

# Protein Content of Single Kernels of Wheat by Near-Infrared Reflectance Spectroscopy

S. R. Delwiche

USDA-ARS, Beltsville Agricultural Research Center, Instrumentation and Sensing Laboratory,  
Building 303, BARC-East, Beltsville, Maryland 20705-2350 U.S.A.

Received 8 October 1996

## ABSTRACT

Protein content is well known to affect the functional properties of processed wheat products. Traditionally performed on aliquots (0.25–2.2 g) from samples ranging in size from 30–40 g (for combustion and Kjeldahl analyses) to several hundred grams (for whole-grain near-infrared analysis), these methods inherently do not provide information on single-kernel protein variability. Inspection procedures by the United States Department of Agriculture for grading and classification of wheat are undergoing change to provide the processor or end user with information on the variability of several single-kernel properties including hardness, moisture, weight, and wheat class. The present study has focused on demonstrating the feasibility of measuring crude protein content of single wheat kernels by near-infrared reflectance. More than 300 commercial wheat samples from the 1992 U.S. harvest, representing five (hard red winter, hard red spring, soft red winter, hard white, and soft white) of the six (durum excluded) market classes were chosen, from which 10 kernels were randomly selected and handled on a single-kernel basis. Handling consisted of reflectance scanning (1100–2498 nm), drying (for moisture compensation), and combustion (for reference protein-content determination). Partial least squares and multiple linear regression models, when applied to samples excluded from calibration, demonstrated standard errors of performance ranging from 0.462 to 0.720% protein depending on the modeling technique, number of classes used to develop the model, and the wheat class tested. The pooling of wheat classes to produce a general model did not diminish model accuracy. Best results were achieved with an 1100–1400-nm region. Model performance worsened as the wavelength region widened or as the minimum wavelength shifted from 1100 nm to higher values.

*Keywords:* wheat, single kernel, protein, near-infrared.

**ABBREVIATIONS USED:** GIPSA = Grain Inspection, Packers and Stockyards Administration; HRS = hard red spring; HRW = hard red winter; HWW = hard white wheat; MLR = multiple linear regression; MSC = multiplicative scatter correction; near-IR = near-infrared; PLS = partial least squares; RMSD = root mean squared difference; SEC = standard error of calibration; SEP = standard error of performance; s.k. = single kernel; SRW = soft red winter; SWW = soft white wheat.

Mention of trademark or proprietary product does not constitute a guarantee or warranty of a product and does not imply its approval to the exclusion of other products that may also be suitable. This article is in the public domain and not copyrightable. It may be freely reprinted with customary crediting of the source.

## INTRODUCTION

In domestic and overseas trade of United States wheat, crude protein content, aside from wheat class and grade, is often the most important factor that determines price. Affecting the suitability of wheat for various products (pan bread, noodles, crackers, flat bread, etc.) and dough characteristics, protein content is used as a measure of grain quality<sup>1</sup>. Owing to its rapid and accurate characteristics, near-infrared (near-IR) reflectance or transmittance spectroscopy is routinely used to measure protein content of wheat on samples that are typically 400–600 g in size. With such sizes, where thousands of kernels

comprise each sample, information on kernel-to-kernel variability of protein content is not attainable. This loss of information is especially acute in the U.S.A. because wheat within a market class is generally not segregated by variety during post-harvest operations. Further, environmental conditions within the Great Plains of North America are widely variant such that wheat quality traits are often more greatly affected by environment than by genotype<sup>2,3</sup>. Knowledge of individual kernel protein contents from samples representative of wheat lots is potentially useful to processors as a gauge of the consistency of raw product. Additionally, such knowledge could be valuable in U.S. official grading procedures in which the wheat must be examined for mixtures of two or more market classes. Provided single-kernel (s.k.) protein analysis can be performed non-destructively, such a procedure could be useful in plant-breeding programs.

In an earlier study<sup>4</sup>, the feasibility of s.k. protein content determination based on near-IR transmittance (850–1048-nm wavelength region) through intact kernels was examined. On a limited set of pure variety samples from six U.S. market classes—hard red winter (HRW), hard red spring (HRS), soft red winter (SRW), hard white wheat (HWW), soft white wheat (SWW), and durum—best near-IR model accuracies, defined as the standard deviation of residuals (modeled value minus reference value), ranged from 0.42 to 0.78% protein content, depending on the wheat class modeled. Although these results were encouraging, the small number of unique samples (five) per class made it difficult to draw general conclusions on the suitability of near-IR techniques for s.k. protein content measurement. The present study has expanded on the previous by incorporating a larger number of unique samples (minimum of 56 samples per class). Additionally, near-IR reflectance (>1100 nm) has been utilized in lieu of transmittance, owing to its easier adaptability to real-time analysis in commercial grading instruments. The specific goals of the experiment reported herein were to examine the accuracies of s.k. near-IR protein content models, to examine the effect of incorporating more than one wheat class in a model, to define the most suitable wavelength region, and to identify the sources of variation under near ideal conditions of kernel presentation.

## EXPERIMENTAL

### Wheat

Wheat samples were obtained from the United States Department of Agriculture (USDA) Grain Inspection, Packers and Stockyards Administration (GIPSA) and were a subset of the Agency's 1992 crop year market survey. Samples for the survey were originally collected from the commercial market stream throughout the wheat-growing regions of the U.S.A., inclusive of the Great Plains, Pacific Northwest, Eastern Seaboard, California, and Michigan. Wheat class for each sample was confirmed by the GIPSA Board of Appeals and Review (the U.S. supreme arbitrator for grain classification). From a total of more than 3000 samples in the survey, 318 samples were drawn at random for this study. All U.S. market classes, excluding durum, were represented, with the exact number per class as follows: HRW = 72, HRS = 62, SRW = 58, HWW = 56, and SWW = 70. Origin by state within the U.S.A. (16 states total) was known for 316 of the 318 samples. The origin by state, as well as the means and standard deviations for protein content and near-IR hardness of the bulk samples are summarized in an earlier publication<sup>5</sup>. From each sample, 10 kernels were drawn at random. Broken, shriveled, or cracked kernels were discarded and replaced. Thus, the total size of the set of kernels for protein analysis was 3180.

### Equipment

#### *Near-infrared reflectance spectrophotometer*

A reflectance detector assembly for an NIRSystems model 6500 scanning monochromator (Silver Spring, MD, U.S.A.) was customized to enable the collection of spectra from individual kernels, as described previously<sup>5</sup>. Briefly, each kernel was placed, crease side down, on the end of a revolving blackened tube within an enclosed cavity. Reflected energy from the kernel was referenced to readings from a pressed disk of a 1:200 (w/w) mixture of carbon black and polytetrafluoroethylene whose absolute reflectance was nominally 15%. The scan range was 1100–2498 nm at 2-nm increments, though the lower wavelength portion (1100–1798 nm) was preferred during modeling, owing to (1) a non-linear response for reflectance at the higher wavelengths and (2) a desire to mimic more closely the

operating range of an indium gallium arsenide detector for which diode array spectrometers (which offer the advantage of speed) are commercially available. Each kernel's spectrum was the average of 32 successive scans (i.e. grating oscillations) and stored to computer file in  $\log_{10}(1/R)$  format. Immediately before the scanning of each kernel, its mass was recorded (0.01 mg resolution) for later use in moisture-content determination. Afterward, dry kernel mass (130°C for 19 h<sup>6</sup>) was measured.

