Rapid Chemometric Method for the Determination of Oleic and Linoleic Acid in Sunflower Seeds by ATR-FTIR Spectroscopy

Chalermkwan Sukkasem [a], Thitiporn Machikowa* [a], Waraporn Tanthanuch [b] and Sodchoi Wonprasaid [a]

[a] Institute of Agricultural Technology, Suranaree University of Technology, Nakhon Ratchasima 30000, Thailand.
[b] Synchrotron Light Research Institute (Public Organization), Nakhon Ratchasima 30000, Thailand.

*Author for correspondence; e-mail: machiko@sut.ac.th

Received: 18 October 2013
Accepted: 7 October 2014

ABSTRACT

Screening of sunflower seeds in a breeding program to produce high oleic acid varieties requires a fast, simple and accurate analytical method. Experiments were carried out to: (i) investigate the potential of attenuated total reflectance-fourier transform infrared (ATR-FTIR) spectroscopy for rapid determination of fatty acid in sunflower seeds; and (ii) evaluate the calibration models for determination of oleic and linoleic acid contents in sunflower seeds by ATR-FTIR. Spectral data manipulation was performed using 5 different processing treatments for monitoring the calibration models, including two pretreatments (1st and 2nd derivative), two preprocessing strategies (Multiplicative Scatter Correction (MSC), vector and min/max normalization), and spectral ranges (3050-2800 and 1800-900 cm⁻¹). FTIR spectra combined with gas chromatography (GC) values of oleic and linoleic acid contents were used for predictive models generated by partial least square (PLS) regression. The results showed that treatments 1 and 2, the combinations of 2nd derivative with MSC or vector normalization in spectral ranges 3050-2800 and 1800-900 cm⁻¹, presented the lowest errors and produced the highest coefficients of determination of oleic and linoleic acid prediction models. It can be concluded that the ATR-FTIR method combined with chemometric analysis provided results comparable to GC analysis but they are more rapid and simple.

Keywords: Fourier transformed infrared, attenuated total reflectance, PLS, fatty acid composition, Helianthus annuus L.

1. INTRODUCTION

Sunflower oil is a high quality vegetable oil due to its high content of unsaturated fatty acids (approximately 89%). Among unsaturated fatty acids, linoleic acid (omega-6 group oil) is dominant in sunflower oil followed by oleic acid (omega-9 group oil), which are classified as polyunsaturated (PUFA) and monounsaturated fatty acids (MUFA),
respectively. The oil being rich in MUFA, low in PUFA and free of trans fatty acids provide a high level of oxidative stability, making it tolerant to high cooking temperatures which is required for food preparation and industrial applications. High oleic sunflower oils are classified as having MUFA levels of 80% and above. The development of sunflower types with high oleic acid and low linoleic acid content is possible through breeding programs. Recently, breeding programs in Thailand have been started to generate sunflower lines with high oleic acid content. The success of the program depends on many factors, especially the effectiveness of the seed screening method.

Analysis of oleic acid and other fatty acids in sunflower seeds can be accomplished using several methods. Gas chromatography with a flame ionization detector (GC-FID) has been proposed for the quantification of fatty acid composition. This method is reliable and reproducible, but is costly and time consuming (extraction and derivatization in sample preparation). Fourier transform infrared (FTIR) spectroscopy, both near infrared (NIR) and mid-infrared (MIR), has been used in the determination of fatty acid content as an interesting alternative to GC. This method has been proven to be useful for determining the fatty acid content of vegetable oil and quantification of free fatty acids in edible oils [1-3]. However, these measurements required the extraction of oil from samples before evaluation. Attenuated total reflectance FTIR (ATR-FTIR) is a rapid method for fatty acid analysis that does not require oil extraction and samples can be analyzed directly on the ATR surface [4]. The method has been used to discriminate between different oil content and to estimate the fatty acid composition of vegetable oils [5-7]. It has been applied to determine the relative PUFA, MUFA, free fatty acid and trans fatty acid content in vegetable oils and oil seeds [4, 6, 7, 8]. In addition, this technique has been also used for the monitoring and quantitative prediction of free fatty acids and trans fatty acids in food samples [9], and total fatty acid in poultry feed [10]. The method provided results comparable to official procedures, with the advantages of being inexpensive and rapid. However, no results have been reported on the use of ATR-FTIR for determination of the fatty acid quantitative mean in sunflower seeds samples. The aim of this study was to take advantage of the useful features of the ATR-FTIR method for the direct determination of fatty acid in sunflower seed samples.

