

# Single Wheat Kernel Analysis by Near-Infrared Transmittance: Protein Content

STEPHEN R. DELWICHE<sup>1</sup>**ABSTRACT**

Cereal Chem. 72(1):11-16

Regulatory agencies within the USDA have expressed an interest in single kernel analysis of grain for determination of basic intrinsic properties such as protein content. A study was undertaken to determine the feasibility of measuring protein content of intact wheat kernels using near-infrared transmittance spectrophotometry. Six wheat classes were examined: hard red winter, hard red spring, hard white, soft red winter, soft white, and durum. Five samples per class were used, with each sample represented by 96 randomly drawn kernels. Within each class, single kernel spectra (850–1,050 nm) were calibrated to single kernel protein content using four samples; the remaining sample was set aside for model validation. The chemometric method of partial least squares (PLS) analysis was used. For reference values, the protein content ( $N \times 5.7$ ) of each

kernel was measured by combustion. Typical model accuracies of  $r^2$  in the range of 0.85–0.93 and standard errors of prediction of 0.4–0.9% protein were determined on the validation sets. The number of PLS factors required for optimal model performance ranged from 6 to 14, dependent on wheat class and the type of mathematical transformation on the spectra before modeling. Repeatability measurements on kernels that were reloaded into the kernel clamp indicated that the variance of model-determined protein content caused by kernel positioning was largest for the smaller wheat kernels. Improvement in model performance could probably be attained by the averaging of transmittances taken from several positions on each kernel.

The direct relationship between protein content in wheat and the quality of flour for bakery products is well known (Wrigley and Bietz 1988). Along with wheat hardness and class, protein content is one of the most important factors for determining end use. Often, trade premiums are offered on high protein wheats. Until recently, nearly all grain quality analyses were performed on samples consisting of hundreds, if not thousands, of kernels, most often in ground form. In so doing, information was lost on the characteristics of the individual kernel. With the implementation of the Grain Quality Improvement Act in 1986 (OTA 1989), in which grading tests were mandated to address end-use quality, the USDA's Federal Grain Inspection Service (FGIS) has been seeking new methods for grading. Ideally, these methods should be fast, objective, capable of characterizing end-use functionality, and in cases where mixture detection is important, capable of operating on single kernels of grain. Plant breeders desire new techniques for single kernel analyses (e.g., hardness, protein, oil, moisture content, and viability) that are nondestructive in order to circumvent the existing time-consuming and expensive performance tests performed on new genetic lines.

Near-infrared (NIR) transmittance spectrophotometry is a possible means to meet these criteria in single kernel quality assessment. Previous research has demonstrated the capability of this instrumentation in the measurement of moisture in corn (Finney and Norris 1978), moisture in soybeans (Lamb and Hurburgh 1991), oil in corn (Orman and Schumann 1992), and hardness in wheat (Delwiche 1993). The present study explores the possibility of using such technology to determine the overall protein content of individual wheat kernels. The objective has been to develop an NIR transmittance technique, with associated chemometrics, that is capable of estimating single kernel protein content for a given class of wheat. Six wheat classes were examined: hard white wheat, hard red winter, hard red spring, soft red winter, soft white wheat, and durum (HWW, HRW, HRS, SRW, SWW, and DUR, respectively).

**MATERIALS AND METHODS****Wheat**

Wheat samples from the six classes were supplied by the FGIS Technical Center in Kansas City, MO. These samples were drawn from the protein standard reference stock that the Technical Center distributes to the FGIS field offices for standardizing near-infrared (formerly, ground-sample reflectance; currently, bulk-sample transmittance) instruments. Each wheat class was represented by five samples (1–5) selected for a broad range in protein content (Table I). Within a class, sample 1 had the lowest protein content (with the exception of the SRW class, in which sample 2 had the lowest) and sample 5 had the highest. Samples were grown in the geographical regions of the continental United States that are typically attributed to each class (central for HRW and HRS, eastern-central for SRW, northern-central for DUR, western for HWW, and northwestern for SWW). All samples were from the 1988–1992 harvests, with two classes (HWW and SWW) containing samples from one year, two (HRS and SRW) from two years, and two (HRW and DUR) from three years. Generally, each sample was of one commercial variety, though not specified. At the Kansas City depository, the entire contents of each sample (typically several hundred kilograms) were thoroughly mixed to ensure uniformity among subsamples. The stock was kept in cold storage (0–5°C) by FGIS. Likewise, samples were held at this temperature at the Beltsville, MD, laboratory until the humidity conditioning process.

**Nitrogen Analyzer**

A nitrogen and food protein determinator (model FP428, LECO, St. Joseph, MI) was used to determine the protein content ( $N \times 5.7$ ) of individual wheat kernels. This system bases the determination of nitrogen on the thermal conductivity of the combustion gases, having first removed water and carbon dioxide gases and converted the oxides of nitrogen to pure nitrogen gas. High purity ethylenediamine tetraacetic acid (nitrogen content 9.59% by weight) was used to calibrate the analyzer. An electronic balance with readability to 0.01 mg was used in conjunction with the analyzer.

**NIR Spectrophotometer**

Single kernel transmittance spectra were collected on a computerized Cary model 14 prism grating spectrophotometer. Each kernel was individually placed in a spring clamp such that the crease was facing the front of the clamp and the germ was always on the same side of the clamp. Within the clamp, the kernel

<sup>1</sup>Instrumentation and Sensing Laboratory, Beltsville Agricultural Research Center, USDA-ARS, Beltsville, MD.

