JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Article

pubs.acs.org/JAFC

Measurement of Single Soybean Seed Attributes by Near-Infrared Technologies. A Comparative Study

Lidia Esteve Agelet*

Department of Agriculture and Biosystems Engineering, Iowa State University, Ames, Iowa 50011, United States

Paul R. Armstrong

Research Engineer, Engineering and Wind Erosion Research Unit, Center for Grain and Animal Health Research, Agricultural Research Service, U.S. Department of Agriculture, 1515 College Avenue, Manhattan, Kansas 66502, United States

Ignacio Romagosa Clariana

Department of Crop and Forest Science, ETSEA, Universitat de Lleida, 25006 Lleida, Spain

Charles R. Hurburgh

Department of Agriculture and Biosystems Engineering, Iowa State University, Ames, Iowa 50011, United States

ABSTRACT: Four near-infrared spectrophotometers, and their associated spectral collection methods, were tested and compared for measuring three soybean single-seed attributes: weight (g), protein (%), and oil (%). Using partial least-squares (PLS) and four preprocessing methods, the attribute that was significantly most easily predicted was seed weight (RPD > 3 on average) and protein the least. The performance of all instruments differed from each other. Performances for oil and protein predictions were correlated with the instrument sampling system, with the best predictions using spectra taken from more than one seed angle. This was facilitated by the seed spinning or tumbling during spectral collection as opposed to static sampling methods. From the preprocessing methods utilized, no single one gave the best overall performances but weight measurements were often more successful with raw spectra, whereas protein and oil predictions were often enhanced by SNV and SNV + detrending.

KEYWORDS: *single seed, near-infrared spectroscopy, protein, oil, weight, Glycine max*

INTRODUCTION

For plant-breeding facilities, obtaining the right genetic material involves a careful selection of the best individual traits. During the selection process, a large number of seed lines are often produced and, consequently, a large number of samples must be evaluated; if the heritability of a desired trait is low, only a small fraction from the total may be of interest. Nondestructive selection of single seeds according to the specific attribute of interest would make the breeding process much faster and more efficient. Near-infrared spectroscopy (NIRS) technologies are highly suitable for breeders because they offer the advantage of performing multiple-trait nondestructive analysis in a relatively short time. In general, NIRS has experienced rapid growth in the past 20 years for compositional measurement of agricultural products. This growth has occurred due to better and less expensive hardware implementations coupled with the need for improved quality control of agricultural products and investigations into novel uses.

The near-infrared (NIR) electromagnetic region (700−2500 nm) is absorbed by C—H, O—H, N—H, and C═O bonds, which are prevalent in water and organic compounds such as carbohydrates, proteins, oils, or alcohols. The amount of absorbed light is proportional to the compound concentration in accordance with Beer's law in scattering media such as agricultural compounds. There are two basic or traditional measurement modes in NIRS instrumentation for agricultural analysis: transmittance and reflectance. In transmittance mode, radiation impinges on the sample with a fraction of the radiation being absorbed by the organic compounds, whereas the remaining fraction passes through the sample and is measured by a sensor. Reflectance measurements are based on sensing diffusely reflected radiation that penetrates to only a shallow depth of a few millimeters into the sample. Other modes of measurement such as transflection or interactance are very popular in fruit quality and medical sciences. Measurements by transflection are mainly, but not exclusively, done on liquid samples. Samples are placed on sampling cells with a backing reflective plate such that NIR radiation penetrates the sample and is reflected by the plate. Interactance measurements

Received: March 26, 2012
Revised: July 23, 2012
Accepted: July 25, 2012
Published: July 25, 2012

dx.doi.org/10.1021/jf3012807 | J. Agric. Food Chem. 2012, 60, 8314−8322

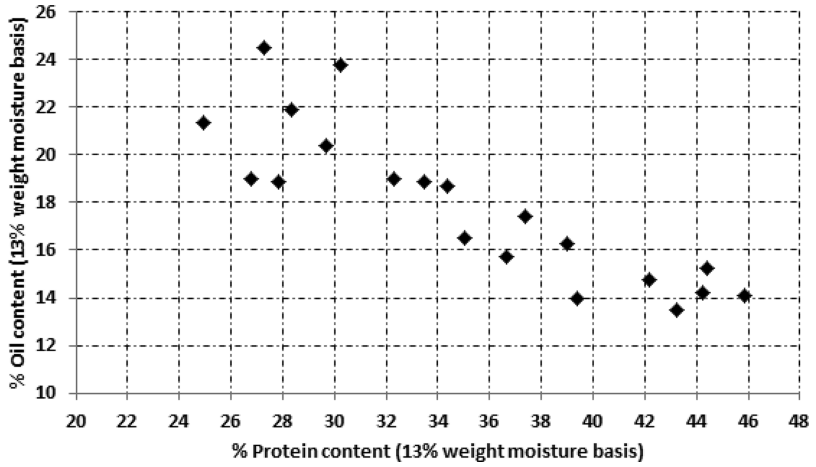

Figure 1. Plot of bulk compositions of the 20 selected samples (percentage protein and oil, 13% weight moisture basis).

are made with the detector parallel to the radiation source, avoiding the collection of specular reflectance. The signal from any of the measurement modes is then mathematically transformed to absorbance spectra to improve its correlation with the concentration of the analyzed compound.