The objectives of this study were to: (i) investigate the potential of ATR-FTIR spectroscopy for rapid determination of oleic acid and linoleic acid in sunflower seeds; and (ii) evaluate difference ways to select the calibration set for the determination of fatty acid in sunflower seeds by ATR-FTIR spectroscopy. The hypothesis of this experiment was that ATR-FTIR could be used as a simple and accurate analytical tool to determine the oleic and linoleic acid in sunflower seeds for rapid screening in breeding programs.

2. MATERIALS AND METHODS

2.1 Sunflower Materials

Two hybrid cultivars of sunflower, high oleic acid (Pacific 22) and low oleic acid content in seeds (Pacific 77), were cultivated in the experimental field at the Suranaree University of Technology, Nakhon Ratchasima, Thailand. At maturity stage (120 days after planted), 20 plants from each cultivar were randomly sampled. Seed samples were ground into powder using a grinder and passed through an 80-mesh sieve. To eliminate moisture interference, the samples were dried at
60 °C for 24 h before chemical analysis and FTIR sampling. After grinding, the samples were divided into two subsamples. One subsample was used for GC analysis (reference analysis) and another sub-sample was used to collect FTIR spectra.

2.2 Gas Chromatography Analysis

The ground sunflower seed was analyzed for oleic and linoleic acid contents as their corresponding methyl esters by GC analysis [11]. The method involves extraction of the lipids with organic solvents, addition of an internal standard, and methylation to prepare fatty acid methyl ester (FAME) for GC analysis. For oil extraction, 1 g of ground, sieved material was placed into a vial, and 5 ml diethyl ether was added. The vial was shaken periodically over 5 h, and then the solvent was evaporated. The fatty acid composition of the oil was analyzed by methyl esterification [12]. About 1 ml of C₁₇ internal standard (2.0 mg/ml in hexane) and 2 ml of BF₃ in methanol was added and heated at 100 °C with shaking. Briefly, a 2.0 ml sample was injected into a GC equipped with a flame ionization detector (FID). The carrier gas was nitrogen at a flow of 1.5 mL min⁻¹. The oven, injector, and FID were held at 195, 275, and 250 °C, respectively. Each sample was analyzed in duplicated. Identification of oleic acid methyl ester was achieved by comparison with authentic reference standards. The results were expressed as percentages according to the total integrated area.

2.3 ATR-FTIR Measurement

ATR-FTIR measurement was performed at the off-line IR spectroscopy facility, at the Synchrotron Light Research Institute (Public Organization), Thailand. Data acquisition was performed using a Bruker TENSOR 27 spectrometer (Bruker Optics Inc., Ettlingen, Germany) equipped with a nitrogen cooled MCT (HgCdTe) detector. Approximately 0.10 g of the grinding sunflower seeds was placed on the surface of Pike “Miracle” single reflection ATR accessory equipped with a germanium composite ATR crystal. FTIR spectra acquisitions were performed by collecting 32 scans, with 4 cm⁻¹ resolution over a measurement range of 4000-700 cm⁻¹. The spectra were collected with the OPUS software 6.5 (Bruker Optics, Inc.).

