1991

Moisture Determination in Single Soybean Seeds by Near-Infrared Transmittance

D. T. Lamb
_Iowa State University_

Charles R. Hurburgh Jr.
_Iowa State University, tatry@iastate.edu_

Follow this and additional works at: http://lib.dr.iastate.edu/abe_eng_pubs
Part of the Agriculture Commons, and the Bioresource and Agricultural Engineering Commons

The complete bibliographic information for this item can be found at http://lib.dr.iastate.edu/abe_eng_pubs/427. For information on how to cite this item, please visit http://lib.dr.iastate.edu/howtocite.html.

This Article is brought to you for free and open access by the Agricultural and Biosystems Engineering at Iowa State University Digital Repository. It has been accepted for inclusion in Agricultural and Biosystems Engineering Publications by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
Moisture Determination in Single Soybean Seeds by Near-Infrared Transmittance

Abstract
A calibration set of 140 soybeans from seven different varieties, ranging in moisture from 5 to 22%, wet basis, was used to calibrate a spectrophotometer for predicting moisture of single soybean seeds. Near-infrared absorbance (A) of individual soybean seeds was measured over the spectral region from 800 to 1100 nm by 0.5 nm increments. The axis of the soybean seed parallel to the incident light beam was measured as an estimate of optical pathlength. Three mathematical techniques were used to develop calibration equations: linear correlation with a difference in absorbance (AA), stepwise multiple linear regression (MLR), and partial least squares (PLS). A validation set contained 100 soybeans from 5 different varieties, ranging in moisture from 5 to 20%, wet basis. The standard error of prediction (SEP) for equations using absorbance data only was 0.88% for AA, 0.82% for MLR, and 0.81% for PLS. The SEP for equations using pathlength and absorbance data was 0.73% for AA, 0.69% for MLR, and 0.65% for PLS.

Keywords
Near-infrared transmission Soybean, Moisture

Disciplines
Agriculture | Bioresource and Agricultural Engineering

Comments
This article is from Transactions of the ASAE (1991): 2123–2129. Posted with permission.
MOISTURE DETERMINATION IN SINGLE SOYBEAN SEEDS BY NEAR-INFRARED TRANSMITTANCE

D. T. Lamb, C. R. Hurburgh, Jr.
MEMBER
ASAE

ABSTRACT
A calibration set of 140 soybeans from seven different varieties, ranging in moisture from 5 to 22%, wet basis, was used to calibrate a spectrophotometer for predicting moisture of single soybean seeds. Near-infrared absorbance (A) of individual soybean seeds was measured over the spectral region from 800 to 1100 nm by 0.5 nm increments. The axis of the soybean seed parallel to the incident light beam was measured as an estimate of optical pathlength. Three mathematical techniques were used to develop calibration equations: linear correlation with a difference in absorbance (ΔA), stepwise multiple linear regression (MLR), and partial least squares (PLS). A validation set contained 100 soybeans from 5 different varieties, ranging in moisture from 5 to 20%, wet basis. The standard error of prediction (SEP) for equations using absorbance data only was 0.88% for ΔA, 0.82% for MLR, and 0.81% for PLS. The SEP for equations using pathlength and absorbance data was 0.73% for ΔA, 0.69% for MLR, and 0.65% for PLS. Keywords. Near-infrared transmission, Soybean, Moisture.

INTRODUCTION
Near-infrared analysis is a valuable tool for proximate analysis of grains. Near-infrared reflectance (NIR) is in common use for the analysis of ground grain samples (Hurburgh, 1988) and has been investigated for whole-grain samples (Stermer et al., 1977; Tkachuk, 1987). Near-infrared transmittance (NIT) of whole-grain samples is becoming increasingly popular (Williams et al., 1985).

Whole-grain NIT is advantageous because of the sample preservation and simple sample preparation. Nondestructive testing provides for replication, additional tests, and storage of the sample for future analysis. Plant breeders could use a nondestructive, single-seed NIT to isolate individual seeds for genetic improvement without the loss of seed stock.

