter-specific hybrid progenies. The seed coat colour also influences the absorbance or reflectance of the NIR. Thus expectable calibrations for any seed component should contain both yellow and brown seeded genotypes (Daun *et al*., 1994). Apart from these, the environmental condition under which the crop is grown has a profound influence on the various seed components. Thus, calibration sets composed of seeds obtained from different environments (year of harvest) for the same genotype are also essential to take into account the environmental variations (Garrido-Varo *et al*., 1998). The present study was undertaken to access the accuracy of NIRS to estimate the oil, moisture and de – oiled meal protein content from whole seeds of *B. juncea*, *B. napus* and *B. rapa* genotypes and to develop calibrations that are independent of seed coat colour, year of harvest and species to estimate the three components.

Material and Methods

Plant material

The seeds of various genotypes of the three *Brassica* species were obtained from self-pollinated populations maintained at the field research station of the oilseed improvement group, The Energy and Resources Institute (TERI), New Delhi, India. These included yellow or brown seeded Indian and exotic genotypes. The seed samples for genotypes were harvested from one/ two consecutive seasons (Table 1). The cleaned seed samples, free from inert material, soil or plant debris constituted the calibration set.

Table 1. Details of Brassica germplasm used for development of calibration equations

Species	Germplasm and its origin				Seed coat colour (No. of accessions)*			Year wise seed samples (no. of seed samples)**		Grand Total
	Origin	No. of cultivars	No. of breeder germplasm collections (Indian)	Total No. of geno-types	Yellow seeded	Brown seeded	Yellow and brown mixed seeds	Two years	One year	Total no. of seed samples used for calibration development
B. juncea	Indian	8	30	55	7	43	5	22	33	77
	Exotic	17								
B. napus	Indian	6	40	54	0	54	0	12	42	66
	Exotic	8								
B. rapa	Indian	10	5	23	16	7	0	3	20	26
	Exotic	8								

* Accessions with either yellow or brown or mixed seed coat colour out of the total no. of genotypes for eg 7 genotypes out of total 55 have yellow seed coat colour for the *B. juncea* samples.

** Genotypes with seed samples of two or one years harvest eg. Out of the total 55 of *B. juncea* seed samples, the seeds of 22 genotypes were available for two years and for the rest of the genotypes, the seed samples were available for one year.

Wet chemistry analysis (reference methods)

The moisture content (% seed weight) was determined by oven drying as per Hanson (Geladi *et al*., 1985) for which the seed samples weighing 500 to 800 mg were kept for 20 to 24 hrs at 80 ^{0}C. The protein content (% de-oiled meal weight) was estimated by determining the Nitrogen content using the micro Kjeldahl N distillation and subsequent multiplication of the nitrogen content value by 6.25, the universally accepted protein content estimation factor (Golbiowski *et al*., 2005). The oil content (% seed weight) was estimated by solvent extraction (Golebiowski, 2004) and by low resolution pulsed wave Nuclear Magnetic Resonance (NMR) using a Bruker PC-20 (20 mhz) analyser[16*]. Oil content from both the reference methods for various *Brassica* samples in two replicates was compared by single factor analysis of variance (ANOVA) [17]. The standard error of laboratory (SEL) was also calculated for the two methods used as per Windham et al[18]: $[3_i\ [3(X_{ij} - X_j)^2/ (R-1)]/ N]^{0.5}$ where X_{ij} is the $_j$th replicate of an $_i$th sample, X_j is the mean value of all replicates of the $_i$th samples by both reference methods, R is the number of replicates and N is the number of samples. Based upon the statistical comparison of the two methods, NMR was chosen as the reference method for calibration development. The five replicate reference method data for each seed sample was used for chemo-metric purposes.

NIR scanning

The samples were scanned on a monochromator NIRSystems model 5000 (NIRSystems, Inc., Silver Springs, USA) equipped with a sample transport module. The instrument was operated at a constant temperature (25 ± 2 ^{o}C) with 40 – 45 % relative humidity. The standard sample cup (part number IH-0307 and IH-0308) was fitted with a special adapter (part number IH- 0337) to reduce the sample volume to ¼ of the filled sample cup, thus reducing the sample size to approximately 300 mg of intact seed. The seed samples were acclimatized for about 24 hrs, to the ambient conditions of instrument use mentioned above, prior to NIR scanning. Five different, 300 mg seed sample spectrums were obtained for each seed sample (grand total, Table 1), and each reflectance spectrum (log 1/R) was obtained by averaging 32 scans from 1100 to 2500 nm recorded at 2 nm intervals with smoothing (R refers to reflectance), using air (empty sample cup) as the reference. Four types of spectral data sets were obtained, three corresponding to the three *Brassica* species and one for all the species combined.