2.4 Multivariate Data Analysis

An experimental protocol was followed for the calibration and validation steps to develop and validate a FTIR model. At least thirty different FTIR spectra of each fatty acid concentration were used for calibration. Principal Component Analysis (PCA) was performed on the entire FTIR spectral data set of the differences lipid content sample in order to select a high quality data set in which outlier data was eliminated. The Unscrambler software (version 9.2, CAMO Software AS, Oslo, Norway) was used for PCA analysis employing the combined selected spectral ranges. PCA was performed on second derivative spectra using the Savitzky-Golay algorithm with nine points of smoothing and normalized with Extended Multiplicative Signal Correction (EMSC).

Partial least square (PLS) calibration models, including cross- and test-set validation, were obtained with the QUANT 2 module of the OPUS 6.5 software. The PLS analysis was applied for developing the calibration models to determine the oleic and linoleic acid contents of sunflower seed samples based on two spectral regions (3050-2800 cm⁻¹ and 1800-900 cm⁻¹). Spectral data manipulation was performed using 5 different treatments (Table 1, 2). The performance of the calibrations was assessed
by linear regression and evaluated by testing the samples of known fatty acid composition. The calibration accuracy was described by the multiple coefficients of determination ($R^2$) and the overall errors between predicted and reference GC values. For validation within the whole calibration set, cross validations were evaluated based on the root mean square errors of cross validation (RMSECV) for samples in the calibration set. Spectral outliers that obviously did not fit the prediction model were removed from the sample sets. After removal, the full-cross validation was repeated. The root mean square error of prediction (RMSEP) and actual values were compared to test the regression slope. Therefore, the calibration and validation model error was calculated through RMSECV and RMSEP, finding minimum values for both cases. Multivariate calibration step was carried out by comparing the oleic and linoleic acid contents predicted by PLS model to the values obtained from GC analysis.

Table 1. Processing and statistical indices for prediction of oleic acid content of sunflower at concentration range 15.6-64.9%.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Data processing</th>
<th>Spectral range (cm⁻¹)</th>
<th>Spectra (n)</th>
<th>$R^2$</th>
<th>Slope</th>
<th>RMSECV</th>
<th>Test set Spectra (n)</th>
<th>RMSEP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2nd derivative, 13 point Smooth, MSC</td>
<td>3050-2800, 1800-900</td>
<td>290</td>
<td>0.971</td>
<td>0.997</td>
<td>2.48</td>
<td>118</td>
<td>1.08</td>
</tr>
<tr>
<td>2</td>
<td>2nd derivative, 13 point Smooth, Vector normalization</td>
<td>3050-2800, 1800-900</td>
<td>290</td>
<td>0.972</td>
<td>0.997</td>
<td>2.42</td>
<td>118</td>
<td>1.49</td>
</tr>
<tr>
<td>3</td>
<td>2nd derivative, 13 point Smooth, Min/max normalization</td>
<td>3050-2800</td>
<td>290</td>
<td>0.958</td>
<td>0.997</td>
<td>2.96</td>
<td>118</td>
<td>1.57</td>
</tr>
<tr>
<td>4</td>
<td>1st derivative, 13 point Smooth, MSC</td>
<td>3050-2800, 1800-900</td>
<td>266</td>
<td>0.965</td>
<td>0.995</td>
<td>2.58</td>
<td>118</td>
<td>2.74</td>
</tr>
<tr>
<td>5</td>
<td>1st derivative, 13 point Smooth, Vector normalization</td>
<td>3050-2800, 1800-900</td>
<td>261</td>
<td>0.972</td>
<td>0.996</td>
<td>2.28</td>
<td>118</td>
<td>2.60</td>
</tr>
</tbody>
</table>

