

Development of Nondestructive Screening Methods for Single Kernel Characterization of Wheat

Jesper Pram Nielsen,^{1,2} Dorthe Kjær Pedersen,¹ and Lars Munck¹

ABSTRACT

Cereal Chem. 80(3):274–280

The development of nondestructive screening methods for single seed protein, vitreousness, density, and hardness index has been studied for single kernels of European wheat. A single kernel procedure was applied involving, image analysis, near-infrared transmittance (NIT) spectroscopy, laboratory density determination, single kernel characterization system (SKCS), and finally Kjeldahl protein determination on the crushed single kernels. Single kernel NIT spectroscopy showed excellent ability to determine protein content, and some ability for determination of single kernel vitreousness. Nondestructive determination of single kernel density, either based on NIT spectroscopy or based on image

analysis and kernel weight, needs to be further improved for practical use. The use of SKCS hardness index as a true single kernel hardness reference in a NIT prediction model resulted in a poor predictability. However, by applying an averaging approach, in which single seed replicate measurements are mathematically simulated, a very good NIT prediction model was achieved. This suggests that the single seed NIT spectra contain hardness information, but that a single seed hardness method with higher accuracy is needed to achieve a good NIT prediction model for single kernel hardness.

Protein content, kernel density in terms of test weight, and kernel vitreousness by visual inspection are normally used in the miller's quality evaluation of wheat for milling. Protein content largely determines the end-use quality, and premiums are often offered on high protein wheat. Test weight reflects kernel size and density, and should be above a certain level to secure a good flour yield. The vitreousness is used for evaluation of millability, even though the relationship between vitreousness and hardness is not straightforward. Vitreousness and hardness affects the milling processing of wheat, including tempering of the grains, flour yield, and the end-use properties such as particle size distributions and the amount of damaged starch. Grain hardness is mainly determined by the degree of adhesion between the starch granules and the protein matrix, with a tight adhesion of the starch granules in the hard wheat and a weaker adhesion in soft wheat. Even though wheat can be divided into genetically soft and hard classes, a substantial variation in texture is seen within the two classes, and the apparent vitreousness of the wheat is therefore used by the millers in their evaluation of millability.

Wheat quality evaluation has traditionally been performed on bulk samples, which implies that the characteristics of the individual kernels within the sample are lost, and thereby the opportunity to evaluate sample homogeneity. In seed sorting and grading by size, form and density for better and more uniform quality, the single seed is the functional unit to be investigated. New developments in instrumentation have made single kernel characterization possible, and for some quality parameters rapid enough to become a valuable tool for homogeneity evaluation in the cereal industry. The single kernel characterization system (SKCS) (4100, Perten Instruments, Reno, NV) is an example of such an instrument for rapid, albeit destructive, measurement of single kernel hardness, weight, diameter, and moisture content (Martin et al 1993). The single kernel measurements are normally conducted on 300 single kernels in a bulk sample to classify the sample into soft, hard, or mixed wheat.

One of the limitations of destructive single seed analysis is that several readings on the same kernels are impossible. It therefore becomes difficult to differentiate between instrument variability and kernel-to-kernel variability. By using nondestructive single

seed analyses these problems could be circumvented. Additionally, fast and nondestructive single kernel quality analyses would be valuable tools in plant breeding for quality selection in early generations and for single kernel quality evaluation within the heads.

Near-infrared spectroscopy on single kernels fulfils these requirements, and the technique has been used for several single kernel applications. Near-infrared transmittance (NIT) spectroscopy has been reported for determination of oil in maize (Orman and Schumann 1992) and meadowfoam (Patrick and Jolliff 1997), protein in wheat (Delwiche 1995) and soybeans (Abe et al 2000), and for wheat hardness (Delwiche 1993). NIR spectroscopy has similarly been applied for wheat classification (Delwiche and Massie 1996), for determination of single seed protein (Delwiche 1998; Delwiche and Hruschka 2000), for differentiation between vitreous and nonvitreous durum wheat kernels (Dowell 2000), and for assessment of heat-damaged wheat kernels (Wang et al 2001).

Image analysis is another method for fast nondestructive characterization of kernels. Image analysis has been used for discrimination between kernels of different species (Chtoui et al 1996), discrimination between wheat classes and cultivars (Zayas et al 1986) and, used in combination with physical measurements, for cultivar identification (Zayas et al 1996). Berman et al (1996) used the method for screening of flour milling yield in wheat breeding.

This investigation involves a combination of image analysis, NIT spectroscopy, hardness analysis (SKCS), protein analysis, as well as a simple laboratory density analysis applied on single kernels of European wheats. The report includes a survey of the use of nondestructive screening methods for prediction of single kernel protein, vitreousness, density, and hardness.

MATERIALS AND METHODS

Samples

Bulk samples of 43 different wheat cultivars or mixtures of cultivars in common use were collected from two different locations in Denmark (Jutland and Funen), representing both genetically hard and soft cultivars. To select fully developed kernels, the samples were screened on a 2.2-mm screen and the fractions >2.2 mm were stored separately in plastic bags. Five kernels were chosen randomly from each of the 86 bulk samples to make up the calibration set (430 kernels in total). Another 10 kernels from each of 11 of the 86 bulk samples (11 cultivars from Funen) were selected as the test set (110 kernels in total). Because the measurements of the kernels in the calibration set showed no significant

¹ Centre for Advanced Food Studies, Food Technology, Dept. of Dairy and Food Science, The Royal Veterinary and Agricultural University, Rolighedsvej 30, DK-1958 Frederiksberg C., Denmark.