Low-noise transmittance measurements of whole grain can be made in the spectral region from 800 to 1100 nm (Norris, 1983). This region has a water absorption peak at 970 nm and a weak water absorption band at 845-850 nm (Curcio and Petty, 1951) making it a possible region for moisture prediction.

Norris and Hart (1965) used NIT to predict the moisture content of individual, intact peanuts and lima beans. They predicted moisture by using a linear correlation with the difference in optical density at 970 and 900 nm. Finney and Norris (1978) used NIT to predict the moisture content of single kernels of corn by using a linear correlation with the difference in optical density at 942 and 931 nm. The standard error of the estimate was 2.78% moisture, wet basis (w.b.). Norris (1983) used an average transmission spectrum to predict moisture content of single sunflower seeds from ostensibly uniform sample lots. Moisture was predicted with a standard error of prediction of 0.36% by using a ratio of second derivatives. Oil content of individual soybeans was predicted by an identical method with a standard error of prediction of 0.50%. No information could be found on the moisture prediction of individual soybean seeds by NIT. Bulk, whole-grain near-infrared analysis tests for soybean protein and oil have a standard error of prediction of 0.6% and 0.5%, respectively (Federal Grain Inspection Service, USDA, 1989).

It is vital that NIT analysis be able to predict moisture content accurately so as to determine the moisture basis for reporting protein and oil content. Protein and oil content are intrinsic characteristics of soybean seed dry matter and determine the value of the commodity (Brumm and Hurburgh, 1990). Moisture content of a soybean can be changed by drying or rewetting. Plant breeders usually report constituent percentages on a dry-matter basis (0% moisture), but the U.S. national grain inspection system reports soybean constituent percentages on a 13% moisture basis (Federal Grain Inspection Service, USDA, 1989).

Typically, NIT measures composition of bulk samples. In certain applications, notably plant breeding, single-seed analysis would be valuable. Of the compositional factors, moisture is the easiest to measure and has the strongest spectral response. Therefore, moisture prediction of single soybean seeds was chosen as a precursor to protein and oil prediction.

OBJECTIVE
The objective of this study was to establish the feasibility of predicting individual soybean seed moisture content with near-infrared transmittance analysis (NIT).
MATERIALS AND METHODS

SAMPLES

Soybeans were obtained from the Iowa State University Grain Quality Laboratory. The calibration set contained soybeans from seven varieties from the 1989 Iowa Soybean Yield Test (Iowa State University, 1989). The validation set contained soybeans of five varieties from the American Soybean Association national soybean quality survey for 1989. Both sets of samples were received at approximately 5% moisture (w.b.). The composition of the sample sets is shown in Table 1.

LABORATORY ANALYSIS

The samples were hand-picked to remove all foreign material, splits, and damaged beans. The cleaned sample was sized for the spectrophotometer mount with a 6.4-mm (16/64 in.) round-hole hand screen. The beans remaining on top of the screen were retained. The sized sample was divided into four equal subsamples with a Boemr divider.

To obtain a range of moistures, the subsamples were rewetted by using a humidifier. Three subsamples were placed in mesh baskets and located in the path of the humid air for various lengths of time. The original and rewetted subsamples were stored in sealed mason jars for 10 days at 4.4° C (40° F). The jars were occasionally tumbled for equilibrium. Oven moistures of three whole-grain samples, approximately 10 g each, were averaged to obtain the moisture content of the subsample (Hartwig and Hurburgh, 1989). The final range of moistures is shown in Table 2.

From each subsample, five beans were selected for analysis. The calibration data set contained 140 beans (seven samples, four moisture groups, five beans per group). The validation set contained 100 beans. The axis of the bean perpendicular to the plane that separated the two cotyledons was measured to the nearest 0.03 mm (0.001 in.) with a micrometer caliper. Each bean was weighed to the nearest 0.0001 g, and its spectral absorbance was immediately recorded. The bean was then dried in an air-oven at 130° C (266° F) for 3 h and reweighed. Moisture content was calculated from the weight change as a percentage (w.b.). Typically, bulk-sample air-oven methods have published precisions of about 0.2 percentage points, although there is no way to verify that on single seeds.