Calibration and validation procedures

The chemometric software NSAS™ version 3 (NIRSystems Inc. Silver Springs, USA) was used for spectral manipulation and calibration development. The original reflectance spectrums were corrected prior to calibration by applying 2nd derivative transformation. For scatter and base line shift correction, various gap width and segment sizes [2,0; 2,2; 2,3; 2,4 and 3,3 respectively] were used for the 2nd derivative transformation of the log (1/ R) spectrums. These spectral manipulations were first attempted for the *B. juncea* samples only. The Partial Least Square (PLS) calibrations for various modified 2nd derivative spectral data and reference method data for oil and moisture content were made using the spectral information from 1100 to 2500 nm. Cross validation was used to avoid over fitting of data. The Standard Error of Cross Validation (SECV), Coefficient of determination (R^2) and the Standard Error of Calibration (SEC) were used to compare the validity of the calibration derived from various 2nd derivative spectral treatments (Batten, 1998). The most suitable spectral manipulation was then used for transformation of the *B. napus*, *B. Rapa, B. juncea* and combined species spectral data. Further, calibrations using PLS regression were made for the three species for oil, moisture and de-oiled meal protein content. The various calibration equations developed for the three seed components, from spectral data of different species and combined species spectral data, were compared using the SECV, R^2, SEC, the Standard Error of Performance (SEP) and the Standard Deviation (SD)[19].

Validation in terms of routine NIRS analysis followed by reference method confirmation of the selected best calibrations was done using 30 unknown samples (samples not included in the calibration set)

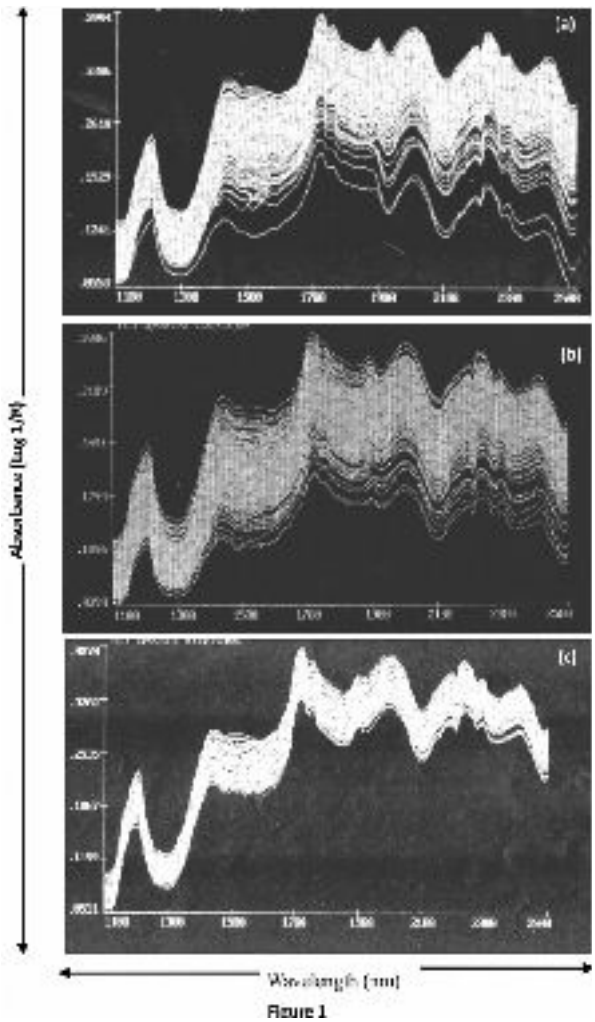
comprising of 10 genotypes of each of the three species. The r^2, SEC and single factor analysis of variance were used for comparison of the reference method and the NIRS results.

Results and Discussion

Spectrum attributes and development of calibrations equations

The original Log1/R seed spectrums for the three species showed sharp rise in absorption readings in four segments; 1124 to 1300 nm, 1592 to 1868 nm, 1920 to 2104 nm and 2208 to 2356 nm (Figure 1). However, the peaks for absorption readings were diffused at longer wavelengths specifically 2104 to 2500 nm. A considerable base line shift was also observed between the spectrums of the three species. The four high absorption regions (Figure 1) correspond closely to the high NIR absorption regions of *B. napus* (Canola) oil *viz*. 1150 to 1250 nm, 1650 to 1800 nm and 2250 to 2380 nm (Hansen, 1985) and are in agreement with various other reports that have described the location of NIR regions specific to vegetable oils (Kumar *et al*., 2003, Mika *et al*., 2003, Miller *et al*., 1962). Although theoretically, the 1100 to 1800 nm region carries sufficient information as it covers the protein (1735 nm) and water (1450 nm) absorption region (Pallot *et al*., 1999, Panford and deMan, 1990, Rittiron *et al*., 2004). However, such information related specifically to *Brassica* seed protein and moisture corresponding NIR absorption is not available.