² Corresponding author. E-mail: jpn@kvl.dk.

differences between the two locations, we chose to use test set kernels from only one of the locations. The single kernels were put through a sequence of measuring steps. The kernels were analyzed one-by-one, with their identity retained during the measurement procedure.

GrainCheck

Grain morphology was measured by digital image analysis (GrainCheck 310, FossTecator, Höganäs, Sweden). The instrument was used for single kernel characterization by manually placing each kernel under the RGB camera from which the kernels were imaged and from which several morphological and color characteristics were automatically assessed. In this investigation, nine kernel characteristics were registered from the instrument and used in the data analysis. They were kernel width, kernel length, roundness, area, volume, red reflectance, green reflectance, blue reflectance, and total light reflectance.

NIT Spectra

After the digital image analysis, the single kernels were moved to a food and feed analyzer (Infratec 1255, FossTecator). Each kernel was placed in a single seed sample cassette with slots for 23 single kernels, and NIT spectra at 850–1,050 nm were automatically recorded. Spectra were recorded three times on each kernel, and the average of the three spectra was used. The position of the kernels in the sample cassette was manually changed between each of the three measurements. The time required for scanning (single scan) 23 single kernels in the cassette was ≈90 sec.

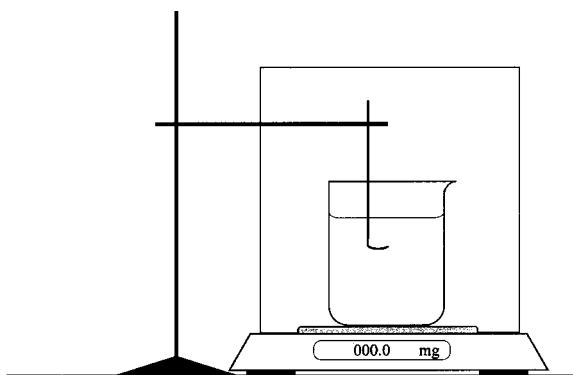


Fig. 1. Method for determination of single kernel volume.

Single Kernel Density

A laboratory single kernel density measurement was developed and applied to the 110 test set kernels before the SKCS analysis. The kernels were individually weighed to the nearest 0.1 mg (Mettler/Toledo scale, Type AB204). When immersing a wheat kernel in water, the weight of the displaced water divided by the density of the water equals the kernel volume. This measurement was made by using the equipment shown in Fig. 1, which was specially designed for the purpose. A beaker containing water at 20°C was placed on the scale. A single kernel holder (modified sample spoon) was mounted on a rack outside the scale chamber (without touching the scale) with the kernel holder end immersed in the water. The scale was tared and the kernels (one at a time) were placed in the holder using a needle. The weight of the water displaced by the volume of the kernel was recorded immediately after to avoid too much water uptake by the kernel. After the analysis, the kernels were dried for 16 hr at 30°C, and checked to have returned to the same weight as before the volume measurements.

Having determined the kernel volume from the weight of the displaced water, the single kernel density (g/cm^3) is subsequently calculated by dividing the kernel weight (g) by the volume (cm^3). Before the single seed analyses, the volume method was tested on 10 glass beads differing slightly in volume. The average deviation between the real volume and the volume determined using the method shown here was 0.0004 cm^3 for an average of 0.0142 cm^3 , (error of 2.8%).

SKCS Analysis

The kernels were subsequently analyzed using a single kernel characterization system (SKCS) (4100 Perten). The SKCS measures a single kernel hardness index (HI), moisture content (%), diameter

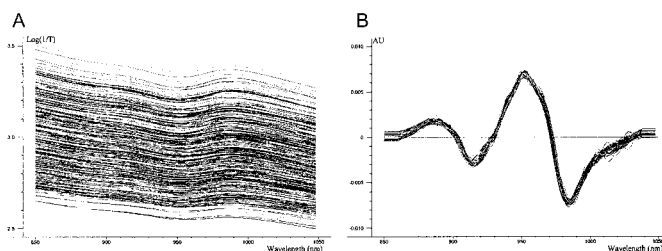


Fig. 2. Single seed near-infrared transmittance (NIT) spectra of 523 wheat kernels shown as raw spectra (A) and corrected spectra using second derivative followed by multiplicative scatter correction (MSC) (B).

