AAPS PharmSciTech 2007; 8 (4) Article 96 (http://www.aapspharmscitech.org).

Themed Issue: Process Analytical Technology
Guest Editor - Ajaz Hussain

Determination of Figures of Merit for Near-Infrared and Raman Spectrometry by Net Analyte Signal Analysis for a 4-Component Solid Dosage System

Received: May 3, 2007; Final Revision Received: May 29, 2007; Accepted: May 31, 2007; Published: Novemebr 9, 2007

Steven M. Short,[1] Robert P. Cogdill,[1] and Carl A. Anderson[1]

[1]Duquesne University Center for Pharmaceutical Technology, 410A Mellon Hall, 600 Forbes Avenue, Pittsburgh, PA

ABSTRACT

Process analytical technology has elevated the role of sensors in pharmaceutical manufacturing. Often the ideal technology must be selected from many suitable candidates based on limited data. Net analyte signal (NAS) theory provides an effective platform for method characterization based on multivariate figures of merit (FOM). The objective of this work was to demonstrate that these tools can be used to characterize the performance of 2 dissimilar analyzers based on different underlying spectroscopic principles for the analysis of pharmaceutical compacts. A fully balanced, 4-constituent mixture design composed of anhydrous theophylline, lactose monohydrate, microcrystalline cellulose, and starch was generated; it consisted of 29 design points. Six 13-mm tablets were produced from each mixture at 5 compaction levels and were analyzed by near-infrared and Raman spectroscopy. Partial least squares regression and NAS analyses were performed for each component, which allowed for the computation of FOM. Based on the calibration error statistics, both instruments were capable of accurately modeling all constituents. The results of this work indicate that these statistical tools are a suitable platform for comparing dissimilar analyzers and illustrate the complexity of technology selection.

KEYWORDS: Near-infrared, Raman, partial least squares, analyte signal, calibration, tablet.

INTRODUCTION

Several technologies may be suitable for a given analytical measurement application. Ultimately, a decision must be made as to which device will be deployed. Common methods for comparison of instruments based upon different fundamental principles are not well established. While instrument selection can be based on many different aspects of a technology (cost, performance, infrastructure, etc), this work will focus only on method performance characterization for 2 sample technologies.

Process analytic measurement applications in pharmaceutical science include qualitative and quantitative identification of compounds, examination of phase transformations and polymorphs, and investigations of manufacturing and process development. Near-infrared (NIR) and Raman spectroscopy are both sensitive to, and capable of accurately predicting, these phenomena.[1,2] As NIR spectroscopy is based on optical absorption and Raman spectroscopy on the measurement of inelastic scattering, comparison of the parallel performance of these 2 classes of instrumentation is indirect. More often than not, the instrument eventually deployed is selected based on limited criteria that do not consider all aspects of performance.

Calibrations are frequently evaluated using the coefficient of determination (R^2) and/or estimation of prediction error.[3-7] These measures do not directly consider issues such as precision[5,6] and signal-to-noise (S/N) ratio.[8,9] The effectiveness of a method should be judged based on a complete assemblage of indicators that describe all aspects of performance and are generalized across technological platforms. Net analyte signal (NAS) theory provides a mechanism for directly determining figures of merit (FOM) for the performance of a method.

NAS theory is the concept of separating relevant signals for a particular component of interest from the remaining interfering elements present within the spectra.[10] Lorber[11] is widely acknowledged as the originator of multivariate NAS theory; Brown,[12] however, makes the clarification that Morgan[13] published on a similar topic prior to Lorber. NAS provides a tool for calculating multivariate FOM; prior to Morgan/Lorber, techniques similar to NAS were applied to data from only univariate methods.[14] For a more detailed discussion of FOM for univariate calibration, please refer to Olivieri et al and their references.[14]

Multivariate NAS was first implemented using pure component projection[11] and classical regression.[15] Implementation of NAS via pure component or classical regression methods is potentially cumbersome[10]; this problem was solved by Lorber et al[10] by using inverse regression instead.[15] Both classical and inverse regression mathematics are suitable for

Corresponding Author: Carl A. Anderson, Duquesne University Center for Pharmaceutical Technology, 410A Mellon Hall, 600 Forbes Avenue, Pittsburgh, PA 15282. Tel: (412) 396-1102; Fax: (412) 396-4660; E-mail: andersonca@duq.edu

the determination of NAS; however, the remaining descriptions employ inverse regression.

Mathematically, NAS is defined as the portion of signal unique to the constituent being considered and is orthogonal to all other factors present in the data.[10,11] NAS, therefore, involves the signal directly useful for quantification.[10] A mixture spectrum (r) extracted from a spectral matrix containing multiple constituents can be resolved into the following:

$$\mathrm{r} = \mathrm{r}^{*} + \mathrm{r}^{\perp} \qquad (1)$$

where $\mathbf{r}^*$ and $\mathbf{r}^{\perp}$ are mutually orthogonal components representing the NAS vector for the particular component of interest and the vector of interferences, respectively.[10] It is well understood that controllable and uncontrollable errors influence the performance of any analytical technique. This error (ε), which varies for each sample acquired, can also be partitioned into its respective mutually orthogonal components: the part that is orthogonal to the interferences (ε^*), and the portion that lies within the interference space ($\varepsilon^{\perp}$).[16] It is important to understand that the estimated r* will not lie exactly in the true direction of the NAS vector as a result of ε^*.[16] ε^* is the portion of total stochastic error (ε) that contributes to imprecision.

Multiple algorithms for calculating NAS have been reported in the literature,[17,18] with the method of computation and resulting output differing. In this work, the method proposed by Bro and Andersen is used.[19] All equations use **X** to represent spectral matrices; bold characters other than **X** (eg, **x**, **y**, and **NAS**) represent vectors; small, italicized characters (eg, *x* and *y*) represent scalars. Additionally, the notation $\|\mathbf{x}\|$ signifies the Euclidean norm of **x** and the superscript *T* indicates the transposition of a **vector**.

