JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Article

pubs.acs.org/JAFC

Characterization of Near-Infrared Spectral Variance in the Authentication of Skim and Nonfat Dry Milk Powder Collection Using ANOVA-PCA, Pooled-ANOVA, and Partial Least-Squares Regression

James M. Harnly,*[,†] Peter B. Harrington,*[,‡] Lucy L. Botros,[§] Joseph Jablonski,[∥] Claire Chang,[∥] Marti Mamula Bergana,[⊥] Paul Wehling,[#] Gerard Downey,[▽] Alan R. Potts,[§] and Jeffrey C. Moore[§]

[†]Agricultural Research Service, Beltsville Human Nutrition Research Center, Food Composition and Methods Development Laboratory, U.S. Department of Agriculture, Building 161, BARC-East, Beltsville, Maryland 20817, United States

[‡]Center for Intelligent Chemical Instrumentation, Department of Chemistry and Biochemistry, Ohio University, Clippinger Laboratories, Athens, Ohio 45701-2979, United States

[§]U.S. Pharmacopeial Convention, 12601 Twinbrook Parkway, Rockville, Maryland 20852-1790, United States

[∥]Center for Food Safety and Applied Nutrition, Division of Food Processing Science and Technology, U.S. Food and Drug Administration, 6502 South Archer Road, Bedford Park, Illinois 60501, United States

[⊥]Abbott Nutrition, Division of Abbott Laboratories, 3300 Stelzer Road, Columbus, Ohio 43219, United States

[#]General Mills, Inc., 330 University Avenue SE, Minneapolis, Minnesota 55414, United States

[▽]Teagasc Food Research Centre Ashtown, Ashtown, Dublin 15, Ireland

S *Supporting Information*

ABSTRACT: Forty-one samples of skim milk powder (SMP) and nonfat dry milk (NFDM) from 8 suppliers, 13 production sites, and 3 processing temperatures were analyzed by NIR diffuse reflectance spectrometry over a period of 3 days. NIR reflectance spectra (1700−2500 nm) were converted to pseudoabsorbance and examined using (a) analysis of variance-principal component analysis (ANOVA-PCA), (b) pooled-ANOVA based on data submatrices, and (c) partial least-squares regression (PLSR) coupled with pooled-ANOVA. ANOVA-PCA score plots showed clear separation of the samples with respect to milk class (SMP or NFDM), day of analysis, production site, processing temperature, and individual samples. Pooled-ANOVA provided statistical levels of significance for the separation of the averages, some of which were many orders of magnitude below 10^{-3}. PLSR showed that the correlation with Certificate of Analysis (COA) concentrations varied from a weak coefficient of determination (R^2) of 0.32 for moisture to moderate R^2 values of 0.61 for fat and 0.78 for protein for this multinational study. In this study, pooled-ANOVA was applied for the first time to PLS modeling and demonstrated that even though the calibration models may not be precise, the contribution of the protein peaks in the NIR spectra accounted for the largest proportion of the variation despite the inherent imprecision of the COA values.

KEYWORDS: *pooled-ANOVA, milk powder, near-infrared spectroscopy, partial least-squares regression, ANOVA-PCA*

INTRODUCTION

Skim milk powder (SMP) and nonfat dry milk (NFDM) are important foods/ingredients in the international food supply.[1] In 2011, the United States Pharmacopeial Convention (USP) initiated a project to characterize compositional and functional variations of these powders and develop nontargeted methods for detection of adulteration. One phase of this project analyzed 41 samples collected from 8 national and 3 international suppliers at 13 production sites (Table 1) using diffuse reflectance NIR.[1] The variance of the data was initially characterized using principal component analysis and varimax rotation[2] of the principal components. This Article provides the second step of the analysis by characterizing the NIR variance using analysis of variance-principal components analysis (ANOVA-PCA),[3] pooled-ANOVA,[4,5] and partial least-squares regression (PLSR).[6]

It has become customary to use chemometric methods to simplify and glean information from complex data sets. However, classical statistical methods, such as ANOVA, can also be applied to complex data. ANOVA can be used to compare the difference among the means of multiple groups, and in a factorial mode to isolate and quantify the variance associated with different experimental factors and their interactions. Pooled-ANOVA extends ANOVA from a single dependent variable to a pooled set of multiple variables.[4,5] Thus, ANOVA tests the differences of means among groups, and pooled-ANOVA tests the differences among two or more vectors of means by comparing the pooled variance of the variables.

Harrington et al.[3] used ANOVA to decompose the original sample matrix of mass spectrometry (MS) data into a series of

Received: March 27, 2014
Revised: June 26, 2014
Accepted: July 10, 2014
Published: July 10, 2014

DOI: 10.1021/jf5013727
J. Agric. Food Chem. 2014, 62, 8060−8067

Table 1. Sample Provenance

supplier[a]	1	1	1	2	2	2	2	3	4	5	6	7	8
prod. site[b]	US1	US2	US2	US3	US4	US5	US6	I1	I2	I3	US7	?	US8
milk class[c]	N	N	S	S	S	S	S	S	S	S	N	N	N
samples[d]	5	10	2	3	3	4	2	2	1	2	2	2	3
proc. type[e]	L,H	L,M	L,?	L,M	L,M	L,H	M	M	M	M	L,H	?	L
no. of analyses[f]	30	60	12	18	18	24	12	12	6	12	12	12	18

[a]Eight suppliers. [b]Production site per supplier in the United States (US), internationally (I), or unknown (?). [c]N = nonfat dry milk and S = skim milk powder. [d]Number of samples collected from each production site. [e]Process type: L = low, M = medium, H = high, or ? = unknown heat. [f]Six subsamples per sample (two subsamples/sample/day for 3 days).

submatrices, each associated with the means of different experimental factors. The final residual matrix was then added to each means matrix and analyzed by principal component analysis (PCA). This combined approach of ANOVA-PCA provides a scores plot for each experimental factor. The separation of scores of the different levels of the factor indicates significance. The separations are defined by 95% confidence ellipses derived from the Hotelling T^2 statistic. ANOVA-PCA provides information about the separation of the groups of objects with respect to the experimental factors.