MSC: Multiplicative Scatter Correction, RMSECV: Root mean square error of cross validation, RMSEP: Root mean square error of prediction.
Table 2. Data Processing and statistical indices for prediction of linoleic acid content of sunflower at concentration range 4.8-39.9%.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Data processing</th>
<th>Spectral range (cm⁻¹)</th>
<th>Spectra (n)</th>
<th>( R^2 )</th>
<th>Slope</th>
<th>RMSECV</th>
<th>Test set Spectra(n)</th>
<th>RMSEP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2nd derivative, 9 point Smooth, MSC</td>
<td>3050-2800, 1800-900</td>
<td>301</td>
<td>0.990</td>
<td>0.997</td>
<td>1.13</td>
<td>121</td>
<td>0.51</td>
</tr>
<tr>
<td>2</td>
<td>2nd derivative, 9 point Smooth, Vector normalization</td>
<td>3050-2800, 1800-900</td>
<td>301</td>
<td>0.990</td>
<td>0.997</td>
<td>1.12</td>
<td>121</td>
<td>0.55</td>
</tr>
<tr>
<td>3</td>
<td>2nd derivative, 9 point Smooth, Min/max normalization</td>
<td>3050-2800</td>
<td>301</td>
<td>0.987</td>
<td>0.996</td>
<td>1.30</td>
<td>121</td>
<td>0.80</td>
</tr>
<tr>
<td>4</td>
<td>1st derivative, 9 point Smooth, MSC</td>
<td>3050-2800, 1800-900</td>
<td>292</td>
<td>0.987</td>
<td>0.996</td>
<td>1.28</td>
<td>121</td>
<td>1.24</td>
</tr>
<tr>
<td>5</td>
<td>1st derivative, 9 point Smooth, Vector normalization</td>
<td>3050-2800, 1800-900</td>
<td>292</td>
<td>0.985</td>
<td>0.995</td>
<td>1.38</td>
<td>121</td>
<td>1.05</td>
</tr>
</tbody>
</table>

MSC: Multiplicative Scatter Correction, RMSECV: Root mean square error of cross validation, RMSEP: Root mean square error of prediction.

3. RESULTS AND DISCUSSION

3.1 Gas Chromatography

The fatty acid profiles of sunflower seed samples were obtained by GC analysis. The results found that oleic and linoleic acids were predominant among unsaturated fatty acid. Oleic acid (C18:1) of two cultivars varied from 15.6 to 64.9% and linoleic acid (C18:2) ranged between 4.8-39.9%.

3.2 FTIR Spectra and Selection of The Spectral Range

The collected ATR-FTIR spectra were recorded between 4000-700 cm⁻¹ as shown in Figure 1a and 1b. Sunflower seeds containing different oleic and linoleic content show different absorption intensity and spectral patterns. The FTIR spectral similarity was observed in the lipid region. However, the region of interest for fatty acid composition was between 3050-900 cm⁻¹. There were two dominant regions related to fatty acids as shown in Figure 1. The first lipid region, the bands mainly associated with lipids are located between 3050 and 2800 cm⁻¹. This region corresponds to asymmetric and symmetric stretches of CH₃ and CH₂ of the triglycerides fatty acid chains [3]. In this range, the band at 3006 cm⁻¹ results from the C-H stretching vibration of HC=CH groups were observed which could be a useful indicator of the different degrees of unsaturation in acyl chains of phospholipids [13]. The second
lipid region of typical vibration absorption bands located between 1800 and 1350 cm\(^{-1}\). The absorption band of carbonyl stretches of carboxyl at 1747 cm\(^{-1}\) exhibited the presence of lipids \([14, 15]\), assigned to C=O stretching vibration of ester groups in triacylglycerols \([5, 16]\). The increased intensity of CH\(_2\) bending vibration of lipids at 1458 cm\(^{-1}\) could also be observed along higher oleic content \([17]\). These results are in agreement with other studies which reported the IR regions of edible oil for unsaturated fatty acid featured around 2900 cm\(^{-1}\), as well as oleic acid and linoleic acid profiles located between 3033-2400 and 2260-700 cm\(^{-1}\), respectively \([7, 10, 18]\).