### Combustion nitrogen analyzer

Nitrogen content of each kernel was determined by combustion, using a Leco model FP-428 nitrogen analyzer (St Joseph, MI, U.S.A.). High purity ethylenediamine tetraacetic acid was used to calibrate the instrument daily. Kernel protein content ( $N \times 5.7$ ) was mathematically corrected to 12% moisture (wet basis) using knowledge of the dry mass.

### Protein content modeling

Based upon the protein content of the initial bulk samples (determined by GIPSA using near-IR reflectance of ground material) all samples within a wheat class were ranked from lowest to highest protein content. From each wheat class, three sets, termed calibration, validation, and test, were formed as follows: calibration set = odd-numbered kernels from the 10 lowest protein samples, the 10 highest protein samples and the 10 samples exactly at midrange in protein content (total = 150 kernels); validation set = same as calibration set, except even-numbered kernels were selected (150 kernels); test set = all remaining kernels (260 to 420 kernels, dependent on class). The purpose of the validation sets was to evaluate the performance of a near-IR model on a set of kernels possessing intrinsic properties that were as close to those of the calibration set as possible. The test sets, consisting of kernels from samples not used in calibration, served as estimators of the near-IR models' performances. In addition to the single classes, calibration sets from HRW, HRS, and SRW were pooled to form a 'RED' calibration set; likewise, pooled RED validation and test sets were also formed. Similarly, HWW and SWW classes were pooled to form calibration, validation, and test 'WHITE' sets, then all five classes were pooled to form equivalently three 'ALL' sets. The

means and standard deviations of s.k. protein content (by combustion), dry mass, and moisture content (at near-IR scanning) for each set within single or pooled class are listed in Table I.

Single-kernel protein content models were developed using partial least squares (PLS) analysis or multiple linear regression (MLR). Both forms of analyses are used extensively in near-IR modeling, with abundant description in the literature<sup>7-9</sup>. PLS modeling was performed with commercial software<sup>10</sup>, and in-house software was used for MLR modeling. The PLS analyses were assumed to provide a lower bound in error for the mathematically simpler MLR analyses. Despite the larger error, the latter technique might be the preferred choice for commercialization, owing to its adaptation to less-expensive equipment (e.g. fixed filter vs. scanning devices).

To reduce kernel-to-kernel spectral variation caused by variation in kernel size, all spectra in PLS analyses were multiplicatively scatter-corrected<sup>11</sup> to the mean spectrum of the appropriate calibration set. Preliminary PLS analyses (results not shown) indicated that multiplicative scatter correction (MSC) greatly improved model performance. A one-sample-out cross validation was used during each PLS calibration to select the optimal number of factors for that model. The optimal number was selected by *F* test of the cross validation residual squared errors<sup>10</sup>. MLR analyses consisted of a stepwise search of all wavelengths within a 1100–1798-nm region to produce the highest  $R^2$  value for a given number of terms (up to eight),

$$P = \beta_0 + \beta_1 f_1 + \dots + \beta_i f_i$$

where  $P$  is protein content,  $f_1, \dots, f_i$  are the  $\log(1/R)$  values,  $i = 1, 2, \dots, 8$ , and  $\beta_0, \beta_1, \dots, \beta_i$  are regression coefficients.

Near-IR model performance is reported as the multiple coefficient of determination ( $R^2$ ) and standard error of calibration (SEC) of each calibration. It is also reported as standard error of performance (SEP) and bias for each validation set and similarly reported for each test set. Mathematical definitions for SEC, SEP, and bias are defined elsewhere<sup>9</sup>.

### Error analysis

The overall error of the near-IR models was further partitioned into that which could be attributed to the reference method and to scan repeatability,

**Table I** Means and standard deviations (s.d.) of protein content, dry mass and moisture content of single wheat kernels

Wheat class <sup>a</sup>	Set <sup>b</sup>	n	Protein content		Dry mass mg		Moisture content	
			[% , 12% moisture basis]				[% wet basis]	
HRW	C	150	10.86	(2.57)	29.78	(6.40)	8.76	(0.35)
	V	150	10.71	(2.57)	28.06	(6.73)	8.77	(0.33)
	T	420	10.50	(2.28)	28.78	(6.48)	8.86	(0.30)
	A	720	10.62	(2.41)	28.84	(6.53)	8.82	(0.32)
HRS	C	150	13.26	(2.85)	30.27	(7.40)	8.61	(0.34)
	V	150	12.80	(2.76)	30.50	(6.89)	8.64	(0.33)
	T	320	13.25	(2.45)	30.32	(7.46)	8.63	(0.30)
	A	620	13.15	(2.63)	30.36	(7.30)	8.63	(0.32)
SRW	C	150	8.85	(1.58)	29.24	(6.95)	8.94	(0.33)
	V	150	8.99	(1.79)	28.23	(7.64)	8.95	(0.35)
	T	280	8.70	(1.70)	27.94	(7.70)	9.06	(0.26)
	A	580	8.81	(1.70)	28.35	(7.50)	9.00	(0.31)
HWW	C	150	11.93	(2.57)	38.76	(9.75)	8.51	(0.26)
	V	150	12.02	(2.24)	37.99	(9.29)	8.51	(0.28)
	T	260	12.05	(1.89)	38.14	(10.12)	8.38	(0.32)
	A	560	12.01	(2.18)	38.27	(9.80)	8.45	(0.30)
SWW	C	150	10.38	(2.70)	33.74	(9.88)	8.90	(0.33)
	V	150	10.38	(2.71)	34.06	(9.13)	8.91	(0.34)
	T	400	9.89	(2.19)	35.32	(9.72)	8.81	(0.29)
	A	700	10.10	(2.43)	34.71	(9.64)	8.85	(0.31)
RED	C	450	10.99	(3.00)	29.76	(6.92)	8.77	(0.37)
	V	450	10.84	(2.86)	28.93	(7.17)	8.79	(0.36)
	T	1020	10.87	(2.82)	29.03	(7.20)	8.84	(0.33)
	A	1920	10.89	(2.87)	29.18	(7.13)	8.81	(0.35)
WHITE	C	300	11.16	(2.74)	36.25	(10.12)	8.70	(0.35)
	V	300	11.20	(2.62)	36.02	(9.40)	8.71	(0.37)
	T	660	10.74	(2.33)	36.43	(9.97)	8.64	(0.36)
	A	1260	10.95	(2.51)	36.29	(9.87)	8.67	(0.36)
ALL	C	750	11.06	(2.90)	32.36	(8.93)	8.74	(0.36)
	V	750	10.98	(2.77)	31.77	(8.84)	8.76	(0.36)
	T	1680	10.82	(2.64)	31.94	(9.14)	8.76	(0.36)
	A	3180	10.91	(2.73)	32.00	(9.02)	8.76	(0.36)

<sup>a</sup>HRW=hard red winter, HRS=hard red spring, SRW=soft red winter, HWW=hard white wheat, SWW=soft white wheat, RED=HRW+HRS+SRW, WHITE=HWW+SWW, ALL=all wheat classes.