Mention of company or trade name is for purpose of description only and does not imply endorsement by the U.S. Department of Agriculture.

was sandwiched between opposing pieces of butyl rubber, which formed an elliptical aperture that was approximately half the cross-sectional area of a typical wheat kernel. Light from a tungsten lamp (24 V DC, 150W) was chopped and then dispersed by a prism and grating before being passed through a 1.7-mm slit (6 nm nominal bandpass) and focused to a size (11 × 3 mm) slightly larger than the clamp's aperture. A silicon detector oriented normal to the path of incident light was positioned 10 mm from the far side of the kernel. These procedures are further discussed and illustrated in Delwiche (1993).

### Experimentation

For each sample, ~30 g of seed was placed in a polystyrene dish contained in a desiccator fixed at 53% rh (25°C) (Greenspan 1977) by an aqueous solution of magnesium nitrate at saturation. This humidity corresponded to an approximate moisture content of 12% (wb) in wheat (ASAE 1983), which is commonly the basis moisture content for trade in the United States. Samples were allowed to equilibrate for at least three weeks before removal from the desiccator.

At random, 96 kernels per sample were selected and individually

scanned in the spectrophotometer. Transmittance readings were collected from 740 nm to 1,139 nm and recorded to RAM every 0.2 nm at a scan speed of 10 nm/sec. Each recorded value was the average of 1,024 pairs of light and dark readings digitized with a 16-bit analog-to-digital converter operating at 100 kHz. Energy readings were referenced to corresponding wavelength readings through a 0.5-mm thick Teflon sheet and metal screen assembly with an average optical density of 1.3 *A* over the wavelength region. A running mean smooth of 1-nm half-width was performed on the values in RAM, whereupon every fifth value was kept, converted to log(1/T), and stored on external disk. Each stored spectrum consisted of 400 successive log(1/T) values in 1-nm increments. Upon scanning, each kernel was immediately weighed to the nearest 0.01 mg. Because of the length of time required for each scan (~3 min), each kernel was scanned only once.

For single kernel protein contents to be expressed on a constant moisture basis, the dry weight of each kernel was determined by oven drying at 130°C for 19 hr in accordance with standard procedure (Hart et al 1959, ASAE 1983). Empty 9-mm bullet casings (one per kernel) were used as drying containers; 9.52-mm (3/8 in.) diameter ball bearings were placed on the casings before

TABLE I  
Means and Standard Deviations (SD) for Single Wheat Kernels

| Sample <sup>a</sup> | No. of Kernels | Dry Kernel Weight (mg) |      | Moisture (% wb) |      | Protein (%) <sup>b</sup> |                              |
|---------------------|----------------|------------------------|------|-----------------|------|--------------------------|------------------------------|
|                     |                | Mean                   | SD   | Mean            | SD   | FGIS <sup>c</sup> Mean   | Combustion <sup>d</sup> Mean |
| HWW 1               | 96             | 45.10                  | 9.56 | 10.31           | 0.21 | 11.24                    | 11.63                        |
| 2                   | 96             | 42.53                  | 8.49 | 10.60           | 0.25 | 12.06                    | 11.78                        |
| 3                   | 96             | 46.05                  | 9.00 | 10.26           | 0.13 | 12.87                    | 12.70                        |
| 4                   | 96             | 44.28                  | 8.40 | 10.17           | 0.13 | 13.26                    | 13.29                        |
| 5                   | 96             | 44.52                  | 8.37 | 10.35           | 0.21 | 13.70                    | 13.92                        |
| Average             | 480            | 44.50                  | 8.76 | 10.34           | 0.19 | 12.63                    | 12.66                        |
| HRW 1               | 95             | 30.35                  | 4.00 | 11.31           | 0.34 | 11.59                    | 11.24                        |
| 2                   | 96             | 29.65                  | 4.00 | 11.10           | 0.24 | 12.27                    | 11.65                        |
| 3                   | 96             | 24.13                  | 4.36 | 12.16           | 0.19 | 13.51                    | 12.94                        |
| 4                   | 96             | 30.21                  | 3.80 | 10.97           | 0.19 | 14.46                    | 13.54                        |
| 5                   | 86             | 25.84                  | 3.63 | 10.69           | 0.19 | 15.84                    | 15.97                        |
| Average             | 469            | 28.04                  | 3.96 | 11.25           | 0.23 | 13.53                    | 13.07                        |
| HRS 1               | 96             | 34.69                  | 6.32 | 11.07           | 0.23 | 12.33                    | 12.48                        |
| 2                   | 96             | 32.55                  | 5.39 | 11.92           | 0.23 | 14.48                    | 13.63                        |
| 3                   | 96             | 24.15                  | 3.93 | 12.17           | 0.18 | 14.50                    | 13.21                        |
| 4                   | 96             | 35.76                  | 5.89 | 11.62           | 0.25 | 15.38                    | 14.83                        |
| 5                   | 96             | 33.10                  | 4.77 | 11.50           | 0.21 | 16.71                    | 16.55                        |
| Average             | 480            | 32.05                  | 5.26 | 11.66           | 0.22 | 14.68                    | 14.14                        |
| SRW 1               | 96             | 26.66                  | 4.73 | 11.22           | 0.19 | 9.08                     | 8.63                         |
| 2                   | 96             | 28.40                  | 3.35 | 12.05           | 0.16 | 8.02                     | 7.86                         |
| 3                   | 96             | 26.89                  | 5.67 | 12.06           | 0.20 | 10.46                    | 10.49                        |
| 4                   | 96             | 26.94                  | 4.95 | 10.98           | 0.28 | 11.09                    | 10.38                        |
| 5                   | 96             | 29.00                  | 6.40 | 10.60           | 0.25 | 11.55                    | 10.65                        |
| Average             | 480            | 27.58                  | 5.19 | 11.38           | 0.22 | 10.04                    | 9.60                         |
| SWW 1               | 96             | 37.43                  | 5.81 | 10.75           | 0.22 | 7.90                     | 8.60                         |
| 2                   | 96             | 47.99                  | 7.98 | 10.89           | 0.24 | 8.92                     | 9.16                         |
| 3                   | 96             | 47.30                  | 8.81 | 10.88           | 0.22 | 10.04                    | 9.88                         |
| 4                   | 96             | 45.95                  | 7.50 | 11.06           | 0.16 | 10.38                    | 9.93                         |
| 5                   | 96             | 33.82                  | 7.12 | 10.65           | 0.17 | 12.05                    | 11.71                        |
| Average             | 480            | 42.50                  | 7.44 | 10.85           | 0.20 | 9.86                     | 9.86                         |
| DUR 1               | 96             | 44.07                  | 7.47 | 10.64           | 0.26 | 13.28                    | 13.18                        |
| 2                   | 96             | 46.86                  | 8.04 | 10.34           | 0.22 | 13.87                    | 13.74                        |
| 3                   | 96             | 44.27                  | 6.51 | 10.83           | 0.19 | 14.20                    | 13.13                        |
| 4                   | 96             | 35.50                  | 5.77 | 10.58           | 0.25 | 15.22                    | 14.29                        |
| 5                   | 96             | 44.34                  | 6.02 | 10.14           | 0.25 | 16.47                    | 15.90                        |
| Average             | 480            | 43.01                  | 6.76 | 10.51           | 0.23 | 14.61                    | 14.05                        |