Harnly[4,5] showed that the sum of the squared residuals of the variables of these submatrices derived from ANOVA-PCA could be pooled into a single scalar value and the mean squared statistic could be tested as it is done with ANOVA. The submatrices allowed computation of the percentage of the total variance for each experimental factor, the level of significance of the variance, and the variance associated with the interaction of the factors. The number of variables is omitted from calculation because it is difficult to determine the true rank or degrees of freedom from underdetermined data sets. Thus, the sum of squares across the number of variables is treated as a single squared variable, and the F statistic will underestimate significant differences. If a significant difference is observed, it will indeed be significant. Pooled-ANOVA provides a conservative test for differences among the level averages for each factor.

In this study, the previously reported NIR data for SMP and NFDM[1] are subjected to analysis by ANOVA-PCA, pooled-ANOVA, and PLSR coupled with pooled-ANOVA. The percentages of the total variance according to day of analysis, supplier, production site, milk class, process type, between sample variation, and instrumental variation were determined. A third aspect of the study examined the correlation of the NIR spectra with the protein, moisture, and fat content of the samples. Because this study was large with milk powders collected from many sources, it was difficult to achieve a balanced experimental design, and some of the data with respect to the experimental factors were missing, which prevented interaction from being studied for several factors. A more systematic study would attempt to collect an equal number of milk samples with respect to each experimental factor. Collecting milk powders from different sources produced on the same date and using the same process may be impossible because of limitations of the facility and production schedule.

MATERIALS AND METHODS

Milk Powders. As previously reported,[1] 41 milk powders (19 skim milk and 22 nonfat dry milk) produced between August 2008 and May 2012 were acquired from eight suppliers. Certificates of analysis (COA) indicated product origin details including production sites, lot numbers, process type (labeled as high, medium, and low heat), and weight/weight concentrations of moisture, fat, and protein. The sample information is summarized in Table 1.

The data were analyzed with respect to four sample variables, that is, sample class, production site, process type, and individual samples. The supplier was not treated as a variable because the majority of the suppliers (6 out of 8) only had one production site. A fifth instrumental variable was the day of analysis because the samples were each analyzed on three different days.

NIR Analysis. The NIR analysis methodology was previously reported.[1] In brief, Fourier transform (FT) near-infrared spectra were acquired at the U.S. Food and Drug Administration (Center for Food Safety and Applied Nutrition, Division of Food Processing Science and Technology) with a PerkinElmer Frontier FT-NIR system (Waltham, MA) fitted with the NIRA reflectance accessory. Data were acquired from 4000 to 10 000 cm^{-1} (1000−2500 nm). A 99% Spectralon diffuse reflectance standard (Labsphere, cat. no. AS-01160-060, North Sutton, NH) was used to acquire 84 background scans throughout the experiment. NIR spectra of six subsamples for each of the 41 milk powder samples were acquired in a randomized order on three consecutive days of analysis with two subsamples per milk powder being acquired on each day. In addition, on each day of analysis, six replicate measurements were acquired for randomly selected subsamples with the sample remaining on the sampling interface between replicates. Thus, spectral acquisition included 246 unique subsample spectra, 105 additional replicate spectra, and 84 background spectra, yielding a total of 435 spectra.

Data Processing. Raw data files consisted of 6001 reflected intensities (1000−2500 nm) for each of the 435 samples. As previously described,[1] the 84 spectra of the diffuse reflectance standard were used to convert reflectance intensities of the samples to pseudoabsorbance (A_s):

$$A_s = \log_{10}(I_{standard}/I_{sample}) \quad (1)$$

for which I_{sample} is the diffuse reflectance intensity of the samples, and $I_{standard}$ is the reflected intensity of the 99% reflectance standard. In the current study, only 228 of the 246 unique subsample spectra were used. The 105 replicate spectra were not included, and another 18 spectra (3 samples) were omitted due to provision of incomplete information, that is, missing information regarding process type (Table 1). Initial analysis by PCA showed little variation for the variable loadings of the first and second principal components in the region between 1000 and 1700 nm.[1] Consequently, only the region from 1700 to 2500 nm was used in this study.

Preprocessing of the pseudoabsorbance data prior to PCA transformed the data to the first derivative using a cubic polynomial fit of 35 points followed by standard normal variate (SNV) correction. The 20 points from the 2500 nm end of the wavelength range were removed so that the spectra would range from 1700 to 2480 nm. The derivative was calculated before the shorter wavelengths were excised from the data set so that no points were lost from the shorter wavelength ends of the derivatized spectra. Each derivatized spectrum comprised 1850 data points.