Bro and Andersen calculate the NAS vector for a particular component of interest as follows:

$$\hat{\mathrm{NAS}}_i = (\mathrm{x}_i \cdot \mathrm{b}) \cdot \left(\mathrm{b}^T \cdot \mathrm{b}\right)^{-1} \cdot \mathrm{b}^T \qquad (2)$$

where $\mathbf{x}_i$ is a sample spectrum from matrix **X** and **b** is a column vector of the regression coefficients for **X**[19]; principal components regression or partial least squares (PLS) are common regression techniques used to estimate **b**.[20] It should be noted that **X** is corrected for the mean; thus, outputs from computations employing $\hat{\mathrm{NAS}}$ are mean-centered. Results can be rescaled using the vector of means from the centering operation of **X** to the original range. NAS can also be expressed in scalar form, with no loss in information, by the following equation[10]:

$$\hat{\mathrm{NAS}}_i = \left\| \hat{\mathrm{NAS}}_i \right\| \qquad (3)$$

Additional discussion concerning the mathematics behind the determination of NAS can be found elsewhere.[10-13,16-19,21-23]

Materials and Methods

Tablet Production

A fully balanced, 4-constituent mixture design composed of anhydrous theophylline (Lot No 92577, Knoll AG, Ludwigshafen, Germany), Lactose 316 Fast Flo NF Monohydrate (Lot No 8502113061, Hansen Labs, New Berlin, WI), microcrystalline cellulose (MCC; Avicel PH 200, Lot No M427C, FMC BioPolymer, Mechanicsburg, PA), and soluble starch GR (Lot No 39362, EMD Chemicals Inc, Gibbstown, NJ) was generated. The approximate median particle size of the theophylline, lactose, MCC, and starch (reported by documentation from their respective suppliers) was ~90, ~100, ~180, and ~17 μm, respectively. No further analysis or alterations were performed on the materials to determine or alter the particle size distribution. Twenty-nine design points were chosen to cover a wide range in all constituents and to remove any possibility of factor aliasing. Analysis of the mixture covariance matrix (not shown) demonstrated that the design was balanced in all directions, giving equal emphasis to all constituents.

Materials for each design point mixture were dispensed by weight (Data Range, model no AX504DR, Mettler Toledo, Columbus, OH) and were transferred to 25 mL glass scintillation vials. In total, 6000 mg of material was weighed out for each point, and the nominal weights for all constituents were adjusted to the observed mass data to calculate actual concentration. The vials were tumbled for 5-minute cycles on a rotating Jar Mill (United States Stoneware, East Palestine, OH). After every blending period, an NIR reflectance spectrum was acquired through the bottom of the glass (FOSS NIRSystems 5000, FOSS NIRSystems, Inc, Laurel, MD). An ad hoc partial least squares II (PLS-2) calibration, using the constituent concentrations as reference data, was constructed after each blending cycle to assess homogeneity. Mixtures were assumed to be homogeneous when further mixing failed to yield an increase in the calibration's coefficient of determination.

The mixtures from each design point were then subdivided and tableted at 5 levels of compaction force (67.0, 117.3, 167.6, 217.8, and 268.1 MPa) on a Carver Automatic Tablet Press (Model 3887.1SD0A00, Wabash, IN) using flat-faced punches and a 13-mm die. A dwell time of 10 seconds was employed. Six ~800-mg compacts were produced per design point, with the sixth tablet's compaction force randomly selected from 1 of the 5 possible levels. The compaction order was randomized to ensure homoscedasticity of experimental error. Following compaction, the samples were allowed to relax for 15 days prior to spectroscopic analysis.

Data Acquisition, Instrumentation, and Software

NIR reflectance measurements for both sides of each tablet were acquired over the wavelength range of 1100 to 2498 nm at a 2-nm increment averaging 32 scans (FOSS NIR Systems 5000, Vision version 2.00, FOSS NIRSystems, Inc, Laurel, MD). Prior to scanning, the tablets were precisely centered using the positioning iris standard on this instrument. Two ad hoc PLS-2 calibrations, using the constituent concentrations as reference data, were constructed from spectra corresponding to a specific surface of the tablets.

Raman data were measured using a prototype *P*^{h}*AT System* spectrometer with a laser excitation wavelength of 785 nm and a fiber-coupled probe head (HoloGRAMS version 4.0, Kaiser Optical Systems, Inc, Ann Arbor, MI). The *P*^{h}*AT System* samples a spot size of ~6 mm. Two accumulations were acquired per scan employing an integration time of 10 seconds over the range of –64.2 to 1895.7 cm^{-1} at a 0.3 cm^{-1} increment. A dark scan was subtracted and the cosmic ray filter and intensity calibration options were selected. The tablets used in this study were larger than could be accommodated in the tablet holder located in the sample chamber of the *P*^{h}*AT System,* so the tablets had to be manually positioned such that the laser spot was visually centered on the flat face.

All spectral data were analyzed in the Matlab environment (version 7.1, MathWorks, Natick, MA) using the PLS_Toolbox (version 3.0, Eigenvector Research, Inc, Manson, WA) and software developed by the Duquesne University Center for Pharmaceutical Technology.

PLS Analysis

NIR and Raman data were analyzed separately but in an identical fashion. The NIR spectral range and resolution were not altered; however, the Raman spectral range was truncated to 205.5 to 1895.7 cm^{-1} to remove residual Rayleigh line radiation and to reflect the operating range of the analyzer. Prior to calibration, the Raman data were evaluated using a moving-window calibration technique[21] with various window widths to determine whether further wavenumber truncation would be beneficial. PLS regression[24] was used via the SIMPLS algorithm[25] to relate spectroscopic response to concentration for each constituent individually. Since analyte concentration is incorporated into the denominator of some FOM calculations, samples having a corresponding zero concentration for the component being considered were excluded. Therefore, samples included in the actual calibration data sets were unique for each component.

Preprocessing routines, including standard normal variate scaling, detrending, derivatives, and combinations of these, were tested.[2] The most favorable data pretreatment method was selected based on a minimization of "batchwise" cross-validation error, where the batches in this instance were the 5 different compaction levels. The root-mean-standard error (RMSE) for cross-validation (RMSECV) and calibration (RMSEC) were calculated using the following formula:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \quad (4)$$

where y_i is the measured concentration, $\hat{y}_i$ is the predicted concentration, and n is the number of samples for the data set under consideration. According to International Conference on Harmonisation (ICH) guidelines, accuracy expresses the agreeability between reference and predicted values.[26] Therefore, RMSEC and RMSECV are used to report the accuracy of the selected calibration.

Multivariate FOM

The optimal number of latent variables selected during the estimation of each PLS regression vector was determined by minimizing RMSECV. Once established, the regression vector was used to determine the NAS according to Equations 2 and 3. Given the number of chemical constituents and physical factors varying in this design, it was anticipated that no more than 4 latent variables would be required; models with greater rank are feasible but would be increasingly difficult to justify. The NAS vector affords the opportunity to calculate numerous FOM, such as sensitivity, analytical sensitivity, selectivity, and S/N ratio. FOM can be determined for every sample using Equations 5 to 11.