Absorbance spectra were recorded by using a Shimadzu W-160 dual-beam spectrophotometer. The spectrophotometer contained an RS-232 communications port for external computer control and data transfer. The wavelength range from 800 to 1100 nm was scanned, and absorbance values were recorded by 0.5 nm increments. The 601-point curves were recorded by the computer, along with the oven moisture data and bean dimensions.

The 601-point spectral curves were smoothed by using a 5-point (1 nm gap width) moving average to reduce the influence of noise in the spectra:

\[ A_n = \frac{A_{n-2} + A_{n-1} + A_n + A_{n+1} + A_{n+2}}{5} \]  

where \( A_n \) is the absorbance in the \( n \)-th 0.5-nm wavelength increment (\( n = 3 \) to 608). The integer wavelength values were retained, which reduced the spectral data to 299 wavelengths, 801 to 1099 nm in 1 nm increments. The reduced data set was needed to accommodate the capabilities of the statistical packages used in calibration.

The bean was positioned in the sample beam path by a spring-loaded mount (fig. 1). The mount was fabricated from backing material for a near-infrared reflectance sample cup and two black rubber O-rings. The mount had a 6.4-mm (0.25 in.) opening to allow the incident radiation to pass through the bean. The O-rings held the bean in place and prevented radiation from passing around the bean. The bean was positioned with its hilum towards the top of the mount. The plane that separated the two cotyledons was perpendicular to the incident radiation beam. The measurement of the axis of the bean perpendicular to the

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Protein (%)</th>
<th>Oil (%)</th>
<th>Variety</th>
</tr>
</thead>
<tbody>
<tr>
<td>AB0001</td>
<td>32.8</td>
<td>19.7</td>
<td>Variety names from</td>
</tr>
<tr>
<td>AB0021</td>
<td>33.0</td>
<td>20.3</td>
<td>Iowa Soybean Yield</td>
</tr>
<tr>
<td>AB0881</td>
<td>31.4</td>
<td>21.4</td>
<td>Test not available by university policy.</td>
</tr>
<tr>
<td>AB1051</td>
<td>33.5</td>
<td>20.3</td>
<td></td>
</tr>
<tr>
<td>AB0951</td>
<td>32.8</td>
<td>20.3</td>
<td></td>
</tr>
<tr>
<td>AB0521</td>
<td>35.3</td>
<td>19.6</td>
<td></td>
</tr>
<tr>
<td>AB0141</td>
<td>35.8</td>
<td>18.4</td>
<td></td>
</tr>
<tr>
<td>AB0321</td>
<td>33.8</td>
<td>18.6</td>
<td>Golden Harvest H-1308</td>
</tr>
<tr>
<td>AB1052</td>
<td>35.0</td>
<td>18.4</td>
<td>Asgrow 1937</td>
</tr>
<tr>
<td>AB1192</td>
<td>34.6</td>
<td>19.4</td>
<td>Northrup King 42.40</td>
</tr>
<tr>
<td>AB1972</td>
<td>34.3</td>
<td>19.6</td>
<td>Asgrow 2234</td>
</tr>
<tr>
<td>AB2032</td>
<td>31.3</td>
<td>20.6</td>
<td>Pioneer 9271</td>
</tr>
</tbody>
</table>

* Sample ID: ABCCCD, where A = moisture group, B = bean 1-5, CCC = source, D = 1 for calibration, 2 for validation.

† 13% moisture basis as measured on bulk sample.