Outliers were detected by calculating the residual data matrix produced by subtracting the sample means from the corresponding spectra. Using a confidence interval defined by the Hotelling distribution at a 99.99% confidence interval applied to the T^2 statistic obtained from the first 10 principal component scores, three spectra

DOI: 10.1021/jf5013727
J. Agric. Food Chem. 2014, 62, 8060−8067

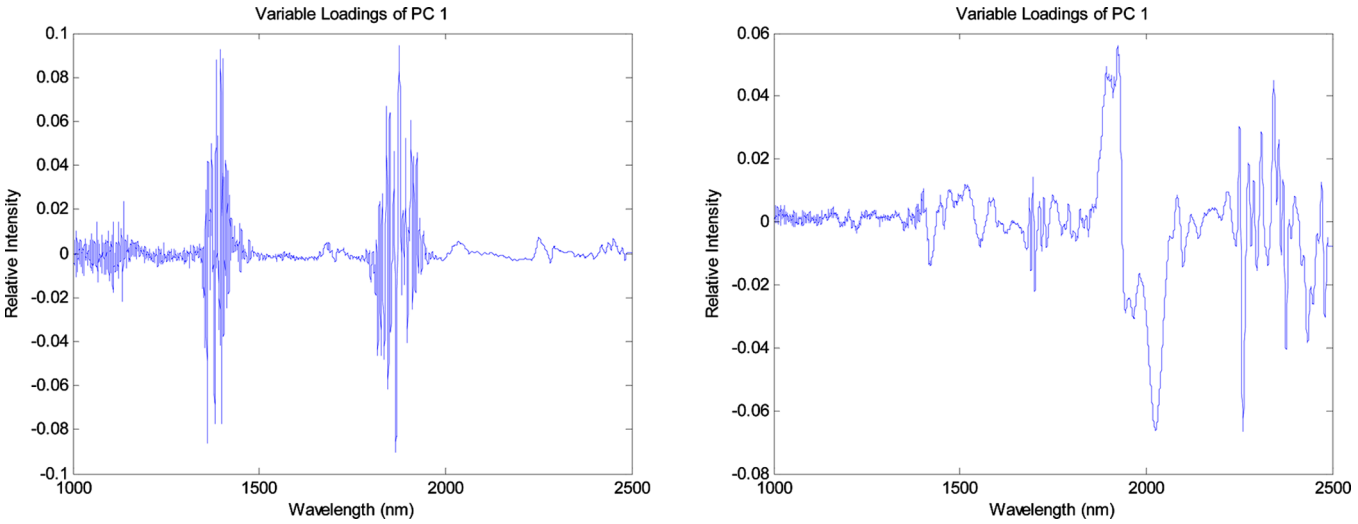

Figure 1. First principal component variable loadings of the SNV corrected first derivative spectra obtained from the single beam spectra (left) and pseudoabsorbance spectra (right).

were rejected because they had scores outside this interval. The number of spectra was reduced using this method from 228 to 225. Consequently, the derivatized and SNV corrected data matrix for the current study consisted of 1850 variables (1700−2480 nm) for 225 spectra (6 subsamples of 38 samples less the 3 outlier spectra). These 225 spectra were subjected to ANOVA-PCA, pooled-ANOVA, and PLSR.

PLSR used the NIPALS algorithm implemented in PLS2 mode (i.e., three dependent variables, moisture, fat, and protein). PLSR used a data set further reduced in size because some of the spectra were incomplete with respect to the COA for moisture, fat, and protein. This data set comprised 144 spectra from 24 samples and 7 production sites.

ANOVA-PCA and Pooled-ANOVA. Both approaches have been previously described.[2−5] In summary, submatrices are constructed for each experimental factor, for experimental factor interaction, and for the factor residuals. For ANOVA-PCA, the final residual matrix is added back to the means matrices for each factor and then subjected to PCA. For pooled-ANOVA, the residuals matrix for each factor and factor interaction is squared and summed to provide a pooled variance. A more detailed description of the calculation is given in Supporting Information Appendix A.

PLSR for Protein, Moisture, and Fat. COA concentrations for moisture, fat, and protein were only available for 24 of the 41 samples, which yielded 144 spectra. The data were preprocessed as described in the Materials and Methods, although it is worthwhile noting that the spectra without numerical derivatization yielded better PLSR prediction results. These results are not included. The PLS-2 nonlinear iterative partial least-squares (NIPALS) algorithm was used as described.[6] The optimum number of principal components was determined by the leave-one-spectrum-out cross validation and the minimum total sum of squared prediction error. Note that the 6 replicates per sample were run at two different times for three sequential days of analysis, so the replicates characterize uncontrolled sources of variation (e.g., instrumental drift with time).

Pooled-ANOVA PLS Regression. The pooled-ANOVA can be applied to the PLS calibration to determine the contribution to the X-block (NIR spectra) of each Y-block property (concentration of moisture, fat, and protein). The procedure is quite simple. For each latent variable or component, the pooled sum of squares for the NIR block of the PLS calculation was apportioned by the squared values of the concentration loadings vector (usually designated as $\boldsymbol{q}$). This calculation allows the variation contributions of each property (i.e., moisture, fat, and protein concentration) to be attributed to the NIR spectra.