<sup>b</sup>C=calibration, V=validation, T=test, A=C+V+T.

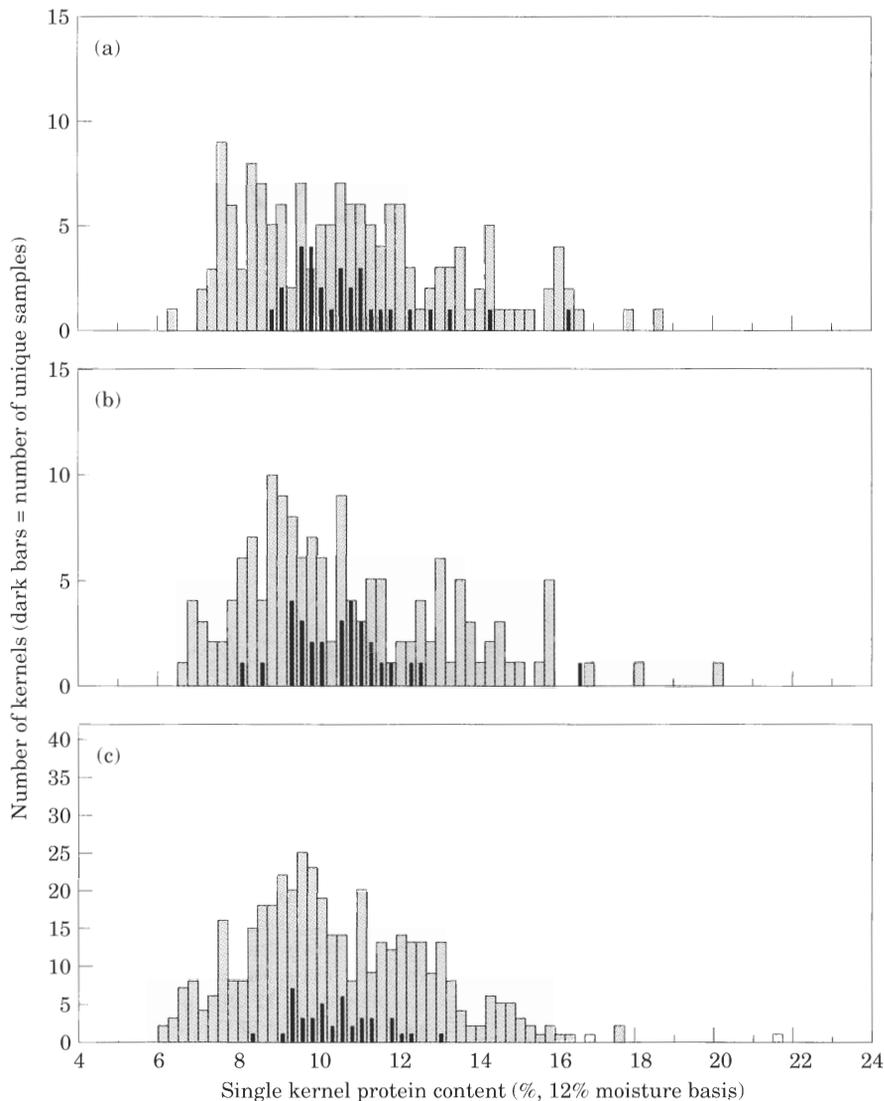
with the remaining value collectively called the chemometric error. Using the equation,

$$SEP^2 = \sigma_{reference}^2 + \sigma_{repeatability}^2 + \sigma_{chemometric}^2 \quad (2)$$

the best estimates of the squared reference method and scan repeatability errors are subtracted from the test set model performance ( $SEP^2$ ) to obtain the squared chemometric error. Reference method error was estimated by periodic combustion measurements of 40-mg portions from a thoroughly mixed 5-g check of ground HRW wheat possessing a nominal protein content of 13.4%. Prior to the combustion of the kernels within a holding tray (80 kernels/tray), a minimum of three portions of check were analyzed for protein content. Likewise,

a minimum of three portions were analyzed at the end of a day, having analyzed one or (more often) two complete trays during the day. The reference method error was defined as the standard deviation of the protein content measurement on 200 portions of check.

Repeatability error was estimated by applying an all-class (i.e. ALL) PLS model to multiple spectra collected on a set of 15 kernels. From an arbitrarily chosen unique sample from each class, a small, medium, and large kernel were selected (avoiding kernels destined for calibration sets). Each kernel was loaded, scanned, and its spectrum stored to disk a total of 25 times. With each reload, care was taken to place the kernel on the shaft in



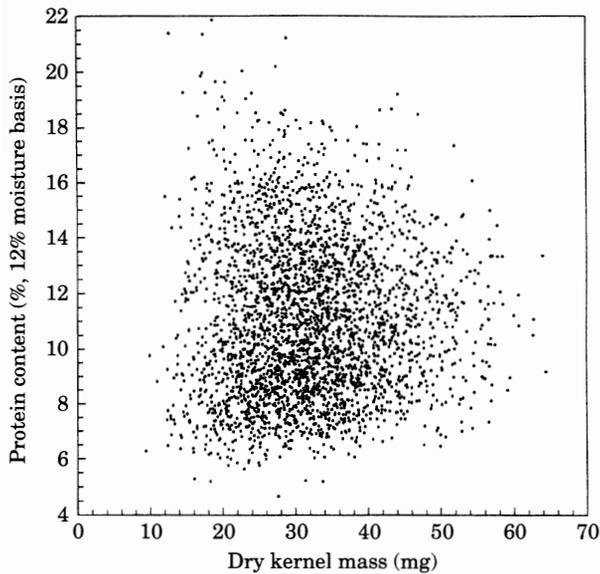
**Figure 1** Single-kernel protein content of hard red winter wheat samples. Shaded bars: distributions of single-kernel protein content (by combustion). (a) calibration set ( $n=150$ ), (b) validation set ( $n=150$ ), (c) test set ( $n=420$ ). Dark bars: distributions of sample protein contents, based on the mass average of the single kernels within each unique sample [No. kernels/sample = 5 for sets (a) and (b), 10 for set (c)].

the same orientation, thus simulating ideal conditions for kernel alignment. Repeatability error was defined as the square root of the average ( $n=15$  kernels) variance ( $n=25$  spectra/kernel) of predicted protein.