3.3 Calibration and Quantitative Analysis

The FTIR spectral data set used for PLS was selected by PCA analysis to eliminate outlier spectra. Different ranges of the spectra were evaluated in order to reach the best prediction results. The best frequency regions (3050-2800 cm\(^{-1}\) and 1800-900 cm\(^{-1}\) as range of the C-H stretching, ester C=O vibration and CH\(_2\) vibration) were selected for calibration between FTIR and GC values for oleic and linoleic acid contents. The overall data set (≤301 spectra) was used for calibration and at least 118 spectra were used for the test set validation. The models for determination of oleic acid and linoleic acid contents were built by PLS regression. Five treatments were considered for the evaluation of the capabilities of ATR-FTIR in order to establish oleic and linoleic acid contents with the use of different spectra ranges, pretreatments and preprocessing techniques (Tables 1 and 2).

3.3.1 PLS models for oleic acid

Table 1 illustrated the correlation between GC and FTIR values for oleic acid in the seed samples from each calibration model. The model for treatments 1, 2 and 5 performed quite well for calibration, yielding good R\(^2\) values (R\(^2\) = 0.971, 0.972 and 0.972,
respectively), whereas treatment 3 (min/max normalization with the spectra range 3050-2800 cm⁻¹) had the lowest $R^2$ (0.958). The RMSECV values were of the magnitude of the standard deviation of the reference method. In this investigation, low RMSECV values were observed in treatments 1, 2 and 5 (2.48, 2.42 and 2.28), whereas highest RMSECV was obtained in treatment 3 (2.96). The predictions were also very satisfactory in terms of RMSEP. The results presented RMSEP values of 1.08 and 1.49 for treatments 1 and 2, respectively. The model of treatments 1, 2 and 3 yielded similar slope regression line values of 0.997, very close to 1, which indicated low bias and absence of systematic errors of the method for estimation. The accuracy of the ATR-FTIR predictions was assessed based on the highest $R^2$ and low RMSECV. It can be seen that treatments 1 and 2 provided similar calibration and prediction performances, these treatments also produced the lowest RMSECV value and produced the highest $R^2$ of all treatments. Therefore, it is concluded that using MSC or vector normalization preprocessing and second derivative treatment with the spectral ranges 3050-2800 and 1800-900 cm⁻¹ provided predictive results comparable to GC analysis.

### 3.3.2 PLS models for linoleic acid

A similar procedure for calibration and prediction using different data processing were carried out to cross the linoleic acid content. Table 2 shows the correlations between measured and predicted values of linoleic acid on validation data set. For treatments 1 and 2, the highest $R^2$ presented similar value (0.990) with the application of second derivative in two spectra ranges (3050-2800 cm⁻¹ and 1800-900 cm⁻¹) with MSC and vector normalization preprocessing, whereas lower $R^2$ values were obtained for treatments 3-5 (0.985-0.987).

From the calibration, the assessment of the error was carried out by calculating the residual RMSECV after comparing the actual concentration with the computed one for each component. Treatments 1 and 2 revealed low RMSECV (1.13 and 1.12) and RMSEP (0.51 and 0.55) values, while higher RMSEP values were observed in treatments 3-5. The slopes of the actual and predicted regression lines depicted in these plots were close to unity (0.995-0.997), the highest slope value (0.997) was found in treatments 1 and 2. As observed, models of treatments 1 and 2 performed well for calibration, yielding very good $R^2$ values and low RMSECV values, predictions were very satisfactory, in terms of RMSEP and bias values. Therefore, the best result for linoleic acid prediction was obtained from treatments 1 and 2.