Figure 1: The original Log 1/R spectra of the whole seeds of *B. juncea*, *B. napus* and *B. rapa*. (a) *B. juncea* seed spectra for 77 samples; total 385 (77×5 replicate) spectra displayed (b) *B. napus* seed spectra for 66 samples; total 325 spectra displayed (c) *B. rapa* seed spectra for 26 samples; total 130 spectra displayed



The oil signal in *Brassica* seeds is considerably weak and in several reports on *Brassica* seed analysis using NIRS, the variability of NIR absorption spectrum within a species, has been related to difference in growing season and site (Hansen,1985) and due to the spectral characteristics of the sample itself (Sato *et al*., 1991, Shenk and Westerhause, 1993). Along with this the spectral properties of the seed are governed by several complex interactions between various factors such as plant genotype, growing environment and agronomic practices (Tiwari *et al*., 1974). This situation warrants the use of full spectrum information for NIRS analysis of biochemical components in *Brassica* seeds, and the need for signal improvement using mathematical procedures such as the 2nd derivative transformations.

2nd derivative transformation is the most exploited spectral data correction tool used for calibration development for various seed components in brassicas. However, the choice of the mathematical treatment (gap width and segment size) is largely dependent on the chemo-metric software being used for regression and the seed

component for which the calibration is being developed (Van Deynze and Pauls, 1994, Velasco and Moellers, 1998, Velasco and Mollers, 2002, Velasco *et al*., 1996, Velasco *et al*., 1995). Selection of specific wavelengths is the common approach till date for calibration development and these calibrations have been reported to be relatively accurate (Velasco *et al*., 1999a, 1998). Alternately, it has been reported that full spectrum information can successfully produce accurate calibrations for several seed components (Velasco *et al*., 1999b, Velasco *et al*., 1998b). In the present study, five 2nd derivative transformations with varying gap width and segment size were used for signal improvement specifically base line, noise and scatter correction (Figure 2), and these resulted in a significant reduction in spectral variation within the species and between species. Such effective reduction in the within – class variation has also been reported earlier by using similar spectral manipulations (Tiwari *et al*., 1974, Velasco *et al*., 1997, Wang *et al*., 1999, Welle *et al*., 2007). However, the transformation of original spectral data may also give rise to artefacts that may lead to misinterpretation of the original absorption values[42]. In this situation the use of different treatment transformed spectral data (for eg. varying the gap width and segment size) for regression analysis with the wet chemistry or reference method data is perhaps the only viable option for choosing the correct spectral manipulation to be applied for accurate routine analysis.

Figure 2: Second derivative transformation of the original Log 1/R spectrum.
a) 2nd derivative transformed *B. juncea* seed spectrums (gap width = 2, segment size = 2 nm)
b) 2nd derivative transformed seed spectrums for the three species combined (gap width = 2, segment size = 2 nm), total 840 spectrums displayed.

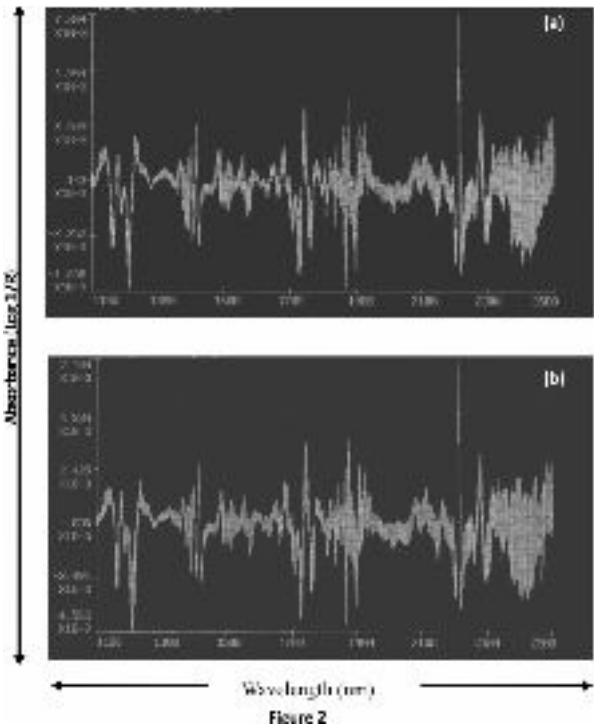


Figure 2

A comparison of the solvent extraction method and the NMR based oil content estimation was done for the seed samples of the three *Brassica* species. The single factor analysis of variance revealed that for the three species, there was no significant difference between the oil content values obtained by the two methods (based on F-test for equality of variance: P> 0.05). The SEL for oil content estimated by the two methods was also low for each of the three species (*B. juncea* = 1.63 %, *B. napus* = 0.59 % and *B. rapa* = 0.48 %). For this reason, NMR was chosen as the reference method for development of calibrations.