<sup>a</sup>HWW = hard white wheat; HRW = hard red winter; HRS = hard red spring; SRW = soft red winter; SWW = soft white wheat; DUR = durum.

<sup>b</sup>12% mc.

<sup>c</sup>Values furnished by USDA Federal Grain Inspection Service (FGIS) Feb. 93. Representative of the protein content of the entire stock of each sample.

<sup>d</sup>Total protein mass of *n* kernels/mass (12% mc-adjusted) of *n* kernels.

placement in a desiccator containing molecular sieve (type 2A) for a 30-min cool-down period. Upon recording the dry weights, the kernels were kept at room conditions (typically for several days) before each was loaded in the nitrogen analyzer. Immediately before being consumed by the analyzer, each kernel was weighed one last time. Kernel protein weight could then be determined as the final kernel weight times the analyzer's value for protein content. Two values for protein content were subsequently calculated: **Pro**<sub>12</sub>, the percent protein by weight of the kernel adjusted to 12% moisture; and **Pro**, the percent protein by weight of the kernel at the time of scanning. Average weight, moisture content, and protein content of each sample is summarized in Table I. Both FGIS-determined (by Kjeldahl analysis) and the nitrogen analyzer mass average-determined (**Pro**<sub>12</sub>) protein contents are listed in this table.

### Repeatability Kernels

Three additional kernels from the stock of HWW 2 were analyzed. These kernels, having dry weights (24, 41, and 57 mg, respectively) representative of small, medium, and large sizes for this class, were used to determine the repeatability of the spectral measurement. Each of the three kernels was repeatedly loaded into the spring clamp (orientation same as described earlier), scanned, and removed from the clamp 32 times. Afterward, the kernels were dried and analyzed for protein in the same manner as all other kernels.

### Model Development

Separate models were developed for each wheat class. Within each class, spectra and accompanying protein contents of the kernels from the samples 1, 2, 4, and 5 formed the calibration set. The nominally intermediate protein sample (3) was set aside for model validation. Chemometric modeling was performed on mean-centered (Galactic 1992) spectra using partial least squares (PLS) analysis. Mathematics of the PLS modeling process and the related process of principal component regression (PCR) are briefly described in Delwiche (1993) and more fully elsewhere (Gunst and Mason 1979, Lindberg et al 1983). During modeling, the wavelength region of interest was reduced to 850–1,050 nm to match the range in commercially available whole grain ana-

lyzers, and because the spectral response of silicon is greatly diminished at wavelengths >1,050 nm.

### Spectral Pretreatment

Before PLS analysis, each spectrum received one of three mathematical transformations: 1) no change to log(1/T) spectra, 2) multiplicative scatter correction (MSC) on log(1/T) spectra, and 3) MSC on second difference log(1/T) spectra.

The MSC procedure, as defined in Martens and Naes (1987), involved the linear regression of each spectrum against the mean spectrum of the calibration set for the purpose of reducing kernel-to-kernel variations in light scatter. A separate regression was performed on every spectrum, and then each was slope- and bias-adjusted to the mean.