## RESULTS AND DISCUSSION

The distributions of s.k. protein content by combustion for the calibration, validation, and test sets of HRW are shown in Figure 1. Similarly shaped histograms (not shown) were observed for the

remaining four classes. Histograms of the five-kernel-per-sample (in calibration and validation sets) and 10-kernel-per-sample (in test set) weighted averages overlay the s.k. histograms. Single-kernel protein content typically ranged from *c.* 6 to 20% for the hard classes and *c.* 5 to 17% for the soft classes. Similar results for sets consisting of one or four certified seed samples per class ( $n=96$  kernels/sample, six classes total) were reported in a previous study<sup>1</sup>. The large range appears plausible upon review of Levi and Anderson<sup>12</sup>, who measured ranges approaching 6% protein for kernels

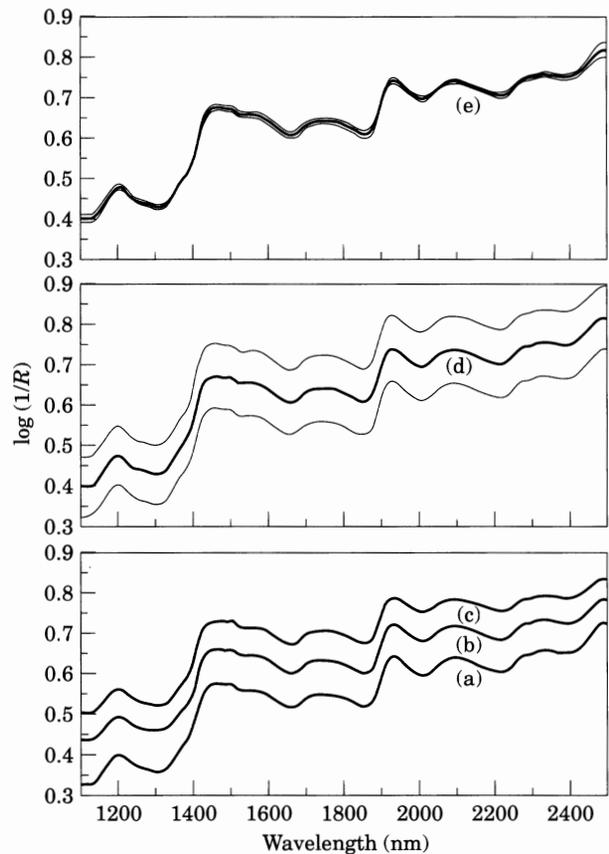


**Figure 2** Demonstration of lack of relationship between single-kernel protein content and kernel mass for all kernels in study ( $n=3180$ ).

within the same wheat spike. In general, the calibration and validation set [e.g. Fig. 1 (a) and (b)] distributions closely matched each other, as to be expected by the manner in which these sets were defined. The test sets [e.g., Fig. 1(c)] tended to be more normally distributed than the calibration and validation sets; however, their ranges were often as wide.

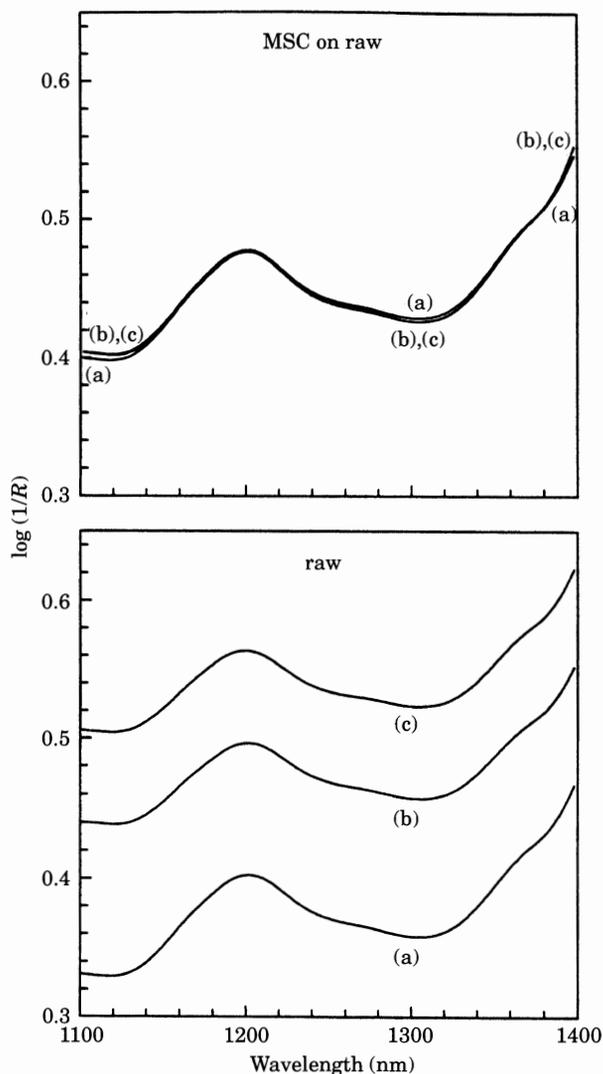
A lack of relationship between kernel protein content and kernel mass is demonstrated in Figure 2, using all kernels. Similar behavior (plots not shown) was noted when each class was isolated. Thus, it is impossible to estimate protein content of sound, mature wheat kernels, strictly based on visual examination.

Typical spectra of single kernels are displayed in the lower graph of Figure 3. Two of three HWW kernels [curves (a) and (b)] from the same unique sample were selected to represent the case of constant protein content (*c.* 12%) and different mass (47 and 26 mg, respectively). Likewise, the case of constant mass (*c.* 25 mg) and different protein content is represented by curves (b) and (c) (12 and 18%, respectively). The most noticeable difference between these curves is the vertical offset between any two curves. Surprisingly, the magnitude of the offset for the constant mass case is nearly as large as that for the constant protein case. To compare the variation in spectral response



**Figure 3** Reflectance spectra of wheat kernels. Lower: typical single-kernel reflectance spectra of three hard white wheat kernels from the same sample. The protein content (pc), dry mass, and moisture content (mc) of each kernel are as follows: (a), 12.09% pc, 46.64 mg, 8.44% mc; (b), 11.73% pc, 25.87 mg, 8.88% mc; (c), 18.26% pc, 25.15 mg, 9.11% mc. Middle: average spectrum (d) of the hard white wheat calibration set ( $n=150$ ), with  $\pm$  one-standard deviation envelope. Upper: average spectrum (e) of the hard white wheat calibration set ( $n=150$ ) pretreated by multiplicative scatter correction, with  $\pm$  one-standard deviation envelope.

of the three kernels to that of the HWW calibration set (which is inclusive of the three kernels), the 150-kernel average spectrum and the  $\pm$  one standard deviation spectra are displayed in the middle graph of Figure 3. In addition to kernel size, spectral variation can also be caused by kernel-to-kernel differences in vitriousness<sup>13</sup> (this cause laying beyond the scope of the present study). By comparing the middle and lower graphs of Figure 3, it is seen that the offsets in either of the cases [(a) vs. (b); (b) vs. (c)] were approximately equivalent to the standard deviation envelope. Also plotted in Figure 3 (upper graph) is the average (with  $\pm$  one standard deviation) spectrum of the 150 multiplicatively



**Figure 4** Demonstration of the multiplicative scatter correction (MSC) normalization procedure. Curves (a), (b), and (c) correspond to the same kernels described in Figure 3. MSC was performed on the spectral data in the 1100–1398 nm region. Before and after normalization conditions are indicated by the lower and upper graphs, respectively.

scatter corrected spectra. Evident from this graph is the reduction in spectral variability that is accomplished by MSC. With MSC applied to strictly the 1100–1398-nm region (Fig. 4), spectral differences caused by mass are faintly apparent, as seen by the slight displacement of (a) and (b) curves in the regions 1100–1150 and 1260–1340 nm. Spectral differences caused by protein content [(b) vs. (c)] are even less apparent.