TABLE I
Mean and Range of Recorded Single Kernel Characteristics

Method	Parameter	Calibration set (n=415)			Test set (n=108)			Total (n=523)		
		Mean	Min	Max	Mean	Min	Max	Mean	Min	Max
GrainCheck	Width (mm)	3,7	2,3	5,0	3,8	2,5	4,7	3,7	2,3	5,0
	Length (mm)	6,2	5,0	7,5	6,1	5,0	7,2	6,2	5,0	7,5
	Roundness (AU) ^a	0,50	0,25	0,83	0,54	0,31	0,82	0,51	0,25	0,83
	Area (mm ²)	16,9	9,4	25,8	17,1	10,2	24,8	17,0	9,4	25,8
	Volumen (mm ³)	40,8	14,6	82,8	42,6	16,6	76,0	41,2	14,6	82,8
	Red	46,4	31,9	62,4	44,0	25,9	60,5	45,9	25,9	62,4
	Green	33,6	22,7	46,5	31,8	17,6	43,9	33,3	17,6	46,5
	Blue	24,4	17,4	34,0	23,2	14,8	31,1	24,1	14,8	34,0
	Intensity	34,8	24,2	47,4	33,0	19,5	44,9	34,4	19,5	47,4
SKCS	Weight (mg)	45,1	24,5	68,0	45,1	24,1	69,3	45,1	24,1	69,3
	Diameter (mm)	2,9	1,7	4,6	3,0	1,7	4,3	2,9	1,7	4,6
	Moisture (%)	11,9	10,4	13,3	11,0	10,0	11,6	11,7	10,0	13,3
	Hardness (HI)	44,0	-21,4	101,5	32,3	-28,8	82,2	41,6	-28,8	101,5
Reference	Protein (% DM)	10,0	6,8	15,2	9,8	7,0	17,0	10,0	6,8	17,0
	Density (g/cm ³)				1,16	0,99	1,25			

^a Values in the range of 0–1. A perfect circle has roundness = 1, while a very narrow elongated object has roundness close to 0.

(mm), and weight (mg). Normally, SKCS analysis is conducted on a small bulk sample (300 kernels), but in this experiment, the single kernels were fed one by one into the vacuum wheel to retain their identity. The normal container for collecting the crushed kernels was removed, and the single kernel grist from the individual kernels was collected in a small container and used without further grinding for determination of single seed protein according to Kjeldahl.

Protein Determination on Single Kernels

Single kernel nitrogen content was finally determined directly by a modified Kjeldahl (1883) method according to Approved Method 46-12 (AACC). The protein content is reported as percent in dry matter calculated using the moisture content measured by the SKCS instrument.

Before the single kernel analysis, the method was tested on samples of 30–40 mg of wheat flour. The analytical error in terms of standard deviation of 20 replicates amounted to 0.16% (percent protein content in dry matter).

Digital imaging data, NIT spectra, SKCS data, and protein content were then recorded for each kernel, and single kernel density was determined on each of the kernels in the test set. A disadvantage of destructive single seed analysis is that if a measurement fails, there is no sample left for a second analysis. Here, a few of the SKCS, protein, and volume analyses failed, and the results are therefore based on a slightly reduced number of kernels. The calibration set consists of 415 out of the original 430 kernels, while the test set of 110 kernels gave valid data for 108 kernels, except for the density measurements where valid results were obtained for only 99 kernels. The mean and range of all the 14 nonspectral single kernel characteristics for the calibration set kernels and the test set kernels are given in Table I.

Data Analysis

Partial least squares regressions (PLSR) (Martens and Næs 1993) were performed using Unscrambler (v. 7.6, CAMO A/S, Norway) to predict a given quality parameter (y) from fast acquirable X data. The multivariate prediction results are presented and discussed as correlation coefficients (r) between predicted and measured values, and prediction error in terms of root mean square error of pre-

diction (RMSEP) for true test set predictions, and root mean square error of cross validation (RMSECV) for cross-validated results. Relative predictions errors (RE %) are calculated by dividing the prediction errors (RMSECV or RMSEP) by the range (max – min value) of a given parameter.

RESULTS AND DISCUSSION

Single Kernel Protein

The statistics of the Kjeldahl protein determination are listed in Table I. The single seed protein content was 6.8–17.0% for all the analyzed kernels and thus, in principle, covers the whole range of end-use requirements from low-protein wheat for crackers to high-protein wheat for breadmaking. To evaluate and utilize this single seed protein variation, a spectroscopic method would be useful. For this purpose, we use single seed NIT spectra recorded on each of the 523 wheat kernels in the spectral region 850–1,050 nm. The NIT spectra of the 523 single wheat kernels are shown in Fig. 2A and B as both raw spectra and scatter-corrected spectra, respectively, applying a combination of second derivative followed by multiplicative scatter correction (MSC) (Geladi et al 1985). This combined scatter correction has been discussed by de Noord (1994) and applied to single seed NIT spectra by Delwiche (1995). The raw spectra show large intensity offsets, as well as less clear multiplicative effects. These scatter effects are probably due to differences in kernel size and texture, together with kernel orientation in the single seed cassette. With respect to the scatter-corrected spectra (Fig. 2B), it is evident that the spectral scatter has been corrected for, and thereby more spectral emphasis could be focused to represent chemical composition (the level of water, starch and protein content in the kernels). Delwiche (1995) showed that the combination of second derivative of the single seed NIT spectra followed by MSC gave the best predictions. Our results are in agreement with this finding because raw spectra, first-derivative spectra, second-derivative, MSC or MSC followed by second-derivative corrected spectra (data not shown) were less efficient in a prediction model. The issue of scatter in single seed NIT spectra, including suggestions for more general and powerful pretransformations is further investigated by Pedersen et al (2002).