Sensitivity characterizes the extent of signal variation as a function of analyte concentration; the higher the sensitivity, the greater the instrumental response to an increase in concentration.[10,11] Sensitivity is calculated as follows[10,14,18]:

$$\hat{\mathrm{SEN}}_i = \frac{\hat{\mathrm{NAS}}_i}{y_i} \quad (5)$$

where $\hat{\mathrm{SEN}}_i$, $\hat{\mathrm{NAS}}_i$, and y_i are the vector of sensitivity for each instrument variable, the NAS vector, and the measured concentration for the *i*th sample, respectively. Sensitivity is reported in units of instrument intensity per concentration. Measured concentrations are autoscaled before being used in Equation 5. It is also possible to express sensitivity as a univariate figure of merit by taking the Euclidean norm of the sensitivity vector:

$$\hat{SEN}_i = \left\| \hat{\mathrm{SEN}}_i \right\| \quad (6)$$

where $\hat{SEN}_i$ is the univariate measure of sensitivity for the *i*th sample.[18,22] Sensitivity is reported in this document as

the mean of the univariate sensitivity values for all samples under consideration.

Sensitivity is applicable to only calibrations constructed on devices operating under the same fundamental principles because it incorporates units of instrument signal (NIR absorbance intensity and Raman scattering intensity). The parameter analytical sensitivity (γ) was developed to provide an impartial assessment of dissimilar analytical techniques.[22] Analytical sensitivity is calculated as follows:

$$\gamma = \frac{\hat{SEN}}{\delta r} \tag{7}$$

where $\hat{SEN}$ is the mean of the sensitivity values under consideration found using Equation 6 and δr is a measure of instrumental noise. This normalization procedure allows direct comparison of the measure of sensitivity associated with NIR and Raman data. Analytical sensitivity has the dimension of concentration^{-1}. For the work presented in this article, δr was determined as the mean standard deviation of the predicted concentrations of 4 tablets broadly varying in constituent concentrations, scanned once a day for 6 consecutive days. Additionally, this figure of merit allows an estimation of the minimum discernible concentration difference given the dynamic range modeled[22] (γ^{-1}); this is referred to as effective resolution.

Predicted values were determined by the following equation[27]:

$$\hat{\mathrm{y}} = \mathrm{X} \cdot \mathrm{b} \tag{8}$$

where **X** is the spectral matrix and **b** is the PLS regression vector, which varies depending on the number of latent variables applied. It should be noted that the concentration data have been previously autoscaled; thus, predicted values need to be rescaled using the mean and standard deviation of the measured concentrations before the accuracy is determined via Equation 4.

Selectivity is a dimensionless univariate measure of the portion of instrumental signal that is not lost because of spectral overlap—in other words, the quantity of signal unaffected by interfering factors—and is restricted to a value between 0 and 1.[10] This statistic is calculated for each sample, using the following equation[10,23]:

$$\hat{SEL}_i = \frac{\left\| \hat{\mathrm{NAS}}_i \right\|}{\|\mathrm{x}_i\|} \tag{9}$$

where $\hat{\mathrm{NAS}}_i$ and x_i are the NAS vector and the original spectrum for the **i**th sample. The magnitude of the selectivity parameter is directly dependent on the degree of spectral interference associated with the particular analyte under consideration. Selectivity is reported in this article as the mean of the selectivity values for all samples under consideration.

The S/N ratio, one of the most important metrics for general comparison of methods, is calculated as follows[10]:

$$\hat{S/N}_i = \frac{\hat{NAS}_i}{\delta r} \tag{10}$$

where $\hat{NAS}_i$ is the scalar representation of the NAS vector. Linear regression was performed between the measured concentration and the univariate NAS values in order to estimate scale (a_1) and offset (a_o) coefficients to transform the NAS value into units of concentration. This enables the S/N ratio to be a dimensionless statistic for this 4-constituent mixture design:

$$\hat{S/N}_i = \frac{a_1 \cdot \left(\hat{NAS}_i \right) + a_o}{\delta r} \tag{10a}$$

The S/N ratio is reported in this article as the mean of the S/N values for all samples under consideration.

Given the wide range of concentrations present within the design, limit of detection (LOD) is a practical figure of merit. LOD can be computed as follows[28]:

$$LOD = \frac{k_D \sigma}{m} \tag{11}$$

where k_D is the statistical confidence factor (here, $k_D = 3$), m is the slope of a univariate classical least squares fit of the predicted and reference data, and σ is defined as δr in Equation 7. Since the predicted vs measured plot was not significantly different from unity, a value of 1.0 was assumed for m in all cases.

Precision Statistics

Precision figures of repeatability and intermediate precision were determined in accordance with ICH guidelines[26] and were reported as the standard deviation of the predicted concentration values (Equation 8) for repeat measurements. Repeatability and intermediate precision values were established using the randomly chosen design point comprising 20% theophylline, 20% lactose, 0% MCC, and 60% starch, compacted at a force of 167.6 MPa. Repeatability was determined without repositioning of the tablet between successive scans, as well as by removing and subsequently recentering the compact before acquiring the next measurement. Six scans for each type of repeatability test were collected one after the other on the same day. Intermediate precision was determined by scanning the tablet once a day for 6 consecutive days.