Based on an 1100–1398-nm wavelength region,

statistics for the optimal PLS models are summarized in Table II for each wheat class. Two kernels, one each from HRW and HRS, were deemed as outliers during initial calibration runs (having modeled values that were different than the reference values by more than 2% protein units) and therefore removed from subsequent analyses. With the exception of the SRW model,  $R^2$  values exceeded 0.96 and the SEC ranged from 0.379 (HRW) to 0.518% protein (HRS). The number of PLS factors determined as optimal during the cross-validation procedure ranged from 7 (HRS) to 10 (HRW and SRW). When kernels from classes of like color were pooled (RED and WHITE, Table II), the  $R^2$  and SEC values were similar to the single-class models, although additional factors were required. Likewise, the pooling of all (designated as ALL) wheat classes resulted in a model with  $R^2$  and SEC values within the range of values for the individual class models.

The results of applying each single-class model to the validation and test sets of the appropriate class are summarized in Table III. Also included are the results of applying the RED, WHITE, and ALL models to the appropriate single-class and pooled-class groups. Single-class model SEPs ranged from 0.418 to 0.589% protein for the validation sets. Similarly, the range was 0.462 to 0.590% protein for the test sets. Normalized with respect to the standard deviations of reference protein contents for all kernels within each class (Table I), the relative predicted deviation ( $RPD = SEP_{\text{test set}}/SD_{\text{all}}$ ) ranged from 2.88 for SRW to 4.72 for HWW, falling within the range sufficient for screening in plant-breeding programs<sup>14</sup>. In Figure 5, modeled (10-factor single-class model) protein content is plotted against the corresponding reference values in the HRW test set. (Plots for the other wheat classes demonstrated similar behavior and are therefore omitted.) The greatest absolute bias for the single-class models, expressed with sign restored, was  $-0.042\%$  protein for the validation sets and  $0.138\%$  protein for the test sets. When the RED or WHITE models were applied to the appropriate single-class validation and test sets, the SEPs were smaller than their corresponding one-class values in all cases but one (HWW test). When a comparison of two correlated variances<sup>15</sup> was performed on the test set to determine the level of significant difference between the single-class SEP and the appropriate RED or WHITE SEP, 1% levels of significance were determined for HRW, HRS, and HWW, a 5%

**Table II** Calibration equation statistics for partial least squares (PLS) models

Modeled classes <sup>a</sup>	<i>n</i>	No. PLS factors	R <sup>2</sup>	SEC <sup>b</sup>
HRW	149	10	0.979	0.379
HRS	149	7	0.968	0.518
SRW	150	10	0.901	0.514
HWW	150	9	0.978	0.389
SWW	150	9	0.971	0.473
RED	448	11	0.976	0.466
WHITE	300	11	0.978	0.409
ALL	748	11	0.971	0.494

<sup>a</sup>HRW=hard red winter, HRS=hard red spring, SRW=soft red winter, HWW=hard white wheat, SWW=soft white wheat, RED=HRW + HRS + SRW, WHITE=HWW + SWW, ALL=all wheat classes.

<sup>b</sup>SEC=standard error of calibration, defined in text (units are % protein, 12% moisture basis).

**Table III** Summary of near-infrared partial least squares (PLS) model performances

Wheat class <sup>a</sup>	Model applied <sup>b</sup>	Units of percent protein (12% moisture basis)				
		Validation set <sup>c</sup>		Test set <sup>d</sup>		
		SEP	Bias	SEP <sup>e</sup>	Bias	
HRW	HRW	0.518	0.016	1% [0.587]	1% -0.012	
	RED	0.481	-0.093			0.546
	ALL	0.484	-0.180			0.542
HRS	HRS	0.559	0.016	1% [0.579]	1% -0.056	
	RED	0.507	0.026			0.476
	ALL	0.537	-0.038			0.506
SRW	SRW	0.589	-0.034	ns [0.590]	ns 0.021	
	RED	0.593	0.010			0.584
	ALL	0.565	-0.043			0.574
HWW	HWW	0.418	-0.042	1% [0.462]	ns -0.092	
	WHITE	0.410	-0.002			0.508
	ALL	0.420	0.104			0.468
SWW	SWW	0.558	-0.037	5% [0.587]	1% 0.138	
	WHITE	0.555	-0.041			0.554
	ALL	0.565	0.125			0.491
RED	RED	0.530	-0.019	0.548	-0.029	
WHITE	WHITE	0.487	-0.021	0.537	0.012	
ALL	ALL	0.527	-0.006	0.543	-0.003	

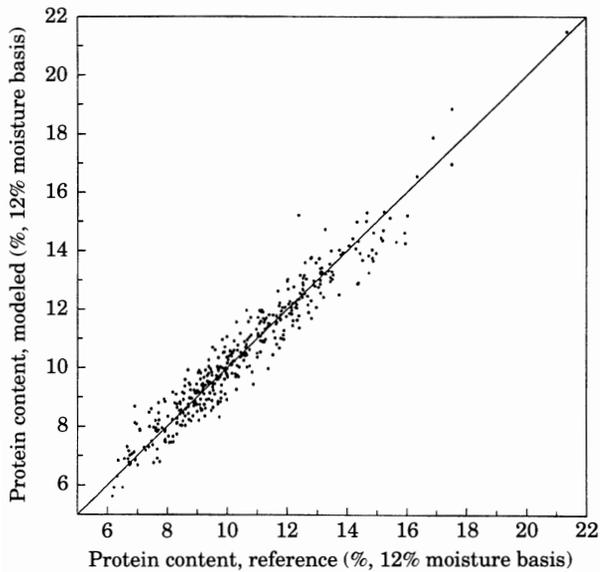
<sup>a</sup>HRW=hard red winter, HRS=hard red spring, SRW=soft red winter, HWW=hard white wheat, SWW=soft white wheat, RED=HRW + HRS + SRW, WHITE=HWW + SWW, ALL=all wheat classes.

<sup>b</sup>Applied models are those which are summarized in Table II.