TABLE II
Correlations Coefficients (r) Between Protein Content, Density, GrainCheck Vitreousness and SKCS Hardness

	Correlation Coefficient (r)
Protein content vs. GrainCheck vitreousness ^a	(-) 0.63
Protein content vs. density ^b	0.65
Protein content vs. SKCS hardness ^a	0.38
SKCS hardness vs. GrainCheck vitreousness ^a	(-) 0.55
SKCS hardness vs. density ^b	0.34
Vitreousness vs. density ^b	(-) 0.53

^a 523 kernels.

^b 99 kernels.

TABLE III
Correlation Coefficients (r) and Prediction Errors (RMSECV) of Replicate Simulation by Averaging Kernels for NIT Prediction Models for Kjeldahl Protein Content and SKCS Hardness

No. of Kernels	Protein Model		SKCS HI Model	
	r	RMSECV	r	RMSECV
1	0.93	0.58	0.70	18.6
2	0.96	0.44	0.80	15.8
4	0.98	0.32	0.85	13.7
8	0.98	0.31	0.90	11.9
16	0.98	0.31	0.91	11.0
32	0.99	0.31	0.93	10.4

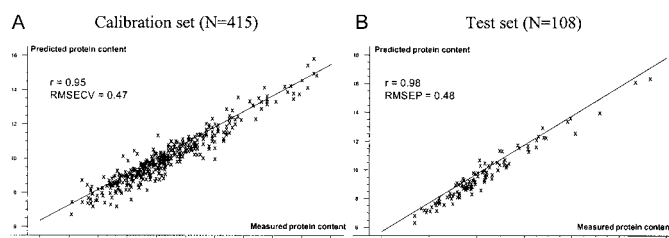


Fig. 3. Predicted vs. measured plot of a partial least squares regression (PLSR) five component regression model for single seed protein using scatter-corrected near-infrared transmittance (NIT) spectra for the calibration set (A) and the subsequent prediction of the test set kernels (B).

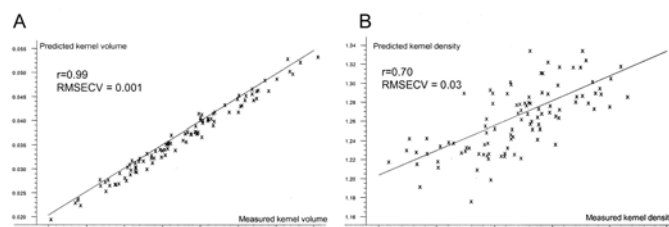


Fig. 4. A, Predicted vs. measured plot of a partial least squares regression (PLSR) model for kernel volume using the nine GrainCheck variables plus single kernel weight. B, Predicted vs. measured plot of a PLSR model for kernel density using the nine GrainCheck variables plus single kernel weight.

A prediction model for protein content was developed based on single seed NIT spectra corrected by the second derivative followed by MSC. The cross-validated calibration model using 5 PLSR components including 415 single kernel spectra is shown in Fig. 3A. This calibration model is used for independent prediction of the 108 test set kernels (Fig. 3B). The relatively low number of PLSR components (5) as compared with other PLSR models in the near infrared range implies a simple and thus robust model. The prediction error (RMSEP) of 0.48% protein when tested independently on 108 new kernels also indicates a good and robust calibration model. Our results for single seed protein determination are comparable to results reported earlier using near infrared transmittance (850–1,050 nm) (Delwiche 1995) and near infrared reflectance (1,100–2,498 nm) (Delwiche 1998).

Single Kernel Vitreousness

Kernel vitreousness is normally determined by visual inspection, where vitreous kernels appear glassy and translucent whereas nonvitreous kernels appear starchy and opaque. Vitreousness is mainly controlled by nitrogen availability in the field as well as temperature during grain filling (Pomeranz and Williams 1990). Vitreous kernels are often harder and have higher protein content. In this investigation, we apply RGB image analysis by the GrainCheck instrument to provide a fast and objective analysis of vitreousness. As a pretest to the current investigation, we analyzed vitreous and nonvitreous kernels (selected by visual inspection) on the image analyzer (GrainCheck). Among the registered color data it was found that especially the red color reflectance differentiated well between vitreous and nonvitreous kernels. The red reflectance from GrainCheck was therefore selected as a quantitative measurement of vitreousness and denoted “GrainCheck vitreousness”. The more vitreous the kernel, the lower the red reflectance and vice versa, i.e., the higher the number, the more nonvitreous the kernel appears. A single seed correlation coefficient of -0.63 (Table II) between protein content and GrainCheck vitreousness shows that the kernels with high protein kernels tend to be more vitreous.

A PLSR model (not shown) was computed using the raw NIT spectra for the prediction of GrainCheck vitreousness to see whether the NIT spectra contained information regarding the GrainCheck vitreousness. The correlation coefficient between measured and predicted GrainCheck vitreousness was 0.76 with a prediction

error of 4.5 AU. A subsequent test of this model on the 108 test kernels confirmed the calibration results ($r = 0.76$, RMSEP = 4.6 AU). Even though the NIT model is based on six PLSR components, most of the spectral NIT information is simply based on the level of absorbance. This can be concluded, because the first score from a principal component analysis (not shown) on the raw NIT spectra (Fig. 2), mainly representing differences in optical densities (offset), correlates well ($r = 0.71$) with GrainCheck vitreousness. The raw NIT spectra thus contain information regarding the GrainCheck vitreousness.