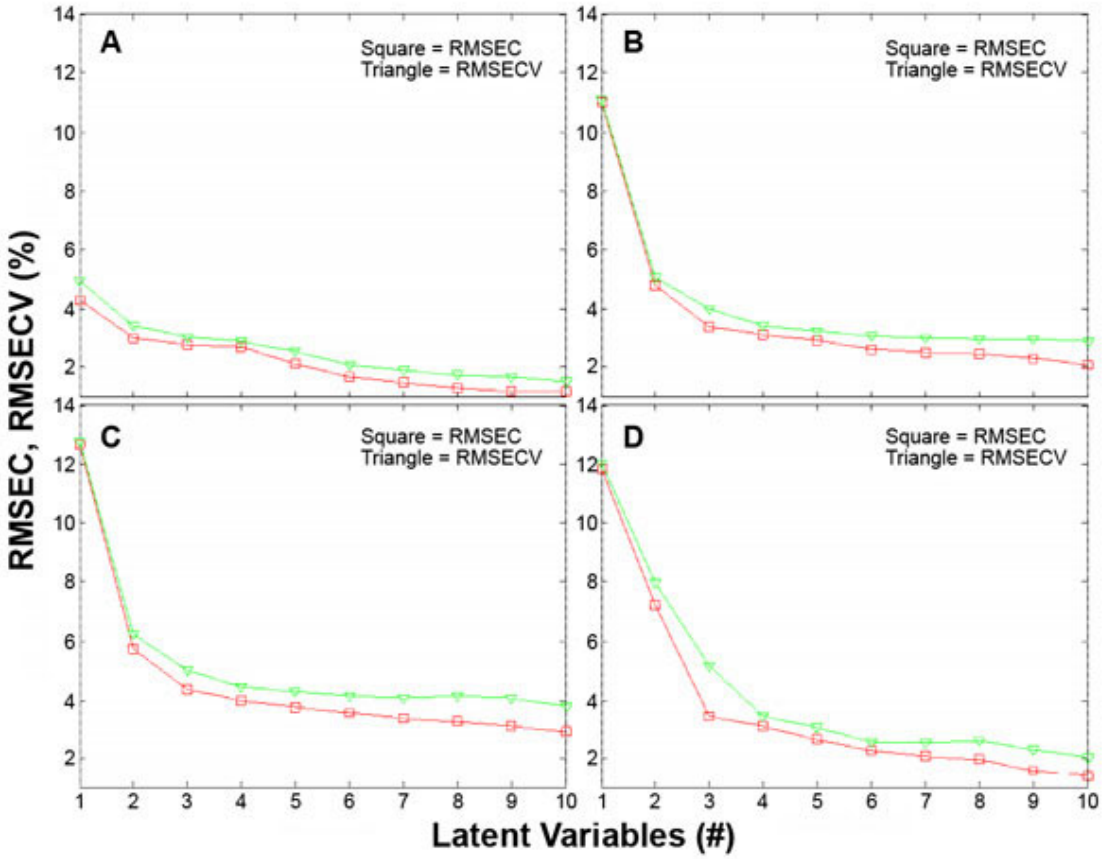

Figure 1. Plot of RMSEC (squares) and RMSECV (triangles) vs the number of partial least squares factors used to model near-infrared data for theophylline (a), lactose (b), microcrystalline cellulose (c), and starch (d), respectively. RMSEC indicates root-mean-standard error for calibration; RMSECV, root-mean-standard error for cross-validation.

Results and Discussion

NIR Analysis

The RMSEC and RMSECV values were plotted against the number of latent variables selected for PLS modeling (Figure 1). Savitsky-Golay first-derivative preprocessing[29] (11-point smoothing and second-order polynomial fit) was chosen based on the minimization of RMSECV. It was independently determined that to adequately model the NIR absorbance data, theophylline required 3 latent variables, while 4 were required for lactose, MCC, and starch. Model rank was chosen as the point where a rapid decline in the incremental variance captured was observed, with an awareness of the expected feasible limit of dimensionality based on the factors varying within the design. Without derivative preprocessing, an additional latent variable would have been required to compensate for the variation in compact density. It is speculated that the derivative preprocessing most effectively suppressed the physical effect of compaction, which has been shown to have a significant effect on the spectral baseline.[1,2]

Figure 2 displays the regression vectors in addition to the pure component spectra theophylline, lactose, MCC, and starch; note that the pure component spectra and the PLS regression vectors were scaled to facilitate viewing. To gather the pure component scans, powder for each constituent was placed in a glass scintillation vial and spectra were acquired directly through the bottom of the glass; each pure component spectrum represents a mean of 3 scans. As expected, each regression vector is highly correlated with its associated pure component scan. The goodness-of-fit seen in the predicted vs measured concentration values for the 4 constituents demonstrates the linearity of the PLS models implemented (Figure 3a).

As an aside, it is important to observe that all predicted values herein are independent of any NAS calculations performed. Certain NAS techniques allow the determination of predicted concentrations using the following equation[10,18]:

$$\hat{y}_i = \frac{\|\hat{\mathrm{NAS}}_i\|}{\|\hat{\mathrm{SEN}}_i\|} \qquad (12)$$

where $\hat{NAS}_i$, $\hat{SEN}_i$, and y_i are as defined previously. Considering Equation 5 for the calculation of the sensitivity vector, the NAS method employed within this article forbids the use of Equation 12 because it forces $\hat{y}_i = y_i$.

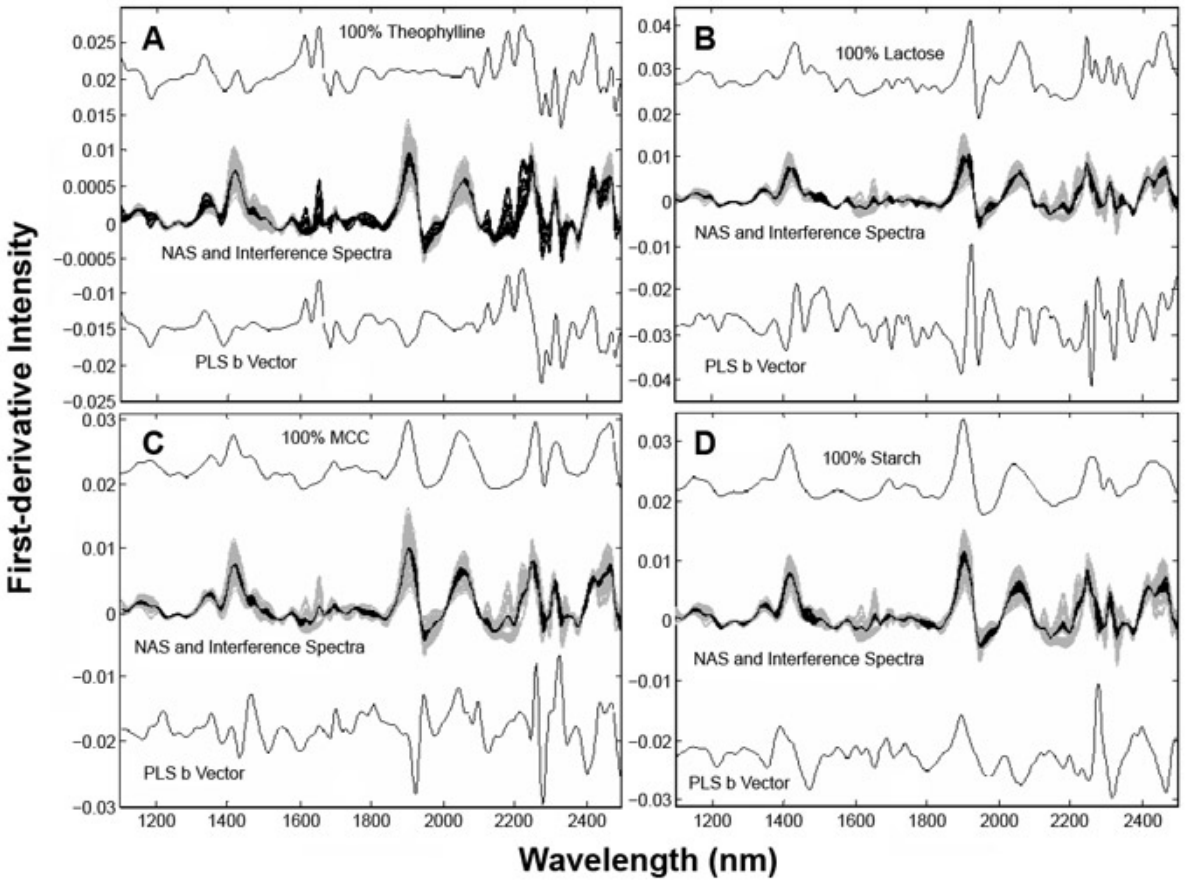

Figure 2. Near-infrared pure component spectra (upper solid lines), PLS regression vectors (lower solid lines), NAS (black), and interference (gray) vectors for each calibration sample, for theophylline (a), lactose (b), and MCC (c), and starch (d), respectively. NAS indicates net analyte signal; PLS, partial least squares; MCC, microcrystalline cellulose.