Both reflectance and transmittance modes have been utilized for NIRS single-seed analysis. One of the challenges for these analyses is the small and variable size of the seeds. The shape, size, and inhomogeneity of seeds produce NIR light scattering (i.e., radiation that enters the sample surface and emerges at many angles), which are translated in multiplicative effects on the absorbance spectra.[1] Cogdill et al.[2] found changes in corn kernel size affected the light path length and calibration performance when measuring oil in corn kernels. Reflectance measurements may not be affected by sample path length, but kernel size and the surface exposed to radiation contribute to differences in focal distance, changes in directions of reflected light, and loss of radiation on the background.[3] Furthermore, higher wavelengths sometimes involved in this measurement mode may not produce linear responses between spectra and compound concentrations in single-kernel analysis.[4]

Some research has reported NIRS calibrations for predicting the major attributes of single soybeans. The first NIRS study analyzing single soybeans determined moisture by transmittance (standard error of prediction (SEP) = 0.65−0.69%).[5] Armstrong later obtained better cross-validation results with the USDA light tube reflectance instrument (SECV = 0.32%), which scans the bean while tumbling.[6] Armstrong also developed PLS protein calibrations (SECV = 0.99%), with results on the same order as those of Tajuddin et al.[7] (SEP = 1.32%), who used a transmittance instrument and large-diameter soybeans (>6 mm). However, older research by Abe et al.[8] reported a protein calibration with an even lower SEP (0.67%). Those results were achieved utilizing the average of two transmittance spectra taken at two measurement points. Baianu et al.[9] obtained lower cross-validation errors for protein compared to Armstrong's when using a Fourier transform (FT) reflectance instrument (SECV = 0.77%), but results using a diode array reflectance instrument were of the same order (SECV = 1.10%).[6,7] Delwiche et al.[10] developed protein and inorganic phosphorus calibrations using absolute units (RMSEP = 13.93 g/kg and 568.6 mg/kg, respectively). However, those calibrations had low predictive ability (RPD = 1.20).

Oil content has been also successfully measured in single soybeans. Tajuddin et al.[7] reported SEPs ranging from 1.32 to 1.57%. Baianu et al.[9] again obtained lower cross-validation prediction errors with the FT-NIR instrument than those from the diode array reflectance (SECV = 0.20 and 0.50%, respectively). From their results it was suggested that FT instruments showed significant advantages for predicting single-seed attributes due to the narrow light beam and accuracy in wavelength measurement.

Despite all available literature reporting successful soybean seed calibrations, comparing the performance among instruments and technologies is not possible because of the variety of factors involved. First, samples involved in the calibration and validation sets were not the same. Research was developed on calibration and validation sets with a different concentration range and standard deviation of the compound to be measured, so even if the SEP among studies were the same, the predictive ability of the calibrations would not be the same according to the RPD (standard deviation of the validation set divided by the standard error of prediction).[11] The numbers of samples and seeds, seed size, or crop years are also well-known factors that influence the robustness and calibration performance. Second, statistics utilized to report the validation results are not comparable. Statistics such as the RPD ratio express the calibration predictive ability and account for the concentration range in the validation set, making calibration performances among different instruments and sample sets more comparable. However, the SEP and the coefficient of determination (R^2) were the most common statistics to report calibration precision and explained variability, respectively. Early research erroneously focused on R^2 values, which are dependent on the compound range. On the other hand, the validation of the calibrations was often not performed utilizing a completely independent validation set (seeds from samples not contained in the calibration set and belonging to a different crop year), or cross-validation statistics were reported instead, which may lead to overoptimistic results.

Finally, the moisture basis on which results are expressed must also be taken into account (i.e., dry weight, 13% weight moisture basis, or as-is weight) as well as the different reference methods employed for oil (hexane extraction and nuclear magnetic resonance (NMR)) and protein quantification (combustion, biuret, and NMR). Pulsed low-resolution NMR gives better repeatability than Soxhlet or supercritical fluid

Table 1. Instrument Specifications[a]

instrument	technology	spectral range	working spectral range	sampling interval	resolution	sampling system
A	reflectance Si and InGaAs diode array	350−2500 nm	900−2250 nm (1) 900−1650 nm (2)	1 nm	3 nm, 10 nm	static
B	reflectance monochromator Si and lead sulfide detectors	400−2498 nm	900−2400 nm (1) 900−1650 nm (2)	2 nm	10 nm	rotating cell
C	transmittance Fourier transform	6000−11520 cm^{-1} 868.1−1666.7 nm	6060−10528 cm^{-1} 950.75−1650.17 nm	4 cm^{-1}	8 cm^{-1}[b]	static
CDI NIR256-1.7T1 (light tube)	reflectance InGaAs diode array	904−1686 nm	955−1635 nm	1 nm	3 nm	random tumbling

[a]For instruments covering up to 2500 nm, calibrations with both the entire NIR region (1) and excluding the combination band region (2) were developed. [b]Boxcar apodization.

extraction, as pointed out by Robertson and Windham[12] and Cogdill et al.,[2] who targeted the reference method as one of the sources of error for their corn kernel oil calibration. Those factors add variable sources of error to the published research.

In this study we tested the performance of four NIR instruments with different measurement methods to develop single soybean seed calibrations for oil, protein, and weight using the same seeds. The validation was carried with an independent sample set. Two of the instruments are well-known for product analysis, one is utilized for pharmaceutical analysis, and the fourth instrument was specifically built for single-seed analysis. We compared the instruments with their most suitable modules for single-seed analysis, in terms of their calibration performances for all of the attributes (oil, protein, and weight), utilizing raw and preprocessed spectra. Four common spectra preprocessing methods and combinations of those were tested. Two of the instruments, which extended up to the combination band spectral region (2500 nm), were considered as two independent instruments after the development of calibrations utilizing the entire available NIR spectral range (900−2500 nm) and up to the first overtone region (900−1650 nm), to test for significant changes of performances between these ranges. Baianu et al.[9] obtained the best results with FT-NIR instruments covering up to the combination band region (2400 nm), but Delwiche[4] suggested that higher wavelengths show possible nonlinearities when measuring single seeds, so we wanted to test if there was a significant improvement when omitted.

