

Calibration of NIRS for the Estimation of Fatty Acids in *Brassica Juncea*

Balraj Kaur¹ · Manjeet K. Sangha¹ · Gurpreet Kaur²

Received: 14 August 2015 / Revised: 15 December 2015 / Accepted: 30 January 2016 / Published online: 2 March 2016
© AOCs 2016

Abstract This study was done to test whether near infrared reflectance spectroscopy (NIRS) could be used as a quick substitute for measuring quality characteristics, i.e., a fatty acid profile over gas liquid chromatography. Near infrared reflectance spectroscopy is a multi-trait technique. In the present study, the fatty acid profiles of seeds of 200 genotypes of *Brassica juncea* were analysed by gas liquid chromatography. Near infrared reflectance spectra of intact seeds of the same samples were collected (400–2500 nm) on a NIR systems Model 6500 spectrophotometer. The spectra were subjected to scanning, mathematical processing, and statistical analysis using Win ISI software. Data were scored to remove redundancy. Spectra were treated as outliers with $H > 3.0$ (global H) and similar samples with $H < 0.6$ (neighbourhood H). Cross validation of the spectra was done using a modified partial least-square method to develop the calibration equation. The calibration equation had good R^2 values for oleic acid ($R^2 = 0.91$), linoleic acid ($R^2 = 0.83$), and erucic acid ($R^2 = 0.88$). The internal validation was done to test the goodness of fit of the developed equation. The equation provided reliable estimations of these traits in internal validation with R^2 values of 0.77, 0.68, 0.81 for the above quoted fatty acids, respectively. The external validation results also showed higher R^2 values for oleic acid (0.89), linoleic acid (0.69), and erucic acid (0.90). The equation was less reliable for linolenic acid, which had R^2 values of 0.53 in cross validation, 0.25

for internal validation, and 0.20 for external validation. The results indicated that NIRS could be used to rapidly determine oleic acid, linoleic acid, and erucic acid in intact *B. juncea* seeds.

Keywords *Brassica juncea* · Near infrared reflectance spectroscopy · Fatty acid · Calibration equation · Validation

Introduction

The importance of oilseed *Brassic*as in human and animal nutrition and also as a crop with industrial application, is well recognized. Over the past three decades the production of oilseed *Brassic*as has increased to an extent that it is now one of the important sources of vegetable oil in the world. The modification of fatty acid composition of seed oil to develop new genotypes having alternative oil characteristics has been an important objective in quality breeding in rapeseed-mustard [1, 2]. The functional and nutritional value of vegetable oil is dependent on the type of constituent fatty acids present in the oil (triacylglycerols). The development of high oleic/low linolenic acid cultivars for use in frying applications and selection for low and very low (zero) content in saturated fatty acids for certain markets is a prerequisite [3]. On the other hand, high erucic acid oilseed rape cultivars have received renewed interest for industrial purposes, e.g., the lubricant industry [4]. The utilization of *Brassic*a oil for fuel has also been commercialized, as a substitute for traditional fossil diesel fuel, to reduce green house gas emissions [5]. The development of cultivars for different target usages demand extensive evaluation of germplasm for traits of interest utilising standard analytical methods, which are labour intensive, time consuming, sometimes expensive, and hazardous too.

✉ Manjeet K. Sangha
manjeet_kaursangha@yahoo.com

¹ Department of Biochemistry, Punjab Agricultural University, Ludhiana 141004, India

² Department of Plant Breeding and Genetics, Punjab Agricultural University, Ludhiana 141004, India

Near-infrared reflectance spectroscopy (NIRS) is an alternative powerful tool for analysis of chemical and physical properties without sample preparation and has been applied for the analysis of quality characteristics in food and agricultural commodities [6–8]. The technology is non-destructive, non-polluting, rapid, accurate, and less labour intensive using simple sample preparation methods [9]. To predict the concentration of various parameters, prediction equations are required and are based on the calibration data sets, which encompass the wide range of sample characteristics previously documented based on wet chemistry techniques. Multivariate calibration techniques are often employed to extract information of importance from NIRS analysis. Standardization of NIRS is mandatory for all the laboratories that install it and require one time accurate wet analysis for development of a calibration equation for simultaneous estimation of various biochemical quality traits. Standardization requires initial labour input, but once standardized, the NIRS system is efficient, fast, and safe. The aim of this study was to develop a calibration equation for estimation of different fatty acids in *Brassica juncea* germplasm for use in various *Brassica* spp. breeding programmes.