Following estimation of the PLS regression vectors for each constituent, the portion of the NIR signal related to only the component being analyzed was determined for all calibration samples (Equation 2). Figure 2 exhibits the NAS vector along with the corresponding interference spectrum for each calibration sample for theophylline, lactose, MCC, and starch. This unique plotting scheme directly illustrates the contrast between analyte and interference signals. Furthermore, this

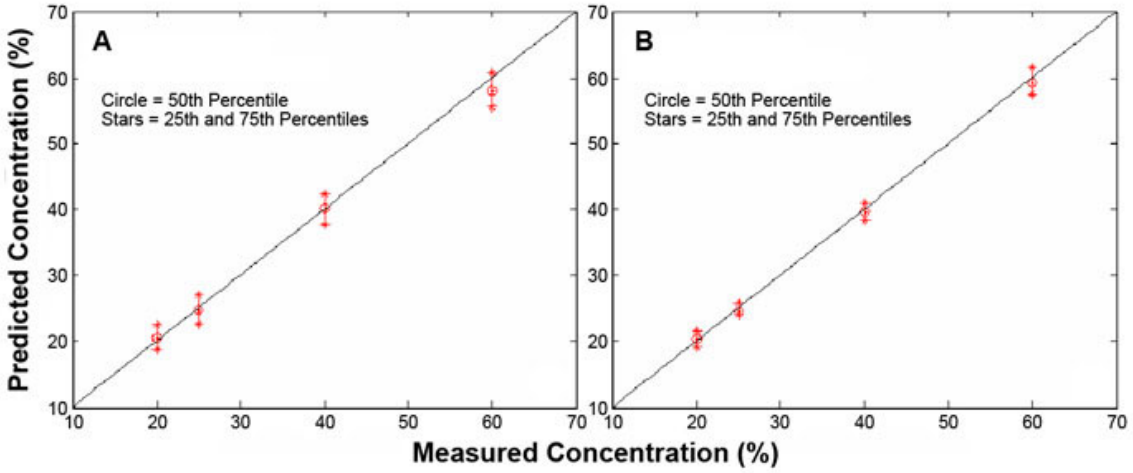

Figure 3. Predicted vs measured concentration plot for near-infrared (a) and Raman (b) data. Circles represent the 50th percentile, while the upper and lower stars represent the 25th and 75th percentiles, respectively. The unity line is shown in black.

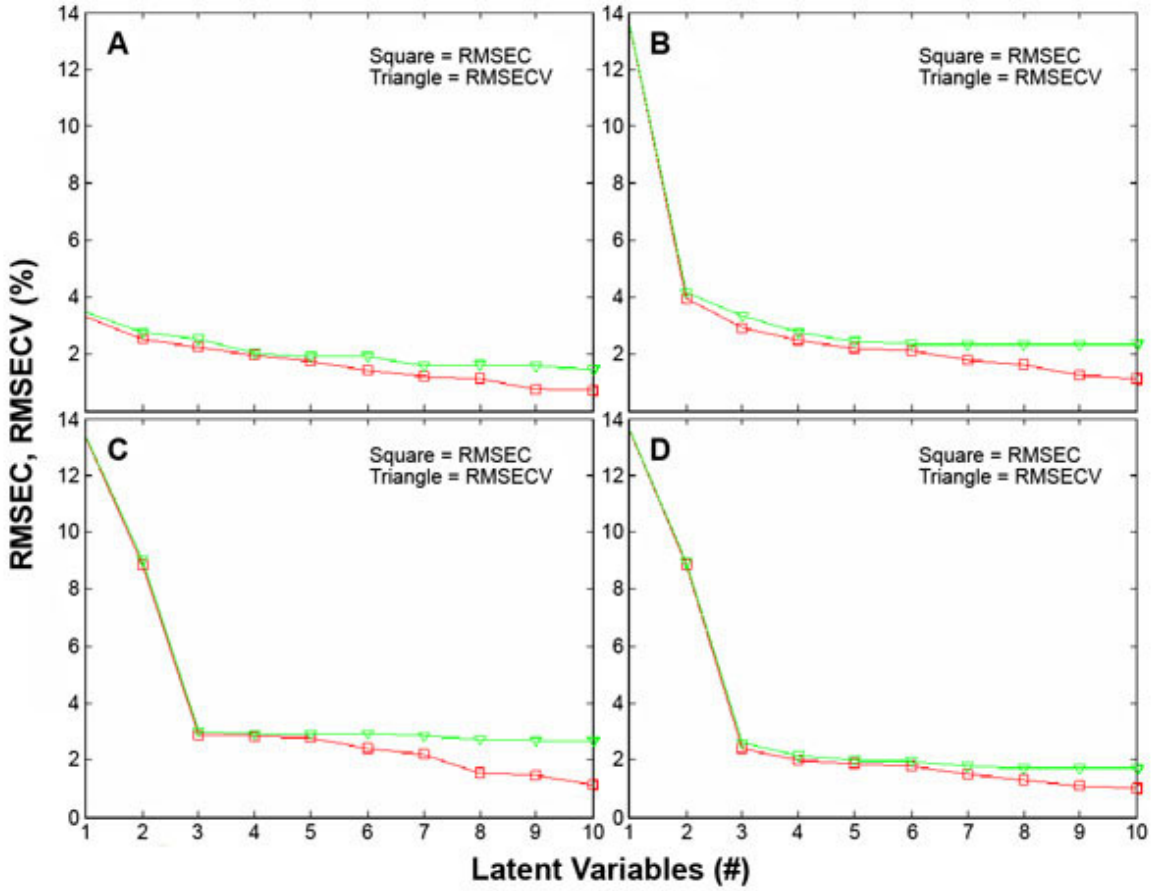

Figure 4. Plot of RMSEC (squares) and RMSECV (triangles) vs the number of partial least squares factors used to model Raman data for theophylline (a), lactose (b), microcrystalline cellulose (c), and starch (d), respectively. RMSEC indicates root-mean-standard error for calibration; RMSECV, root-mean-standard error for cross-validation.

graphically illustrates the ability of multivariate calibration to achieve selectivity. For example, in Figure 2a, a great deal of spectral variance can be observed around 1500 nm. However, this variation is not attributed to the presence of theophylline; rather, it is the result of the interfering components. This effect is evident in the relatively small range in intensity for the NAS spectra in comparison to the larger intensity range for the interference spectra. Conversely, much of the spectral variation around 1650 nm is due to the variance in theophylline concentration. Similar phenomena can be seen for the other 3 components (Figure 2).