A PLS model comprises a set of components that span the covariance of the X- and Y-blocks of data. Prediction is accomplished by iteratively removing the contribution of each PLS component or latent variable to the X-block and adding the contribution to the Y-block for the number of components chosen in the model.

$$t_{i,j} = \boldsymbol{x}_{i,j-1}\boldsymbol{w}_j \tag{2}$$

for which $x_{i,j}$ is the NIR spectrum of object i that has been centered with respect to the mean of the calibration set of data. The spectrum is modified for each iteration j of the PLS calculation, and $\boldsymbol{w}_j$ is the jth PLS component. The contribution to the NIR spectrum is calculated as

$$\hat{\boldsymbol{x}}_{i,j} = t_{i,j}\boldsymbol{p}_j \tag{3}$$

for which $\hat{\boldsymbol{x}}_{i,j}$ is a component of the reconstructed NIR spectrum i and p_j is a row vector from the PLS regression for component j. This contribution is iteratively removed or peeled from the NIR spectrum as follows.

$$\boldsymbol{x}_{i,j} = \boldsymbol{x}_{i,j-1} - \hat{\mathbf{x}}_{i,j} \tag{4}$$

The properties are iteratively estimated as

$$\hat{\boldsymbol{y}}_{i,j} = \hat{\boldsymbol{y}}_{i,j-1} + t_{i,j}b_j\boldsymbol{q}_j \tag{5}$$

for $\hat{\boldsymbol{y}}_{i,0}$ is the average of the properties and the estimate is built up with respect to PLS component number. The b_j is the inner relation, which can be considered a correlation coefficient between the X- and Y-blocks for each component. The property loadings $\boldsymbol{q}_j$ for component j are a row vector.

The pooled sums of squares of the X-block or NIR spectra with respect to the properties of the Y-block or concentrations are calculated as follows:

$$\boldsymbol{ss}_{\boldsymbol{x}_{i,j}} = \boldsymbol{q}_j^2 \sum_{k=1}^{n} \hat{\mathrm{x}}_{i,j,k}{}^2 \tag{6}$$

The sum of squares for the reconstructed component is calculated across the n variables of the spectrum or object i. The $\boldsymbol{ss}_{\boldsymbol{x}_{i,j}}$ is a row vector that contains the sums of squares of the X-block variance with respect to the PLS component j and the properties for the calibration, in this case, the concentrations of moisture, fat, and protein.

$$\boldsymbol{SS}_{\boldsymbol{x}} = \sum_{i=1}^{m}\sum_{j=1}^{r} \boldsymbol{ss}_{\boldsymbol{x}_{i,j}} \tag{7}$$

DOI: 10.1021/jf5013727
J. Agric. Food Chem. 2014, 62, 8060−8067

Table 2. Pooled-ANOVA of NIR Pseudo-Absorbance Data

treatment	df	SS	MS	F	p-value	RSS
milk class	1	221	221	130	$\ll 10^{-3}$	8.7%
day	2	21	11	6.2	3×10^{-3}	0.8%
processing type	2	223	111	65	$\ll 10^{-3}$	8.8%
production site	10	980	98.0	57	$\ll 10^{-3}$	38.8%
sample	37	665	18.0	10	$\ll 10^{-3}$	26.3%
milk × processing	2	108	54.2	31	$\ll 10^{-3}$	4.3%
processing × day	4	5	1	0.7	0.59	0.2%
milk × day	2	2	1	0.7	0.51	0.1%
site × day	20	53	2.7	1.5	0.07	2.1%
residual error	144	247	1.7	9.8%		

The $ss_{x_{ij}}$ is summed across the m objects and r PLS components that were used for calibration, and then the pooled-ANOVA is applied to the total ss_x as described previously.

RESULTS AND DISCUSSION

Selection of Spectral Region. The SNV corrected spectra, first derivative with SNV spectra, and variable loadings of the first principal component for both diffuse reflectance and pseudoabsorbance modes are given in Figure 1S of the Supporting Information. Figure 1 gives the variable loadings obtained from the diffuse reflectance spectra, which have two major water bands at 1340−1440 and 1800−1950 nm. The variable loadings obtained from the pseudoabsorbance spectra are given on the right. The water bands of these variable loadings are attenuated when compared to those of the diffuse reflectance spectra. However, the pseudoabsorbance calculation did not completely remove the moisture bands as will become apparent later in the subsequent results. The reference standards were only run every sixth analysis and did not account for any moisture in the milk powder but only for ambient moisture in the NIR beam path. There was no attempt to interpolate a reference spectrum for the spectra acquired between the two reference spectra. Thus, depending on the direction of the drift, the reference standards may have over- or under-corrected the sample pseudoabsorbance spectra.

Pooled-ANOVA. Table 2 reports the pooled-ANOVA results for the analysis of the NIR pseudoabsorbance data matrix comprising 38 samples (i.e., 225 spectra) each with 1850 variables (1700−2500 nm). The residuals matrix for each experimental factor and factor interaction were squared to provide the factor variance (see Supporting Information Appendix A). Numbers in bold designate cases associated with significant differences (i.e., less than 10^{-2} of a type I error). In other words, bold designees indicate if the average levels of the experimental factors came from the same distribution, which is the Null Hypothesis. The likelihood of obtaining this result would be equal to or less than the corresponding p-value. The relative sum of squares (RSS) gives the relative contribution of each factor with respect to the total sum of squares. The sample residuals account for 10% of the RSS as indicated by the residual error (bottom row) of Table 2. The sum of the pairwise factor interactions accounted for less than 7% of the RSS. Rows 6−9 of Table 2 list the interactions between pairs of factors.

At the 99% confidence level, there are significant differences for milk class, day of analysis, production site, process type, and samples (Table 2). It should be noted that pooled-ANOVA, and ANOVA in general, tests for a difference among the averages of each level. Thus, significant differences for factors reported in Table 2 do not indicate that all of these averages are different, but that at least one average of a level differs from the other averages. Further matrix calculations can provide information for specific levels. However, visual methods such as ANOVA-PCA are easier to interpret.

ANOVA-PCA. The submatrices used for the pooled-ANOVA calculations in the previous section were also used to compute reconstituted matrices for each factor prior to PCA (see Appendix A in the Supporting Information). The same experimental factors evaluated by pooled-ANOVA in Table 2 are described by ANOVA-PCA in the PCA scores and variable loadings given in Figure 2. However, pooled-ANOVA takes advantage of the power of the Central Limit Theorem to determine if the average spectra are varying, while ANOVA-PCA compares the separation of data points with respect to the individual spectra.