Various 2nd derivative spectral manipulations were evaluated for development of reliable calibrations, initially for oil and moisture content estimation in *B. juncea* samples. The cross validation statistics for oil content estimation implicated that though the SECV (2.26 %) was higher for 2, 2 (gap, segment size) 2nd derivative transformed spectral data, the R^2 was higher (0.91) and the SEC was lower (1.4 %) as compared to the rest of the spectral manipulations. Similarly for moisture content also the same 2nd derivative spectral treatment (2, 2: gap, segment size) was observed to be most appropriate (Table 2). Based on these observations it was concluded that calibrations based on the 2nd derivative spectral transformation using segment size of two and gap width of 2 nm was most appropriate for reliable estimation of these seed constituents. Therefore, this transformation was then applied to the original log (1/R) spectrum of the remaining two

Table 2. Calibration statistics for oil and moisture content of *B. juncea* samples using various mathematical treatments for 2nd derivative transformation of 385 seed samples spectra with the reference method values of oil (ranging from 21.9 to 45.9%) and moisture content (ranging from 5.2 to 9.1%).

Mathematical treatment		SECV Ú		R^{2} €		SEC ›	
Gap width (nm)	Segment size (number)	Oil	Moisture	Oil	Moisture	Oil	Moisture
2	0	1.57	0.50	0.79	0.60	2.11	0.55
2	2	2.26	0.16	0.91	0.83	1.40	0.37
2	3	1.70	0.14	0.88	0.70	1.61	0.43
2	4	1.70	0.16	0.88	0.76	1.61	0.43
3	3	1.63	0.11	0.83	0.65	1.92	0.50

1.

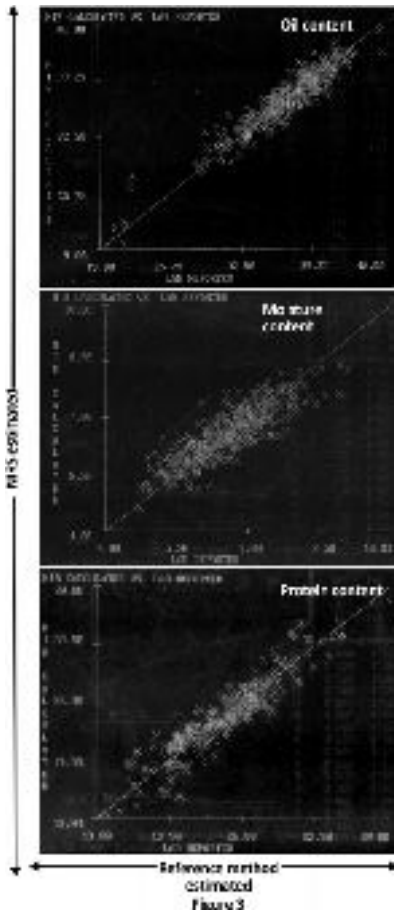
2.

3.SECV = Standard error of Cross Validation.

4. SEC = Standard error of Calibration

5. R^2 = Coefficient of determination

Figure 3: Reference method Vs. NIRS oil, moisture and protein content estimated by the *B. juncea*, *B. rapa* and *B. napus* combined spectra calibrations.



species and the joint spectral data file of the three species. The PLS calibrations were then worked out for the three seed constituents (Figure 3).

Oil content estimation

The *B. rapa* calibration equation for oil content estimation was observed to be superior as compared to the rest of the two species; *B. juncea* and *B. napus*, and the combined calibration for the three species based upon the low SECV (1.48), SEC (0.30) and the high R^2 value (0.99). However, based on SEP/SD ratio comparison, the combined species calibration (SEP/SD=0.89) was observed to be relatively more accurate as compared to the rest of the calibrations for the three species (Table 3).

Velasco et al[41] have reported cross validation statistics for single seed estimation of oil content in *B. napus* (SECV = 1.98, R^2 = 0.94, SEP = 1.87 and SEP/SD = 0.39). Our results of the combined species calibration show similar or better performance of NIR based oil content estimation (SECV = 1.30, R^2 = 0.94, SEP= 1.81 and SEP/SD = 0.89). Similarly for *B. napus* bulk samples (120g), Daun et al[22] have reported R^2 of 0.99 and SEP/SD = 0.19 for the developed calibrations, however, the sample size used for the calibration development and for routine estimation is four times higher as compared to that reported by us (300 mg). Similar to our strategy of including more than one seasons harvest seeds in the calibration set, Greenwood *et al.* (1999) have reported the development of oil content estimation calibrations in *B. napus* (SECV

= 0.77, SEC = 0.61 and R^2 = 0.98) using similar equipment but different chemometric software. The R^2 value (greater than 0.91) for the seed oil content estimation using NIRS in the present study is higher than that reported for rapeseed-mustard (R^2 = 0.87) (Williams and Norris, 1987). To the best of our knowledge, the calibration of NIRS for oil content estimation for *B. rapa* or for a combined calibration for the three species has not been reported as yet.