The values of $R^2$, RMSECV and slope were useful for distinguishing the accuracy of using the FTIR method compared to the GC in analysis of fatty acid contents. In addition, low bias and absence of systematic errors were demonstrated by the slopes of the actual vs. predicted regression lines. This investigation indicated that the prediction results of treatments 1 and 2 were very good yielded, because RMSECV was close to 0, the slope and $R^2$ values of both fatty acid contents were close to 1. The PLS analysis for FTIR spectra of sunflower seeds in this experiment could be used to evaluate oleic acid and linoleic acid contents in the samples and provided good prediction for these fatty acid contents. For pretreatment, second derivative treatment was yielded better results than first derivative treatment in both fatty acid models. The spectral ranges 3050-2800 cm⁻¹ and 1800-900 cm⁻¹ were found to have good
predictive capability demonstrated by the low RMSECV values and values for both the slope and $R^2$ of both fatty acid contents being close to 1. However, RMSECV and RMSEP of treatment 3 were higher than other treatments due to the lack of the 1800-900 cm$^{-1}$ spectral range, which contained C-H stretching and C=O stretching [9]. It can be seen that the higher $R^2$ and lower RMSECV for linoleic acid model, when compared to oleic acid model, could in part be explained by the fact that for linoleic acid the reference method is also more accurate.

### 3.3.3 Prediction of oleic acid and linoleic acid contents

The predictive ability of the model achieved from treatment 1 (second derivative and MSC) of both oleic acid and linoleic acid contents was tested with five triplicate analysis of unknown samples. The predicted oleic acid and linoleic acid contents from the PLS method were compared to GC analysis. Table 3 shows the oleic acid and linoleic acid contents with deviation of 5 validation samples. The predicted oleic acid content in the test samples varied between 19.41-48.96%, and the actual measured values were 20.10-49.70%. In this case, the mean differences of oleic acid of five validation samples ranged from 1.49-3.43%. The predicted linoleic acid contents were 13.55-31.27%, and the actual measured values were from 13.93 to 30.72%. The mean differences of linoleic acid of samples ranged from 0.66-2.73% were smaller than mean differences of oleic acid. The comparable results determined using FTIR and GC indicates that ATR–FTIR can be used as an alternative method for determination of oleic and linoleic acid contents in sunflower seeds.

![Table 3. GC measurement and FTIR predicted values of oleic and linoleic acid content in sunflower seeds.](attachment:table3.png)

<table>
<thead>
<tr>
<th>Sample</th>
<th>Oleic acid (%)</th>
<th>Linoleic acid (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GC</td>
<td>FTIR Prediction</td>
</tr>
<tr>
<td>1</td>
<td>20.10</td>
<td>19.41</td>
</tr>
<tr>
<td>2</td>
<td>29.91</td>
<td>30.57</td>
</tr>
<tr>
<td>3</td>
<td>31.70</td>
<td>31.20</td>
</tr>
<tr>
<td>4</td>
<td>44.69</td>
<td>43.58</td>
</tr>
<tr>
<td>5</td>
<td>49.70</td>
<td>48.96</td>
</tr>
</tbody>
</table>

CONCLUSION

The application of ATR-FTIR spectroscopy for fatty acid content determination provided results statistically similar to GC analysis. The analytical procedure of this method is simple compared to the GC reference method in terms of sample handling and no toxic chemical reagents are needed. Furthermore, the scanning of ATR-FTIR spectra and the analysis time of each sample is less than 5 minutes. This method would be very effective for determination of seed sample fatty acid in breeding programs, when there are large populations. The ATR-FTIR technique using the spectral ranges between 3050-2800 cm$^{-1}$ and 1800-900 cm$^{-1}$ with second derivative pretreatment MSC or
vector normalization preprocessing is suitable for the screening of oleic and linoleic acid contents in seed samples. The prediction models can determine the concentrations of oleic and linoleic acid with reasonable accuracy. Therefore, the ATR-FTIR technique can be used as an alternative method in plant breeding programs for fatty acid determination in seed samples since it is accurate, rapid and simple.

ACKNOWLEDGEMENTS

This research was supported by Suranaree University of Technology research fund. The authors gratefully acknowledge the Synchrotron Light Research Institute (Public Organization), Nakhon Ratchasima, Thailand, for time allocation of using off-line ATR-FTIR spectroscopy.

REFERENCES