In the top row of Figure 2, the effect of milk class is clearly significant. Skim milk powder (SMP) is compared to nonfat dry milk (NFDM). The RSS reported in Table 2 for milk type accounted for 8.7% of the variation in the data set. The two milk types were significantly different (i.e., p-value $\ll 10^{-3}$), and this estimate underestimates the significance because the degrees of freedom from the 1850 NIR spectral variables are neglected.

The second row of Figure 2 gives plots for the effect of the day to day variation. In the score plots, the 95% confidence ellipses for the 3 days of the analysis all overlap with respect to the residual error. However, the third day of analysis, C, clearly has a different mean that is characterized by a shift with respect to the second principal component. The variable loadings for this component indicate that moisture peaks are primarily responsible for the difference. The pooled-ANOVA result (Table 2) is at a level of significance of 0.003, indicating that the averages are different even though the fraction of variance explained by the RSS is a relatively small 0.8%.

The third row corresponds to the three process types labeled as low (LH), medium (MH), and high (HH) heat. The medium and high temperature processes are more different from each other with the low temperature appearing between the two. This trend can be explained by the process differences that arise in milk condensation and dehydration. The pooled-ANOVA results indicate a RSS of 8.8% and similar p-value that was far below 10^{-3}.

The fourth row corresponds to the production site's geographical location, which accounts for the largest source of variation with an RSS of 38.8% and a p-value also far less than 10^{-3}. The score plots indicate that the production sites form nicely defined clusters in the data space.

DOI: 10.1021/jf5013727
J. Agric. Food Chem. 2014, 62, 8060−8067

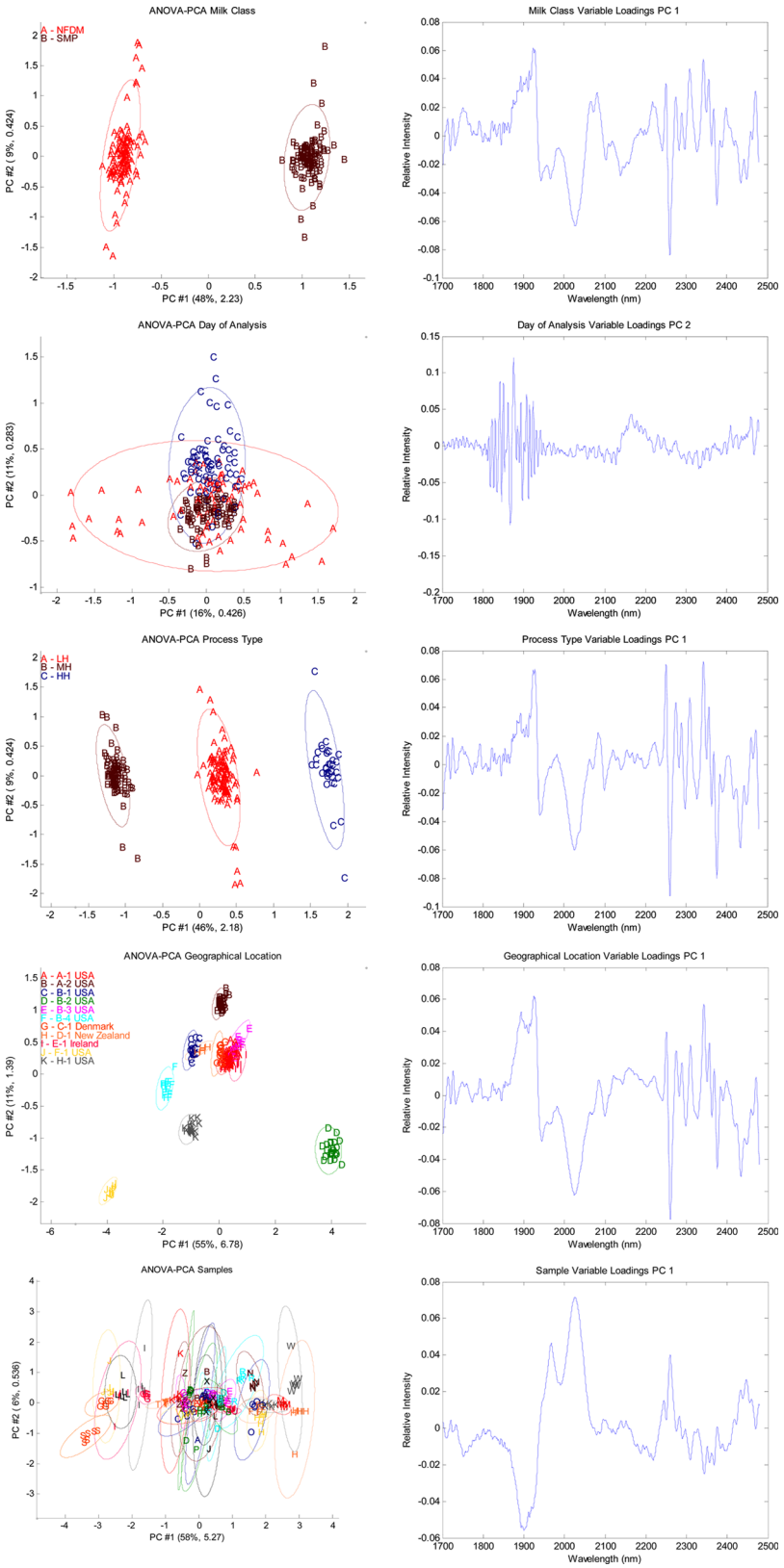

Figure 2. (Left) ANOVA-PCA score plots with 95% confidence ellipses. (Right) Corresponding variable loadings. (Row 1) Skim milk (SMP) and nonfat dry milk (NFDM). (Row 2) Day of analysis; for this case, the variable loading for the second principal component is plotted. (Row 3) Low, medium, and high processing temperatures. (Row 4) Production site of milk supplier. (Row 5) Thirty-eight samples.