Materials and Methods

The plant material (seeds of 200 *B. juncea* genotypes) was procured from Oilseed Section, Department of Plant Breeding and Genetics, PAU, Ludhiana.

Fatty Acid Analysis Using Gas Liquid Chromatography

The selected samples were analysed for fatty acid composition by gas liquid chromatography (GLC) [10] on a M/s Nucon Engineers AIMIL Gas Chromatograph (solid state) model 5700 series equipped with a flame ionization detector fitted with 6 % butane diol succinate (BDS) on Chromosorb WAW/DWCS column, 6 feet in length \times 1/4" outer diameter. For this about 50 mg of ground seed was thoroughly mixed with 1 ml petroleum ether (60–80 °C). To this was added 1.5 ml sodium ethoxide (0.02 M NaOH in 99.5 % of ethanol). The tubes were vortexed and kept at room temperature for 30–40 min. Added 1.5 ml of 8 % sodium chloride solution in water and the tubes were vortexed again. After half an hour the two layers got separated. The upper petroleum ether layer was transferred into another tube and allowed to evaporate and finally dissolved in 10 μ l of petroleum ether and 2 μ l was then injected into GLC using a microsyringe (Hamilton).

Scanning Samples for NIRS Analysis

A NIR Systems Model 6500 spectrophotometer (Foss-NIR Systems, Inc., Silver Spring, MD, USA) equipped with a

transport module was used to perform the NIRS analyses [11]. Intact seeds from samples of *Brassica juncea* were used for wet analysis and were also analyzed on NIRS using a small ring cup and reflectance spectra (log 1/R) from the 400 to 2500 nm that were recorded.

Selecting Samples for NIRS Analysis

The NIRS manipulation for scanning, mathematical processing, and statistical analysis was performed with the WinISI II software (Windows version 4.0, Foss and Infrasoft International LLC, USA). The distance between a sample and its neighbour was measured as the Mahalanobis distance (H distance) called the neighborhood H . The score algorithm ranks spectra according to the Mahalanobis distance (H distance) from the average spectrum, and provides spectral boundaries to eliminate outliers with $H > 3.0$ (global H) and similar samples with $H < 0.6$ (neighbourhood H). Therefore, the final number of samples for calibration and validation was variable based on the cutoff point of H distance, depending on the spectral and chemical variability of samples in the population used for NIRS estimation.

Developing the NIRS Equation

After the removal of outliers by scoring, the calibration set was used to calibrate and cross-validate the equation derived, and internal cross validation (using the same set of samples) was used to avoid overfitting of the equation. The external validation was done to test the goodness of fit of the developed equation using random samples not included in calibration sample set.

Results and Discussion

The range, mean, and standard deviation of the parameters analysed by reference methods are summarised in Table 1 for *B. juncea*. The quality exhibited by the total

Table 1 Fatty acid composition of *Brassica juncea* seeds (wet chemistry data)

Parameters	Range	Mean	SD
Fatty acid profile (%)			
Oleic acid	6.4–49.9	24.8	13.6
Linoleic acid	10.9–40.8	24.8	7.8
Linolenic acid	3.9–36.3	13.0	3.3
Erucic acid	0.4–57.2	32.2	17.8
Total No. of samples	200		

