Determination of total oil content in corn seeds using near infrared spectroscopy and moving window partial least squares regression

Reunreudee Kaewcheenchai,[a,*] Sumaporn Kasemsumran,[b] Vichien Keeratinijakal[a] and Warunee Thanapase[b]

[a]*Department of Agronomy, Faculty of Agriculture, Kasetsart University, Bangkok 10900 Thailand. E-mail: g5161123@ku.ac.th*

[b]*Kasetsart Agricultural and Agro-Industrial Product Improvement Institute (KAPI), Kasetsart University, Bangkok 10900 Thailand*

Introduction

The major projected use of high oil corn is for animal feed and human food. In recent years, the oil of corn has gained such a wide commercial utilisation that under the present market conditions, it has become the most valuable constituent of the grain. It is proposed to meet this demand by breeding corn for increased total oil content. The breeding of a new corn variety requires methods that are non-destructive and rapid, for selection from generation to generation. Analysis for total oil content in corn seeds is important in a high-oil corn breeding program, but the conventional method (soxhlet extraction) is time-consuming and expensive. Therefore, NIR spectroscopy becomes a promising method for the early screening of quality traits. The objective of this study is to compare the accuracy of NIR calibration models for total oil content prediction in corn seeds that were developed by partial least squares regression (PLSR), in combination with moving window partial least squares regression (MWPLSR).

Materials and methods

Samples

A total of 302 samples of corn, in which the total oil content ranged from 2.18 to 10.82 $\%ww^{-1}$ were used in this study. Of these, 102 samples were collected from seed storage of the Suwan Wajokkasikit Field Crops Research Station, and another 200 samples derived from population of the corn breeding program, in which the varieties Suwan04L-336 and Suwan04L-337 were used as base population. The samples were divided into two sets; 222 samples for the calibration set and 80 samples for a prediction set.

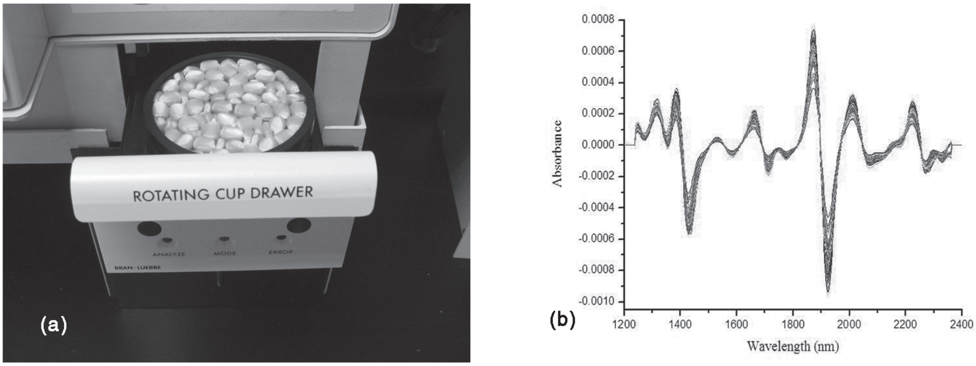

Figure 1. (a) Whole seed samples packed in a rotating cup and (b) second derivative spectra of corn samples.

Spectral acquisition

Seed samples were analysed as bulk whole seeds. The NIR spectra were collected in the region of 1200–2400 nm at 2 nm intervals with a rotating cup (Figure 1A) by the InfraAlyzer 500, reflectance mode (BRAN+LUEBBE, Norderatedt, Germany).

The spectral data obtained from measurement of two replications were averaged and used for calibration and prediction.

Reference analysis

Corn oil was extracted from the seeds by soxhlet extraction with petroleum ether 40:60, for about 16 hours.[1] Composition of the spectral files are summarised in Table 1.

Data analysis

All spectra were studied with various pretreatments, such as second derivative (2nd derivative, multiplicative scatter correction (MSC) and their combinations, before developing calibration models. Calibration equations were developed using PLSR and MWPLSR with the Unscrambler 9.8 (Camo, Oslo, Norway), and an in-house written program (run by Matlab ver. 6.1, The MathWorks, USA), respectively.

Table 1. Statistical characteristics of percentage of total oil content in the calibration and prediction sample sets.