MATERIALS AND METHODS

Sample Sets. Twenty soybean samples from 1993, 1994, 2008, and 2009 crop years were selected from the Grain Quality Laboratory (Iowa State University, Ames, IA, USA). The bulk sample composition of oil and protein covered a wide compositional range (Figure 1). Twenty-four seeds were randomly picked from each sample and scanned by the instruments (480 seeds total). The seeds were then placed into 10 48-well plates and sent for oil analysis at the USDA NCAUR in Peoria, IL, USA. Total oil and moisture, expressed as mgs mass, was determined by NMR (AOCS method Ak 4-95 (09)) using an mq20 minpec (Bruker Optik GmbH, Ettlingen, Germany). Individual seed mass and the corresponding oil and moisture masses were used to convert oil values to a percentage dry matter basis. Seeds were then measured for protein content by the Soil Testing Laboratory at Iowa State University with a Truspec CN analyzer (Leco Corp., St. Joseph, MI, USA) using the combustion method (AOCS method Ba 4e-93 (09)). Twenty-two of the initial 24 seeds per sample were analyzed (440 seeds total); the remaining 2 were retained for future work.

Instruments. Three commercial NIR spectrometers (referred to as A, B, and C) were used in this study to compare the prediction ability of single-seed soybean content. The manufacturer of each instrument is omitted as the study does not intend to compare instrument models, but different instrument conformations and sampling methods in single-seed analyses. The fourth instrument included in the study, the light tube CDI NIR256, was a noncommercial instrument built by the U.S. Department of Agriculture, Center for Grain and Animal Health Research (USDA, CGHAR) Manhattan, KS, USA. Table 1 shows the specifications of the four spectrometers. The spectral working range was obtained after removal of visibly noisy wavelength extremes and adjusting to a standard range (900−1650 nm) when appropriate. Because two of the instruments, A and B, cover the vis range up to the NIR combination band region (400−2500 nm), calibrations were developed for both first- and second-overtone range (up to 1650 nm) and the whole NIR range (up to 2400 nm) to test for performance differences from each spectral range. The two instruments use different detector types to cover the different NIR regions; therefore, there is an impact on the spectra at the wavelengths of detector change. Hence, the two data points on each side of the wavelength crossover (1000 and 1974 nm for instrument A and 1100 for instrument B) were removed.

The specific mechanisms used to collect spectra for each spectrometer are described as follows. For instrument A, a high-intensity light accessory with an independent halogen light was utilized and connected to the instrument through an optic fiber collecting probe provided by the company. The probe collected the diffuse reflectance signal at an angle designed to avoid specular reflection. Individual seeds were always placed on the same spot of the sampling glass surface, defined by a sample tray (maximum diameter of 12 mm). Spectralon was utilized as reference material every 20 min. Three consecutive spectra were taken and averaged for each seed without the seed's being moved from its initial position.

For instrument B, a small ring rotatory cup module was utilized to analyze each seed. Seeds were individually placed on the center of the cup, positioned in a small indenture on a blank circular supporting material. The material, provided by the manufacturer, served to hold the seed during cup rotation. Twenty-five spectra were taken for each rotating seed and co-added by the instrument.

For instrument C, designed for pharmaceutical products, a tablet measurement cell accessory with 10 iris aperture wells was used to collect transmission spectra. The hole where the light beam passed through the seed was always smaller than the seed surface, which helps reduce possible scattering effects that would be otherwise produced by seed edges. Sixty-four spectra were collected and averaged by the instrument software for each seed.

The light tube CDI NIR256 was specifically designed to analyze single seeds. Spectra of individual seeds are taken as they fall by gravity through an illuminated borosilicate light tube. The specifics of this

design are explained by Armstrong[6] and Tallada et al.,[13] Three spectra were taken from each seed; that is, the seed was dropped three times through the light tube, and spectra were averaged after mean centering to remove the significant offset differences between spectra.

Spectral Data Preprocessing. Raw and preprocessed spectra with four common pretreatments were utilized for calibration development: raw spectra, standard normal variate (SNV), seven-point window average smoothing (smt), SNV with second-degree polynomial detrending (SNV + detr), and SNV with second-degree polynomial detrending and seven-point average smoothing (SNV + detr + smt).

Calibration and Validation. Spectra obtained from the 1993 crop year samples (3 samples, 72 seeds for both oil and weight, 54 for protein after 11 seeds were discarded during protein combustion measurements) were kept as an independent validation set because these samples had compound ranges equivalent to the calibration set ranges, and thus no prediction extrapolation would occur. The remaining samples (17) were used as the calibration set (408 seeds for oil and weight and 374 for protein). Table 2 shows the descriptive statistics of each set. For protein, the seeds that had protein content >55% were removed as the reference laboratory flagged them as possible outliers as well as outliers in the NIR calibration modeling. Calibration models were developed using partial least squares (PLS-1) regression for each combination of instrument, preprocessing, and compound (for information regarding PLS, refer to Naes et al.[14]). A total of 90 calibrations were developed. Spectral data, raw and preprocessed, were mean-centered prior to PLS software modeling. The modeling software used was The Unscrambler v. 9.8 (Camo Software AS, Oslo, Norway). Outliers were detected by visual inspection of spectral plots, by inspection of the residual versus leverage plots, and by the built-in outlier detection system employed by the software, which is based on a combination of limits for sample leverage, residuals, and explained variance. The optimal number of latent variables for the model was determined by 15 segment cross-validation as suggested by the software. The software selects latent variables according to either the first local minimum from the residual variance curve or last largest significant decrease of variance from the previous model developed with one less latent variable from what was determined by the *F* test.

