



Prediction of crude protein and oil content of soybeans using Raman spectroscopy[☆]

Hoonsoo Lee^a, Byoung-Kwan Cho^{a,*}, Moon S. Kim^b, Wang-Hee Lee^a, Jagdish Tewari^c, Hanhong Bae^d, Soo-In Sohn^e, Hee-Youn Chi^f

^a Department of Biosystems Machinery Engineering, Chungnam National University, Daejeon, South Korea

^b Environmental Microbial and Food Safety Laboratory, USDA-ARS, Beltsville, MD 20705, USA

^c Fiber Science, Cornell University, Ithaca, NY, USA

^d School of Biotechnology, Yeungnam University, Gyeongsan, South Korea

^e Biosafety Division, National Academy of Agricultural Science, RDA, Suwon, South Korea

^f SBLab Co., Suwon, South Korea

ARTICLE INFO

Article history:

Received 9 January 2013

Received in revised form 22 April 2013

Accepted 23 April 2013

Available online 15 May 2013

Keywords:

Raman spectroscopy

Oil content

Protein content

Soybean

Partial least squares analysis

ABSTRACT

While conventional chemical analysis methods for food nutrients require time-consuming, labor-intensive, and invasive pretreatment procedures, Raman spectroscopy can be used to measure a variety of food components rapidly and non-destructively without supervision from experts once the instrument has been calibrated. The purpose of this study was to develop an optimal prediction model for determining the protein and oil contents of soybeans using a dispersive Raman spectroscopy method. In general, the crude oil content of soybeans is chemically determined using the Soxhlet extraction method, while the semimicro-Kjeldahl method and an auto protein analyzer have been used to assess crude protein content. In the present study, Raman spectra were measured in the 200–1800 cm⁻¹ wavenumber range and partial least squares (PLS) analysis methods were used to develop optimal models for predicting the crude protein and oil contents of soybeans. The resultant PLS models that used the effective wavenumber regions determined by intermediate PLS (iPLS) method were better than those models developed using the entire wavenumber range under investigation. The R_p^2 and SEP of the optimal PLS model for crude protein content were 0.916 and 0.636%, respectively. Likewise, the R_p^2 and SEP for crude oil content were 0.872 and 0.759%, respectively. The result suggests that the conventional Raman techniques investigated in this study can be applied to the prediction of soybean crude protein and oil content.

© 2013 The Authors. Published by Elsevier B.V. All rights reserved.

1. Introduction

In 2011, approximately 261 million tons of soybeans were produced worldwide. This represents approximately 10.1% of the total cereal production in 2010 [1]. Soybeans are an important food source for humans and animals and contain considerably more protein and oil than many other crops [2]. The protein and oil content of soybeans are considered to be important quality indicators and are announced immediately before packaging and distribution. Thus, the classification of soybeans could lead consumers to selectively take nutritionally high valued diets. Kjeldahl and Soxhlet extraction methods are widely used as standard methods of

measuring the crude protein and oil content of crops. However, these methods have several disadvantages, as they are time-consuming, labor-intensive, and destructive procedures that hamper the fast and economical quality evaluation of mass-produced soybeans. Therefore, the demand for rapid, robust, and nondestructive measurement methods for the main nutrient components of crops, such as the protein, oil, carbohydrate, water, and ash content of soybeans has recently increased.

Since the 1970s, near infrared reflectance spectroscopy has been widely used to measure the quality of agriculture and food materials. This noninvasive spectroscopic technique can rapidly provide physical and chemical information about specimens. Near infrared diffuse reflectance spectroscopy and near infrared transmittance spectroscopy have been used to predict the fatty acid composition of soybeans [3,4]. Baye et al. used near-infrared spectroscopy to predict maize seed composition [5] and Choung et al. employed near infrared reflectance spectroscopy to determine the protein and oil content of soybeans [6]. However, combination and overtone absorption bands in the near infrared region that are caused

[☆] This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial-No Derivative Works License, which permits non-commercial use, distribution, and reproduction in any medium, provided the original author and source are credited.

* Corresponding author. Tel.: +82 42 821 6715; fax: +82 42 823 6246.

E-mail addresses: chobk@cnu.ac.kr, bcx195@gmail.com (B.-K. Cho).

by O–H, C–H, and N–H functional groups result in broad overlapping spectra and the sensitivity of these techniques is limited to approximately 0.1% for trace components. Hence, a novel technique is necessary to improve accuracy and precision. As an alternative, Raman spectroscopy technique has emerged as a novel technology that could compensate for the disadvantage of NIR spectroscopy such as low sensitivity for minor components, providing high accuracy and precision [7]. Unlike near infrared spectroscopy, Raman spectroscopy provides detailed information about molecular vibrations. Extremely high sensitivity has been demonstrated for Raman spectroscopy in various fields [8]. This method involves the analysis of Raman scattering, which is observed in samples when they are illuminated by a strong light source. There are two types of Raman spectroscopy techniques that are of interest for the evaluation of agricultural commodities. One is dispersive Raman spectroscopy and the other is Fourier-transform (FT) Raman spectroscopy. Dispersive Raman spectroscopy typically uses various visible lasers, such as 473 nm, 532 nm and 633 nm lasers, and 785 nm near infrared (NIR) laser. In contrast, FT-Raman spectroscopy has used 1024 nm laser generally to minimize strong fluorescence signal from agro-product [9], preventing it from directly measuring the nutrient content of agricultural products. For this reason, many researchers have turned to interferometer-based NIR FT-Raman systems that cause less fluorescence for agriculture products than the dispersive Raman systems. The NIR FT-Raman spectroscopy has been used to predict the protein and apparent amylose contents of rice, determine milk fat content, monitor and quantify ethyl esters in soybean oil, and distinguish between oil and fat [10–13]. However, the large size of FT-Raman system due to interferometer is not suitable for recent demands on compact Raman system in the field of agriculture.