Each sample was taken from a different lot of milk powders. The fifth row of Figure 2 corresponds to the individual sample level, and, according to its RSS, the samples contribute 26.3% of the variance in the data set, making it the second largest contributor. The sample scores that form tight clusters indicate that NIR spectroscopy of the dried milk powders has the

DOI: 10.1021/jf5013727
J. Agric. Food Chem. 2014, 62, 8060−8067

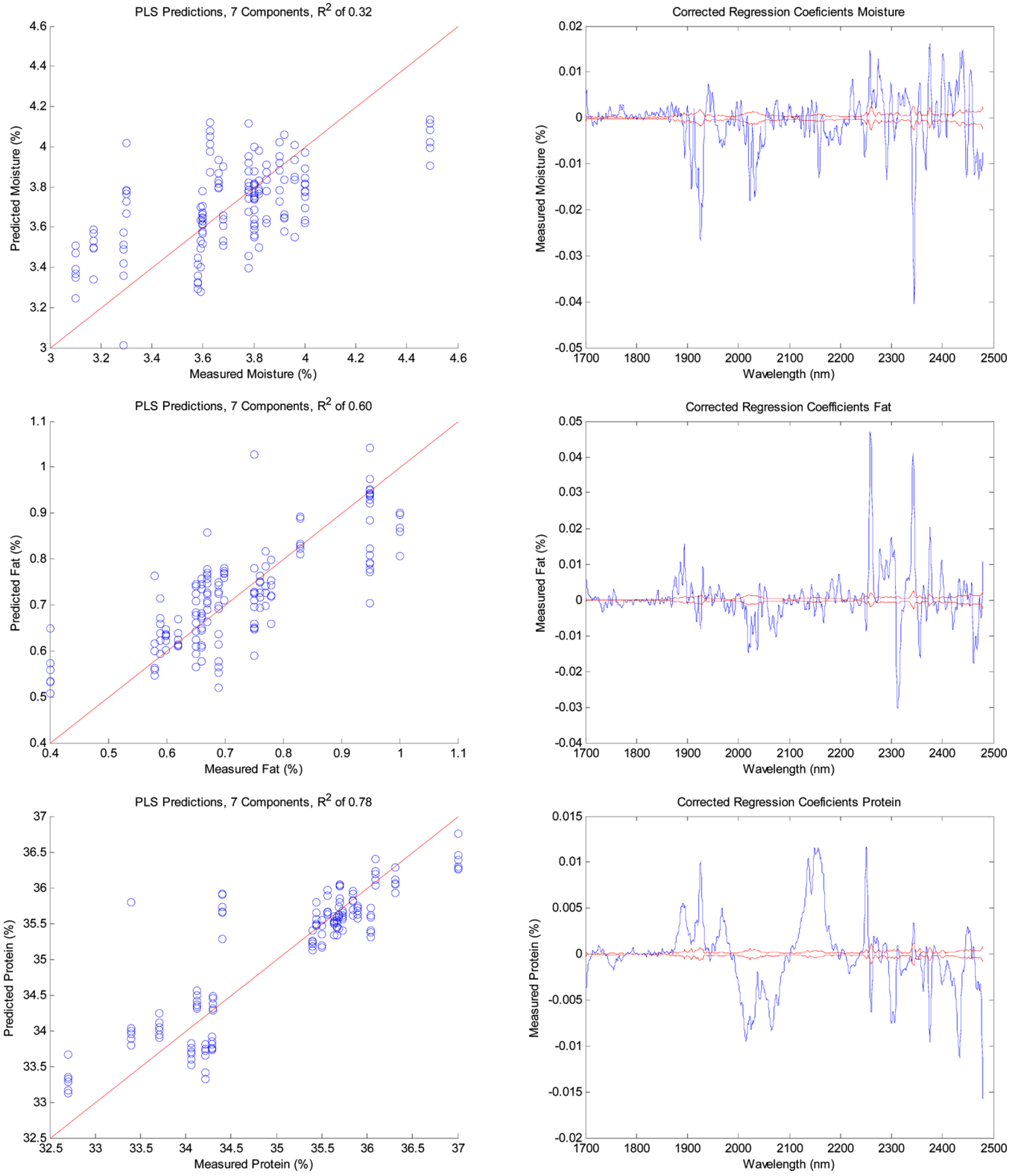

Figure 3. Partial least-squares regression cross-validation prediction results using seven latent variables (left). The corresponding corrected average regression coefficients and their 95% confidence intervals (right) for moisture (top), fat (middle), and protein (bottom).

potential to discriminate among the different lots of milk powders.

Interaction between factors was also examined; the figures are available in the Supporting Information (Figure 2S). However, not all interactions could be studied because the experimental design had confounded some of these factors. When plotting the interaction scores, the main effect averages were added back to the data as well as the interaction means for the ANOVA-PCA. The addition of the main effect averages to the interaction averages allows the interplay of factors to be displayed in Supporting Information Figure 2S. The top row/left side of Supporting Information Figure 2S gives the scores

DOI: 10.1021/jf5013727
J. Agric. Food Chem. 2014, 62, 8060−8067

for the interaction of process type with milk class. One can see that six subgroups are all resolved. On the right side of the top row, the variable loadings for the first principal component are given, which indicate the features of the NIR spectrum that are prominent for this interaction.