TABLE 2. Air-oven moisture contents for calibration and validation sets

| Sample ID| Average moisture (%) by group*
|----------|-----------------
| Calibration | 1  | 2  | 3  | 4  |
| AB0901    | 4.8 | 5.4 | 7.5 | 11.2 |
| AB0021    | 5.6 | 7.6 | 15.3 | 16.6 |
| AB0881    | 5.0 | 5.6 | 8.2 | 15.8 |
| AB1051    | 5.0 | 6.0 | 9.2 | 18.6 |
| AB0951    | 5.4 | 7.8 | 10.6 | 18.4 |
| AB0521    | 5.1 | 5.6 | 8.6 | 22.6 |
| AB0141    | 5.2 | 6.0 | 8.1 | 18.6 |
| Validation|     |     |     |     |
| AB0532    | 5.8 | 11.5 | 8.4 | 14.7 |
| AB1052    | 5.7 | 9.8 | 9.2 | 15.1 |
| AB1192    | 6.3 | 10.7 | 10.8 | 19.9 |
| AB1972    | 5.8 | 6.5 | 7.4 | 9.8 |
| AB2032    | 5.8 | 9.3 | 10.3 | 12.3 |

* Average of triplicate whole grain samples (Hartwig and Hurburgh, 1989).
† Sample ID: ABCCCD, where A = moisture group, B = bean 1-5, CCC = source, D = 1 for calibration, 2 for validation.
plane separating the cotyledons was used to estimate the optical pathlength (t).

**ABSORBANCE MATHEMATICS**

The dual-beam spectrophotometer is a monochromator that is computer controlled to scan through a range of wavelengths at preset intervals. A beam splitter divides the output of the monochromator into two beams, a reference beam and a sample beam. Absorbance is defined as:

\[
A = \log \left( \frac{I_r}{I_s} \right) \tag{2}
\]

where \(I_r\) is the intensity of the reference beam, equal in intensity to the radiation incident on the sample. \(I_s\) is the intensity of the beam that has passed through the sample. Transmittance (T) is the ratio of radiation passing through the sample to the radiation incident on the sample:

\[
T = \frac{I_s}{I_r} \tag{3}
\]

Absorbance, in terms of transmittance is:

\[
A = \log \left( \frac{1}{T} \right). \tag{4}
\]

Absorbance is also called optical density (OD). Because the absorbance of the bean exceeded the normal scale for the spectrophotometer, an aperture was placed in the reference beam path. A baseline shift was produced that allowed the spectrophotometer to measure absorbance. This was needed because the optical density of the seed reduced \(I_s\) to an intensity small enough that the absorbance was off scale for the spectrophotometer. The aperture reduced \(I_r\) and, therefore, shifted the spectrophotometer scale.

The absorbance reading for the empty seed mount and aperture, \(A_1\), is:

\[
A_1 = \log \left( \frac{I_{aperture}}{I_{mount}} \right). \tag{5}
\]

\(I_{aperture}\) is the intensity of the beam after passing through the aperture; \(I_{mount}\) is the intensity of the beam after passing through the seed mount. The absorbance reading for the seed mount with seed and aperture, \(A_2\), is:

\[
A_2 = \log \left( \frac{I_{aperture}}{I_{mount} + seed} \right). \tag{6}
\]

\(I_{mount} + seed\) is the intensity of the beam after passing through the seed held in the mount. \(A_3\) is the absorbance of the soybean seed held by the mount:

\[
A_3 = A_2 - A_1 = \log \left( \frac{I_{mount} + seed}{I_{mount}} \right). \tag{7}
\]

Absorbance is also called optical density (OD).

Beer's law relates absorbance and constituent concentration:

\[
A = \varepsilon c t \tag{9}
\]

where \(A\) is the absorbance, \(\varepsilon\) is the molar absorptivity, \(c\) is the molar concentration, and \(t\) is the optical pathlength. In a complex medium such as a whole seed, optical pathlength is not easily defined. Therefore, Beer's law, using measured geometrics, is only an approximation of light diffusion. Solving for concentration:

\[
c = \frac{1}{\varepsilon} \left( \frac{1}{t} \right) A. \tag{10}
\]

The concentration of a sample can be determined by a reference method. Absorbance and pathlength can be measured, leaving molar absorptivity as the only unknown. Ideally, the value of \(\varepsilon\) can be solved algebraically, and only one wavelength absorbance would be needed. But, because of interference from other constituent concentrations and absorbance overlap of constituents, the final calibration equation needs multiple absorbance values. If absorbance is correlated strongly with pathlength and pathlength is only weakly (or not at all) related to the constituent concentration, then absorbance alone can be used to determine a calibration equation.