Single Kernel Density

Kernel density is an important parameter in the milling industry, which is normally determined on bulk samples as test weight. The test weight measurement is greatly influenced by kernel packing, kernel size, and kernel density, without differentiation between those factors. Utilizing differences in kernel density by grading for a better and more uniform quality on for example gravity tables, the link between single kernel density and other single kernel quality parameters is essential, to predict whether a given sample is worthwhile sorting for density. For instance, there should be a link between single kernel density and single kernel protein to be able to sort for higher protein content by indirectly sorting for density.

The single kernel density in the test set of 99 kernels ranges from 0.99 g/cm^3 to 1.25 g/cm^3 . This material shows a correlation coefficient of 0.65 (Table II) between protein content and density and a correlation coefficient of -0.53 between GrainCheck vitreousness and density was seen. A single seed correlation coefficient of 0.65 between protein and density would probably be too low to be able to sort for protein by use of density grading on a gravity table.

The Archimedes procedure developed and used for single seed volume analysis in this investigation is rather tedious and it was of interest to investigate whether the much more rapidly acquirable NIT or GrainCheck data could be used for good volume and density determinations. The GrainCheck provides a calculated value of kernel volume based on a 2D-image. Densities derived from these calculated volumes gave, however, a poor correlation ($r = 0.07$) to the real densities based on Archimedes. This low correlation is most likely due to the approximation of a 3D-volume based on a 2D-image, which even if it gives a correlation coefficient of 0.9 to the real volume (Archimedes) is not sufficiently accurate to provide the basis for an accurate measurement of single kernel density.

A second approach, in which the nine GrainCheck variables (Table I) plus the kernel weight were used as **X** in a PLSR model, gave a good prediction of the single kernel volume (Fig. 4A). This combination of image analysis data with kernel weight gives an excellent, rapidly acquirable estimate of the single kernel volume

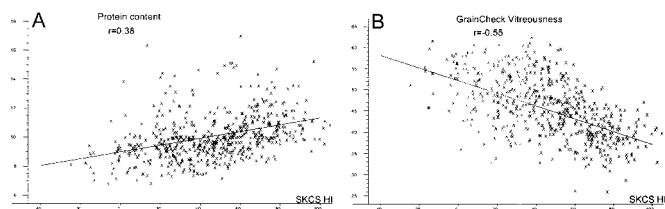


Fig. 5. Scatter plots of single seed single kernel characterization system (SKCS) hardness vs. protein content (A) and single seed SKCS hardness vs. GrainCheck vitreousness (B). $n = 523$.

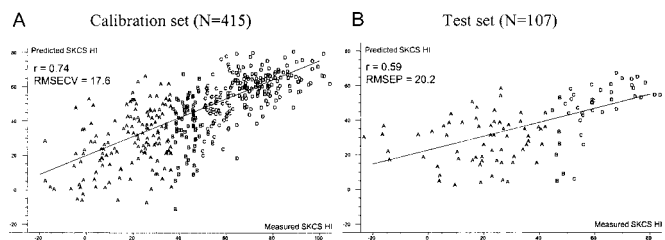


Fig. 6. Predicted vs. measured plot of a partial least squares regression (PLSR) six component regression model for single seed single kernel characterization system (SKCS) hardness using raw near-infrared transmittance (NIT) spectra for the calibration set (A) and the subsequent prediction of the test set kernels (B).

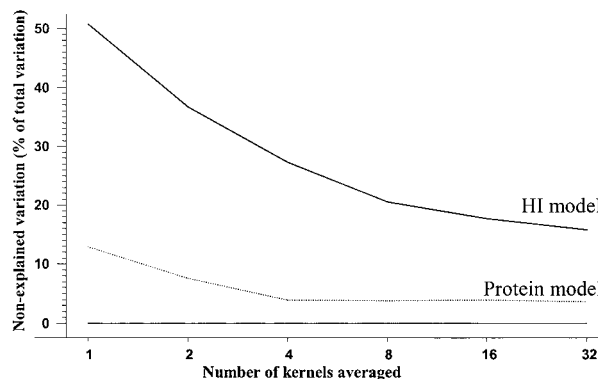


Fig. 7. Plot of nonexplained variation in percent of total variation vs. levels of averaging for the near-infrared transmittance (NIT) prediction model for hardness (solid line) and protein content (dotted line).

($r = 0.99$, RMSECV = 0.001 cm^3) using full cross-validation ($n = 99$). The subsequent calculation of the single kernel density based on this predicted volume provides a considerably better estimate of kernel density, but still only a correlation coefficient of 0.68, as compared to the 0.07 above, with a prediction error of 0.04 g/cm^3 (plot not shown).

Thirdly, by directly using the nine GrainCheck variables plus the kernel weight for PLSR prediction of density, the results can be improved slightly, giving $r = 0.70$ and a lower prediction error (RMSECV = 0.03 g/cm^3) (Fig. 4B).

