APPLICATION OF NIR REFLECTANCE SPECTROSCOPY ON DETERMINATION OF MOISTURE CONTENT OF IN-SHELL PEANUTS: A NON-DESTRUCTIVE ANALYSIS

J. Sundaram, C. V. K. Kandala, C. L. Butts, W. R. Windham

ABSTRACT: NIR spectroscopy was used to measure the moisture content (MC) of Virginia and Valencia type in-shell peanuts. Peanuts were conditioned to various moisture levels between 7% and 26% (wet basis), and the MC was verified using the standard oven method. Sample from the various moisture levels were separated into two groups, as calibration and validation. NIR absorption spectral data from 400 nm to 2500 nm were collected using peanuts within the calibration and validation sample sets. Measurements were obtained on 30 replicates within each moisture level. Partial least squares (PLS) analysis was performed on the calibration set, and models were developed using the raw spectral data and its derivative function data. The standard error of calibration (SEC) and $R^2$ of the calibration models were calculated to select the best calibration model for each peanut market type. Both Valencia and Virginia types gave an $R^2$ value of 0.99 for the derivative spectral data treatment as well as for the raw data. The selected models were used to predict the moisture content of peanuts in the validation sample set. Predicted and reference moisture contents were compared. Relative percent deviation (RPD) and standard error of prediction (SEP) were calculated to validate the goodness of fit of the prediction model. The raw reflectance spectra model gave an RPD of 5.55 with a corresponding SEP of 0.97 for Valencia type peanuts, which is an indicator that the model is good for quality control and analysis. For Virginia type peanuts, the derivative reflectance spectra model gave the highest RPD value of 5.75 and the lowest SEP of 0.771. Thus, these two models were selected for the respective peanut types as the best models for prediction of moisture content.

Keywords. In-shell peanuts, NIR spectroscopy, Partial least squares, Pretreatment, Relative percent deviation, Standard error of prediction.

Accurate measurement of peanut moisture content (MC) is critical in marketing, storing, and processing. Typically, peanuts are marketed after partial drying in the windrow followed by mechanical drying until the MC is below 10.5% (moisture content is expressed in % wet basis throughout this article). During the grading process, peanuts must be shelled and the grading procedure nearly completed before the moisture content can be determined. If the moisture content is greater than 10.5%, then the load is rejected (USDA, 2000), returned to the dryer to complete drying, and then regraded. Blankenship et al. (2001) and Lamb et al. (2003) showed that peanuts could be accurately graded at MC up to 18% and eliminated the need to regrade the peanuts after re-drying. However, this process did not eliminate the need to ensure that the MC of the peanuts was below 10% prior to placing the peanuts in storage. Storing peanuts with excessive MC increases the risk of microbial growth and aflatoxin contamination during storage. Therefore, measurement of peanuts MC is important for storage and processing.

Meters currently used to measure peanut MC are based on capacitance measurements. During drying, a sample of about 500 g of in-shell peanuts is periodically collected from a trailer, shelled, and then the peanut kernels are placed in the sample holder of a moisture meter. The moisture meter measures the capacitance of the sample holder with the sample, which is a function of the dielectric properties of the kernels, and displays it as the average MC. These moisture samples are usually discarded and cumulatively result in a large loss of edible peanuts. In-shell measurement of kernel moisture would result in considerable savings in both time and money during grading and drying.

Current capacitance moisture meters require the peanuts to be shelled before the measurement due to limited sample cup size and the absence of in-shell calibrations. Single-kernel devices provide valuable information regarding the variability of MC within a sample, but the methods are destructive and time-consuming when performed on large samples. Techniques using near-infrared (NIR) spectroscopy for food quality measurements are becoming more popular in food processing and quality inspection of agricultural commodities. NIR spectroscopy has several advantages over conventional physical and chemical analytical methods of food quality analysis. It is rapid, nondestructive, and provides
more information about the components and the structures present in food and food products. It can be used to measure more than one parameter simultaneously.