Raman Analysis

The RMSEC and RMSECV were plotted against the number of PLS latent variables used in the model (Figure 4). Savitsky-Golay first-derivative preprocessing[29] (33-point smoothing and second-order polynomial fit) was also selected based on minimization of RMSECV. Presently, there are no known published reports identifying any consistent correlation between variation in Raman spectra and tablet hardness, implying that Raman spectra are insensitive to compact hardness variation. The data collected in this study are in agreement with this conclusion; no discernible pattern was observed relating Raman intensity and tablet compaction force (Figure 5). Hence, the role of derivative preprocessing was apparently not to mitigate any spectral effect of hardness variation; rather, it suppressed the baseline effect present in the spectra. Theophylline and lactose each required 4 latent variables, while 3 were required for MCC and starch.

Figure 6 displays the scaled regression vectors in addition to the scaled pure component spectra for theophylline, lactose, MCC, and starch. Raman pure component scans were gathered in the same manner as the NIR. Again, it was anticipated that the PLS regression vectors would include information pertaining to the component, which was confirmed by the similarities between the pure component scans and the regression vectors for all 4 constituents. Less dispersion in predicted values was observed around each concentration level, which is in agreement with the higher-calibration R^2 statistics for the Raman calibration (Figure 3).

Following the construction of the PLS regression vectors for each constituent, the $\hat{\mathrm{NAS}}_i$ was determined for each

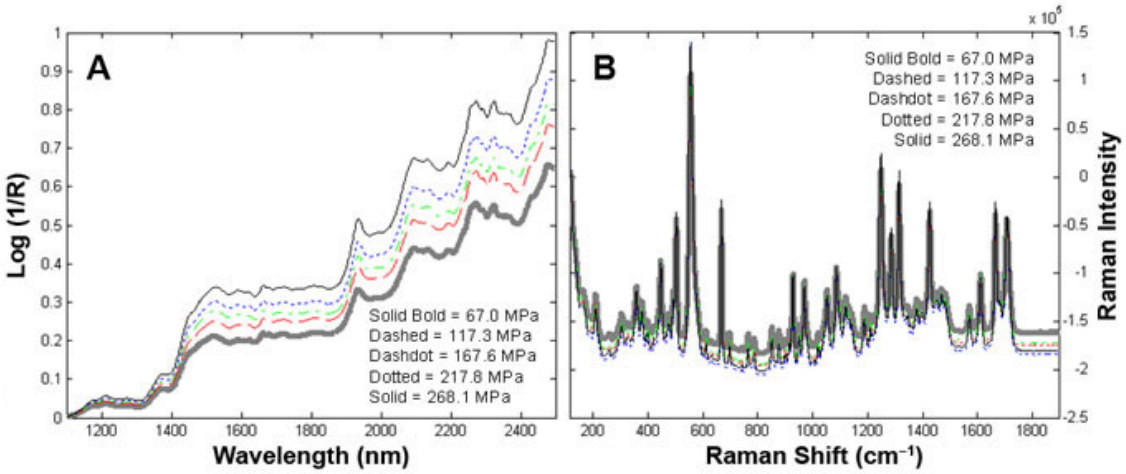

Figure 5. Near-infrared (a) and Raman (b) spectra of the same design point (40% theophylline, 40% lactose, 20% microcrystalline cellulose, and 0% starch) compacted at 67.0, 117.3, 167.6, 217.8, and 268.1 MPa. For each spectrum, the value for the first variable was subtracted to facilitate viewing.

calibration sample (Equation 2). Figure 6 depicts the NAS vector and the interference spectrum for each calibration sample for theophylline, lactose, MCC, and starch. Although 205.5 to 1895.7 cm^{-1} was used during calibration, a reduced range was plotted to highlight the contrast between NAS and interference spectra. As was observed for NIR, the patterns demonstrate the selectivity of multivariate calibration for each component.

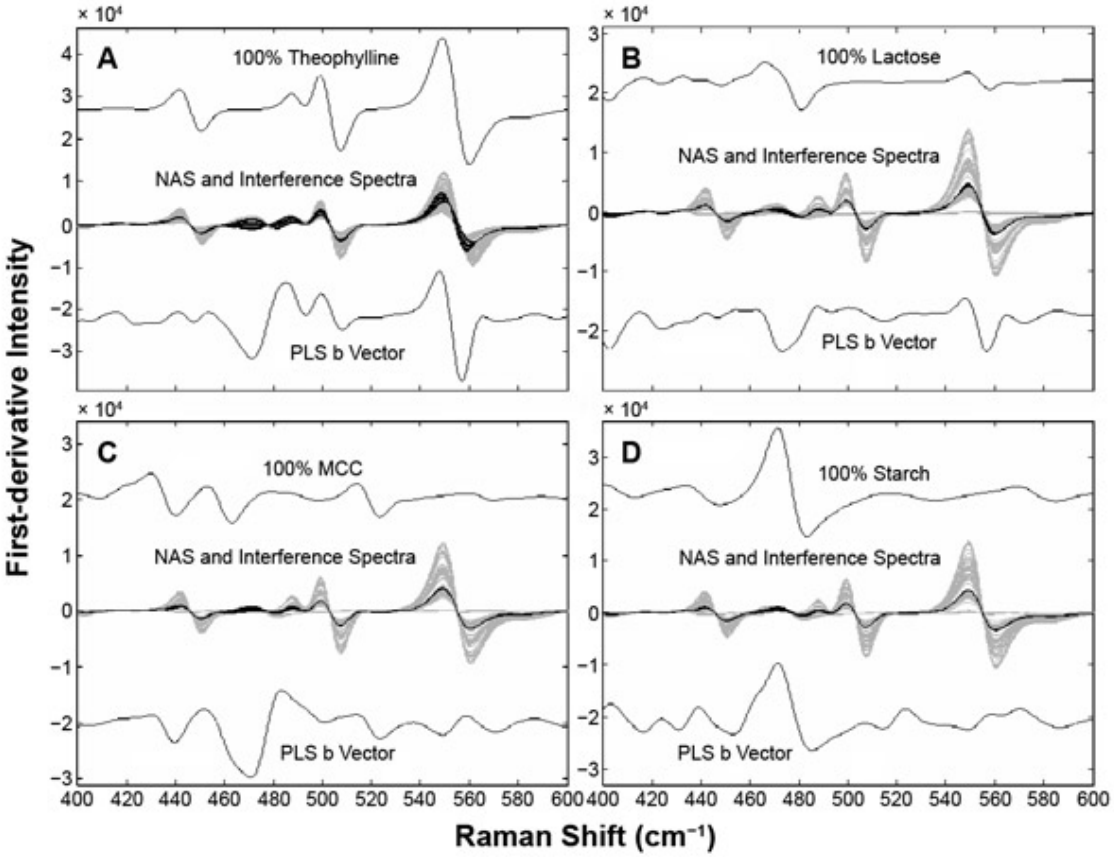

Figure 6. Raman pure component spectra (upper solid lines), PLS regression vectors (lower solid lines), net analyte signal (black), and interference (gray) vectors for each calibration sample, for theophylline (a), lactose (b), MCC (c), and starch (d), respectively. NAS indicates net analyte signal; PLS, partial least squares; MCC, microcrystalline cellulose.