The second difference was determined by:

$$g(\lambda_i) = k[f(\lambda_{i+j}) - 2f(\lambda_i) + f(\lambda_{i-j})]$$

where  $g(\lambda)$  is the second difference spectrum,  $f(\lambda)$  is the spectrum before transformation,  $k$  is a scaling constant,  $i$  is the wavelength index, and  $j = 6$  is the gap (i.e., the half-width, expressed in points) of the finite difference window. The magnitude of  $j$  corresponded to the bandpass of the instrument (~6 nm). A second difference transformation eliminates spectrum-to-spectrum baseline differences, which generally is advantageous for reducing the effect of light scatter caused by unequal kernel weights and differences in vitreosity.

### Modeling Procedure

The regression procedure in PCR or PLS models is performed on the scores or eigenvalues of each spectrum, once the spectra are decomposed into their principal components (eigenvectors), weighted by protein value in PLS, nonweighted in PCR. In the present study, models consisting of 1–20 principal components were developed using **Pro** as the dependent variable. A commercial software package, GRAMS/386 (Galactic Industries, Salem, NH), running in the Microsoft Windows operating system, was used to develop the models.

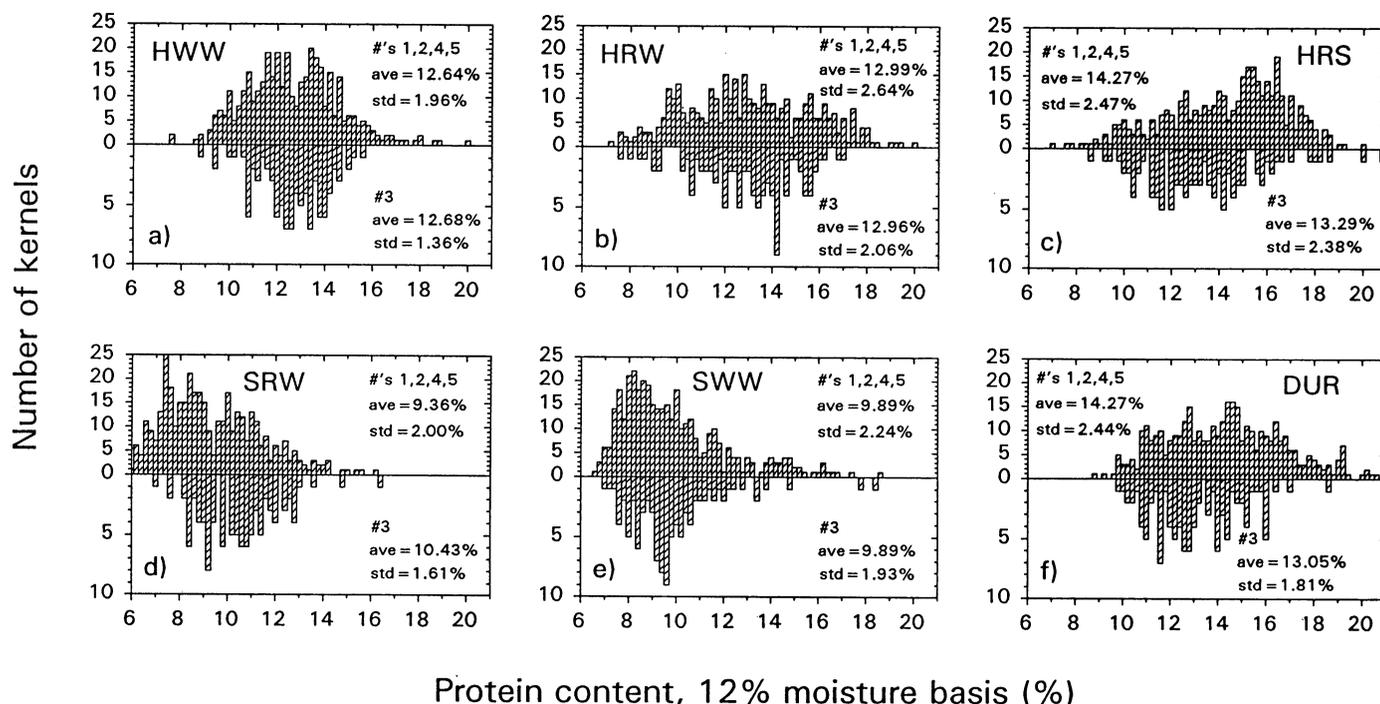


Fig. 1. Distributions of single kernel protein contents (determined by reference method) for six wheat classes: a) hard white wheat (HWW); b) hard red winter (HRW); c) hard red spring (HRS); d) soft red winter (SRW); e) soft white wheat (SWW); and f) durum (DUR). For each class, samples (#s) 1, 2, 4, and 5 ( $n = 384$ , except for HRW  $n = 373$ ) used in model development. Sample 3 ( $n = 96$ ) reserved for validation.

## Reported Statistics

Within a wheat class, each model was subsequently applied to the corresponding 96-kernel sample 3 of the class. Model performance was defined in terms of the multivariate coefficient of determination ( $r^2$ ), the standard error of prediction (SEP) between measured and modeled protein values, and the bias between the average for **Pro** by reference measurement and **Pro** by the model. (These terms are mathematically defined in Hruschka [1987].)