NIR spectroscopy is well-suited for measurement of water, since the water O-H group overtone and combination bands are pronounced in the NIR region of the spectrum. Water is a strong absorber in the NIR region, and samples with high moisture content are strongly dominated by the signature of water. NIR is readily adaptable, and the low intensities of NIR absorptions permit direct measurement of water over wide concentration ranges in solid samples. Spectra measured using NIR spectroscopy contain absorbance bands that are mainly due to three chemical bonds: C-H, which is usually for fats and oil; O-H, which is found in water; and N-H, which is found in protein (Cozzolino et al., 2008). Other chemical bonds may appear in overtone bands in the NIR region, but they are generally too weak to consider for analysis in complex food systems such as peanuts, which contain water, oil, fat, protein, etc. NIR spectroscopy may be applied with minimal sample preparation, and it has been used successfully in many other crops, including soybean (Nimaiyar et al., 2004), sunflower seed (Pérez-Vich et al., 1998), rapeseed (Velasco and Becker, 1998), canola (Daun et al., 1994), flaxseed (Bhatty, 1991), and for oil analysis. In addition, NIR has been used to determine the oleic and linoleic fatty acid concentrations of oil in individual peanut kernels (Tillman et al., 2006). The primary objective of this research is to develop calibration models to predict the MC of the peanut kernels using NIR reflectance on in-shell peanuts.

MATERIALS AND METHODS

SAMPLE PREPARATION

Approximately 50 kg of Virginia and Valencia type peanut pods were obtained from the 2008 and 2007 crops, respectively. The initial MC of the whole peanut pods was determined using the time/temperature protocol specified in the standard oven method (ASAE Standards, 1982). Sample size was reduced to about 100 g of each market type due to limited quantities. About 100 g of peanut sample was placed in a small aluminum pan, and this pan was kept in an air oven at 130 °C for 6 h. At the end of the drying time, the peanut samples were removed from the oven, allowed to reach room temperature in a desiccator, and weighed. The percentage MC of the peanut sample was calculated from the weight of the sample before and after drying, and expressed as percentage MC on wet basis (w.b.).

The initial MC of both Virginia and Valencia peanuts was about 6%. These peanut lots were divided into 12 Valencia and 15 Virginia sublots, and each was placed in a separate vapor-sealed, airtight plastic container. Appropriate quantities of water were added to each container to raise the moisture level in increments of 2% of the MC of the peanuts. The containers were sealed and allowed to equilibrate at 4 °C for one week. The containers were periodically rotated during this period to allow uniform moisture distribution. This resulted in 12 and 15 moisture levels for Valencia and Virginia type peanuts, respectively. Final MC ranged from 7% to 26%. After a week, the containers were removed from cold storage and allowed to equilibrate to room temperature before taking the NIR spectroscopy measurements. The final MC of each subplot after the equilibration was measured, using the standard oven method described above, for three 100 g subsamples from each subplot. The average of the three replicates was taken as the MC of each subplot (reference value), and each subplot was labeled accordingly.

NIR SPECTROSCOPY MEASUREMENTS

The peanut pod samples, after conditioning, were separated into calibration and validation groups (table 1). NIR spectral measurements were made using a scanning monochromator (model 6500, FOSS NIRSystems, Inc., Laurel, Md.). Spectral data were collected using Vision software (version 1.0, FOSS NIRSystems, Inc., Laurel, Md.). The calibration groups of in-shell Valencia and Virginia peanuts had seven and eight different moisture levels, respectively. Thirty replicates of each moisture level of each peanut type were scanned. Each replicate sample consisted of 100 to 200 g of peanut pods depending on the amount required to adequately fill the sample holder. The room temperature during the measurements varied from 21 °C to 23 °C. NIR light was allowed to fall on the bottom of the sample holder containing the in-shell peanuts, where it penetrated and interacted with the samples. The reflected energy was measured over the wavelength range of 400 nm to 2500 nm. This reflected light spectrum, which carried absorption information, was collected.

DATA ANALYSIS

The NIR spectral data were analyzed using multivariate data analysis software (Unscrambler, version 9.7, CAMO Software, Inc., Woodbridge, N.J.). Absorption values of the spectra between 400 nm and 2500 nm with 0.5 nm intervals were taken as independent variables, and the MC of the peanuts was taken as the dependent variable for the analysis. These absorption spectra were converted into reflection spectra using the Unscrambler (version 9.7) and saved as a separate file. Both the absorption and reflection spectra were then used for development of the MC prediction models. Absorption spectra were obtained from the energy absorbed by the sample at a given wavelength. Measurement of absorption spectra is a widely used technique in chemical analysis. Reflection spectra are generally obtained from the light energy that is scattered and reflected by the sample. This reflected light energy contains information about the sample components. Using data from the calibration dataset, partial least

Table 1. Moisture content (MC) levels of calibration and validation groups of in-shell peanuts.