In a final approach, it was investigated whether the NIT spectra contained information, which could be used for prediction of single kernel density. For a PLSR model using the raw NIT spectra for the prediction of the kernel density, the correlation between measured and predicted density gave 0.63 with a cross-validated prediction error of 0.035 g/cm^3 . An attempt to combine GrainCheck and NIT data for an improved prediction of kernel density was not successful.

Single Kernel Hardness

We have now provided data on the single kernel basis for protein content, kernel density, and apparent vitreousness, the tools normally used by the miller for wheat quality evaluation. Hardness is also used for classification of wheats and its quality in relation to different end uses. It was of interest to investigate to what extent hardness added any further information to the structural characterization of wheat in addition to kernel vitreousness and density. In this investigation, each kernel was fed separately into the SKCS to retain its identity and thereby explore the link between SKCS HI and other single kernel quality parameters. The range in SKCS HI for the analyzed kernels is shown in Table I. Figure 5 shows a scatter plot of single seed SKCS HI versus a) the protein content and b) the GrainCheck vitreousness. A low correlation ($r = 0.38$) between protein content and SKCS HI indicates that the SKCS HI is nearly independent of the kernel protein content in this wheat material. This is surprising, as it is often assumed that high protein wheat kernels tend to be harder. The low correlation between single kernel Kjeldahl protein content and SKCS hardness might be explained by the fact that the kernels originate from a range of genotypes, and that the link between seed protein and seed hardness is seen in some genotypes but not in others. The low number of kernels (10) within each cultivar in this experiment, however, does not allow for investigation of the correlations within each of the cultivars.

A higher, yet still low, correlation ($r = -0.55$) is seen between the GrainCheck vitreousness and the SKCS HI (Fig. 5B). Table II summarizes the correlations between protein content, density, GrainCheck vitreousness, and SKCS HI. Only a small portion of the SKCS HI information seems to be explained in protein content, vitreousness, or density as seen by the relatively low correlations.

In bulk, NIT has been successfully applied for prediction of texture in wheat. Williams (1991) concluded that a bulk NIT measurement was capable of predicting whole-wheat kernel texture with

precision equal to that of the particle size index (PSI) method and slightly better than the NIR method. Delwiche (1993) reported on the use of single kernel NIT measurements for hardness determination. When calibrating single seed NIT spectra against bulk hardness data, Delwiche found that NIT spectra of single seeds had some ability to determine wheat hardness.

Here, we attempt to develop a PLSR model between single seed NIT spectra and true single seed hardness data (SKCS hardness index). In general, we achieve better prediction models for kernel hardness using the raw NIT spectra than with scatter corrected spectra, which agrees with the findings of Delwiche (1993). A prediction model (6 PLSR components) for SKCS HI based on the raw single seed NIT spectra using segmented cross-validation was performed on the calibration kernels. A reasonable calibration is achieved ($r = 0.74$, RMSECV = 17.6 HI) as shown in Fig. 6A. This calibration was subsequently used for HI prediction of 107 of the original 108 test set kernels (Fig. 6B). A low correlation coefficient of 0.59 and a high prediction error of 20.2 HI was achieved. This prediction error corresponds to 20% of the hardness range and thus limits the practical use. In Fig. 6A and B, the samples are labeled according to the hardness groups, where A is soft ($\text{HI} < 33$), B is semi-soft ($33 < \text{HI} < 46$), C is semi-hard ($46 < \text{HI} < 59$), and D is hard ($\text{HI} > 59$). It is apparent that the soft kernels (A) give a more scattered picture in the plots, which means that the hardness index of these kernels are more difficult to predict. However, an exclusion of the soft A kernels did not improve the results.

Various aspects have been considered when interpreting the reason for the relatively poor NIT prediction of SKCS HI of this investigation. First, there might not be a link between single seed NIT spectra and single seed kernel hardness but, as mentioned above, earlier reports have demonstrated the use of NIT spectroscopy on whole-wheat kernels for hardness determination. Second, irrelevant noise in the NIT spectra (X) and the SKCS hardness data (y) might impair the model. Our single seed NIT spectra are averages of three spectra recorded on each kernel. As shown earlier, these spectra correlate very well with kernel protein, so the quality of the NIT spectra seems to be satisfactory. On the other hand, the single seed HI, as determined by the SKCS, might be too inaccurate and thereby problematic as y-values in a NIT prediction model. Because the SKCS HI measurement is destructive, multiple HI readings on the same kernel are not possible and an average of replicate readings is thereby impossible to obtain. This essential condition also makes it difficult to quantify the uncertainty of the instrument measurement. If it was possible to prepare a uniform set of kernel shaped particles from a polymer material, it could be an opportunity to estimate the single kernel uncertainty of the SKCS HI measurements. In the current investigation, a possible way to investigate this problem of uncertainty is to mathematically simulate replicate measurements by averaging across single kernels that are nearly identical. First, we have applied such an averaging approach for the NIT model to protein content where we are certain of both the NIT spectra and the Kjeldahl

TABLE IV
Summary of Nondestructive Screening Methods on Single Kernels

Data (X)	Parameter (y)	r^a	RMSEP ^b	RE ^c
NIT 850–1,050 nm (scatter corrected)	Protein	0.98	0.48	4.7%
NIT 850–1,050 nm (raw)	Vitreousness ^d	0.76	4.6	12.6%
NIT 850–1–050 nm (raw)	Density	0.63	0.035 ^e	13.4%
GrainCheck data plus kernel weight	Volume	0.99	0.001 ^e	2.9%
GrainCheck data plus kernel weight	Density	0.70	0.030 ^e	11.5%
NIT 850–1,050 nm (raw)	Hardness	0.59	20.2	15.5%

^a Correlation coefficient (r) between measured and predicted.