However, for the pooled-ANOVA results reported in Table 2, the main effect averages were not included in the sum of squares calculation. The RSS is smaller than those of the main effects with a value of 4.3%; yet this interaction is significant with a p-value far less than 10^{-3}. This result is expected because the process type should affect the milk classes differently since the milk classes differ in chemical and physical composition. NIR spectrometry is affected by the powder morphology, particle size, and chemical composition, all of which may differ among the milk classes, processes, and production sites.

The interactions between milk class and process type with the day of analysis are illustrated by the ANOVA-PCA score plots in the second row of Supporting Information Figure 2S. The confidence intervals of the days are not resolved, although we can see the third day is shifted a bit off center. The respective RSS values for the interaction between day and milk class and between day and process type are 0.2% and 0.1%, respectively. For these cases, the pooled-ANOVA did not detect a significant difference among the means as reported in Table 2.

The bottom row of Supporting Information Figure 2S gives the ANOVA-PCA score plots of the interaction of production site with day of analysis. Because of the large number (3×11) of labels for this interaction, the same set of scores is plotted with the day label (left) and the production site label (right). In Table 2, the corresponding RSS is 2.1%, which was not significant at a 95% level of confidence. For this interaction term, the p-value was 7%. This result occurs because different samples may have various interactions with moisture because of their varied chemical and physical compositions.

Influence of Fat, Protein, and Moisture. An earlier report by this team of investigators implicated the involvement of fats, proteins, and moisture as an influence on the sample variability.[1] Varimax rotations provided variable loadings that correlated with the bands of all three families of compounds. For the partial least-squares regression (PLSR) study, spectra from the 24 samples that had complete COA concentrations for moisture, fat, and protein were used. The concentrations were weight to weight percentages. No spectra from this set of samples were detected as outliers using the same method that was applied to the ANOVA-PCA, so 144 spectra (24 samples × 6 replicates) were available for calibration by PLS regression. Because the two replicates on each day of analysis were blocked with respect to time, a leave one spectrum out cross-validation was used for model building. Each replicate characterized variations with respect to time and sampling by the NIR spectrometer. The squared prediction errors were summed across the fat, moisture, and protein concentrations and plotted with respect to the number of latent variables used by the PLS model. A minimum total prediction error was obtained with seven latent variables or PLS components, so this model was used for the subsequent calculations and figures. Data are not shown.

The prediction results using cross validation of the seven-component PLS model are given in Figure 3. During the cross-validation procedure, each PLS model was converted to a linear set of regression coefficients. These regression coefficients were stored and used to calculate average regression coefficients with 95% confidence intervals that are also given in Figure 3. The coefficient of determination (R^2) values were calculated with respect to the cross-validated prediction results and are included with each figure along with a red reference line indicating ideal behavior.

Table 3 reports the pooled-ANOVA results derived from the PLS model. The sum of squares for the NIR (X-Block) are calculated for each component as described by eqs 2−7. From the PLS calibration, the contributions of each concentration can be calculated as a sum of squares for the NIR spectra. In addition, the residual sum of squares (i.e., variation that is not accounted for by the calibration model) can also be calculated for in the spectra. These sums of squares were totaled for the seven components to furnish the top half of Table 3.

Table 3. Pooled-ANOVA Partial Least-Squares Regression Analysis

Near-Infrared Spectra Block							
treatment	df	SS	MS	F	p-value		RSS
moisture	6	438	72.7	52	$\ll 10^{-3}$		25.7%
fat	6	178	29.7	21	$\ll 10^{-3}$		10.5%
protein	6	897	149.5	107	$\ll 10^{-3}$		52.6%
residuals	137	192	1.4				11.2%
Concentration Block							
treatment	df	SS	MS	F	p-value	RMSEP	R^2
moisture	6	46	17.7	10.8	$\ll 10^{-3}$	0.24%	0.32
residual	137	97	0.7				
fat	6	87	14.4	35.0	$\ll 10^{-3}$	0.09%	0.61
residual	137	97	0.4				
protein	6	112	18.7	82.2	$\ll 10^{-3}$	0.48%	0.78
residual	137	97	0.2				

The lower half of Table 3 was obtained from the moisture, fat, and protein concentration predictions using the leave-one-spectrum-out cross validation. The sum of squares of the regression was obtained by subtracting the sum of the squared errors from the sum of the squared mean-centered measured concentrations for moisture, fat, or protein.

Although the coefficients of variation (R^2) are not large, it is important to remember that the calibration did not comprise prepared standards over a wide range of concentrations; instead, the concentration ranges used for the calibration were those that were measured from each lot of milk powder, which happened to vary over a narrow range.

Note that, although the contribution to the NIR data for moisture is relatively large as compared to fat, the R^2 value of 0.32 is quite low. This trend indicates that the moisture bands in the spectra do not correlate well with the COA moisture concentrations. The samples were stored under ambient conditions, and the moisture content most likely varied during storage. In addition, the COA moisture concentration was measured for the lot of milk powder and not for the individual sample. Last, the range of moisture values was relatively small at 1.8−4.5%, which typically can result in smaller coefficients of determination. NIR spectrometry is sensitive to moisture, and variations of ambient moisture during the measurement may not have been sufficiently corrected by the reference spectrum. According to the RSS, moisture was responsible for one-quarter of the total variation of the NIR spectral data. However, it is important to remember that these data were obtained from a multinational study of many diverse milk powder samples, so the lack of correlation with the moisture concentrations could

DOI: 10.1021/jf5013727
J. Agric. Food Chem. 2014, 62, 8060−8067

be attributed to the different chemical and physical properties of this diverse set of samples.