**CALIBRATION MATHEMATICS**

Three mathematical techniques were used to produce prediction equations: linear correlation with the difference in absorbance at two wavelengths (\(\Delta A\)) (Hruschka, 1987; Finney and Norris, 1978; Norris and Hart, 1965), stepwise multiple linear regression (MLR) (Hruschka, 1987), and partial least squares (PLS) (Martens and Naes, 1987).

\(\Delta A\) is limited to two absorbance wavelengths and involves two steps, the selection of wavelengths and the linear regression of the absorbance difference against moisture content. The wavelengths were selected by an iterative technique that converged to the pair of wavelengths with the largest absolute correlation with moisture. The wavelength absorbance with the largest
absolute correlation with moisture, $A_A$, was selected. All remaining absorbances were subtracted from $A_A$, and the difference with the largest absolute correlation to moisture ($A_A - A_B$) was determined. $A_A$ was replaced by $A_B$, the differences were again calculated, and a new $A_B$ was found. The process continued until the selection converged to an $A_A$ and $A_B$ that selected each other. An interactive program was written to perform the wavelength selection.

The prediction equation for $\Delta A$, absorbance values only was:

$$M = b_0 + b_1 (A_A - A_B)$$  \hspace{1cm} (11)

where $b_0$ is the intercept, $b_1$ is a regression constant, and $M$ is moisture (%). Pathlength, $\ell$, was included in $\Delta A$ by dividing all wavelengths by $\ell$, per Beer's law. The prediction equation for $\Delta A$ with $\ell$ included was:

$$M = b_0 + (b_1 / \ell) (A_A - A_B).$$  \hspace{1cm} (12)

MLR is not limited to a set number of absorbance wavelengths. Absorbance wavelengths to include in the prediction model were selected by a stepwise technique (SAS Institute Inc., 1985). The stepwise technique selected only wavelengths that contributed most to the accuracy of the model prediction at a preset level of statistical significance.

The prediction equation for MLR, absorbance values only was:

$$M = b_0 + b_1 (A_1) + b_2 (A_2) + \ldots + b_n (A_n)$$  \hspace{1cm} (13)

where $b_0$ is the intercept, and $b_1, \ldots, b_n$ are regression constants for the selected absorbance wavelengths $A_1, A_2, \ldots, A_n$. Pathlength was included in MLR as suggested by Beer's Law:

$$M = b_0 + b_1 (A_1 / \ell) + b_2 (A_2 / \ell) + \ldots + b_n (A_n / \ell).$$  \hspace{1cm} (14)

PLS with orthogonal regression factors was developed by Martens and Naes (1987). Two programs were written to implement the technique: one to calculate the calibration factors and a companion program to predict sample moisture content by using the calibration factors. Pathlength was included in PLS by considering the factors to be $A_n / \ell$.

### Statistical Methods

Root mean square errors (RMSE) for calibration and validation were calculated as:

$$\text{RMSE} = \left( \sum d_i^2 / (n - 1) \right)^{1/2}$$  \hspace{1cm} (15)

where

$$d_i = (\text{Oven}_i - \text{Predicted}_i)$$  \hspace{1cm} (16)

$\text{Oven}_i = \text{moisture determined by air-oven method}$ for the method for the $i$th sample, for the $i$th sample,

$\text{Predicted}_i = \text{moisture predicted by calibration equation for the } i\text{th sample}$, $n = \text{number of samples}$.

The optimum number of regression constants was selected by observing the RMSE for the calibration and validation sets. The RMSE for calibration decreases for increasing number of constants as the MLR method tracks more and more detail of the calibration set. The RMSE for the validation set should decrease to a minimum, then increase as wavelengths are no longer useful for constituent prediction, being related only to characteristics of the specific calibration set. The MLR equation that produced the minimum RMSE for validation was selected for use.