^b Average prediction error.

^c Relative error (RE); RMSECV or RMSECV divided by the range (max-min values, %).

^d Determined using GrainCheck.

^e Models validated using cross-validation and RMSEP should be RMSECV.

protein content determinations. Because this method requires a great number of samples, we use all the 523 analyzed kernels. The NIT spectra and corresponding protein content values are sorted according to protein content. As a start, a PLSR model is developed on the basis of all the 523 calibration kernels. Then the sorted data are averaged across two kernels. Because the kernel data are sorted according to protein content, the two-kernel average is an average, which might be taken as an average of two duplicated analyses on one kernel. A subsequent PLSR model is then developed for the 262 averaged data objects (averaged kernels). This procedure is repeated another four times in which PLSR models are developed averaging across 1 ($n = 523$), 2 ($n = 262$), 4 ($n = 131$), 8 ($n = 66$), 16 ($n = 33$), and 32 ($n = 17$) kernels, respectively. The percent of nonexplained y variation of the total variation is calculated for each model. The trend of nonexplained variation of the protein data for the different PLSR models can then be evaluated (Fig. 7, dotted line). In an ideal situation (i.e., with no noise in the NIT spectra and with determinations of Kjeldahl protein content without any errors) together with a perfect description of the protein content by the NIT spectra, a horizontal line at an ordinate value of 0 would have appeared. In a situation in which we only have model error (i.e., not perfect description of the protein content by the NIT spectra but still with no noise in the NIT spectra and Kjeldahl protein content measurements) we would expect a horizontal line at a certain level above an ordinate value of 0. The decrease in nonexplained variation when averaging (moving from left to right in the plot) represents the noise and errors in the NIT spectra and in the Kjeldahl protein content determinations, reaching a horizontal level representing only model error as mentioned above. As seen from the dotted line, $\approx 13\%$ of the protein data variation is not explained by the NIT PLSR model using all kernel data (original single kernel data), but already after averaging over four kernels (2^2), a nearly horizontal line is appearing at $\approx 4\%$ nonexplained variation. This means that after four simulated replicates, nearly all data noise and errors have been eliminated.

The exact same strategy was applied to the NIT model for SKCS HI (only 522 out of the original 523 kernels had valid data and were used). The results are shown in Fig. 7 (solid line). It is evident that the nonexplained variation in the HI model is considerably higher than for the protein model. As much as 50% of the HI data variation is not explained by the NIT PLSR model using all kernel data, and even after averaging 32 kernels (2^5), the curve is still declining slightly, reaching a level $\approx 15\%$ nonexplained variation. When comparing the two models which are based on the exact same NIT spectra, it is apparent that the decrease in nonexplained variation when averaging is much smaller for the protein model compared to the HI model, thus indicating considerably higher measurement errors in the HI measurement. Table III summarizes the averaging approach in terms of correlation coefficients (r) and RMSECV for the protein content and SKCS HI prediction models. By averaging 32 times, a good prediction model for HI is developed reaching a correlation coefficient of 0.93 and a prediction error of 10.4 HI, which corresponds to 10% of the range. This good model suggests that the raw NIT spectra can be used for single seed prediction of SKCS HI. However, the results also show that the single SKCS HI values are not sufficiently accurate to be used as reference values in a NIT-based prediction model.

CONCLUSIONS

By applying a single kernel procedure in which the non-destructive analyses are conducted before the destructive analyses, several single kernel characteristics can be linked directly to the same functional unit, the single seed, to be used in cereal processing and breeding. In this investigation, the development of nondestructive screening methods for single seed protein content,

vitreousness, density, and SKCS hardness index for the same set of kernels has been studied by applying this type of procedure.

The results of the nondestructive prediction models for single kernel protein, vitreousness, hardness, volume, and density are summarized in Table IV. NIT spectroscopy, in combination with multivariate analysis, shows excellent ability to determine protein content, and only shows some ability for determination of single kernel vitreousness. It is concluded that the nondestructive determination of kernel density, on the other hand, either based on NIT spectroscopy or a combination of kernel weight and image analysis, needs further improvement for practical use.

The use of a true single seed hardness determination, in terms of SKCS HI, as reference values in a NIT prediction model resulted in poor predictability. However, the results shown in Fig. 7 and Table III suggest that raw NIT spectra actually contain more information about kernel texture than the poor prediction model in Fig. 6 suggests. It seems that a single seed reference method for hardness determination with greater accuracy is needed to achieve a good and useful NIT prediction model. If this is possible, there seems to be a potential for the development of a model, which would allow the use of raw NIT spectra for a nondestructive single seed hardness analysis.