<sup>c</sup>Number of validation kernels, *n*, varies with wheat class as follows: HRW=150, HRS=150, SRW=150, HWW=150, SWW=150, RED=450, WHITE=300, ALL=750.

<sup>d</sup>Number of test kernels, *n*, varies with wheat class as follows: HRW=420, HRS=320, SRW=280, HWW=260, SWW=400, RED=1020, WHITE=660, ALL=1680.

<sup>e</sup>Values appearing beside brackets indicate the statistical levels of significance for a comparison of two correlated variances test performed on the squares of the value identified by the termini of the brackets (ns=not significant at the 5% level).



**Figure 5** Typical single-kernel protein content model results. For the plot shown, the single-class 10-factor partial least squares model for hard red winter (HRW) wheat (performance summarized in Tables II and III) was applied to the HRW test set ( $n=420$ ).

level for SWW, while no significant difference was determined for SRW. Compared to the single-class models, the application of the ALL model to each class resulted in a smaller SEP for all validation and test sets except both sets of HWW and the validation set of SWW. Applying the same comparison of correlated variances procedure to the test set SEPs from the single- and ALL-class models, 1% levels of significance were determined for HRW, HRS, and SWW. No significant differences were determined for the SEPs of HRS and HWW. Thus, in all cases but one (WHITE model on HWW), the multiple-class models caused an improvement in, or at worst, no significant change in model error.

When the single-class validation sets were pooled to form the appropriate RED, WHITE, and ALL groups (Table III), the SEPs ranged from 0.487 to 0.530% protein. Similarly, the SEP range for the test sets of these same groups was 0.537 to 0.548% protein, with an RPD range of 4.67 to 5.24 (the latter value being greater than the threshold value of 5 for accuracies sufficient for quality control<sup>14</sup>). The greatest absolute bias, with sign restored, was  $-0.021\%$  protein for the validation sets and  $-0.029\%$  protein for the test sets. Comparing these biases with those determined when the multiple-class models (i.e. RED, WHITE,

ALL) were applied to the single-class validation and test sets, the decrease in absolute bias (gained from the pooling of sets) was the mathematical result of averaging biases of opposite sign. A weak ( $r=0.113$ ) though significant correlation existed between the PLS ALL model s.k. protein residuals (i.e. model-reference) and s.k. dry mass for the ALL test group. Assuming a direct relation between kernel size (not measured) and mass, it can be concluded that model error was not largely affected by kernel size.

The results of repeated measurements of s.k. spectra to determine the protein contents of a small, medium, and large kernel from each of the five wheat classes are summarized in Table IV. Because of the good overall performance of the ALL model, this model was applied to all repeatability kernels. Expressed as the standard deviation of the modeled protein content for each kernel, the repeatability ranged from 0.163 to 0.461% protein. With the exception of SRW, repeatability was comparable among wheat classes. Pooling all repeatability kernels, only a weak negative correlation was found between repeatability and kernel mass ( $r=-0.598$ ), while the correlation between repeatability and reference protein content ( $r=-0.243$ ) was not significant.

### Choice of wavelength region

The HRW calibration set was arbitrarily selected for investigating the effect of wavelength region on model performance. Figure 6 contains the results of cross validations on wavelength regions of width varying between 300 and 1300 nm, with the lower end of each region varying between 1100 and 2100 nm. Each trial is displayed on the graph as the number of factors deemed as optimal for the combination of region width (*x*-axis) and starting wavelength (italicized value). The *y* value, RMSD, is the root mean squared difference between modeled and reference values of s.k. protein content of the removed samples during a one-sample-out cross validation. Models which used 1100 nm as the starting wavelength were superior to other models that used higher starting wavelengths. At the 1100-nm starting wavelength, as the width of the wavelength region increased (by 100-nm steps), the RMSD increased slightly. Additionally, more factors were required as the region increased, presumably because of a non-linear spectral response (i.e. absorption as a function of

**Table IV** Repeatability statistics from applying the 11-Factor 'ALL' PLS model to  $n=25$  loads and scans of each of three sizes of kernels per wheat class

Wheat class <sup>a</sup>	Dry mass (mg)	Protein content (% , 12% moisture basis)		
		Reference	Model mean	Model standard deviation
HRW	23.66	6.70	6.45	0.295
	31.53	11.79	11.85	0.187
	50.81	14.45	13.64	0.211
HRS	22.59	8.10	8.70	0.248
	33.10	13.08	13.57	0.223
	38.83	15.55	16.09	0.295
SRW	18.14	9.03	8.82	0.461
	27.31	7.04	6.50	0.331
	33.73	8.26	7.00	0.184
HWW	34.73	10.34	10.36	0.181
	41.73	8.61	8.76	0.240
	50.32	12.61	12.21	0.163
SWW	32.81	9.70	10.41	0.190
	43.34	9.37	9.83	0.170
	50.20	9.30	10.10	0.236

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

species concentration) at the longer wavelengths. The 300-nm wide wavelength region with starting wavelength = 1100 nm represented the best combination of low RMSD and small number of PLS factors, hence the reason for its use in models involving HRW or the other wheat classes.

### Error analysis

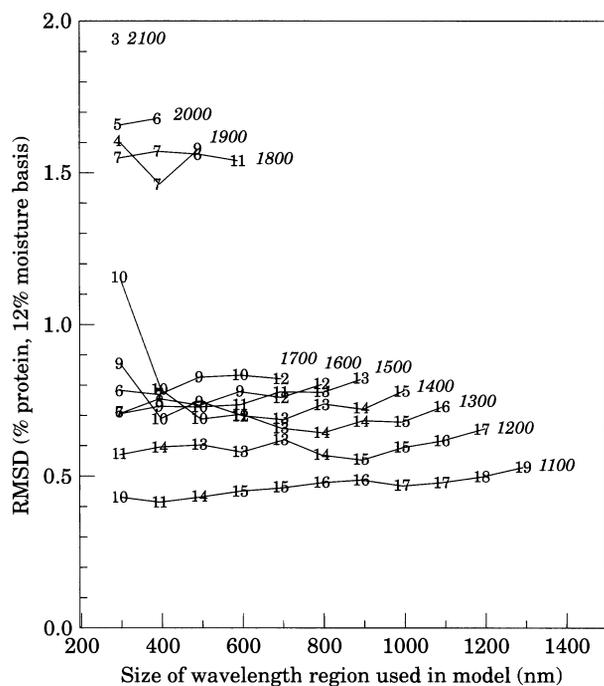
Referring to Equation 2, the squared values for the reference method error and scan repeatability error were as follows:  $\sigma_{reference}^2 = 0.0614$  (% protein)<sup>2</sup>,  $\sigma_{repeatability}^2 = 0.0639$  (% protein)<sup>2</sup>. Subtracting these values from the squared SEP of the ALL model applied to all test kernels, the chemometric error,  $\sigma_{chemometrics}$  was estimated to be 0.411% protein. Thus, under the ideal conditions of negligible reference method error and negligible kernel orientation error, the collective error would be less than one-half percent protein. Possible sources of error that contribute to the chemometric error include (1) uncorrectable spectral non-linearities caused by kernel size and shape, (2) stray light within the detection assembly, (3) varietal differences in the nitrogen-to-protein ratio caused by differences in the relative proportions of their amino acids<sup>16</sup> and (4) error in the s.k. moisture analysis (probably small, e.g. for an 8–16% protein range and kernels at nominally 9% moisture, an

error of  $\pm 0.5\%$  moisture transforms to an absolute error of  $<0.09\%$  protein).