Table 2. Descriptive Statistics of Calibration and Validation Sets

set	compound	*n*	min	max	SD	mean
calibration	oil (%)	408	12.0	29.7	4.0	19.9
	protein (%)	377	16.0	52.4	7.4	37.7
	weight (g)	408	0.068	0.277	0.034	0.148
validation	oil (%)	72	14.0	22.7	2.2	18.2
	protein (%)	56	33.2	44.2	2.1	39.0
	weight (g)	71	0.070	0.266	0.043	0.138

Experimental Design. To compare the effects of instrument, attribute, and preprocessing methods on calibration performance, a factorial ANOVA least-squares was performed. The dependent variable and a standardized indicator of calibration performance was the RPD ratio.[11] This ratio was also utilized by Kovalenko et al.[15] to compare instrument and calibration methods to predict amino acids in bulk soybeans. Calibration models with RPDs < 1.4 are not usable; RPDs between 1.4 and 1.7 are usable for rough screening; RPDs between 1.7 and 2.42 are usable for screening; and RPDS between 2.42 and 3 are usable for most applications with caution; RPDS between 3 and 5 are usable for most applications, and RPDs > 5 are usable for quantitative purposes.[11] The three main factors were the instrument (I), attribute (A), and preprocessing method (P). All three were considered to be fixed effects, and second-degree interactions were initially included in the model. The third-degree interaction (I × A × P) was pulled out as the error term (E) because the model did not have repetitions. The least-squares fit had the following form:

$$\mathrm{RPD} = \mathrm{intercept} + \mathrm{I} + \mathrm{A} + \mathrm{P} + (\mathrm{I} \times \mathrm{C}) + (\mathrm{C} \times \mathrm{P}) + (\mathrm{I} \times \mathrm{P}) + \mathrm{E} \quad (1)$$

The Sidak method was utilized for pairwise comparison of means. The ANOVA analysis was performed using IBM SPSS Statistics 17.0 software (Armonk, NY, USA).

RESULTS AND DISCUSSION

Calibrations and Overall Observations. Tables 3−8 summarize the best two calibrations per each instrument and compound, based on the best independent validation predictions. Because bias values refer to the systematic error and can be easily removed by subtraction from the predicted values, we did not focus on these statistic values as much as on the SEP (random error) and RPD (calibration predictive ability). For all models and instruments the heaviest seed (0.280 g) and three seeds with protein below 20% were removed because they could not be modeled properly, probably because the relationship was not linear at such a low concentration range.

Instrument A performance statistics were the poorest overall (Tables 3 and 4). For oil, most of the calibrations were suitable for rough screening (RPDs > 1.40), but protein calibrations were not usable (RPD of 1.00, on average), and were similar to what Delwiche et al.[10] obtained when developing protein calibrations with absolute reference units (mg/g). The best calibrations were achieved for weight. There were some problems in the modeling such as the presence of clustering in the latent variable score plots for all attributes. The origin of the clusters could not be determined, but it was probably due to light scattering from slight sample positioning changes and seed shape differences, as seeds from the same sample tend to group in the same cluster. Preprocessing with SNV accentuated this problem, as shown in the oil calibration score plot with SNV preprocessing (Figure 2). SNV is proven to minimize scattering effects, but probably the combination and interaction of several factors producing scattering cannot be successfully addressed by SNV. The characteristics of these results suggest that the use

Table 3. Statistics of the Two Best Calibrations for Instrument A, Entire NIR Range (900−2250 nm)

		calibration					independent validation		
	preprocessing	*n*	LVs	SEC (%)	R^2 (%)	SECV (%)	SEP (%)	RPD	bias (%)
oil (%)	raw	397	14	0.69	96.83	1.38	1.44	1.52	1.03
	SNV + Detr + smt	403	12	1.00	93.21	1.35	1.46	1.50	0.75
protein (%)	raw	365	11	1.68	94.32	2.49	1.86	1.10	−0.40
	SNV + Detr	353	10	1.74	93.93	2.46	1.91	1.07	−0.52
weight (g)	raw	399	9	0.014	83.45	0.015	0.011	4.095	0.005
	smt	399	9	0.014	82.00	0.015	0.011	4.095	−0.005

Table 4. Statistics of the Two Best Calibrations for Instrument A, Short Working Range (900−1650 nm)

		calibration					independent validation		
	preprocessing	*n*	LVs	SEC (%)	R^2 (%)	SECV (%)	SEP (%)	RPD	bias (%)
oil (%)	raw	400	9	1.1	91.88	1.31	1.47	1.49	0.98
	smt	401	9	1.19	90.48	1.32	1.5	1.46	−1.05
protein (%)	smt	365	10	1.73	93.71	1.97	1.56	1.31	−0.06
	SNV	360	9	1.7	94.31	2.04	1.56	1.31	0.31
weight (g)	raw	400	10	0.012	86.16	0.015	0.014	3.07	−0.004
	smt	397	10	0.013	84.25	0.015	0.013	3.31	−0.004

Table 5. Statistics of the Two Best Calibrations for the USDA Light Tube Instrument

		calibration					independent validation		
	preprocessing	*n*	LVs	SEC (%)	R^2 (%)	SECV (%)	SEP (%)	RPD	bias (%)
oil (%)	SNV + detr	345	6	0.54	97.99	0.56	0.56	4.46	0.39
	SNV + detr + smt	357	6	0.55	97.99	0.57	0.55	4.55	0.43
protein (%)	SNV	330	5	1.06	97.97	1.10	0.65	3.23	−0.14
	SNV + detr + smt	317	4	1.12	97.78	1.15	0.64	3.28	0.05
weight (g)	raw	356	9	0.010	91.18	0.01	0.010	4.34	0.002
	smt	365	9	0.010	90.59	0.011	0.010	4.39	0.003

Table 8. Statistics of the Two Best Calibrations for Instrument C

		calibration					independent validation		
	preprocessing	*n*	LVs	SEC (%)	R^2 (%)	SECV (%)	SEP (%)	RPD	bias (%)
oil (%)	raw	411	3	1.10	92.11	1.11	0.78	2.81	0.98
	SNV	400	1	0.99	93.59	0.99	0.77	2.84	0.73
protein (%)	SNV	356	3	1.66	94.4	1.68	1.19	1.72	0.56
	SNV + detr	360	3	1.97	92.09	2.01	1.25	1.64	0.78
weight (g)	raw	414	11	0.016	78.15	0.016	0.015	2.87	0.012
	smt	414	10	0.016	77.48	0.017	0.016	2.69	0.013