Moisture content estimation

Similar to the oil content estimation, the calibration equation developed for *B. rapa* seed moisture content was observed to be most accurate (SECV = 0.20, R^2 = 0.98, SEC = 0.11) as compared to the rest of the calibrations developed for other species, or for combined spectral data for the three species. However, the combined species calibration was observed to have lowest SEP/SD ratio (0.88) as compared to the rest of the calibrations (Table 3). Although the seed moisture content is an important contributor to the total seed weight (4 to 9%, our results) and is important from commercial value and storage point of view (Barnes *et al*., 1989), the seed moisture content estimation using NIRS has been reported recently only for *B. napus* (Bandel *et al*., 1991) .

Protein content estimation

Among the three species, the calibration equation developed for *B. juncea* de-oiled meal protein content was observed to be most accurate (SECV = 3.77, R^2 = 0.99, SEC = 0.31) as compared to the rest of the calibrations developed for other species or for combined spectral data for the three species. However, the combined species calibration was

Table 3. Calibration statistics for oil moisture and protein content in Brassica samples with 2nd derivative spectral transformation using gap width of 2nm and segment size of two

Oil content

Species	No. of spectra	Reference method values (%) Mean ± std deviation	Range	SECV	R^2	SEC▲	SEP	SD
B. juncea	385	36.2 ± 3.4	21.9-45.9	2.26	0.91	1.40	1.30	1.38
B. napus	325	37.4 ± 3.1	26.7-42.9	2.63	0.91	1.35	1.21	1.50
B. rapa	130	38.0 ± 4.3	20.5-44.4	1.48	0.99	0.30	0.29	0.28
Combined *B. juncea*, *B. napus* and *B. rapa*	840	36.9 ± 3.5	20.5-45.9	1.30	0.94	1.18	1.81	2.02

Moisture content

Species	No. of spectra	Reference method values (%) Mean ± std deviation	Range	SECV	R^2	SEC	SEP	SD
B. juncea	385	7.0 ± 0.6	5.2-9.1	0.16	0.83	0.37	0.35	0.36
B. napus	325	6.2 ± 0.7	4.2-8.7	0.18	0.85	0.38	0.37	0.38
B. rapa	130	5.7 ± 0.6	4.6-8.8	0.20	0.98	0.11	0.10	0.11
Combined *B. juncea*, *B. napus* and *B. rapa*	840	6.5 ± 0.8	4.2-9.1	0.12	0.87	0.39	0.53	0.61

De-oiled meal protein content

Species	No. of spectra	Reference method values (%) Mean ± std deviation	Range	SECV	R^2	SEC	SEP	SD
B. juncea	385	43.8 ± 2.6	36.0-49.1	3.77	0.99	0.31	0.31	0.29
B. napus	325	39.6 ± 6.8	22.7-47.7	18.25	0.99	0.46	0.41	0.42
B. rapa	130	38.3 ± 4.4	29.0-48.9	7.81	0.79	3.01	2.64	2.67
Combined *B. juncea*, *B. napus* and *B. rapa*	840	40.8 ± 5.4	22.7-49.1	12.19	0.91	2.18	1.99	2.12

1. SECV = Standard error of Cross Validation, ▲ SEC = Standard error of Calibration, R^2 = Coefficient of determination, SEP = Standard error of performance, SD = Standard Deviation.

observed to have lowest SEP/SD ratio (0.93) as compared to the rest of the calibrations (Table 3). Our results are similar to Velasco and Mollers[44] who have reported a SECV = 0.87, R^2 = 0.96, SEC = 0.81 and SEP/SD ratio = 0.28 for single seed estimation of protein content and Daun et al[22] [SEP/SD = 0.15] for bulk samples of *B. napus*. However, our results of combined species calibration present a higher SECV (12.19) primarily because of high protein content variation in the samples of *B. napus* used in the calibration set. The SEP/SD ratio obtained by us (0.93) however, suggests that the developed calibration is practically accurate for protein content estimation.