### MLR analyses

From a stepwise search of wavelengths throughout the 1000–1398 nm region, the best single- and multiple-class eight-term models are summarized in Table V. MSC or other forms of spectral pretreatment were not performed on the  $\log_{10}(1/R)$  data. Using the same class structure for modeling as in the PLS analyses (Table II), all models produced an  $R^2$  value of 0.900 or greater. The range in SEC for the five single-class models was 0.513–0.623% protein. For the RED, WHITE, and ALL models, the SECs were very similar, ranging from 0.626 to 0.648% protein. With the exception of the SRW model, the SEC of each MLR model was greater than that of the corresponding PLS model. However, considering the simpler design requirements for an instrument which utilizes eight discrete filters as opposed to a scanning instrument, the sacrifice in model performance may be acceptable.

Consistent with the PLS analyses, the lower wavelengths, with most less than 1200 nm, were preferably selected by the step-wise analysis, despite allowing an 1100–1798 nm search region. The highest wavelength selected by any of the MLR models was 1504 nm. Generally, the three



**Figure 6** Sensitivity of partial least squares (PLS) model performance to wavelength region chosen. In this example, one-sample out cross validation has been performed on the hard red winter calibration set ( $n = 149$ ). RMSD = root mean square difference (modeled–reference) of protein content, accumulated for each kernel when rotated out during cross validation. Italicized numbers refer to the starting wavelength in each region whose width is identified by the  $x$ -axis. Numerical values used as plot symbols represent the optimal numbers of PLS factors identified during cross validation.

lowest wavelengths selected (*c.* 1106, 1138, and 1156 nm) were the same to within  $\pm 4$  nm among all models. Comparing these wavelengths with absorption bands attributable to protein, starch, lipid, and water<sup>17</sup>, it is reasoned that the first wavelength is for baseline correction, while the second and third coincide with protein- and water-absorption bands, respectively, or their interaction. The next three wavelengths (*c.* 1170, 1186, and 1200 nm) were also consistent among all models, albeit with a greater envelope. The 1170- and 1186-nm wavelengths coincide with protein-absorption bands, and the 1200-nm wavelength coincides with the peak of a broad band that is typically apparent in carbohydrates. The two highest wavelengths were most variable, though even with these, several models had common wavelengths (e.g. 1306–1318 nm for the seventh wavelength of the HRS, SWW, and WHITE models; 1500–1504 nm for the eighth wavelength of the

HRW, SRW, HWW, RED, and ALL models). Interpretation of these wavelengths is difficult; however, for the RED and ALL models, the 1486–1488 nm and 1500–1502 nm wavelengths possess coefficients of equal magnitude but opposite sign, such that when combined, these wavelengths are sensitive to the slope of a protein band whose peak occurs at 1496 nm.

When the MLR single-class models were applied to the validation and test sets (Table VI), SEPs ranged from 0.566 to 0.758% protein for the validation sets and from 0.643 to 0.720% protein for the test sets. Multiple-class models applied to the same sets produced SEPs ranging from 0.575 to 0.736% protein for the validation sets and 0.590–0.685% protein for the test sets. With the exception of the SRW class, the validation and test performances of multiple-class models were generally equivalent to, if not better than, those of the single-class models. Generally, MLR models produced SEPs 0.1–0.2% protein units higher than the corresponding SEPs of the PLS models.

## SUMMARY AND CONCLUSIONS

A method was developed for measuring the protein content of individual wheat kernels non-destructively. Based on the diffuse reflectance of near-IR radiation (1100–2498 nm collected, 1100–1504 nm used) from revolving single kernels, reasonably accurate PLS or MLR models ( $R^2 = 0.900$ – $0.971$ ) were developed for various combinations of five commercial U.S. wheat classes. When applied to test sets consisting of kernels from samples omitted from calibration, these models produced a SEP that ranged from 0.462 to 0.720% protein (dependent on modeling technique, number of wheat classes used to develop the model, and the wheat class tested). Additionally, the following conclusions are drawn: (1) models developed from the pooling of wheat kernels from more than one wheat class generally performed as well as, and often better than, models developed from one class; (2) the most essential spectral region for s.k. protein content analysis is 1100–1400 nm. Lengthening this region (on the 1400 side) or, more detrimentally, shifting the region to higher wavelengths resulted in models with diminished accuracy; and (3) an eight wavelength MLR model, developed from all five classes, produced test-set SEPs ranging from 0.590 to 0.685% protein, which were 0.1–0.2% protein units higher than those

**Table V** Calibration equation statistics for eight-term multiple linear regression models

Modeled classes <sup>a</sup>	Wavelengths, nm (coefficients)	n	R <sup>2</sup>	SEC <sup>b</sup>
HRW	1108 (-489.2), 1138 (2474.9), 1156 (-4540.9), 1176 (2390.1), 1184 (2785.7), 1198 (-2595.9), 1434 (-214.6), 1504 (190.0) (+11.69)	149	0.960	0.525
HRS	1104 (-420.6), 1144 (2856.0), 1154 (-4284.7), 1184 (4862.3), 1198 (-2913.3), 1280 (-1850.4), 1306 (3225.3), 1328 (-1476.5) (+16.76)	149	0.953	0.623
SRW	1104 (-738.3), 1138 (2218.0), 1152 (-2511.0), 1178 (1573.0), 1192 (2195.8), 1202 (-2745.0), 1470 (-229.6), 1502 (240.0) (+9.23)	150	0.900	0.513
HWW	1126 (-224.1), 1148 (3498.3), 1156 (-5380.5), 1182 (5174.6), 1198 (-3123.7), 1368 (109.9), 1420 (-158.8), 1502 (108.3) (+9.17)	150	0.952	0.580
SWW	1104 (-828.6), 1134 (2505.0), 1154 (-4989.6), 1168 (3697.5), 1196 (2693.4), 1208 (-3405.2), 1318 (2780.4), 1324 (-2455.0) (+14.50)	150	0.959	0.560
RED	1108 (-937.5), 1138 (2758.0), 1158 (-5267.7), 1168 (3105.6), 1188 (3387.4), 1200 (-3056.9), 1486 (-283.0), 1500 (292.8) (+13.71)	448	0.956	0.627
WHITE	1106 (-482.4), 1136 (1870.0), 1156 (-4446.1), 1168 (2380.0), 1184 (2672.7), 1206 (-2221.2), 1318 (1425.7), 1326 (-1198.7) (+13.28)	300	0.949	0.626
ALL	1106 (-808.0), 1138 (2827.5), 1156 (-4916.1), 1170 (2733.2), 1186 (2860.4), 1200 (-2707.2), 1488 (-206.6), 1502 (216.8) (+12.44)	748	0.950	0.648

<sup>a</sup>HRW = hard red winter, HRS = hard red spring, SRW = soft red winter, HWW = hard white wheat, SWW = soft white wheat, RED = HRW + HRS + SRW, WHITE = HWW + SWW, ALL = all wheat classes.