Fat had the smallest contribution to the NIR data, but its R^2 of 0.61 is more than twice that of the moisture. The range of fat concentrations was 0.01−1.05%. The RSS of fat was 10.5%, which indicates that, although the concentration of fat was less than moisture, it still contributed a significant source of variation to the NIR spectra. Although fat is nonvolatile and fairly inert, it could also be affected by ambient storage; for example, lipid oxidation may alter the spectral profile.

Protein was the largest contributor to the variation of the NIR data and gave the best R^2 of 0.78. The COA protein concentration range was 32.7−37.0%. The results were encouraging when one considers the wide range of milk powder samples and that the concentration values were determined by lot and not by sample. Typically, calibration models would be devised from a single milk source across a controlled range of concentrations.

Supporting Information Figure 3S gives the production sites with respect to the predicted fat and protein concentrations. These predictions are from the cross-validation of the seven-component PLS regression model. This plot suggests that milk samples processed in the different production sites have different fat and protein concentrations. The 95% confidence ellipses indicate that some sites are statistically different.

In this study, the 38 samples analyzed by NIR showed statistically significant variability with respect to the day of analysis, production site, milk class, process type, and individual samples. NIR spectrometry is very sensitive to chemical composition and physical properties of milk powder products and is an excellent tool for judging the quality of product. The lot to lot variation was not studied. However, the individual samples were precisely determined, which indicates that for targeted or untargeted analysis of milk powders, the models would have to be specific for each sample or lot of milk. Perhaps a future study will examine the lot to lot variations of the milk powders.

Significant correlations were found with the COA concentration values even though some of the coefficients of determination R^2 indicated rather weak predictability of models. Many effects could be responsible for the weak PLS models. The range of concentrations was relatively narrow because the observed concentration range was used as opposed to using a standard range of concentrations that would presumably be wider. In addition, the concentrations of fat, moisture, and protein may have changed during storage of the samples. Because these concentrations are relative to the total weight, a change in any one component could also affect the concentration of the others; for example, the adsorption of moisture by the dried milk sample would lower its relative concentrations of protein and fat.

Pooled-ANOVA was combined with PLS regression for the first time. Although the PLS calibration models were not impressively accurate, the pooled-ANOVA applied to PLS regression models demonstrated that the correlated peaks of the NIR spectra were important contributors to the variances of the data set with the protein bands accounting for more than the majority and the moisture bands one-quarter of the total variation of the data set.

This study indicates that for targeted or untargeted analysis of adulterants in milk powders by near-infrared spectrometry, a global model will probably not succeed because the variations arising from milk class, process type, and production site were all significant. Instead localized models that are specific for milk class, process type, and production site would be expected to have a better opportunity for success. Furthermore, local standardization of moisture, fat, and protein would improve calibration models for these important properties. The significance of the variation with respect to the day of analysis indicates that a better approach for collecting reference spectra may be beneficial, especially with respect to controlling variations that arise from ambient moisture.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Implementation of pooled-ANOVA in Excel, and additional figures. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Authors

*Tel.: (301) 504-8569. E-mail: james.harnly@ars.usda.gov.
*Tel.: (740) 994-0265. E-mail: peter.harrington@ohio.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Walter Hauck and Markus Lipp from USP are thanked for their valuable comments. One of the authors (JMH) acknowledges support or the Agricultural Research Service of the U.S. Department of Agriculture and the Office of Dietary Supplements of the National Institutes of Health.

■ REFERENCES

(1) Botros, L. L.; Jablonski, J.; Chang, C.; Bergana, M. M.; Wehling, P.; Harnly, J. M.; Downey, G.; Harrington, P.; Potts, A. R.; Moore, J. C. Exploring authentic skim and nonfat dry milk powder variance for the development of nontargeted adulterant detection methods using near-infrared spectroscopy and chemometrics. *J. Agric. Food Chem.* **2013**, *61*, 9810−9818.

(2) Kaiser, H. F. The Varimax criterion for analytic rotation in factor analysis. *Psychometrika* **1958**, *23*, 187−200.

(3) Harrington, P. D.; Vieira, N. E.; Espinoza, J.; Nien, J. K.; Romero, R.; Yergey, A. L. Analysis of variance-principal component analysis: A soft tool for proteomic discovery. *Anal. Chim. Acta* **2005**, *544*, 118−127.

(4) Luthria, D. L.; Lin, L. Z.; Robbins, R. J.; Finley, J. W.; Banuelos, G. S.; Harnly, J. M. Discriminating between cultivars and treatments of broccoli using mass spectral fingerprinting and analysis of variance-principal component analysis. *J. Agric. Food Chem.* **2008**, *56*, 9819−9827.

(5) Luthria, D. L.; Mukhopadhyay, S.; Robbins, R. J.; Finley, J. W.; Banuelos, G. S.; Harnly, J. A. UV spectral fingerprinting and analysis of variance-principal component analysis: a useful tool for characterizing sources of variance in plant materials. *J. Agric. Food Chem.* **2008**, *56*, 5457−5462.

(6) Geladi, P.; Kowalski, B. R. Partial least-squares regression: A tutorial. *Anal. Chim. Acta* **1986**, *185*, 1−17.

DOI: 10.1021/jf5013727
J. Agric. Food Chem. 2014, 62, 8060−8067