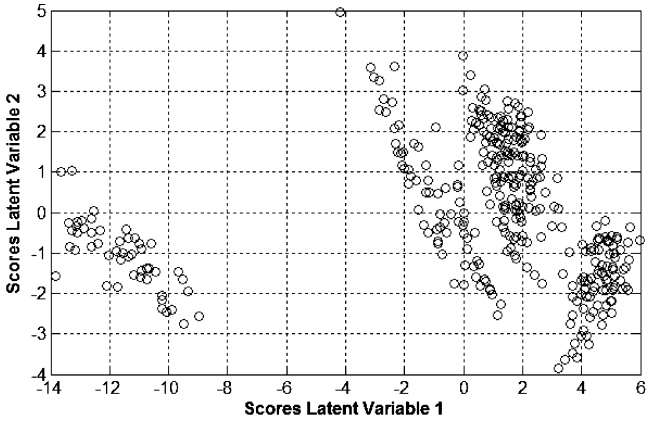

Figure 2. Score plots of LV1 versus LV2 from instrument A whole range PLS oil calibration and SNV preprocessed spectra, showing clustering.

of local models or several PLS submodels based on seed weight would be more suitable than a single PLS model.

Each seed spectrum, which resulted from averaging three spectra, was visually noisy. That was particularly true at higher wavelengths and shown in their respective regression coefficients, indicating that those calibrations may not be robust. Regression coefficients were similar to overfitting a model, and reducing the number of latent variables in one or two factors did not bring any improvement. Reducing the working wavelength range (Table 3) did not improve calibration performance and still led to noisy regression coefficients, but the ability for predicting weight decreased. This could be due to the fact that there is more scattering at lower wavelengths, a phenomenon correlated with sample size. From these results, we can conclude that the use of the accessory for static measurements with the single measuring fiber is not suitable for analyzing spherical seeds such as soybeans: It is extremely sensitive to seed position and size.

For relative comparison, the USDA light tube (Table 5), also a diode array instrument, did not show such noisy regression coefficients, gave better results for weight, and outperformed all instruments in protein and oil predictions with the highest RPD values (above 4 and 3, respectively). The SEP values for protein were close to those obtained by Abe et al.[8] (SEP = 0.67% pts), who measured the seeds from two different angles with a transmittance instrument. None of the preprocessing methods made a large difference for instrument A calibration statistics, whetrsd the USDA light tube had better performances using SNV and SNV + detrending for oil and protein. However, SNV

dx.doi.org/10.1021/jf3012807 | J. Agric. Food Chem. 2012, 60, 8314−8322

Table 6. Statistics of the Two Best Calibrations for Instrument B, Entire NIR Range (900−2400 nm)

		calibration					independent validation		
	preprocessing	*n*	LVs	SEC (%)	R^2 (%)	SECV (%)	SEP (%)	RPD	bias (%)
oil (%)	raw	394	8	0.87	94.99	0.92	0.87	2.52	0.22
	smt	394	8	0.89	94.87	0.93	0.87	2.52	0.22
protein (%)	smt	359	9	1.88	92.80	1.96	1.04	1.97	−0.49
	SNV	325	8	1.75	93.61	1.86	1.01	2.03	0.00
weight (g)	raw	391	7	0.006	96.50	0.006	0.008	5.38	0.001
	smt	392	7	0.006	96.55	0.007	0.008	5.38	0.001

Table 7. Statistics of the Two Best Calibrations for Instrument B, Short NIR Range (900−1650 nm)

		calibration					independent validation		
	preprocessing	*n*	LVs	SEC (%)	R^2 (%)	SECV (%)	SEP (%)	RPD	bias (%)
oil (%)	SNV	348	7	0.69	96.69	0.72	1.00	2.19	0.201
	SNV + detr	342	3	1.00	92.77	1.01	0.98	2.23	−0.09
protein (%)	SNV + detr	302	8	1.39	95.78	1.47	0.88	2.33	−0.20
	SNV + detr + smt	307	8	1.42	95.69	1.50	0.92	2.23	−0.23
weight (g)	raw	349	7	0.007	95.95	0.007	0.006	7.82	0.001
	smt	351	7	0.007	95.93	0.007	0.006	7.82	0.001

greatly diminished weight predictions for both instruments as these preprocessing methods reduce the light-scattering effects in the spectrum, phenomena attributed to seed size/weight. With regard to weight predictions, the statistics are only slightly better than the ones from instrument A with the static measurements (RPD around 4). Light scattering generated from seed size or shape may be diminished when combining the seed tumbling motion and the combination of two measurements taken from two opposite sides of the seed.

Instrument B, using the 900−2400 nm range, gave better oil calibration performance compared to the 900−1650 nm range (RPD of 2.5 vs 2.23), but this was reversed for protein (2.03 vs 2.33) (Tables 6 and 7). Working with the shorter range did make a noticeable difference in predicting weight; SEP improved from 0.008 to 0.006 g (Table 6). When the entire NIR region was studied, seed weight calibrations were not influenced by preprocessing. Although raw and smoothed spectra calibrations are reported in Table 6, the same validation results were achieved with three preprocessing methods left. Regardless of the wavelength range, instrument B was the best for predicting seed weight, with RPDs > 5. This could be due to the fact that the seed was held by the spinning cell blank material provided by the manufacturer, and it was scanned together with the seed. This material, although not interfering in predicting sample attributes, will have its own spectrum. The correlation of seed area with weight (probably stronger at shorter wavelengths) may have helped the sampling method and instrument to outperform others for seed weight prediction.