<table>
<thead>
<tr>
<th>Peanut Type</th>
<th>Calibration Group MC (%)</th>
<th>Validation Group MC (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Virginia</td>
<td>6.9 7.6</td>
<td>10.8 11.3</td>
</tr>
<tr>
<td></td>
<td>14.9 14.1</td>
<td>15.8 15.0</td>
</tr>
<tr>
<td></td>
<td>19.4 15.4</td>
<td>22.3 18.8</td>
</tr>
<tr>
<td></td>
<td>23.7 22.8</td>
<td>26.7</td>
</tr>
<tr>
<td>Valencia</td>
<td>6.2 7.4</td>
<td>10.1 12.2</td>
</tr>
<tr>
<td></td>
<td>13.1 14.6</td>
<td>15.7 18.8</td>
</tr>
<tr>
<td></td>
<td>18.3 20.9</td>
<td>19.3 21.7</td>
</tr>
</tbody>
</table>

184 TRANSACTIONS OF THE ASABE
squares (PLS) regression analysis was conducted to develop an empirical equation to predict the peanut MC. Before analyzing the data, certain mathematical pretreatments, such as a derivative function, were applied to the raw spectral data for better spectral resolution. The derivative of the absorption and reflection spectral data with respect to wavelength was calculated. The derivative computation was used to remove any baseline shifting that may have occurred and resolve overlapping peaks. After applying the derivative treatment, the modified wavelength spectra showed more details by diminishing the low-frequency signals and enhancing the high-frequency signals of the sample composition concentration. Derivative analysis was limited to first order, because at higher-order derivatives, sensitivity may be lost in the characteristic peaks.

PLS analysis was performed on both the raw spectral data and the derivative data to determine the best calibration model based on the standard error of calibration (SEC) and the coefficient of multiple determinations ($R^2$). SEC is calculated as:

$$SEC = \left( \frac{1}{n-p-1} \sum_{i=1}^{n} e_i^2 \right)^{1/2}$$  \hspace{1cm} (1)

where $n$ is the number of observations, $p$ is the number of variables in the regression equation with which the calibration is performed, and $e_i$ is the difference between the observed value and the reference value for the $i$th observation.

SEC is the estimation of the calibration procedure, and $R^2$ allows determination of the amount of variation in the data predicted by the calibration models. The pretreatment and its PLS model that yielded the best fit were used to predict the peanut MC from the spectral data of the validation group of peanuts. Goodness of fit was evaluated based on the standard error of prediction (SEP), which was obtained by comparing the reference (oven measured) MC with the predicted MC:

$$SEP = \left( \frac{1}{n-1} \sum_{i=1}^{n} (e_i - \bar{e})^2 \right)^{1/2}$$  \hspace{1cm} (2)

where $n$ is the number of observations, $e_i$ is the difference between the predicted MC and the MC determined by the reference method for the $i$th sample, and $\bar{e}$ is the mean of $e_i$ for all of the samples.

The ratio of the standard deviation of the predicted MC values to the SEP, which is the residual predictive deviation (RPD), was also used to evaluate the goodness of fit. RPD values usually range from 1 to 10. Higher values indicate a stronger calibration model for accurately predicting the composition of unknown samples. RPD values of 1 or less indicate that the model predicts the same values only by random chance. RPD values in the range of 3.1 to 4.9 are good for screening purposes, and values in the range 5 to 6.4 are good for quality control purposes (Fearn, 2002; Williams, 2001). Bias values, which correspond to the average difference between the standard reference and predicted values, were also considered in model selection.

**RESULTS AND DISCUSSION**

Strong NIR absorption bands near 1400-1440 nm and 1900-1950 nm have often been applied to quantitative analysis of moisture content in foods. Changes in absorption bands at the 1400 nm and 1900 nm regions are caused by the presence of water in food (Buning-Pfaue, 2003).

Figure 1 shows the NIR absorption spectra of Virginia (fig. 1a) and Valencia (fig. 1b) type in-shell peanuts.

Figure 1. NIR absorption spectra of Virginia (a) and Valencia (b) type in-shell peanuts.
2500 nm, a combination of one or more overtone bands accounts for the spectral absorption of water.