## RESULTS AND DISCUSSION

The distributions of single kernel **Pro**<sub>12</sub> of the combined samples 1, 2, 4, and 5 and of sample 3 are shown in Figure 1 for the classes HWW, HRW, HRS, SRW, SWW, and DUR. Within each class, the range of **Pro**<sub>12</sub> for either the one- or four-sample groups was typically 10 percentage units, with standard deviations typically being 1–2 and 2–3 percentage units for these groups, respectively. Though such ranges are large, they do seem reasonable upon review of Levi and Anderson (1950), who measured ranges of up to 6% protein for kernels within the same wheat spike.

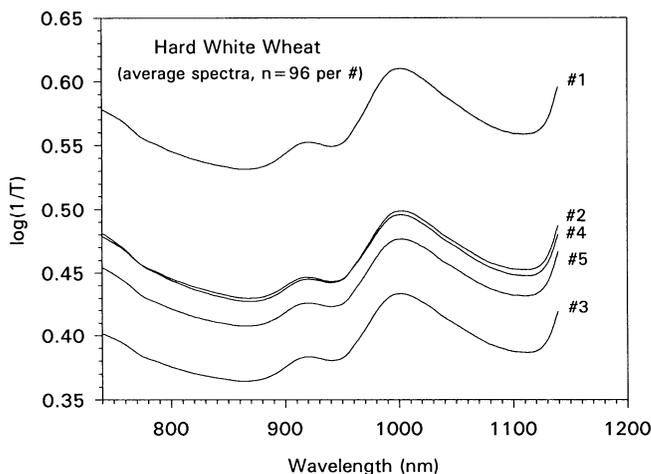


Fig. 2. Average  $\log(1/T)$  spectra of the hard white wheat (HWW) samples. Transmittance values referenced to a standard consisting of a metal screen and Teflon assembly at 1.3  $\mu$ .

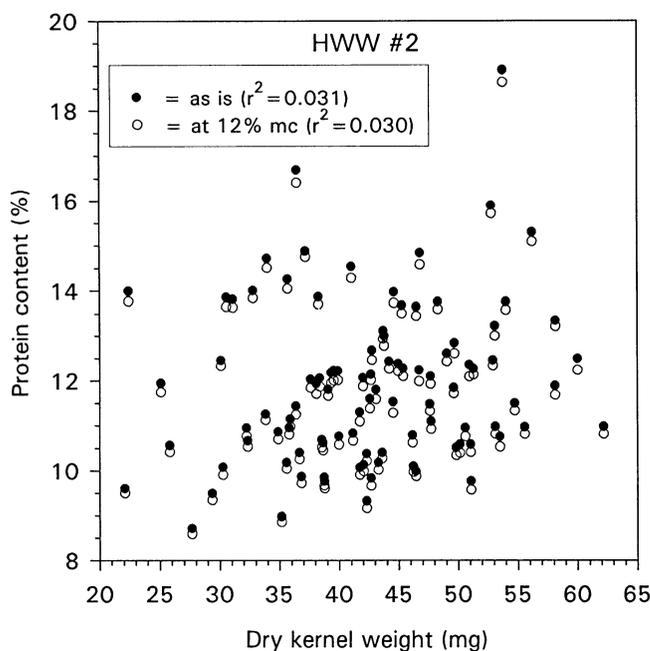


Fig. 3. Lack of relationship between single kernel protein content and dry kernel weight in hard white wheat (HWW) sample 2.

Typical average spectra of the HWW samples are shown in Figure 2. Samples in this class were numbered in order of increasing protein content (Table I), but the same order is not present in the magnitude of their  $\log(1/T)$  spectra. Most noticeable with the average spectra is the large vertical offset between these curves, indicating that, irrespective of protein content, some of the kernels were less translucent to light transmittance. Each spectrum contained two broad carbohydrate absorption bands, one centered at  $\sim 920$  nm and a larger one at 1,000 nm. Not apparent in these  $\log(1/T)$  curves were regions of spectral absorption due to protein and water (1,018 nm and 958 nm, respectively) (Williams and Norris 1987). However, such bands usually become noticeable upon transformation to the second difference, hence the reason for the transformation.

By way of example with HWW sample 2, single kernel protein content was not correlated to kernel mass (Fig. 3). Lack of correlation between **Pro** and dry weight or **Pro**<sub>12</sub> and dry weight was typical of all samples and classes. Likewise, **Pro** and moisture content were uncorrelated for all samples and classes. Consequently, estimation of a kernel protein content is not possible by visual examination of size, translucency, or by gravimetric analysis.

Performance of the PLS-based models of **Pro** for all classes is summarized in Table II. The values listed for  $r^2$ , SEP, and BIAS are associated with the number of PLS factors that yielded the best combination of high  $r^2$  and low SEP for each of the spectral pretreatment cases.