<sup>b</sup>SEC = standard error of calibration, units of % protein (12% moisture basis).

**Table VI** Summary of eight-term multiple linear regression (MLR) model performances

Wheat class <sup>a</sup>	Model applied <sup>b</sup>	Units of percent protein (12% moisture basis)			
		Validation set <sup>c</sup>		Test set <sup>d</sup>	
		SEP	Bias	SEP	Bias
HRW	HRW	0.660	0.010	0.644	0.051
	RED	0.632	-0.041	0.622	-0.081
	ALL	0.640	-0.110	0.623	-0.111
HRS	HRS	0.691	0.118	0.715	-0.042
	RED	0.688	0.040	0.682	-0.133
	ALL	0.691	-0.060	0.685	-0.267
SRW	SRW	0.675	0.079	0.618	0.120
	RED	0.728	0.148	0.683	0.169
	ALL	0.736	0.044	0.664	0.140
HWW	HWW	0.566	0.120	0.643	-0.022
	WHITE	0.585	0.059	0.620	-0.017
	ALL	0.575	0.233	0.590	0.193
SWW	SWW	0.758	-0.053	0.720	0.165
	WHITE	0.711	-0.011	0.662	0.200
	ALL	0.686	0.085	0.623	0.246
RED	RED	0.687	0.049	0.669	-0.029
WHITE	WHITE	0.651	0.024	0.654	0.114
ALL	ALL	0.677	0.038	0.666	0.033

<sup>a</sup>HRW = hard red winter, HRS = hard red wpring, SRW = soft red winter, HWW = hard white wheat, SWW = soft white wheat, RED = HRW + HRS + SRW, WHITE = HWW + SWW, ALL = all wheat classes.

<sup>b</sup>Applied models are those which are summarized in Table V.

<sup>c</sup>Number of validation kernels,  $n$ , is defined as follows: HRW = HRS = SRW = HWW = SWW = 150, RED = 450, WHITE = 300, ALL = 750.

<sup>d</sup>Number of test kernels,  $n$ , varies with wheat class as follows: HRW = 420, HRS = 320, SRW = 280, HWW = 260, SWW = 400, RED = 1020, WHITE = 600, ALL = 1680.

produced by a PLS model. Despite the diminished accuracy, the MLR model might be preferable from the standpoint of instrument simplicity. Regardless of modeling technique employed, incorporation of samples from more than one crop year into a calibration set is advised for the purpose of improving model robustness.

### Acknowledgments

Dr Aarion Johnson, USDA-GIPSA, Kansas City, Missouri, is thanked for providing the wheat samples and survey information. Joyce Shaffer, USDA-ARS, Beltsville, Maryland, collected and managed the spectral and reference data. Karl Norris, Beltsville, MD, is thanked for early discussions on chemometric procedures.

### REFERENCES

1. Wrigley, C.W. and Beitz, J.A. Proteins and amino acids. In 'Wheat: Chemistry and Technology, Volume one,

- third edition' (Y. Pomeranz, ed.), American Association of Cereal Chemists, St Paul, MN (1988) pp 159-285.
2. Peterson, C.J., Graybosch, R.A., Baenziger, P.S. and Grombacher, A.W. Genotype and environment effects on quality characteristics of hard red winter wheat. *Crop Science* **32** (1992) 98-103.
3. Graybosch, R.A., Peterson, C.J., Baenziger, P.S. and Shelton, D.R. Environmental modification of hard red winter wheat flour protein composition. *Journal of Cereal Science* **22** (1995) 45-51.
4. Delwiche, S.R. Single wheat kernel analysis by near-infrared transmittance: protein content. *Cereal Chemistry* **72** (1995) 11-16.
5. Delwiche, S.R. and Massie, D.R. Classification of wheat by visible and near-infrared reflectance from single kernels. *Cereal Chemistry* **73** (1996) 399-405.
6. American Society of Agricultural Engineers. Moisture measurement—grain and seeds. In 'American Society of Agricultural-Engineers Yearbook of Standards', American Society of Agricultural Engineers, St Joseph, MI (1983) **S352.1**.
7. Geladi, P. and Kowalski, B.R. Partial least squares regression: a tutorial. *Analytica Chimica Acta* **185** (1986) 1-17.
8. Lindberg, W., Persson, J.A. and Wold, S. Partial least-squares method for spectrofluorometric analysis of mixtures of humic acid and ligninsulfonate. *Analytical Chemistry* **55** (1983) 643-648.

9. Hruschka, W.R. Data analysis: wavelength selection methods. In 'Near-Infrared Technology in the Agricultural and Food Industries', (P.C. Williams and K.H. Norris, eds), American Association of Cereal Chemists, St Paul, MN (1987) pp 35-54.
10. GALACTIC. Add-on application software. PLSplus version 2.1G for GRAMS/386 (Rev. 07/29/92). Galactic Industries: Salem, NH (1992) 94 pp.
11. Martens, H. and Naes, T. Multivariate calibration by data compression. In 'Near-Infrared Technology in the Agricultural and Food Industries', (P.C. Williams and K.H. Norris, eds), American Association of Cereal Chemists, St Paul, MN (1987) p 73.
12. Levi, I. and Anderson, J.A. Variations in protein contents of plants, heads, spikelets, and individual kernels of wheat. *Canadian Journal of Research, Section F* **28** (1950) 71-81.
13. Delwiche, S.R. Measurement of single kernel wheat hardness using near-infrared transmittance. *Transactions of the American Society of Agricultural Engineers* **36** (1993) 1431-1437.
14. Williams, P.C. and Sobering, D.C. Comparison of commercial near infrared transmittance and reflectance instruments for analysis of whole grains and seeds. *Journal of Near Infrared Spectroscopy* **1** (1993) 25-32.
15. Snedecor, G.W. and Cochran, W.G. Statistical Methods, sixth edition. Iowa State University Press, Ames, Iowa (1967) pp 195-197.
16. Tkachuk, R. Nitrogen-to-protein conversion factors for cereals and oilseed meals. *Cereal Chemistry* **46** (1969) 419-423.
17. Williams, P.C. and Norris, K.H. Qualitative applications of near-infrared reflectance spectroscopy. In 'Near-Infrared Technology in the Agricultural and Food Industries', (P.C. Williams and K.H. Norris, eds), American Association of Cereal Chemists, St Paul, MN (1987) pp 241-246.