Figure 2 shows the derivative spectra of in-shell Virginia (fig. 2a) and Valencia (fig. 2b) peanuts. The first-derivative spectra of both peanut types have a trough corresponding to each MC peak in the original absorption spectra. In this figure, absorption bands indicate MC at 940 nm, 1134 nm, and 2030 nm, other than the peaks shown in the raw NIR absorption spectra (fig. 1). The first-derivative curves at different MC levels show slightly different peak positions for each level. For example, for the Valencia peanuts (fig. 2a) at 21.7% MC, the peak positions are 1890 nm, 1393 nm, and 1134 nm, whereas at 13.1% MC, the peak positions are 1905 nm, 1413 nm, and 1154 nm. At lower moisture levels, the absorption spectra shifted about 20 nm higher. Since the derivative curves show modified wavelength spectra with more details by enhancing the high frequencies, the shifts in peaks are more clearly seen. Water absorption bands are influenced by the effects of solutes in water, such as ions and organic monomers, as well as hydrogen bonds, with which the water bonds are strongly associated. The shift of absorption peaks to lower or higher wavelengths is related to the hydration potential of the solutes and the strength of the hydrogen bonds. In the peanut samples, strong hydrogen bonds might have been present, which could have influenced the NIR absorption.

In pure water, hydrogen atoms that attach with O-H groups are considered free hydrogen, i.e., not attached with any other molecules. These free-hydrogen O-H groups create stronger absorption peaks at lower wavelengths, such as for pure water at 975 nm (Marvin and Singh, 2004; Inoue et al., 1984). However, O-H groups with hydrogen atoms that are bonded with carbon, nitrogen, or some other atoms, as in food systems, create water absorption peaks at higher wavelengths. Therefore, as the MC in the peanuts decreased, there might have been an increased number of O-H groups with bonded hydrogen atoms (i.e., hydrogen atom bonded with other atoms), as mentioned above, which shifted the water absorption peaks toward higher wavelengths for low-moisture samples.

Figure 3 shows the regression coefficients of the calibration models using absorbance and its derivative, as developed using PLS. The wavelengths corresponding with many of the peaks shown in this figure are related with MC. The figure shows that the regression coefficients were obtained from both positive and negative absorption peaks. The peaks ranging between 1400-1440 nm and 1884-1930 nm contributed significantly to the regression coefficients of the calibration equation, and these wavelengths are the strong absorption bands of water molecules. The $R^2$ and SEC of the calibration regression curves are given in tables 2a and 2b.

Tables 2a and 2b show the fitness measures of the calibration group of in-shell peanuts obtained using absorbance and reflectance spectra, respectively. Valencia type peanuts with a calibration set of 210 gave an $R^2$ of 0.99 for each of the spectral data treatments as well as for the raw data. Similarly, 240 calibration samples of Virginia type in-shell peanuts gave an $R^2$ of 0.99. The calibration model equations were used to predict the MC of the 150 Valencia type and 180 Virginia type samples in the validation sample set. Tables 3a and 3b show the fitness measures of the validation set after prediction. The predicted RPD and SEP obtained for the validation samples are given in these tables.

The best calibration model for Valencia in-shell peanuts was found to be the reflectance derivative spectral data model, based on the lowest SEC (0.731) and bias (-0.007) values (table 2b). Among the fitness measurement parameters shown in table 3b for the validation group of Valencia peanuts, the model with reflectance derivative data gave an unsatisfactory RPD of 2.36 with an SEP of 1.93. The reflectance spectra with no treatment gave a much better RPD of 5.55, which indicates that this model is good for quality control and analysis. The corresponding SEP of 0.977 for this model is also reasonable. Thus, the model with reflectance spectra without applying any treatment was selected as the best model for Valencia in-shell peanuts. RPD values were used in the best model selection according to the description of RPD and its values, as given in Data Analysis section.

In the case of Virginia type in-shell peanuts, the models developed with both absorption spectral data and its derivative were found to be the best, with the lowest SEC and bias values (table 2a). Looking at the fitness measurements of the different models in table 3a and 3b, absorbance with the first-
Table 2a. Fitness measurements of calibration group model developed using absorbance spectra.