SD standard deviation

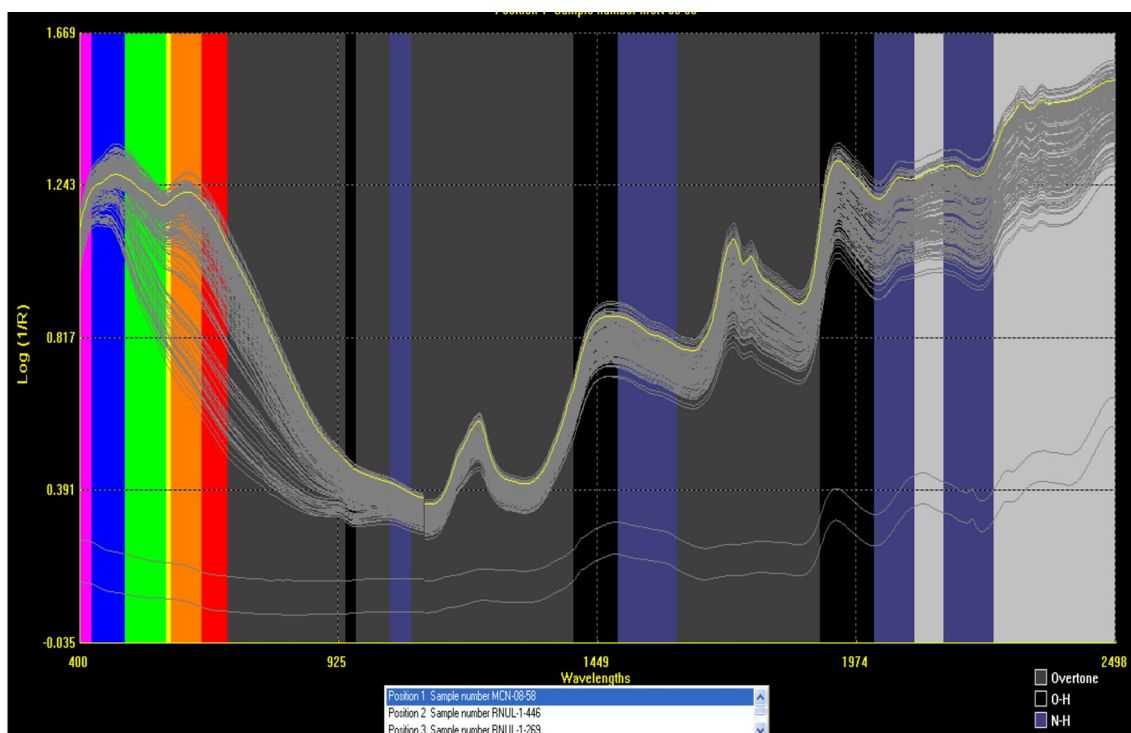


Fig. 1 Near-infrared reflectance (NIR) spectra of intact seeds of *Brassica juncea*

Table 2 Calibration equation statistics for near-infrared reflectance spectroscopy (NIRS) for fatty acid analysis in intact seed samples of *Brassica juncea*

Constituent	N	Mean	SD	SEC	RSQ	SECV	1-VR
Fatty acid profile							
Oleic acid	177	24.31	13.35	3.91	0.91	4.52	0.88
Linoleic acid	180	24.74	7.47	3.02	0.83	3.20	0.81
Linolenic acid	184	12.70	2.21	1.50	0.53	1.77	0.35
Erucic acid	151	31.89	17.41	6.01	0.88	6.76	0.84

N sample number after scoring using cross validation, SD standard deviation, SEC standard error of calibration, RSQ coefficient of determination in calibration, SECV standard error of cross validation, 1-VR coefficient of determination in the cross-validation

Table 3 Internal and external validation statistics for calibration equation for fatty acid analysis in intact seeds of *Brassica juncea*

Fatty acids	Internal validation				External validation	
	Bias	RSQ	SEP(C)	Slope	RSQ	SEP
Oleic acid	0.01	0.77	6.57	0.94	0.89	1.64
Linoleic acid	-0.21	0.68	4.29	0.92	0.69	0.51
Linolenic acid	0.36	0.25	2.96	1.05	0.20	0.98
Erucic acid	0.75	0.81	7.49	0.95	0.90	2.34

RSQ coefficient of determination in calibration, SEP(C) standard error of prediction corrected for bias, SEP standard error of performance

seed samples varied between 6.4–49.9 % in oleic acid, 10.9–40.8 % in linoleic acid, 3.9–36.3 % in linolenic acid, and 0.4–57.2 % in erucic acid. The erucic acid exhibited a

very wide range. The mean value for oleic acid was 24.8 %, linoleic acid 24.8 %, linolenic acid 13.0 %, and erucic acid 32.2 %. Font [12] obtained similar results for *B. juncea*,