Based upon the cross validation statistics discussed above, we observed that the calibrations derived using the combined spectral data of the three species were reliable for the estimation of oil, moisture and protein content. The validation of the developed calibrations in terms of routine NIRS analysis results in comparison to respective primary reference methods was performed. Seed samples of ten genotypes of *B. juncea*, *B. napus* and *B. rapa* (other than those included in the calibration set) were used for this purpose. Each genotype was represented by three randomly selected 300 mg seed samples. Therefore, a total of 90 *Brassica* seed samples were first evaluated for the three seed components by routine NIR analysis and then by the respective reference methods. A strong positive correlation was observed between the NIR estimation and reference method estimation based upon r^2 value (greater than 0.85; Table 4). The low SEC and the non- significant difference as observed by the single factor analysis of variance (P > 0.05) for the three seed components; oil, protein and moisture, analysed by primary reference method and NIRS, establishes the reliability of the NIRS calibrations for the estimation of these seed components (Table 4).

To the best of our knowledge this is the first report of a combined multiple species calibration based on a heterogeneous sample set (132 genotypes of the three *Brassica* species, with 2 years harvest, totalling to 840 seed samples having yellow/ brown

Table 4 Real time analysis validation of *B. juncea*, *B. napus* and *B. rapa* unknown seed samples using calibration developed with 2nd derivative (gap width=2mm and segment size=2) spectral transformation of combined species spectra.

Species	Oil content (%)					Moisture content (%)					Decoiled total protein content (%)				
	NIR	NMR	SEC	r^2	P*	NIR	Oven drying	SEC	r^2	P	NIR	Kjeldahl N estimation	SEC	r^2	P
B. juncea	38.14 ± 1.28[a] (36.50-41.13)**	38.06 ± 1.44 (35.91-40.36)	1.45	0.90	0.79	5.90 ± 0.55 (4.80-7.03)	6.06 ± 0.55 (4.80-7.35)	0.65	0.93	0.20	50.02 ± 3.39 (43.10-55.30)	50.06 ± 3.34 (43.56-59.65)	2.07	0.91	0.95
B. napus	37.68 ± 2.78 (33.70-40.25)	37.76 ± 2.73 (33.97-43.07)	1.68	0.85	0.29	5.87 ± 0.52 (4.91-6.80)	5.83 ± 0.46 (4.66-6.75)	0.47	0.92	0.74	53.30 ± 4.16 (45.53-59.14)	52.95 ± 4.17 (45.50-59.60)	1.17	0.98	0.70
B. rapa	39.91 ± 1.45[a] (37.7-42.33)	39.41 ± 1.67 (35.67-42.05)	1.80	0.91	0.77	5.96 ± 0.46 (4.59-6.61)	5.94 ± 0.56 (5.04-6.85)	0.56	0.90	0.45	39.79 ± 1.78 (37.59-42.80)	39.91 ± 1.76 (37.56-42.80)	1.62	0.89	0.90
Linear regression equations for NIR and reference method															
	Oil content					Moisture content					Protein content				
B. juncea	$y = 10.29 + 0.72x$					$y = 3.67 + 0.04x$					$y = 1.03x - 1.8$				
B. napus	$y = 1.01 + 0.99x$					$y = 2.64 + 0.54x$					$y = 6.73 + 0.88x$				
B. rapa	$y = 1.36 - 0.65x$					$y = 8.73 - 0.55x$					$y = 1.07x - 3.18$				

2. * Each data point represents mean – standard deviation recorded from 30 seed samples obtained from 10 different genotypes that had not been included in the calibration set.
3. ** Values in parenthesis represent range.
4. [c] Indicates P- values obtained from single factor analysis of variance for the reference method and the NIRS estimation.

seed coat colour). The strategy used and the NIRS calibration equations developed in this study permit the simultaneous analysis of oil, protein and moisture content of rapeseed- mustard seeds in a non-destructive, fast, cost effective and reliable way. The developed calibrations are of immense utility for non-destructive and efficient screening of seed components for *Brassica* quality breeding programs.

Acknowledgements

We sincerely thank Dr David Hopkins, Consultant NIRS, Berkley, USA for guiding us regarding the nuances of NIRS calibration and critically reviewing the NIRS calibrations that were developed during the course of this study. We are grateful to Dr S. Nagrajan and Dr D. K. Joshi of Nuclear Research Laboratory, Indian Agricultural Research Institute, New Delhi for providing the NMR facility for oil content estimation. We also thank Mr Bir Singh for his help in wet chemistry analysis. The first author is thankful to TERI and TERI University for providing required infrastructure for carrying out this work as a part of his PhD thesis.

References

Anonymous. 1995. AOAC official method 954.01 for protein content (crude) in animal feed – Kjeldahl method. In: *Official methods of analysis of AOAC international* Cunniff P. (ed) 16th edition, Chapter 4 pp 5.

Anonymous. 2000. AOCS official method Am 2-93. Sampling and analysis of vegetable oil source materials-Determination of oil content in oilseeds. AOCS International, CD-ROM http://www.aoac.org/pubs/omacd_rom_revised.htm

Bandel, VA, Mulford, FR, Ritter, RL Kantzes, GJ, and Hellman, JL. 1991. Fact Sheet- 635: Canola production guidelines. University of Maryland and USDA http://www.agnr.umd.edu/MCE/Publications/PDFs/FS635.pdf

Barnes, RJ, Dhanoa, MS and Lister, SJ. 1989. Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. *Appl Spectrosc* **43(5)**: 772-777.