The standard error of prediction (SEP), used for validation data, was:

$$\text{SEP} = \left( \sum (d_i - \text{Bias})^2 / (n - 1) \right)^{1/2}$$  \hspace{1cm} (17)

where

$$\text{Bias} = \left( \sum (\text{Oven}_i) - \sum (\text{Predicted}_i) \right) / n.$$  \hspace{1cm} (18)

An F-ratio was used to compare root mean square errors (Finney and Norris, 1978):

$$F = \text{RMSE}_1^2 / \text{RMSE}_0^2$$  \hspace{1cm} (19)

where $\text{RMSE}_1$ is the larger of $\text{RMSE}_0$, $\text{RMSE}_1$, $\text{RMSE}_0$ and $\text{RMSE}_1$ were both from calibration or validation data of two different prediction techniques. This test established the statistical equivalence or nonequivalence of the two prediction techniques.

### Results and Discussion

The moisture range of the single soybeans tested is summarized in Table 3. The pathlength statistics are summarized in Table 4. The calibration equations were developed and are summarized in Table 5.

The number of regression constants to include for the MLR prediction equation was determined by the RMSE performance. The RMSE for validation and calibration using the MLR for absorbance data only is shown in

<table>
<thead>
<tr>
<th>TABLE 3. Oven moisture content of single soybean seeds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set</td>
</tr>
<tr>
<td>-----</td>
</tr>
<tr>
<td>Calibration</td>
</tr>
<tr>
<td>Validation</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE 4. Pathlength ($\ell$) of single soybean seeds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set</td>
</tr>
<tr>
<td>-----</td>
</tr>
<tr>
<td>Calibration</td>
</tr>
<tr>
<td>Validation</td>
</tr>
</tbody>
</table>

2126

TRANSACTIONS OF THE ASAE
### TABLE 5. Summary of calibration and validation statistics

<table>
<thead>
<tr>
<th>Method</th>
<th>Calibration</th>
<th>Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RMSE(^*)</td>
<td>RMSE(^*)</td>
</tr>
<tr>
<td></td>
<td>(%)</td>
<td>(%)</td>
</tr>
<tr>
<td>Path length not included</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA(^\ddagger)</td>
<td>0.93(^A) A</td>
<td>0.90(^A) A</td>
</tr>
<tr>
<td>MLR(^\ddagger)</td>
<td>0.82(^A,B) A,B</td>
<td>0.84(^A,B) A,B</td>
</tr>
<tr>
<td>PLSI</td>
<td>0.82(^A,B) A,B</td>
<td>0.81(^A,B,C) A,B,C</td>
</tr>
<tr>
<td>Path length included</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA(^#)</td>
<td>0.76(^B,C) B,C</td>
<td>0.73(^B,C) B,C</td>
</tr>
<tr>
<td>MLR(^**)</td>
<td>0.63(^C,D) C,D</td>
<td>0.69(^B,C) B,C</td>
</tr>
<tr>
<td>PLS(^\ddagger)</td>
<td>0.62(^D) D</td>
<td>0.65(^C) C</td>
</tr>
</tbody>
</table>

* RMSE that are not statistically different (p = 0.01) indicated by like letters, vertical comparisons only.
† Slope of the regression line of residuals vs. oven moisture.
‡ M = 13.43 + 159.74 (A\(_{561\ nm}\) - A\(_{918\ nm}\)) R\(^2\) = 0.97.
§ M = 12.70 + 43.78 (A\(_{516\ nm}\) - 220.36 (A\(_{905\ nm}\)) + 176.59
(A\(_{960\ nm}\)) R\(^2\) = 0.98.
\# 6 factor prediction using absorbance data only.
** M = 13.72 + (1001.01/t) (A\(_{960\ nm}\) - A\(_{918\ nm}\)) R\(^2\) = 0.98.
\** M = 10.90 + 191.26 (A\(_{561\ nm}\)) - 172.47 (A\(_{911\ nm}\)) + 783.84
(A\(_{936\ nm}\)) / t R\(^2\) = 0.99.
\†† Different from 0.00 at p < 0.05.
\‡‡ Different from 0.00 at p < 0.01.