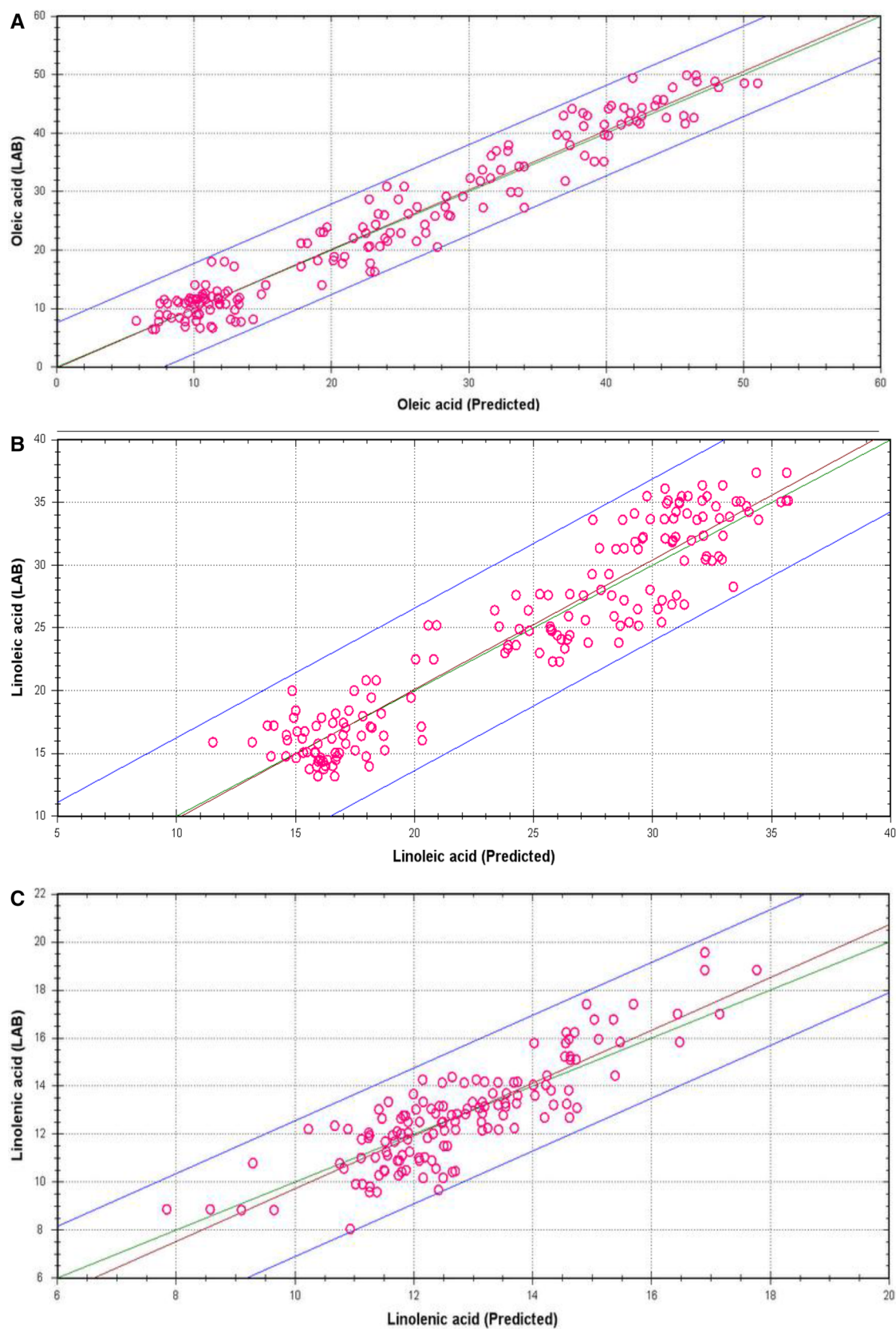


Fig. 2 Correlation plots between predicted values and lab (wet) values for oleic (a), linoleic (b), linolenic (c) and erucic (d) acid (%) for the internal validation of near-infrared reflectance spectroscopy (NIRS) calibration equation for *Brassica juncea*