Table 1. Calibration Statistics and Figures of Merit for NIR and Raman as Determined for Each Constituent*

Data Type		NIR				Raman			
Method		PLS							
Preprocessing		1st Derivative (11, 2, 1)				1st Derivative (33, 2, 1)			
Latent Variables		3	4	4	4	4	4	3	3
Component		Theophylline	Lactose	MCC	Starch	Theophylline	Lactose	MCC	Starch
Accuracy	R^2—cal	0.962	0.951	0.919	0.952	0.981	0.969	0.958	0.972
	R^2—CV	0.962	0.942	0.902	0.941	0.979	0.962	0.955	0.966
	RMSEC (%)	2.7	3.1	4.0	3.1	1.9	2.5	2.9	2.4
	RMSECV (%)	2.8	3.4	4.4	3.4	2.0	2.8	3.0	2.6
Precision	Repeatability—without repositioning (%)	0.01	0.16	0.16	0.02	0.28	0.27	0.09	0.04
	Repeatability—with repositioning (%)	0.07	0.10	0.36	0.46	0.45	0.46	0.45	0.27
	Intermediate (%)	0.11	0.16	0.52	0.66	0.35	0.66	0.36	0.26
Sensitivity (instrument intensity/%)		0.02	0.01	0.01	0.01	12 768.61	5124.58	3732.26	3265.58
Analytical sensitivity (1/%)		126.36	82.47	37.09	31.18	15.13	17.31	11.85	10.53
Effective resolution (%)		0.01	0.01	0.03	0.03	0.07	0.06	0.08	0.09
Selectivity (unitless)		0.59	0.33	0.24	0.27	0.37	0.24	0.18	0.16
Signal-to-noise (unitless)		282.40	189.88	87.23	72.14	34.99	40.69	26.80	23.98
Limit of detection (%)		0.33	0.50	1.08	1.31	2.70	2.32	3.53	3.94

*NIR indicates near-infrared; PLS, partial least squares; MCC, microcrystalline cellulose; cal, calibration; CV, cross-validation; RMSEC, root-mean-standard error for calibration; RMSECV, root-mean-standard error for cross-validation. Preprocessing parameters (11,2,1) correspond to window width, polynomial order, and derivative order, respectively.

FOM Comparison

Table 1 displays the calibration statistics describing the performance of the NIR and Raman spectrometers under investigation for this 4-component solid dosage system. With regard to accuracy, the Raman calibration appears to have an advantage over NIR, as reflected by the lower RMSE as well as the superior coefficient of determination. Among the 4 components, theophylline was modeled the most accurately by both NIR and Raman spectroscopy. This result is most likely attributable to higher sensitivity and selectivity relative to the other components. This topic will be addressed in greater detail in the following paragraphs.

The ICH guidance on method validation for repeatability[26] provides a protocol for partitioning sources of variance (ie, instrumental noise, sample positioning error, instrument drift). In particular, it was noted that repositioning inconsistency was a large contributor to total error for both the NIR and the Raman calibrations. The Raman analyzer sampled approximately 6 mm of the surface (roughly 46% of the tablet surface); in contrast, the off-line NIR instrument sampled nearly the entire compact face. Furthermore, NIR has a greater depth of sampling relative to Raman spectroscopy. Therefore, sample heterogeneity and/or sample presentation effects have a greater influence on Raman spectroscopy's precision. The impact of sampling heterogeneity is reduced as the number of samples analyzed per time period is increased, as a result of averaging. For example, if these methods were implemented in an at-line environment, the difference in precision between Raman spectroscopy and NIR would be expected to decrease. Additionally, more equitable comparisons between these 2 analyzers could be made by implementing a modified sample holder capable of precisely positioning 13 mm tablets, by analyzing tablets of a diameter similar to the diameter of those used for the Raman spectroscopy, or by averaging multiple locations on either side of the tablet.

In some cases, error statistics were inconsistent with expected trends, as shown in Table 1. For example, intermediate precision values calculated using the Raman data were actually lower than repeatability figures for both MCC and starch. This is unexpected, as intermediate precision includes the additional factor of day-to-day instrument drift. This may be indicative of an incomplete estimate of the variance associated with sample repositioning.

While the accuracy and precision data provide a feasible means of comparing these 2 spectrometers, the power of the evaluation can be enhanced by determining additional calibration FOM. Sensitivity for both analytical devices was the largest for theophylline; an increase in its concentration resulted in the greatest response in instrumental intensity. The relative magnitude of peaks in unscaled pure component spectra (not shown) illustrates this effect. It is important to note that the sensitivity values of individual constituents for the 2 instruments cannot be compared.

Analytical sensitivity is used to compare sensitivity across different measurement technologies. This normalized statistic quantifies sensitivity with respect to analytical precision. Although both devices exhibited the greatest sensitivity for theophylline, the NIR device was more sensitive to all 4 constituents (in terms of analytical sensitivity). Error of repositioning had a direct effect on σ in Equation 7, which in turn inflated the denominator, thereby reducing the analytical sensitivity of Raman spectroscopy. The constituent ordering for highest to lowest sensitivity is not identical for the 2 instruments. This emphasizes the importance of pairing the instrument to the analytical task. The effective resolution results reinforce the results reported for sensitivity. Despite the apparent similarity, this statistic should be considered with respect to quality action limits.

Selectivity is important only when adequate sensitivity is available. A lack of selectivity has the effect of suppressing sensitivity. If adequate sensitivity is not available, attempts to improve selectivity are futile. Theophylline, which exhibited high relative sensitivity, also exhibited superior selectivity, which is directly attributable to its inherent pure component orthogonality. In contrast, the collinearity among lactose, MCC, and starch (all carbohydrates) reduced selectivity. For Raman spectroscopy in particular, the sensitivity (Equation 6) and selectivity are lower for these components. This is manifested in the performance-related FOM (LOD, analytical sensitivity, effective resolution, and S/N ratio). An example of the enhanced interference between these 3 components can be seen in the NIR at ~1500 nm, where the NAS signal is quite large for all 3 (Figure 2).