Instrument C did not show any artifacts in the score plots, but did show possible nonlinearities when working with raw spectra: subtle skewing of the data is noticeable in prediction and reference data (Figure 3). Using more or fewer latent variables did not remove this effect, but spectra preprocessing did improve it. Protein calibrations could be used for rough screening (RPDs > 1.4), whereas oil and weight calibrations were more quantitative (Table 8). Observation of the oil

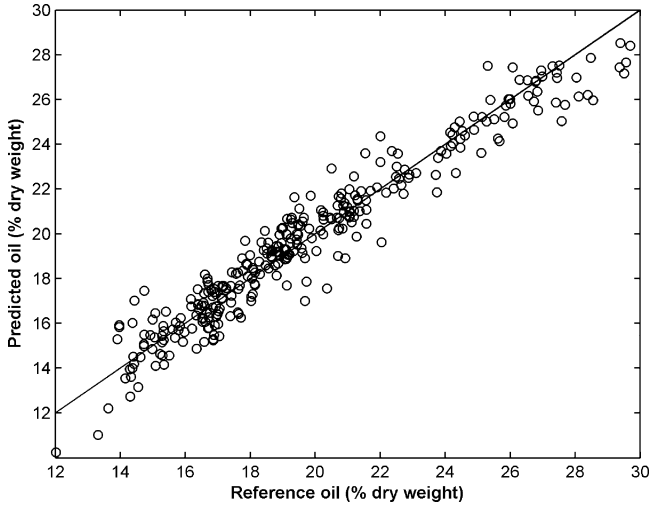

Figure 3. Oil calibration with instrument C raw spectra. Seeds above 24% and below 16% are underpredicted, whereas the range between 17 and 21% is mainly overpredicted.

regression coefficients reveals that higher wavelengths (lower wavenumbers) have more weight in the calibrations (Figure 4). Bulk transmittance instruments for grain analysis often work in the region from 800 to 1100 nm and give very good predictions. As such PLS model calibrations were also developed for instrument C using the transmittance range of 868−1100.35 nm (9088−11520 cm^{-1}) but no improvement was found in predicting oil (SEP = 0.78% and RPD = 2.77, SNV preprocessing), protein (SEP = 1.22% and RPD = 1.68, SNV preprocessing), or weight (SEP = 0.0019 g and RPD = 2.69, raw absorbance data). The number of latent variables needed for oil predictions by instrument C was the lowest of all instruments. For all of the SNV-based preprocessing methods, only one latent variable was required for the calibration (and only three were required with both raw absorbance and

dx.doi.org/10.1021/jf3012807 | J. Agric. Food Chem. 2012, 60, 8314−8322

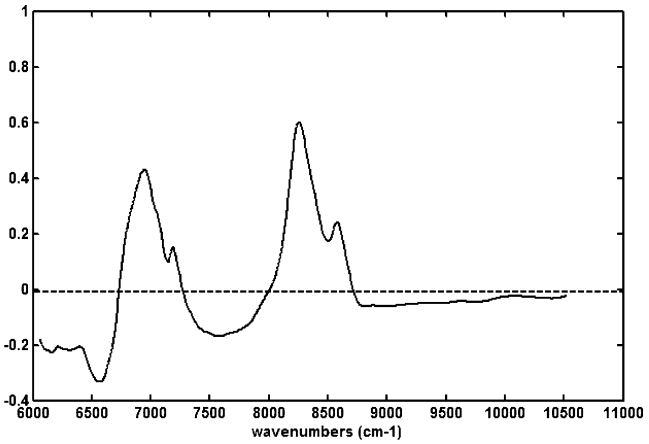

Figure 4. PLS regression coefficients for raw oil calibration, instrument C (FT transmittance).

smoothed data). In practice, this would indicate an inability to model the data. However, both cross-validation and validation statistics showed that the model had predictive ability, and a univariate linear regression between the first latent variable scores and reference values may be enough to obtain a calibration for screening. In our study, the advantage of FT-NIR technology, such as higher signal to noise ratio, did not translate in better performances as Baianu et al.[9] reported, although they had worked with reflectance mode and higher resolutions (4 cm^{-1}). Resolution, however, is not usually a relevant factor when one is working with agricultural samples because their absorption bands are broad. For instance, resolutions of 5 nm or 16 cm^{-1} were found to be suitable for fruit and soybean amino acids, respectively.[9,16] High resolutions may even add spectral artifacts and decrease calibration performances as shown for soybean bulk samples.[17] Further studies could be performed to test the effect of spectral resolution on single-seed analysis.

It is interesting to note for most protein calibrations and all instruments, oil calibrations for instrument C and the USDA light tube, and weight calibrations for instrument A and B, the SECV values were higher than SEP values from the independent sample set. This may occur because protein calibrations were developed from a set with a wide range and standard deviation of the compound of interest, with a distribution not entirely uniform but having fewer samples with high and low concentrations. *N*-Folded cross-validation may split data into random groups with both wide and short reference ranges. This leads to a high likelihood of extrapolation in the cross-validation submodels, cumulating high prediction errors. In that case, SECV was pessimistic. Data artifacts may have also influenced the higher SECV values of the instrument B calibrations.

Comparison of Performance. Table 9 shows the ANOVA results. The RPD, as a dependent variable, was transformed to its inverse (1/RPD) to address a suspicious trend in residuals. With that transformation, a possible violation of the homoscedasticity assumption of ANOVA regression was eliminated. All of the analyses were carried with that transformation, although Figures 5 and 6 use the original RPD values for clearer interpretation of results.

The differences in performances (1/RPD) between the principal factor attributes and instruments were significant at the 0.05 level, so at least one attribute and one instrument

Table 9. Results of Three-Factor ANOVA with Double Interactions[a]

source	degrees of freedom	type III sum of squares	mean squares	*F* value	significance
total	89	4.25			
attributes	2	1.77	0.89	449.60	0.000
preprocessing	4	0.00	0.00	0.51	0.728
instrument	5	1.43	0.29	145.31	0.000
preprocessing × instrument	20	0.02	0.00	0.58	0.907
attributes × preprocessing	8	0.05	0.01	3.22	0.006
attributes × instrument	10	0.88	0.09	44.74	0.000
error	40	0.08	0.00		

[a]The dependent variable was the inverse of the RPD.

differed in achieved performances from the rest. The interactions attribute × preprocessing and attribute × instrument were found to be significant at the 0.05 level. From the Sidak pairwise comparison test, the performance differences between working with the entire NIR range and shorter wavelength range were not significant for either instrument A or instrument B. However, all instruments were different from each other in terms of performances or average RPDs.