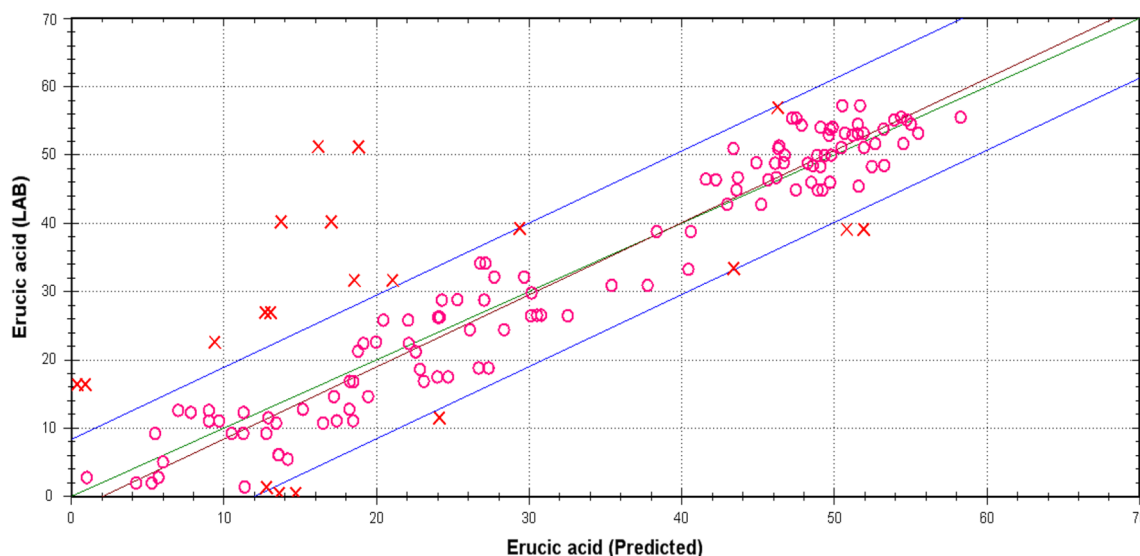


Fig. 2 continued

oleic acid 8.3–50.0 %, linoleic acid 11.2–43.5 %, linolenic acid 5.1–14.0 %, and erucic acid 0.02–56.2 % with mean values of 19.7, 22.8, 9.4, 38.3 %, respectively. The concentrations of the various parameters analyzed by others also reported a good range [13–16].

Reflectance Spectra of *B. juncea* Genotypes and Their Scoring

All the 200 seed samples of *B. juncea* were scanned in ISI SCAN in a NIR spectrophotometer System. The instrument was operated at a constant temperature of 23 ± 2 °C. The spectra were collected over the range 400–2498 nm including the visible and NIR regions. The NIR reflectance spectra of intact seed samples are shown in Fig. 1. The main absorption bands were observed at 1208 nm related to C–H stretching 2nd overtone ($-\text{CH}_2$), 1496 nm related to C–H stretching 1st overtone, 1724 nm related to C–O (oil) and C–H stretching 1st overtone ($-\text{CH}_2$), 1936 nm related to O–H bending 2nd overtone (water), and 2308 nm related to C–H bending 2nd overtone (oil). The information about functional groups in the spectrum were searched from Win ISI software. The overall spectrum showed strong absorption bands related to oil and water, and were similar to those for other oil crops such as perilla, peanut, soybean, and sesame, especially, in the near-infrared region [17].

The reflectance spectra along with laboratory values were then transferred from ISI SCAN to Win ISI 4.0 software as a cal. file where the spectra were subjected to various manipulations for making the calibration equation. Spectra showed the outliers, which were eliminated by scoring. Score file generation showed the ranking of each sample from the spectral mean called global H . The global

H (GH) is the Mahalanobis H distance, which describes the distance of an individual sample from the centre of the population of the calibration samples. It basically seeks spectra, which are most different from other spectra.

Calibration Models for Fatty Acid Composition

In developing a NIRS calibration model for use in estimation of fatty acids of interest in *Brassica* samples, the statistics of calibrations and cross-validations are shown in Table 2, including the standard error of calibration (SEC) and the coefficient of determination in calibration (R^2). The one minus the ratio of unexplained variance divided by variance (1-VR) and the standard error of cross validation (SECV) are also shown in Table 2. The robustness of the equation obtained after calibration was strengthened first by internal cross validation followed by external cross validation. Internal validation was used to avoid overfitting of the equation by selecting the minimum number of PLS terms after the data scans through various cross validation passes in Win ISI software [18]. The MPLS regression model for the whole NIR spectra range (400–2500 nm) using the second derivative transformation with scatter correction (SNVD) of raw reflectance spectra yielded the equation for fatty acid analysis in *B. juncea*. The standard error of calibration for oleic acid was 3.91, the coefficient of determination was 0.91, SECV was 4.52, and 1-VR was 0.88. In the case of linoleic acid, SEC was 3.02, R^2 was 0.83, SECV was 3.20, 1-VR was 0.81, for erucic acid SEC was 6.01, R^2 was 0.88, SECV was 1.77, and 1-VR was 0.84. But for linolenic acid, calibration was not good; MPLS on 184 samples showed SEC 1.50, R^2 was 0.53, SECV was 1.77, and 1-VR was 0.35 as shown in Table 2. These results