For practical use of single seed near infrared spectroscopy as an homogeneity tool, it is important that the measurements are automated, as in the new combined SKCS-NIR instrument (Dowell et al 1999; Delwiche and Hruschka 2000). The Infratec 1255 single seed measurements provides excellent single seed protein data that are much easier to obtain than the traditional Kjeldahl method, but the single seed handling is still not automated and the measurements are quite time-consuming when analyzing high number of kernels. When applied automatically, near-infrared spectroscopy on single seeds, alone or in combination with other automated nondestructive techniques, has a great potential as routine homogeneity analysis. This might not only be limited to protein and hardness, but also for other quality parameters in cereals, as the method is used today on bulk samples.

ACKNOWLEDGMENTS

We are grateful to The Directorate for Food, Fisheries and Agri Business for financial support through the Danish Cereal Network-B4. Karen Thomsen is thanked for development and testing of the single seed Kjeldahl method, and The Danish Institute of Agricultural Sciences for providing the wheat samples. Bo Löfqvist, BoMill AB, Sweden, is thanked for inspiring discussions regarding the manuscript.

LITERATURE CITED

- Abe, H., Kusama, T., and Kawano, S. 1995. Non-destructive determination of protein content in a single kernel of wheat and soybean by near infrared spectroscopy. Pages 457-461 in: *Near Infrared Spectroscopy: The Future Waves*. A. M. C. Davies and P. Williams, eds. NIR Publications: Chichester, UK.
- Berman, M., Bason, M. L., Ellison, F., Peden, G., and Wrigley, C. W. 1996. Image analysis of whole grains to screen for flour-milling yield in wheat breeding. *Cereal Chem.* 73:323-327.
- Chtioui, Y., Bertrand, D., Dattée, Y., and Devaux, M.-F. 1996. Identification of seeds by colour imaging: Comparison of discriminant analysis and artificial neural network. *J. Sci. Food Agric.* 71:433-441.
- de Noord, O. E. 1994. The influence of data preprocessing on the robustness and parsimony of multivariate calibration models. *Chemom Intell Lab Systems* 23:65-70.
- Delwiche, S. R. 1993. Measurement of single-kernel wheat hardness using near-infrared transmittance. *Trans. ASAE* 36(5):1431-1437.
- Delwiche, S. R. 1995. Single wheat kernel analysis by near-infrared transmittance: protein content. *Cereal Chem.* 72:11-16.
- Delwiche, S. R. 1998. Protein content of single kernels of wheat by near-infrared reflectance spectroscopy. *J. Cereal Sci.* 27:241-254.
- Delwiche, S. R., and Hruschka, W. R. 2000. Protein content of bulk wheat from near-infrared reflectance of individual kernels. *Cereal Chem.* 77:86-89.

- Delwiche, S. R., and Massie, D. R. 1996. Classification of wheat by visible and near-infrared reflectance from single kernels. *Cereal Chem.* 73:399-405.
- Dowell, F. E. 2000. Differentiating vitreous and nonvitreous durum wheat kernels by using near-infrared spectroscopy. *Cereal Chem.* 77:155-158.
- Dowell, F. E., Ram, M. S., and Seitz, L. M. 1999. Predicting of scab, vomitoxin, and ergosterol in single wheat kernels using near-infrared spectroscopy. *Cereal Chem.* 76:573-576.
- Geladi, P., MacDougall, D., and Martens, H. 1985. Linearization and scatter-correction for near-infrared reflectance spectra of meat. *Appl. Spectrosc.* 39:491-500.
- Martens, H., and Næs T. 1993. *Multivariate Calibration*. Wiley: New York.
- Martin, C., Rousser, R., and Brabec, D. 1993. Development of a single-kernel wheat characterization system. *Trans. ASAE* 36:1399-1404.
- Orman, B. A., and Schumann, R. A. 1992. Nondestructive single-kernel oil determination of maize by near-infrared transmission spectroscopy. *J. AOCS* 69:1036-1038.
- Patrick, B. E., and Jolliff, G. D. 1997. Nondestructive single-seed oil determination of meadowfoam by near-infrared transmission spectroscopy. *J. AOCS* 74:273-276.
- Pedersen, D. K., Martens, H., Nielsen, J. P., and Engelsens, S. B. 2002. Near infrared absorption and scattering separated by Extended Inverted Signal Correction (EISC). Analysis of NIT spectra of single wheat seeds. *Appl. Spectrosc.* 50:413-419.
- Pomeranz, Y., and Williams, P. C. 1990. Wheat hardness: Its genetic, structural and biochemical background, measurement and significance. *Adv. Cereal Sci. Technol.* 10:471-544.
- Wang, D., Dowell, F. E., and Chung, D. S. 2001. Assessment of heat-damaged wheat kernels using near-infrared spectroscopy. *Cereal Chem.* 78:625-628.
- Williams, P. C. 1991. Prediction of wheat texture in whole grains by near-infrared transmittance. *Cereal Chem.* 68:112-114.
- Zayas, I., Lai, F. S., and Pomeranz, Y. 1986. Discrimination between wheat classes and varieties by image analysis. *Cereal Chem.* 63:52-56.
- Zayas, I., Martin, C. R., Steele, J. L., and Katsevich, A. 1996. Wheat classification using image analysis and crush-force parameters. *Trans. ASAE* 39:2199-2204.

[Received February 4, 2002. Accepted September 27, 2002.]