Item	*n*	Mean	Standard deviation	Minimum	Maximum
Calibration set	222	4.8225	1.1064	2.1792	10.8233
Prediction set	80	6.1544	0.6436	4.5013	8.0925

Table 2. Summary of statistical results were calculated from the NIR spectra that were pretreated with 2nd derivative.

Region	Wavelength(nm)	*F*	Prediction		
			R	*RMSEP*	Bias
Whole region	1200–2400	11	0.4245	1.0151	–0.6829
MWPLS region	1240–1440	12	0.5081	0.6453	0.0150
	1580–1880	11	0.5610	0.5684	–0.1142
	1240–1440,1580–1880	10	0.5152	0.6272	–0.1494

F: the number of factors used; *R*: the correlation; *RMSEC*: root mean square error of calibration; *RMSEP*: root mean square error of prediction, Unit: %ww^{-1}.

Results and discussions

Statistical results for predicting the total oil content in corn seeds are shown in Table 2.

The spectra were pretreated with 2nd derivative [Figure 1(b)] provided the least root mean square error of prediction (*RMSEP*). These pretreatments can reduce the scattering effect and enhance spectral differences, and were selected to develop calibrations with PLSR and MWPLSR. The residual lines for 2nd derivative NIR spectra of corn seeds obtained from MWPLSR are depicted in Figure 2(a).

From the plot of residual lines informative regions can be selected, which, in this work, showed two informative spectral regions of 1240 – 1440 nm and 1580 – 1880 nm. The region of 1580 – 1880 nm selected by MWPLSR provided the lowest root mean square error of prediction (*RMSEP*) of 0.5684 %ww^{-1}. The correlation and bias were 0.8221 and –0.1142 %ww^{-1}, respectively. Scatter plot for the comparison between reference and NIR data is given in Figure 2(b). Further investigation may be necessary for re-calibration to improve the accuracy of the prediction, since the recent model is susceptible to change with climatic or agronomic conditions.

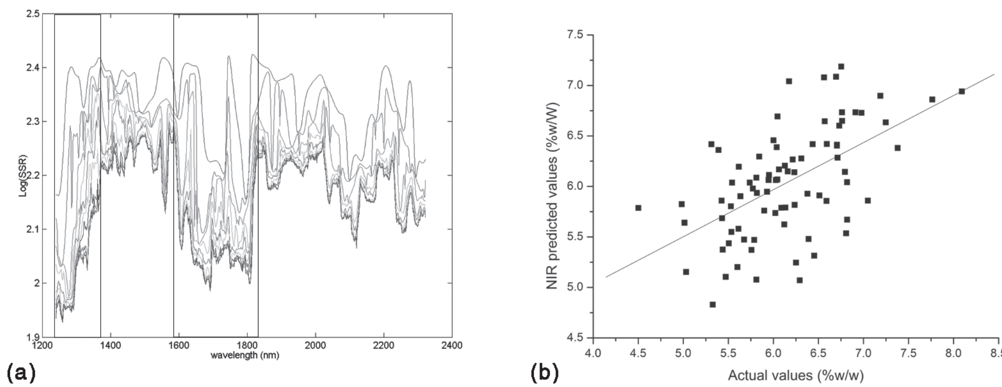

Figure 2. (a) Residual lines for second derivative spectra of corn seeds obtained from MWPLS and (b) scatter plot of actual and NIR predicted oil content in prediction set of corn seeds.

Conclusion

The optimum PLS model was calculated from 2nd derivative pretreated spectra of 1580–1880 nm. This model provided the lowest *RMSEP*. The result shows the potential of NIR spectroscopy in combination with MWPLSR to provide an estimate of total oil content of corn seeds in early generations of breeding programs. This work has demonstrated that NIR spectroscopy has the potential to be an effective technique for use in plant breeding. The technique requires the minimal sample preparation, and permits a large number of samples to be analysed per day.

Acknowledgments

The authors gratefully thank Kasetsart University Research and Development Institute (KURDI), Kasetsart Agricultural for financial support.

Reference

1. AOAC International, 17th Edn, *Official Methods of Analysis of AOAC International*, Ed by W. Horwitz. AOAC, Arlington, USA (2000).