Among the classes, the best performance occurred with the HWW models, in which  $r^2$  was  $\sim 0.9$  and SEP was  $< 0.50\%$  protein for all three spectral pretreatments. High relative performance of the HWW models is thought to be due to the genetic similarity of the five samples (all were the cultivar *Klassic*) and, consequently, uniformity in appearance. In all three pretreatments of the HWW models, the optimal performance was attained in less than 10 factors. As the level of pretreatment became more complex, the number of factors necessary to maintain values for  $r^2$  and SEP comparable to those of less complex pretreatments declined. For a given pretreatment, model performance improved with increase

TABLE II  
Performance Statistics of Partial Least Squares (PLS)  
Models Applied to the Validation Samples<sup>a</sup>

| Class <sup>b</sup> | Pretreatment <sup>c</sup> | Factors <sup>d</sup> | $r^2$ | Units of % Protein |                   |
|--------------------|---------------------------|----------------------|-------|--------------------|-------------------|
|                    |                           |                      |       | SEP <sup>e</sup>   | BIAS <sup>f</sup> |
| HWW                | None                      | 9                    | 0.914 | 0.42               | 0.24              |
|                    | MSC                       | 7                    | 0.882 | 0.48               | 0.17              |
|                    | MSC on 2nd                | 6                    | 0.894 | 0.45               | 0.03              |
| HRW                | None                      | 14                   | 0.851 | 0.83               | 0.02              |
|                    | MSC                       | 8                    | 0.872 | 0.78               | 0.25              |
|                    | MSC on 2nd                | 10                   | 0.889 | 0.83               | 0.46              |
| HRS                | None                      | 14                   | 0.865 | 0.90               | 1.52              |
|                    | MSC                       | 12                   | 0.927 | 0.71               | 1.05              |
|                    | MSC on 2nd                | 6                    | 0.938 | 0.65               | 1.11              |
| SRW                | None                      | 10                   | 0.857 | 0.63               | -0.84             |
|                    | MSC                       | 8                    | 0.904 | 0.60               | -1.03             |
|                    | MSC on 2nd                | 8                    | 0.908 | 0.53               | -0.88             |
| SWW                | None                      | 10                   | 0.782 | 0.94               | 1.04              |
|                    | MSC                       | 10                   | 0.918 | 0.70               | 0.43              |
|                    | MSC on 2nd                | 6                    | 0.925 | 0.65               | 0.15              |
| DUR                | None                      | 8                    | 0.854 | 0.78               | -0.32             |
|                    | MSC                       | 10                   | 0.883 | 0.83               | 0.35              |
|                    | MSC on 2nd                | 8                    | 0.928 | 0.50               | -0.06             |

<sup>a</sup>Separate models developed for each class.  $n = 96$  kernels per validation sample.

<sup>b</sup>See Table I for abbreviations.

<sup>c</sup>Spectral pretreatment was either none, multiplicative scatter correction (MSC), or a MSC on a three-point, 6-nm gap, second central difference (MSC on 2nd).

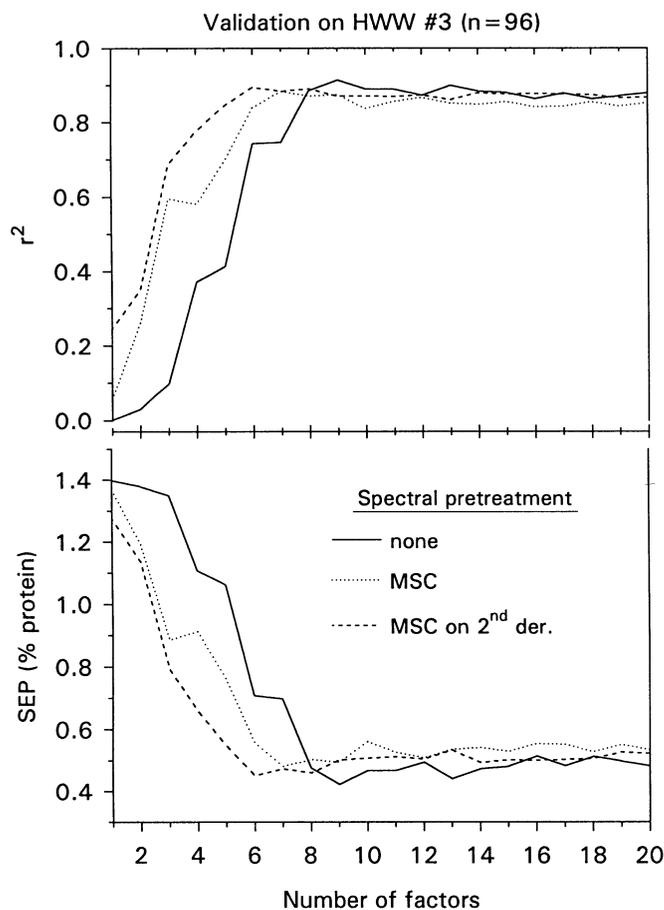
<sup>d</sup>Optimal model conditions as determined on validation samples.

<sup>e</sup>Standard error of prediction.

<sup>f</sup>Bias between the average for **Pro** by reference measurement and **Pro** by the model.

in factors up to ~9, beyond which both  $r^2$  and SEP remained relatively constant (Fig. 4). The second difference pretreatment resulted in higher accuracies at the lower factors (1–8), but made little difference at factors 9–20. Additionally, the HWW model employing the most complex spectral pretreatment produced the least BIAS (Table II), suggesting that kernel-to-kernel baseline differences caused by scatter can be reduced by application of a suitable spectral pretreatment.