Batten, G.D. 1998. Plant analysis using near infrared reflectance spectroscopy: potential and the limitations. *Aus J Experi Agri* **38**: 697–706.

Cowe, IN and McNicols, JW. 1985 The use of principal components in the analysis of near infrared spectra. *Appl Spectroscopy* **39:** 257–266.

Dardanne, P. 1996. Stability of NIR spectroscopy equations. *NIR News* **7:** 8-9.

Daun, JK and Williams, PC. 1995. Use of NIR spectroscopy to determine quality factors in harvest surveys of canola. In: *Proc. 9th Int. Rapeseed Cong*. 4- 7 July 1995. Cambridge, UK, GCIRC : 864-866.

Daun, JK, Clear, KM and Williams, P. 1994. Comparison of three whole seed near-infrared analyzers for measuring quality components of canola seed. *JAOCS*, **71**(10): 1063-1068.

Downey, RK and. Rimmer, SR. 1993. Agronomic improvement in oilseed *Brassicas*. Adv Agron. Academic Press, Inc.

Font, R, Del Rio, M, Dominguez, J, Fernandez-Martinez, JM and De Haro, A. 1999. Using NIRS for determining glucosiolate content in *Brassica juncea* seed. In: 10th Proc GCIRC International Rapeseed Congress, Sept. 26-29, Canberra, Australia and at http://www.regional.org.au/au/gcirc/1/558.htm#P133_9656.

Garrido-Varo, A, Carrete, R and Fernandez-Cabanas, V. 1998. Use of difference near infrared reflectance spectra to extract relevant information from the spectra of agro-food products. *JNIRS* **6**: 89.

Geladi, P, Mac Dougall, D and Martens, H. 1985. Linearization and scatter-correction for NIR reflectance spectra of meat. *Appl Spectrosc* **39**: 491.

Golbiowski, T, Leong, AS and Panozzo, JF. 2005. Near infrared reflectance spectroscopy of oil in intact canola seed (*Brassica napus* L.). II. Association between principal components and oil content. *J Near Infrared Spectrosc* **13(5):** 255-264.

Golebiowski, T. 2004. Near infrared reflectance spectroscopy of oil in intact canola seed (*Brassica napus*, L.). I. Attributes of the intact seed spectrum. *J Near Infrared Spectrosc* **12:** 325 – 330.

Golebiowski, T, Filip, ML, Harman, CF and Pallot, T. 1995. Routine Analysis of Fatty Acids in

Whole Seed of Canola by NIR. *In Proc. 10th Australian Research Assembly on Brassicas,* September 1995, pp 87-89.

Gomez, K and Gomez, AA. 1984. Statistical Procedures for Agriculture Research, 2nd John W. and Sons (ed). IRRI publication, New York, pp130-186.

Greenwood, CF, Allen, JA, Leong, AS, Pallot, TN, Golder, TM and Golebiowski, T. 1999. An investigation of the stability of NIRS calibration for the analysis of oil content in whole seed canola. In: Proc 10th International Rapeseed Congress, Sept. 26-29, Canberra, Australia; and at http://www.regional.org.au/au/gcirc/1/572b.htm#P226_12531

Guo, Q, Wo, W and Massart, DL. 1999. The robust normal variate transform for pattern recognition with near-infrared data. *Anal Chim Acta* **382**: 87.

Hansen, J. 1985. *Practical Manuals for Gene banks: No. 1*, IBPGR Secretariat, International Board for Plant Genetic Resources, Rome, Italy

Kumar, S, Singh, AK, Kumar, M, Yadav, SK, Chauhan, JS and Kumar, PR. 2003. Standadization of near-infrared spectroscopy (NIRS) for determination of seed oil and protein content in rapeseed-mustard. *J Food Sci Tech* **40**: 306-309.

Mika, V, Tillman, P, Koprna, R, Nerusil. P and Kucera. V. 2003. Fast prediction of quality parameters in whole seeds of oilseed rape (*Brassica napus* L.). *Plant Soil Environ* **49 (4)**: 141-145

Miller, RW, VanEtten, CH, McGrew, C, Wolff, IA and Jones, Q. 1962. Amino acid composition of seed meals from forty-one species of Cruciferae. *J Agric Food Chem* **10**: 426-430.

Pallot, TN, Leong, AS, Allen, JA, Golder, TM, Greenwood, CF and Golebiowski, T. 1999. Precission of fatty acid analysis using near-infrared spectroscopy of whole seed Brassica. In: 10th Proc GCIRC International Rapeseed Congress, Sept. 26-29, Canberra, Australia http://www.regional.org.au/au/gcirc/1/573.htm#TopofPage.