From the interaction plot with the original RPD values (Figure 5), we can see that although both instruments A and B did not have significant differences in overall calibration performances for all attributes and preprocessing methods comparing the entire wavelength range with the reduced wavelength range, there were appreciable differences when measuring seed weight (significant instrument × attribute interaction). The reduced NIR range improved the predictive ability of instrument A for measuring weight, whereas the opposite was true for instrument B. This could be related to the sampling method: for instrument A the measurements were statically taken without a background; measurements on instrument B were taken on a spinning seed against a background material that could have contributed to the improvement in weight prediction using lower wavelengths.

Instrument A performed the worst overall. Instrument C, although having the second-best performance in oil prediction, was next after instrument A in worst overall performance. Instrument C was the worst in predicting seed weight, probably because spectra were taken from a small portion of the seed. The best instrument, with larger average RPD across attributes and preprocessing methods, was the USDA light tube, followed by instrument B. According to these results and in agreement with previous publications, sample positioning and sampling method are of major relevance to develop single-soybean calibrations. Although measurements in transmittance mode may be comparable to or better than the reflectance mode for static sampling, to achieve significantly better predictive models the seed must be scanned from as many angles as possible. Factors such as measurement mode, sampling increment, or other instrument characteristics are not as relevant as the sampling method. The results confirm findings by Delwiche[18] on wheat and by Janni et al. on corn kernels;[19] they suggested averaging multiple measurements at different points and homogeneous illumination of the kernel to improve single-seed calibrations. The results indicate that a number of measurements should be taken from different angles and, in

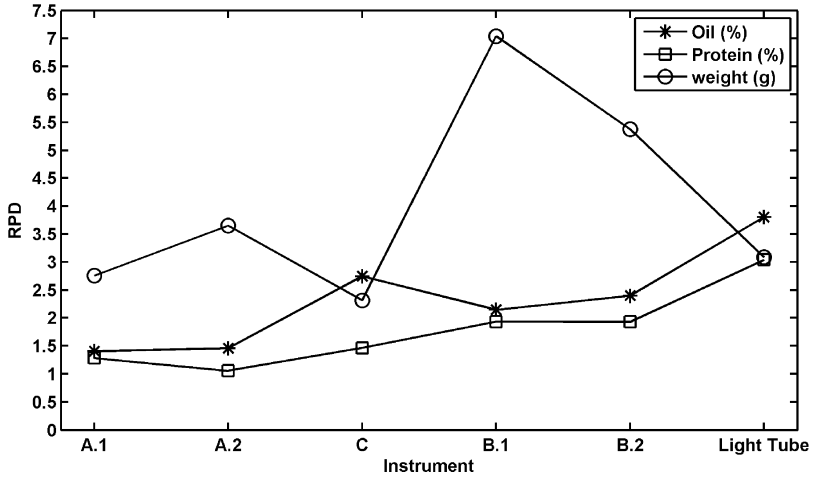

Figure 5. Plot of attribute × instrument interaction with the ANOVA estimated RPD means (A.1 and B.1 are based on calibrations with wavelengths up to 2400 nm; A.2 and B.2 are the same instruments working using the shorter wavelength range).

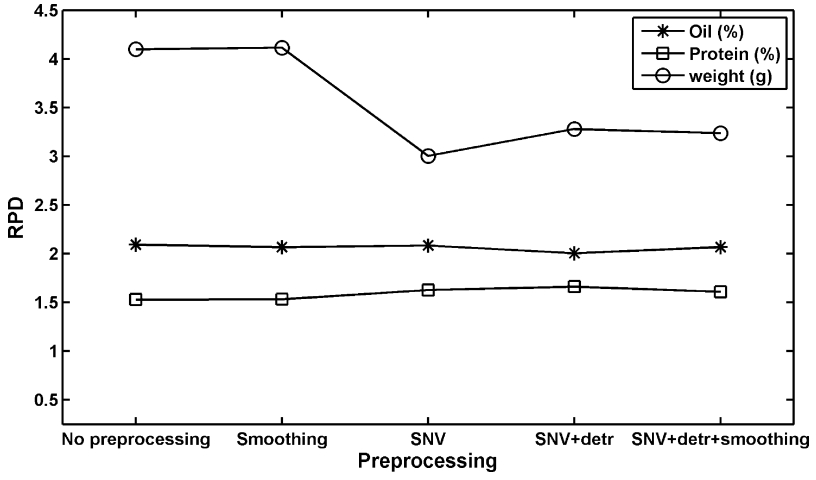

Figure 6. Interaction plot of preprocessing × attributes with the ANOVA estimated RPD means.

the case of working with an instrument with a static sampling system, averaging several spectra from the seed repositioned several times on the sampling surface. The USDA light tube achieves this effect thanks to the complete illumination of the tumbling seed. Instrument B takes the spectra from only one side of the seed, but the spinning cup allows collection of spectral information from many angles compared to static sampling instruments (A and C). However, it is unknown if performing multiple static measurements and co-adding spectra could lead to the same performance as taking the measurements while the seed is in motion.

The overall differences of performance among attributes were significant between each other according to the Sidak pairwise comparison of means. Seed weight predictions were the most accurate overall, followed by oil and protein, respectively. As can be seen from the interaction plot (Figure 6), SNV accompanied with other combinations of preprocessing decrease the predictive ability of weight calibrations. This preprocessing method also affects calibrations developed with absolute units,[20] which suggests that SNV attenuates the effects of seed size and thus negatively affects the prediction of seed weight. Conversely, SNV tends to improve oil and protein prediction for composition expressed in percent thanks to the removal of unnecessary spectral information induced by seed size.