<table>
<thead>
<tr>
<th>Spectral Treatment</th>
<th>R^2</th>
<th>SEC</th>
<th>RMSEC</th>
<th>Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Virginia type peanuts</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No treatment</td>
<td>0.99</td>
<td>0.70</td>
<td>0.653</td>
<td>+0.004</td>
</tr>
<tr>
<td>First derivative</td>
<td>0.99</td>
<td>0.720</td>
<td>0.673</td>
<td>-0.001</td>
</tr>
<tr>
<td>Valencia type peanuts</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No treatment</td>
<td>0.99</td>
<td>0.969</td>
<td>0.898</td>
<td>-0.03</td>
</tr>
<tr>
<td>First derivative</td>
<td>0.99</td>
<td>1.528</td>
<td>1.416</td>
<td>-0.057</td>
</tr>
</tbody>
</table>

Table 2b. Fitness measurements of validation group model developed using reflectance spectra.

<table>
<thead>
<tr>
<th>Spectral Treatment</th>
<th>R^2</th>
<th>SEP</th>
<th>RPD</th>
<th>RMSEC</th>
<th>Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Virginia type peanuts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No treatment</td>
<td>0.87</td>
<td>1.725</td>
<td>2.51</td>
<td>1.793</td>
<td>-0.817</td>
</tr>
<tr>
<td>First derivative</td>
<td>0.97</td>
<td>1.93</td>
<td>2.36</td>
<td>1.735</td>
<td>+0.189</td>
</tr>
<tr>
<td>Valencia type peanuts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No treatment</td>
<td>0.86</td>
<td>1.977</td>
<td>5.55</td>
<td>1.782</td>
<td>+1.55</td>
</tr>
<tr>
<td>First derivative</td>
<td>0.87</td>
<td>1.93</td>
<td>2.36</td>
<td>1.735</td>
<td>+0.189</td>
</tr>
</tbody>
</table>

derivative model gave an RPD of 3.04, which is good for screening purpose only. Reflectance spectral data (table 3b) with derivative treatment had the highest RPD of 5.75 and the lowest SEP of 0.771. Therefore, for Virginia in-shell peanuts, the model developed using reflectance derivative spectral data was selected as the best model for quality control and analysis. For both Virginia and Valencia in-shell peanuts, the absorption derivative model could be used for screening purposes. Screening based on MC would be helpful at the peanut buying point to make grading faster and save time.

draphical comparisons between the reference and NIR-predicted MC values obtained using the various regression equations and validation datasets are given in figures 4 and 5. These figures show the individual predicted MC values obtained from 30 replicates of the validation set in the y-axis and three replicates of reference (oven measured) values in the x-axis. Good correlation between the predicted and reference MC values of both Valencia and Virginia type in-shell peanuts was achieved. Both NIR absorption and reflection data and their derivatives resulted in R^2 values in excess of 88%.
Figure 4. Comparison of NIR-predicted values and reference values measured using air oven for Virginia type in-shell peanuts.

Figure 5. Comparison of NIR-predicted values and reference values measured using air oven for Valencia type in-shell peanuts.
Based on the SEP and RPD values (table 3b), the reflection calibration model was considered a good model for quality control and analysis because the safe storage quality of peanuts is determined mainly by their MC. The reflectance spectral derivative calibration model for in-shell Virginia type peanuts and the reflectance spectral calibration model for in-shell Valencia type peanuts were selected for prediction of unknown moisture samples. Generally, the height and shape of the spectra are dependent on the scattering of light. This scatter is affected by differences in sample particle size and by the reflective nature of the sample surface. Virginia type peanuts differ in particle size and surface color from Valencia type peanuts. This might have changed the spectral baseline of Virginia type peanuts. To remove this baseline shifting, a derivative computation is needed. This could be the reason for the high RPD value and goodness of fit using the reflectance spectral derivative calibration model for Virginia type peanuts.

CONCLUSION

NIR reflectance spectroscopy works well for analysis of the moisture content of in-shell peanuts with minimal sample preparation. The calibration models were obtained using partial least squares regression analysis. NIR measurement is procedurally very simple, considerably reducing the time required for measurement compared to the standard oven method and conventional moisture meters, which require samples to be shelled. The use of NIR spectroscopy as described in this article would result in large savings in time and labor during the drying and storage of peanuts. Regression coefficient peaks corresponded to the respective wavelengths of water. By considering the lowest SEP and the highest RPD values, a model developed with reflectance data was selected as the preferred calibration model for MC prediction of Valencia type peanuts. The reflectance spectral derivative model was preferred for Virginia type peanuts. Calibration tests are underway for runner type peanuts. For commercial application of this technique, validation tests are required for cultivars within each market type from various production regions.

REFERENCES