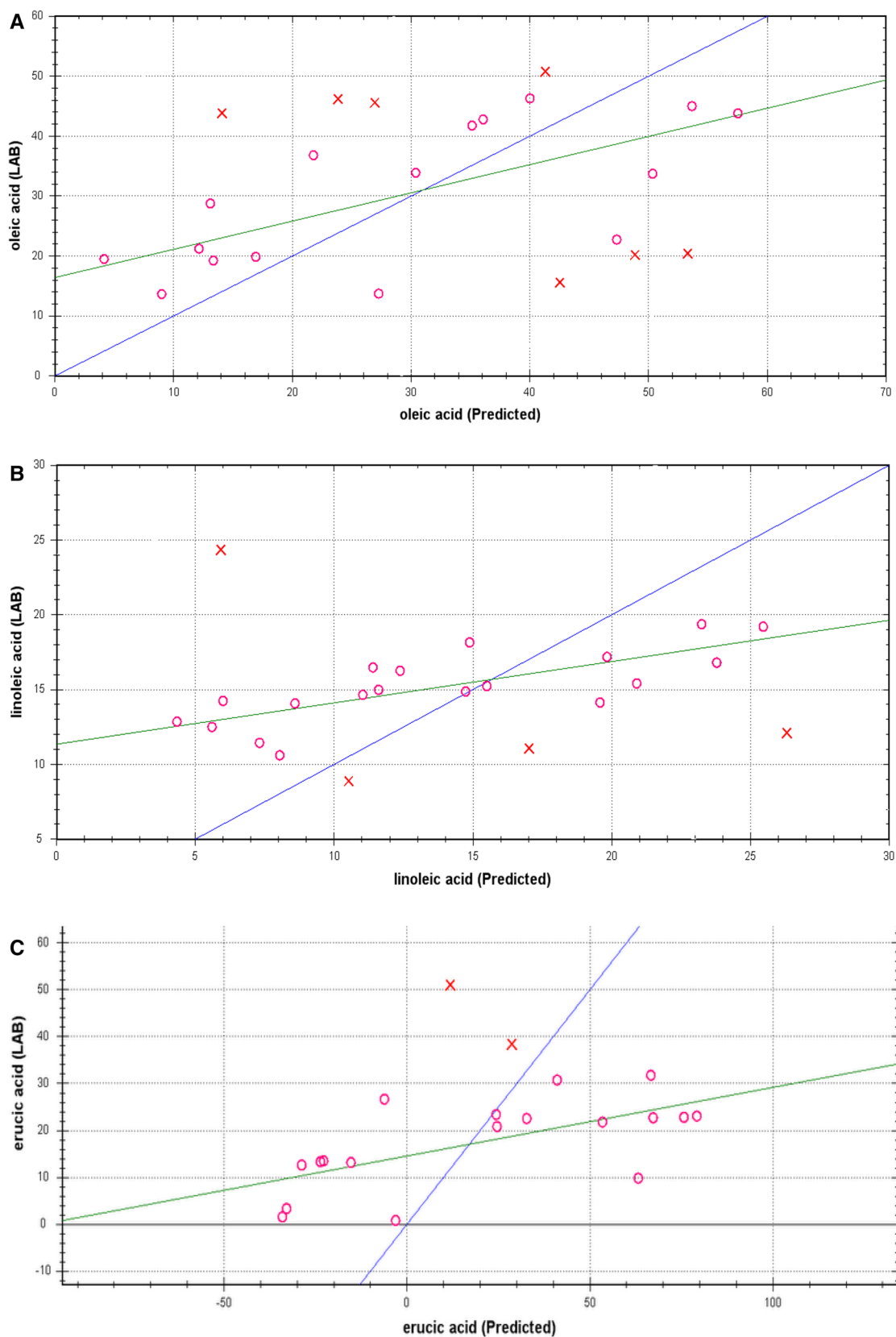


Fig. 3 Correlation plots between predicted values and lab (wet) values for oleic (a), linoleic (b), and erucic (c) acid (%) for the external validation of the near-infrared reflectance spectroscopy (NIRS) calibration equation for *Brassica juncea*