To construct a compact Raman system which satisfies high sensitivity and low fluorescence emission, the strong fluorescence signal must be removed from conventional Raman data, particularly for dispersive Raman system. Recently, various mathematical methods, such as wavelet transformation, Fourier transformation, and polynomial fitting, have been developed to banish the strong fluorescence background signal caused by dispersive Raman spectroscopy [14]. Among the algorithms, the polynomial fitting method is popular because of its rapid and simple nature [15]. This method has been applied to dispersive Raman spectroscopy at 785 nm laser in order to detect lycopene changes in tomatoes [9].

As dispersive Raman data is available by reducing the fluorescence, the objective of this study is to develop an optimal prediction model for determining the protein and oil content of soybeans using data sets obtained by dispersive Raman spectroscopy. We aim at using PLS algorithm to use Raman spectra processed by polynomial fitting method in determination of the optimal model. Because high fluorescence limited the use of dispersive Raman system at the expense of high sensitivity, this study is to invent a customized mathematical method available for Raman spectra so that Raman spectroscopy can be applied to inspect and obtain accurate information for a specific agricultural product.

2. Materials and methods

2.1. Samples and reference analysis

A total of 45 types of soybean sample were purchased from retail stores and 50 types of soybeans were provided by the National Institute of Crop Science in Korea. Of all sample, the color of twenty-five samples were black and remain seventy samples were yellow. The average and standard deviation of the diameter of soybeans used in this experiment were 0.72 cm and 0.12 cm, respectively.

Table 1 illustrates indicate In addition, the range of moisture content in soybean was between 6.13% and 13.44% (**Table 1**). Each sample was milled using a grain grinder and passed through a 250- μm -mesh sieve. Sieved samples were stored in plastic bottles and spectra were obtained using Raman spectroscopy. Crude protein was determined using the semimicro-Kjeldahl method and an auto protein analyzer (Kjeltec 2400 auto-analyzer, Hillerød, Denmark). The ground soybean sample was weighed 1 g accurately and nitrogen to protein conversion factor of 5.71 was used. In order to obtain crude oil content in soybean, weighted 1 g sample were put a dish and were dried during 2 h at 105 °C using drying oven. Then, crude oil content was determined by extraction using diethyl ether and a Soxhlet extractor (Soxtec System HT 1043 extraction unit, Höganäs, Sweden).

2.2. Raman spectroscopy

A 400-mW diode laser with a light source at 785 nm was used for all Raman measurements. The Raman system (Kaiser Optical Systems, Ann Arbor, MI) consisted of a charge-coupled device (CCD) detector and a holographic transmission grating. The spectral range was set to 200 cm^{-1} –1800 cm^{-1} in 0.3 cm^{-1} intervals and the laser point diameter was set to 3 mm. The same amount of each sample (2 g) was placed in a standard 96-hole wall plate and the plate moved automatically to align each sample with the Raman system prior to each measurement. For each sample, the spectrum measurement with 1-s exposure time was repeated 64 times for each sample and then the average spectrum was used for a representative spectrum for each sample.

2.3. Data analysis

The Raman scattering signal was generated by using a strong laser light source. Biological materials, such as agricultural products, may emit strong fluorescence signals that mask the characteristic Raman scattering signal. This problem has been considered to be a challenge for Raman spectroscopy [9]. In the present study, a widely used polynomial fitting method was employed to analyze Raman spectrum data and to correct for fluorescence because this method is efficient and simple [15]. Polynomial fitting involves determining the proper order polynomial for obtaining a baseline through iterative calculation. Various order of polynomial such as 4th, 5th, 8th, 12th, and 16th was test for fitting the spectrum data based on the previous studies that used 5th order polynomial for soybeans [7] and 8th order polynomial equation for lycopene in tomatoes [9]. Finally, a 16th order polynomial equation and the 100th iteration were employed to create the fluorescence correction baseline because of its best prediction than any other polynomials.

The corrected Raman spectra were further subjected to 8 preprocessing methods, including smoothing, mean normalization, maximum normalization, range normalization, multiplicative scatter correction (MSC), standard normal variate (SNV), Savitzky–Golay 1st derivative, and Savitzky–Golay 2nd derivative methods. The preprocessed spectra were used to develop an optimal partial least squares (PLS) model described below for prediction of the protein and oil content of soybeans.

2.4. Partial least squares (PLS) analysis

PLS was the main algorithm used to obtain the prediction models for the crude protein and oil content of soybeans. PLS is a multivariate analysis method that extracts new latent variables from raw spectra. It can compress the large amount of spectral data into a new structure known as latent variables or factors and these latent variables are able to describe the maximum covariance

Download English Version:

<https://daneshyari.com/en/article/7148897>

Download Persian Version:

<https://daneshyari.com/article/7148897>

[Daneshyari.com](https://daneshyari.com)