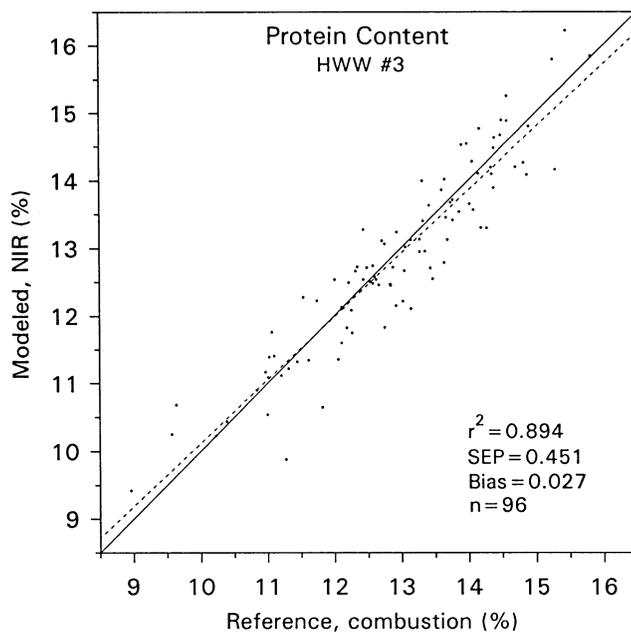
Of the five remaining classes, model performance at a given



**Fig. 4.** Performance of three partial least squares models applied to a validation set of hard white wheat (HWW), with each model using a different spectral pretreatment. MSC = multiplicative scatter correction. MSC on 2nd = MSC on a three-point, 6-nm gap, second central difference. SEP = standard error of prediction.

spectral pretreatment was approximately equivalent between classes. Within each class (excluding DUR), fewer factors were needed at the more complex pretreatments to yield equivalent or superior performance to the case of no spectral pretreatment. In the trials of the durum samples, a MSC pretreatment without second difference produced a model with slightly improved  $r^2$ , but also produced a higher standard error and required two more factors than the model with no spectral pretreatment.

A plot of a modeled (six-factor PLS with MSC on second difference pretreatment) value for **Pro** versus the reference value for HWW 3 is shown in Figure 5. A line at 45° is included that represents the line on which all predictions would lie if the model were perfect. Also included on the plot is the line determined from the linear regression of the modeled values onto the reference values. Failure to find a significant nonzero mean of the difference between modeled and reference values (by Student's  $t$  test), along with an  $r^2 = 0.89$  and low values for SEP and BIAS, collectively indicate that the model was reasonably robust.



**Fig. 5.** Near-infrared (NIR) transmittance spectrophotometry comparison of NIR-predicted and reference single kernel protein contents on a validation set of hard white wheat (HWW). Protein contents not corrected to constant moisture (average =  $10.3 \pm 0.1\%$  mc). Model conditions: 6-factor partial least squares, spectral pretreatment of multiplicative scatter correction on second difference (gap = 6 nm) of  $\log(1/T)$  spectra.

**TABLE III**  
Simplified Analysis of Variance of Protein Measurements

| Sample <sup>a</sup>                | <i>n</i> | Protein Content |                            | Method  |
|------------------------------------|----------|-----------------|----------------------------|---|
|                                    |          | Mean (%)        | Variance (% <sup>2</sup> ) |   |
| Reference method                   |          |                 |                            |   |
| HWW 3                              | 96       | 12.93           | 1.92                       | Combustion, <i>n</i> single kernels   |
| HWW 2                              | 96       | 11.90           | 3.15                       | Same  |
| Ground HWW 2                       | 7        | 12.67           | 0.013                      | Combustion, <i>n</i> 45-mg subsamples of 5 g of ground HWW 2  |
| NIR model on single kernels        |          |                 |                            |   |
| HWW 3                              | 96       | 12.96           | 1.73                       | 6-factor partial least squares model, multiplicative scatter correction on second difference pretreatment |
| HWW 2                              | 96       | 11.96           | 3.11                       | Same  |
| NIR model on repeatability kernels |          |                 |                            |   |
| HWW 2, kernel 1                    | 32       | 11.05           | 0.79                       | Same model applied to <i>n</i> repetitive scans on a small (26 mg) size kernel of HWW 2                   |
| HWW 2, kernel 2                    | 32       | 14.17           | 0.33                       | Same, but on a medium (45 mg) size kernel   |
| HWW 2, kernel 3                    | 32       | 12.28           | 0.13                       | Same, but on a large (63 mg) size kernel  |

<sup>a</sup>See Table I for abbreviations.

## Error Analysis

Table III contains the variances of **Pro** (for both measured and modeled values) for the HWW validation sample 3 and for HWW sample 2 used in calibration, and of **Pro** for the three kernels of the spectrophotometer repeatability set. Also included is the variance associated with repeated reference analyses of a ground portion of HWW 2. Though not intended as a formal analysis of variance (since each kernel of the HWW samples was not loaded and scanned repetitively), this table illustrates the precision of the reference method. The variance of the protein readings (determined by combustion) of subsamples from a 5-g portion of ground HWW 2 was <0.5% of the variance associated with similarly determined single kernel protein readings of HWW 2. Therefore, the kernel-to-kernel protein variability in a single wheat variety (of single harvest), demonstrated in Figure 1 and quantified in Table III, appears to be as large as indicated.

Also in Table III are the means and variances of **Pro** that were determined from the optimal PLS model with full spectral pretreatment (MSC on second difference) applied to HWW 3 and to HWW 2. For both samples, the magnitudes of the means and variances are equivalent to their corresponding values from the reference method determinations. However, such equivalency does not signify that the modeled and reference protein contents match exactly for every kernel, otherwise the  $r^2$  values in Table II would be near unity and the error near zero.