Figure 2. The RMSE for validation decreased for the first three terms and then increased slightly for successive terms. The RMSE for validation using the MLR for absorbance and pathlength data did not reach a statistical minimum in the first six constants (fig. 3). However, the difference between the RMSE for validation and calibration was a minimum after four constants were included. This was also observed for the three-constant MLR for absorbance data only.

The number of factors to include for the PLS prediction was found by observing the RMSE for validation (figs. 4 and 5). The RMSE for validation was at a minimum after the first six factors and then increased for additional factors. This was observed for both the absorbance data alone and the absorbance with pathlength data. This demonstrated that excessive factors were fitting to unique characteristics of the calibration set. The minimum RMSE for validation was also the minimum difference between RMSE's for validation and calibration.

Absorbance was related to pathlength for all wavelengths as shown in figure 6, where all r > 0.6. The correlation of pathlength and moisture content was r = -0.25. Pathlength was apparently shorter for wetter beans. The relationship was weak, and the negative correlation may actually be due to the caliper deforming...
the softer, wetter beans. The correlations of moisture content and pathlength with absorbance for the wavelengths in the MLR absorbance-only equation are shown in Table 6. The correlation of moisture content and pathlength was much smaller than the correlation of absorbance and pathlength.

If the effect of pathlength is a shift of the absorbance spectrum, similar to particle size effects in NIR (Hruschka, 1987), the slope of the line relating absorbance and pathlength should be a constant for all wavelengths. The data in Table 6 indicate that the slope was a constant. Prediction equations using absorbance values only must compensate for the shift in the spectra. AA is not influenced by a shift because it uses a difference of absorbance values. PLS processes the spectrum as a whole and is dependent on the absorbance values and the general correlation with pathlength to compensate for the shift. There will be no influence of a spectral shift on the MLR prediction if the summation of the regression constants $(b_1, b_2, ..., b_n)$ is 0. The summation for the three-term MLR equation was 0.01.

The pathlength data improved the prediction of moisture for all methods. PLS produced marginally the best performance but was never statistically different from the MLR technique. PLS residuals from the validation set are shown in figure 7. The MLR technique required only a few wavelength values in comparison with the complete spectrum utilized by PLS. The MLR or AA approach could easily be implemented with a fixed-filter instrument, as opposed to a monochromator.

### CONCLUSIONS

Moisture prediction equations were successfully developed for single-soybean-seed NIT analysis. Partial least squares (PLS) and multiple-linear regression (MLR) using absorbance and pathlength data provided essentially equal calibrations (RMSE = 0.62%, 0.63%) and validation performance (SEP = 0.65%, 0.69%). Absorbance difference (AA) was the poorest performing technique. Including the pathlength data significantly improved the calibration set statistics for all calibration methods.

NIT appears feasible for estimating moisture content of single soybean seeds. Therefore, individual seed protein and oil content predictions should be investigated.

### TABLE 6. Absorbance wavelength statistics for three term MLR, absorbance only

<table>
<thead>
<tr>
<th>Wavelength (nm)</th>
<th>Correlation coefficient of Moisture (M)</th>
<th>Correlation coefficient of Pathlength (t)</th>
<th>Slope*</th>
</tr>
</thead>
<tbody>
<tr>
<td>816</td>
<td>-0.76</td>
<td>0.62</td>
<td>0.22^A</td>
</tr>
<tr>
<td>905</td>
<td>-0.74</td>
<td>0.64</td>
<td>0.23^A</td>
</tr>
<tr>
<td>960</td>
<td>-0.63</td>
<td>0.70</td>
<td>0.21^A</td>
</tr>
</tbody>
</table>

* Linear regression of A vs. t. Slope in Absorbance/mm. Values not statistically different (p = 0.01) indicated by like letter.
REFERENCES