suggest that the variability for fatty acid composition in the calibration set is one of the main factors determining development of a successful calibration equation. Because in the case of linolenic acid, 1-VR is 0.35 whereas for others it is more than 0.8. There should be a close relation between R^2 and 1-VR as their values should be near about 0.9, whereas the value of SEC and SECV should be as little as it can be. If SEC and SECV are more, they can be corrected by adding more samples. Kim [11] reported R^2 for oleic acid 0.98, linoleic acid 0.90, linolenic acid 0.84, and erucic acid 0.98 with SEC 2.67, 1.00, 0.62, and 2.60, respectively in 249 seed samples of rapeseed. Mailer [19] reported R^2 for oleic acid and linoleic acid as 0.99, 1.0 with 1-VR 0.98, and 0.99, respectively in the case of olive oil. While building a calibration set, the following three parameters should be taken into consideration: the number of samples, the concentration range to be covered by the samples, and the distribution within this range. A large number of samples should be collected to ensure that the model covers a whole variation range and the samples are uniformly distributed within this range (in terms of all possible constituents or other parameters) [20].

Internal and External Validation for Fatty Acid Composition

Although internal cross-validation provides a useful way to estimate the predictive ability of the calibration equation, an external validation was carried out to analyze in more detail the performance of the calibration equations. The results obtained in the development of calibration equations for predicting the concentration of individual fatty acids were very heterogeneous (Table 3). The calibration equations for oleic acid, linoleic acid, and erucic acid showed a good R^2 in validation; Fig. 2 indicates that a reliable estimation of the concentration of these fatty acids can be obtained in samples of *B. juncea* by applying these equations. These values are similar to those reported for calibration equations integrating samples from different species of *Brassica* [4, 21]. The calibration equations for the linolenic acid were characterized by poor validation statistics, with r^2 0.25 (Table 3; Fig. 2) suggesting that the calibration equations for this could be useful only to make preliminary selections or to get an approximate classification of the entries; the SEP (Standard error of performance) was too high to consider them as adequate for reliable analyses. Previous studies developing either single- [22, 23] or multispecies calibration equations [4] for these fatty acids in the genus *Brassica* were reported as considerably lower in SEP values. For example, Velasco [4] obtained a SEP of 2.25 (mean 16.6 %) for C18:2 in a multi-species calibration equation, as compared with a SEP of 3.7 (mean 17.5 %) obtained in single species calibration.

For oleic acid in *B. juncea*, Bias was 0.01, SEP 6.57, slope 0.94, and R^2 0.77. Whereas in case of linoleic acid, Bias was 0.21, SEP 4.29, slope 0.92, R^2 0.68, and for erucic acid Bias 0.75, SEP 7.49, slope 0.95, and R^2 0.81. But for linolenic acid, Bias 0.36, SEP 2.96, slope 1.05, and R^2 0.25. Since, R^2 for calibration (0.91) and validation (0.77), in the case of oleic acid, 0.83 calibration R^2 and 0.68 validation R^2 in linoleic acid, and 0.88 (calibration) and 0.81 (validation) R^2 in erucic acid indicated good correlation and R^2 calculated by calibration and the validation model is close to each other, calibration equations for these parameters are considered as significant for their estimation. For the linolenic acid, the calibration and validation models showed low R^2 values of 0.53 and 0.20 respectively.

The external validation results showed that in the case of *B. juncea* (Table 3; Fig. 3), the equation was valid for oleic acid (R^2 0.89), linoleic acid (R^2 0.69), and erucic acid (R^2 0.90). The external and internal validation results showed good R^2 values for oleic acid, linoleic acid, and erucic acid. Therefore, in conclusion, this NIRS calibration equation for intact seed samples of *B. juncea* can be used for the estimation of the linoleic acid, erucic acid, and oleic acid profile except for linolenic acid.

Conclusion

The study concludes that calibration equation development for individual species of genera increases the accuracy for the estimation of various fatty acids. For good calibration equation development, the values of each quality trait should be uniformly distributed throughout the range of that trait in the population. The calibration equation having higher R^2 values in internal and external cross validation revealed a good calibration equation in comparison to low R^2 values that need to be updated with a new set of samples that cover a wide range in intact seeds of *B. juncea*.