In summary, weight is the best measurable attribute overall for all instruments except C. Oil predictions were generally better than protein, and the best calibrations in terms of predictive ability (RPD) and standard error of prediction (SEP) were obtained from the USDA light tube. There seemed to be a correlation between calibration performance for both oil and protein and for sampling method, that is, static or dynamic (spectra taken from different seed sides or angles). This factor has been shown to be far more relevant than other instrument characteristics such as technology (monochromator, diode array, FT), measurement mode (transmittance, reflectance), sampling interval, or wavelength range. RPD results show the best performanced for protein and oil came from the USDA light tube and the worst, from instrument A. For seed weight, instrument B with the spinning cup outperformed the other instruments.

There was no best preprocessing method or combination overall. SNV and SNV + detrending, however, gave the best results for the light tube and most of the other instruments for protein and oil predictions. In the prediction of weight, raw spectra or smoothing was usually the best option as SNV removes information related to seed size and weight.

AUTHOR INFORMATION

Corresponding Author

*Postal address: 1545 Food Science Building, Iowa State University, Ames, IA 50014. E-mail: esteve.lidia@gmail.com. Phone: (515) 294-8629.

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Kortum, G. *Reflectance Spectroscopy: Principles, Methods, Applications* (translated from German by Lohr, J. E.); Springer: New York, 1969.

(2) Cogdill, R. P.; Hurburgh, C. R., Jr.; Rippke, G. R. Single-kernel maize analysis by near-infrared hyperspectral imaging. *Trans. ASABE* **2004**, *47* (1), 311−320.

(3) Wang, D.; Dowell, F. E.; Lacey, R. E. Single wheat kernel size effects on near infrared reflectance spectra and color classification. *Cereal Chem.* **1999**, *76*, 34−37.

(4) Delwiche, S. R. Protein content of single kernels of wheat by near infrared reflectance spectroscopy. *J. Cereal Sci.* **1998**, *72*, 241−254.

(5) Lamb, D. T.; Hurburgh, C. R., Jr. Moisture determination in single soybean seeds by near-infrared transmittance. *Trans. ASABE* **1991**, *34* (5), 2123−2129.

(6) Armstrong, P. R. Rapid single-kernel nir measurement of grain and oil-seed attributes. *Appl. Eng. Agric* **2006**, *22* (5), 767−772.

(7) Tajuddin, T.; Watanabe, S.; Masuda, R.; Harada, K.; Kawano, S. Application of near infrared transmittance spectroscopy to the estimation of protein and lipid contents in singke seeds of soybean recombinant inbred lines for quantitative trait loci analysis. *J. Near Infrared Spectrosc.* **2002**, *10* (4), 315−325.

(8) Abe, H.; Kusama, T.; Kawano, S.; Iwamoto, M. Non-destructive determination of protein content in a single kernel of wheat and soybean by near-infrared spectroscopy. In *The Future Waves*; Davies, A., Williams, P., Eds.; NIR Publications: Chichester, U.K., 1996.

(9) Baianu, I. C.; You, T.; Costescu, D. M.; Lozano, P. R.; Prisecaru, V.; Nelson, R. L. High-resolution nuclear magnetic resonance and near-infrared determination of soybean oil, protein, and amino acid residues in soybean seeds. In *Oil Extraction and Analysis, Critical Issues and Comparative Studies*; Luthria, D. L., Ed.; AOCS Press: Champaign, IL, 2004.

(10) Delwiche, S. R.; Pordesimo, L. O.; Scaboo, A. M.; Pantalone, V. R. Measurement of inorganic phosphorous in soybeans by near-infrared spectroscopy. *J. Agric. Food Chem.* **2006**, *54*, 6951−6956.

(11) *Near-Infrared Technology in the Agricultural and Food Industries*, 2nd ed.; Williams, P., Norris, K., Eds.; AACC: St. Paul, MN, 2001.

(12) Robertson, J. A.; Windham, W. R. Comparative study of methods of determining oil content of sunflower seed. *J. Am. Oil Chem. Soc.* **1981**, *58* (11), 993−996.

(13) Tallada, J. G.; Palacios-Rojas, N.; Armstrong, P. R. Prediction of maize seed attributes using a rapid single kernel near infrared instrument. *J. Cereal Sci.* **2009**, *50*, 381−387.

(14) Næs, T.; Isaksson, T.; Fearn, T.; Davies, T. *A User-Friendly Guide to MultiVariate Calibration and Classification*; NIR Publications: Chichester, U.K., 2002.

(15) Kovalenko, I. V.; Rippke, G. R.; Hurburgh, C. R., Jr. Determination of amino acid composition of soybeans (*Glycine max*) by near-infrared spectroscopy. *J. Agric. Food Chem.* **2006**, *54*, 3485−3491.

(16) Nicola, B. M.; Beullens, K.; Bobelyn, E.; Peirs, A.; Saeys, W.; Theron, K. I.; Lammertyn, J. Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: a review. *Postharvest Biol. Technol.* **2007**, *46*, 99−118.

(17) Esteve Agelet, L.; Rippke, G. R.; Hurburgh, C. R. Effect of Fourier-transform instrument resolution on soybean calibration performance. *IDRC Conference Poster*; 2010.

(18) Delwiche, S. R. Single wheat kernel analysis by near-infrared transmittance: protein content. *Cereal Chem.* **1995**, *72* (1), 11−16.

(19) Janni, J.; Weinstock, B. A.; Hagen, L.; Wright, S. Novel near-infrared sampling apparatus for single kernel analysis of oil content in maize. *Appl. Spectrosc.* **2008**, *62* (4), 423−426.

(20) Spielbauer, G.; Armstrong, P. R.; Baier, J. W.; Allen, W. B.; Richardson, K.; Shen, B.; Settles, M. High-throughput near −infrared reflectance spectroscopy for predicting quantitative and qualitative composition phenotypes of individual maize kernels. *Cereal Chem.* **2009**, *86* (5), 556−564.