Model-caused variability in single kernel protein is better described by repeated measurements on one kernel (Table III). Such variability is attributed to kernel positioning within the clamp. Of the three kernels repeatedly loaded and scanned, the smallest (26 mg) demonstrated the most variability, while the largest (63 mg) had the least. More freedom of movement within the cavity of the single kernel clamp apparently made kernel alignment more critical for the smaller kernels. Because special care was taken to reload each kernel into the clamp in the same position, further improvements in mechanical alignment are probably impractical. Rather, transmittance measurements at more than one position of the kernel, and an average of such readings, might improve the performance of NIR-based models.

## SUMMARY AND CONCLUSIONS

NIR transmittance measurements (850–1,050 nm) collected on individual kernels of wheat were related by PLS analysis to corresponding single-kernel protein contents measured by combustion. White, red, soft, and hard classes of wheat were examined. Coefficients of determination of ~0.9, with accompanying standard errors of 0.4–0.9% protein, were achieved for models employing 6–14 factors. The exact number was dependent on the class of wheat and the spectra pretreatment. Correction for light scattering effects, by application of a multiplicative scatter correction or a second difference, improved model performance. Further improvements could probably be attained by repetitive scans at multiple orientations of the kernel. Future research will examine the feasibility of using diode array devices or several single diodes as sensing elements, whereby the targeted read-time on each kernel

will be less than 1 sec. Successful results at such a time scale will make it practical to implement single kernel protein analysis in official grading and classification procedures.

## ACKNOWLEDGMENTS

I thank Mary Ann Wickham of USDA/FGIS, Kansas City, MO, for supplying the wheat samples and Joyce Shaffer of USDA/ARS for technical assistance.

## LITERATURE CITED

- ASAE. 1983. Moisture measurement—grain and seeds. S352.1 In: American Society of Agricultural Engineers Yearbook of Standards. The Society: St. Joseph, MI.
- DELWICHE, S. R. 1993. Measurement of single kernel wheat hardness using near-infrared transmittance. *Trans. ASAE* 36:1431-1437.
- FINNEY, E. E., and NORRIS, K. H. 1978. Determination of moisture in corn kernels by near-infrared transmittance measurements. *Trans. ASAE* 21:581-584.
- GALACTIC. 1992. Add-on application software. PLSplus version 2.1G for GRAMS/386. Galactic Industries: Salem, NH.
- GREENSPAN, L. 1977. Humidity fixed points of binary saturated aqueous solutions. *J. Res. Natl. Bur. Stand.* 81A:89-96.
- GUNST, R. F., and MASON, R. L. 1979. Some considerations in the evaluation of alternate prediction equations. *Technometrics* 21:55-63.
- HART, J. R., FEINSTEIN, L., and GOLUBIC, C. 1959. Oven methods for precise measurement of moisture in seeds. USDA Marketing Research Report No. 304. U.S. Government Printing Office: Washington, DC.
- HRUSCHKA, W. R. 1987. Data analysis: Wavelength selection methods. Page 54 in: *Near-Infrared Technology in the Agricultural and Food Industries*. P. C. Williams and K. H. Norris, eds. Am. Assoc. Cereal Chem.: St. Paul, MN.
- LAMB, D. T., and HURBURGH, C. R. 1991. Moisture determination in single soybean seeds by near-infrared transmittance. *Trans. ASAE* 34:2123-2129.
- LEVI, I., and ANDERSON, J. A. 1950. Variations in protein contents of plants, heads, spikelets, and individual kernels of wheat. *Can. J. Res. Sect. F* 28:71-81.
- LINDBERG, W., PERSSON, J. A., and WOLD, S. 1983. Partial least-squares method for spectrofluorimetric analysis of mixtures of humic acid and lignosulfonate. *Anal. Chem.* 55:643-648.
- MARTENS, H., and NAES, T. 1987. Multivariate calibration by data compression. Page 73 in: *Near-Infrared Technology in the Agricultural and Food Industries*. P. C. Williams and K. H. Norris, eds. Am. Assoc. Cereal Chem.: St. Paul, MN.
- ORMAN, B. A., and SCHUMANN, R. A. 1992. Nondestructive single-kernel oil determination of maize by near-infrared transmission spectroscopy. *J. Am. Oil Chem. Soc.* 69:1036-1038.
- OTA. 1989. Enhancing the quality of U.S. grain for international trade. U.S. Congress, Office of Technology Assessment. Pages 189-215 in: OTA-F-399. U.S. Government Printing Office: Washington, DC.
- WILLIAMS, P. C., and NORRIS, K. H. 1987. Qualitative applications of near-infrared reflectance spectroscopy. Page 246 in: *Near-Infrared Technology in the Agricultural and Food Industries*. P. C. Williams and K. H. Norris, eds. Am. Assoc. Cereal Chem.: St. Paul, MN.
- WRIGLEY, C. W., and BIETZ, J. A. 1988. Proteins and amino acids. Pages 159-275 in: *Wheat: Chemistry and Technology*, Vol. 1, 3rd ed. Y. Pomeranz, ed. Am. Assoc. Cereal Chem.: St. Paul, MN.

[Received April 27, 1994. Accepted September 7, 1994.]