References

1. Becker HC, Loiptien H, Robbelen G (1999) Breeding: an overview. In: Gómez-Campo C (ed) *Biology of Brassica Coenospecies*. Elsevier, Amsterdam, pp 413–460
2. Shengwu H, Ovesna J, Kucera L, Kucera V, Vyvadilova M (2003) Evaluation of genetic diversity of *Brassica napus* germplasm from China and Europe assessed by RAPD markers. *Plant Soil Environ* 49(3):106–113
3. Rakow R, Raney JP (2003) Present status and future perspectives of breeding for seed quality in *Brassica* oilseed crops. In: *Proceedings of the 11th international rapeseed congress*, Copenhagen, Denmark, pp 181–85
4. Velasco L, Goffman FD, Becker HC (1998) Variability for the fatty acid composition of the seed oil in a germplasm collection of the genus *Brassica*. *Genet Res Crop Evol* 45:371–382

5. Cardone M, Mazzoncini M, Menini S, Rocco V, Senatore A, Seggiani M, Vitolo S (2003) *Brassica carinata* as an alternative oil crop for the production of biodiesel in Italy: agronomic evaluation, fuel production by transesterification and characterization. *Biomass Bioenerg* 25:623–636
6. Batten G (1998) Plant analysis using near infrared reflectance spectroscopy: the potential and limitations. *Aust J Exp Agric* 38:697–706
7. Huang H, Yu H, Xu H, Ying Y (2008) Near infrared spectroscopy for on/in-line monitoring of quality in foods and beverages. *J Food Eng* 87:303–313
8. Williams P, Norris K (2001) Near infrared technology in the agricultural and food industries, 2nd edn. American Association of Cereal Chemists Inc., St. Paul
9. Workman J, Shenk J (2004) Understanding and using the near-infrared spectrum as an analytical method. In: Roberts CA, Workman J Jr, Reeves JB (eds) *Near-infrared spectroscopy in agriculture*. American Society of Agronomy, Madison, pp 3–10
10. Appleqvist LA (1968) Rapid methods of lipid extraction and fatty acid ester preparation for seed and leaf tissue with special remarks on preventing the accumulation of lipid contaminants. *Ark Kenci* 28:351–370
11. Kim KS, Park SH, Choung MG, Jang YS (2007) Use of near-infrared spectroscopy for estimating fatty acid composition in intact seeds of rapeseed. *J Crop Sci Biotech* 10:15–20
12. Font R, Mercedes DR, Antonio DH (2006) The use of near-infrared spectroscopy (NIRS) in the study of seed quality components in plant breeding programs. *Elsevier Ind Crops Products* 24(3):307–313
13. Elfadl E, Reinbrecht C, Claupeina W (2010) Development of near infrared reflectance spectroscopy (NIRS) calibration model for estimation of oil content in a worldwide safflower germplasm collection. *Int J Plant Prod* 4(4):259–270
14. Mika V, Tillmann P, Koprna R, Nerusil P, Kucera V (2003) Fast prediction of quality parameters in whole seeds of oilseed rape (*Brassica napus* L.). *Plant Soil Environ* 49(4):141–145
15. Hom NH, Becker HC, Mollers C (2007) Non-destructive analysis of rapeseed quality by NIRS of small seed samples and single seeds. *Euphytica* 153:27–34
16. Montes JM, Paul C, Melchinger AE (2007) Quality assessment of rapeseed accessions by means of near-infrared spectroscopy on combine harvesters. *Plant Breed* 126:329–330
17. Kim KS, Park SH, Choung MG (2006) Nondestructive determination of lignans and lignan glycosides in sesame seeds by near infrared reflectance spectroscopy. *J Agric Food Chem* 54:4544–4550
18. Shenk JS, Westerhaus MO (1996) Calibration the ISI way. In: Davis AMC, Williams PC (eds) *Near infrared spectroscopy: the future waves*. NIR Publications, Chichester, pp 198–202
19. Mailer RJ (2004) Rapid evaluation of olive oil quality by NIR reflectance spectroscopy. *JAOCS* 81:823–827
20. Fearn T (1992) Near infrared spectroscopy—bridging the gap between data analysis and NIR applications. Ellis Horwood Limited, New York, pp 61–66
21. Niewietzki O, Tillmann P, Becker HC, Mollers C (2010) A new near-infrared reflectance spectroscopy method for high throughput analysis of oleic acid and linolenic acid content of single seeds in oilseed rape (*Brassica napus*). *J Agric Food Chem* 58:94–100
22. Velasco L, Fernandez-Martínez JM, Haro AD (1997) Determination of the fatty acid composition of the oil in intact-seed mustard by near-infrared reflectance spectroscopy. *JAOCS* 74:1595–1602
23. Velasco L, Becker HC (1998) Estimating the fatty acid composition of the oil in intact-seed rapeseed (*Brassica napus* L.) by near-infrared reflectance spectroscopy. *Euphytica* 101(2):221–229