Panford, JA and deMan, JM. 1990 Determination of oil content by NIR: Influence of fatty acid composition on wavelength selection. *JAOCS* **67**(8): 437.

Rittiron, R, Saranwong, S and Kawano, S. 2004. Useful tips for constructing a near infrared-based quality sorting system for single brown-rice kernels. *JNIRS* **12**: 133

Sato, T, Kawano, S and Iwamoto, M. 1991. Near-infrared spectral patterns of fatty acid analysis from fats and oils. *JAOCS* **68**: 827-833.

Shenk, J and Westerhause, MO. 1993. Analysis of Agriculture and food products by Near-infrared Reflectance Spectroscopy. *Infrasoft International*, Port Matilda, USA.

Tiwari, PN, Gambhir, PN and Rajan, TS. 1974. Rapid and non-destructive determination of seed oil by pulsed NMR technique. *JAOCS* **51**:104-109.

Van Deynze, AE and Pauls, KP. 1994. Seed colour assessment in *Brassica napus* using a near-infrared reflectance spectrometer adapted for visible light measurement. *Euphytica* **76**: 45-51.

Velasco, L and Moellers, C. 1998. Nondestructive Assessment of Sinapic Acid Esters in Brassica Species: II. Evaluation of Germplasm and Identification of Phenotypes with Reduced Levels. *Crop Sci* **38**: 1650-1654.

Velasco, L and Mollers, C. 2002. Non-destructive assessment of protein content in single seed of rapeseed (*Brassica napus* L.) by near-infrared reflectance spectroscopy. *Euphytica* **123**: 89-93.

Velasco, L, Fernandez Martinez, JM and De Haro, A. 1996. An efficient method for screening seed colour in Ethiopian mustard using visible reflectance spectroscopy and multivariate analysis. *Euphytica* **90**: 359-363.

Velasco, L, Fernández-Martínez, JM and De Haro, A.1995. The Applicability of NIRS for Estimating Multiple Seed Quality Components in Ethiopian Mustard. *In: Proc. 9th International Rapeseed Congress*, Cambridge UK. pp 867-869

Velasco, L, Goffman Fernando, D and Becker Heiko, C. 1999a. Development of calibration equations to predict oil content and fatty acid composition

in Brassicaceae germplasm by near-infrared reflectance spectroscopy. *JAOCS* **76**: 25-30.

Velasco, L, Matthaus, B and Moellers, C. 1998. Non-destructive assessment of sinapic acid esters in *Brassica* species: I. Analysis by near-infrared reflectance spectroscopy. *Crop Sci* **38**: 1645-1650.

Velasco, L, Mollers, C and Becker Heiko, C. 1999b. Estimation of seed weight, oil content and fatty acid composition in intact single seeds of rapeseed (*Brassica napus*) by near-infrared reflectance spectroscopy. *Euphytica* **106**:79-85.

Velasco, L, Perez-Vich, B and Feranandez-Martinez, JM. 1998. A rapid and simple approach to identify different sunflower oil types by means of near-infrared reflectance spectroscopy. *JAOCS* **75**(12): 1883

Velasco. L, Fernandez Martinez, JM and De Haro, A. 1997. Determination of the fatty acid composition of the oil in intact-seed mustard by near-infrared reflectance spectroscopy. *JAOCS* **74**: 1595-1602.

Wang, D, Dowell, FE and Lacey, RE. 1999. Single wheat kernel color classification using near-infrared reflectance spectra. *Cereal Chem* **76(1)**: 34

Welle, R, Zahle, K, Hildebrand, C, Kraling, K and Greten, W. 2007. Application of near infrared spectroscopy on- combine for canola breeding. *J Near Infrared Spectrosc* **15(5):** 317–325.

Williams, PC. 1987. Variables affecting near–infrared reflectance spectroscopy analysis. In: *Near – infrared technology in agriculture and food industries.* Williams, P. and Norris K. (ed) American Association of Cereal Chemists Inc. St Pauls, MN, USA: 143 –167.

Williams, PC and Norris, KH. 1.987. Near-infrared Technology in the Agriculture and Food Industries. In: *Near-infrared Technology in the Agriculture and Food Industries*. Williams PC and Norris KH (ed) American Association of Cereal Chemists, St Pauls, MN, USA: 246

Windham, WR, Mertens, DR and Barton, FE. 1989. Protocols for NIRS calibration: simple selection and equation development and validation. In: *Near-infrared Spectroscopy (NIRS): Analysis of forage quality. Agric Handbook No. 643.* Martens *GC*, Shenk JS and Barton FE. (eds), USDA-ARS. US Govt Printing Office, Washington: 96-103.