From the results discussed thus far, the resulting S/N ratio analysis should be straightforward. The NIR calibration included the more usable signal in relation to obstructive noise. In some cases, researchers assume that the coefficient of determination is directly predictive of the S/N ratio.[30] It is for this reason that technology selection criteria are often based upon R^2 and RMSE, since these statistics are frequently generated during calibration, requiring no additional calculations. The results of this work contradict these assumptions. While Raman spectroscopy outperformed NIR in terms of linearity and accuracy, the S/N ratio for NIR measurements was greater. This occurred because error statistics (R^2 or RMSE) are heavily influenced by the experimental design, while S/N ratio is inherent to the method. Further studies are planned in which calibrations will be optimized according to S/N ratio (as opposed to the traditional method of RMSE); the impetus is to address the aforementioned precision issues while simultaneously enhancing sensitivity, selectivity, and LOD.

It is interesting to note that despite the lower relative precision of the Raman measurements, which deflates several of the FOM, no negative effect on the ability of the SIMPLS algorithm to resolve the covariance structure was observed. This is because inverse least squares regression is less affected by precision than by sample leverage in the estimation of the true solution. This supports the notion that calibration quality is not sufficient to fully describe method performance.

CONCLUSION

This study demonstrates that multivariate FOM (determined from NAS theory) can be used to compare calibrations constructed from spectroscopic data collected using 2 analytical instruments detecting different physical phenomena. The observed calibration performance statistics demonstrate that NIR and Raman spectroscopy are both suitable for the quantitative determination of chemical components within this tablet matrix. Beyond error statistics, multivariate FOM provide a clearer assessment of the factors that limit method performance. For all of these reasons, FOM should take a more prominent place among chemometric techniques used in pharmaceutical analytical method development and validation.

ACKNOWLEDGMENTS

Duquesne University Graduate School of Pharmaceutical Sciences in accordance with the Duquesne University Center for Pharmaceutical Science would like to thank Kaiser Optical Systems, Inc (Ann Arbor, MI, www.kosi.com) for its generosity in loaning the *P^{h}AT System* analyzer. The authors would also like to express their gratitude to Ryanne Palermo for her efforts in editing this document.

REFERENCES

1. Bugay DE, Brittain HG. Raman spectroscopy. In: Brittain H, ed. *Spectroscopy of Pharmaceutical Solids.* vol. 160. New York, NY: Taylor & Francis; 2006:271–312.

2. Cogdill RP, Drennen JK. Near-infrared spectroscopy. In: Brittain H, ed. *Spectroscopy of Pharmaceutical Solids.* vol. 160. New York, NY: Taylor & Francis; 2006:313–412.

3. Afseth NK, Segtnan VH, Marquardt BJ, Wold JP. Raman and near-infrared spectroscopy for quantification of fat composition in a complex food model system. *Appl Spectrosc.* 2005;59:1324–1332.

4. Furukawa T, Masahiro W, Siesler HW, Ozaki Y. Discrimination of various poly(propylene) copolymers and prediction of their ethylene content by near-infrared and Raman spectroscopy in combination with chemometric methods. *J Appl Polym Sci.* 2003;87:616–625.

5. Nordon A, Meunier C, McGill CA, Littlejohn D. Comparison of calibration methods for the monitoring of a fluorobenzene batch reaction using low-field 19F NMR, 1H NMR, NIR, and Raman spectrometries. *Appl Spectrosc.* 2002;56:515–520.

6. Nordon A, Mills A, Burn RT, Cusick FM, Littlejohn D. Comparison of non-invasive NIR and Raman spectrometries for determination of alcohol content of spirits. *Anal Chim Acta.* 2005;548:148–158.

7. Qiao Y, van Kempen TATG. Comparison of Raman, mid, and near infrared spectroscopy for predicting the amino acid content in animal meals. *J Anim Sci.* 2004;82:2596–2600.

8. Chung H, Ku M-S. Comparison of near-infrared, infrared, and Raman spectroscopy for the analysis of heavily petroleum products. *Appl Spectrosc.* 2000;54:239–245.

9. Ku M-S, Chung H. Comparison of near-infrared and Raman spectroscopy for the determination of chemical and physical properties of naphtha. *Appl Spectrosc.* 1999;53:557–564.

10. Lorber A, Faber K, Kowalski BR. Net analyte signal calculation in multivariate calibration. *Anal Chem.* 1997;69:1620–1626.

11. Lorber A. Error propagation and figures of merit for quantification by solving matrix equations. *Anal Chem.* 1986;58:1167–1172.

12. Brown CD. Discordance between net analyte signal theory and practical multivariate calibration. *Anal Chem.* 2004;76:4364–4373.

13. Morgan DR. Spectral absorption pattern detection and estimation, I: analytical techniques. *Appl Spectrosc.* 1977;31:404–415.

14. Olivieri AC, Faber NM, Ferre J, Boque R, Kalivas JH, Mark H. Uncertainty estimation and figures of merit for multivariate calibration. *Pure Appl Chem.* 2006;78:633–661.

15. Haaland DM. Classical versus inverse least squares methods in quantitative spectral analyses. *Spectroscopy.* 1987;2:56–57.

16. Boelens HF, Kok WT, de Noord OE, Smilde AK. Performance optimization of spectroscopic process analyzers. *Anal Chem.* 2004;76:2656–2663.

17. Xu L, Schechter I. A calibration method free of optimum factor number selection for automated multivariate analysis. Experimental and theoretical study. *Anal Chem.* 1997;69:3722–3730.

18. Ferre J, Brown SD, Rius FX. Improved calculation of the net analyte signal in inverse multivariate calibration. *J Chemom.* 2001;15:537–553.

19. Bro R, Andersen CM. Theory of net analyte signal vectors in inverse regression. *J Chemom.* 2003;17:646–652.

20. Martens H, Naes T. *Multivariate Calibration*. New York, NY: John Wiley and Sons; 1989.

21. Xu L, Schechter I. Wavelength selection for simultaneous spectroscopic analysis. Experimental and theoretical study. *Anal Chem.* 1996;68:2392–2400.

22. Goicoechea HC, Olivieri AC. Chemometric assisted simultaneous spectrophotometric determination of four-component nasal solutions with a reduced number of calibration samples. *Anal Chem Acta.* 2002;453:289–300.

23. Braga JWB, Poppi RJ. Figures of merit for the determination of the polymorphic purity of carbamazepine by infrared spectroscopy and multivariate calibration. *J Pharm Sci.* 2004;93:2124–2134.

24. Geladi P, Kowalski BR. Partial least-squares regression: a tutorial. *Anal Chem Acta.* 1986;185:1–17.

25. De Jong S. SIMPLS: an alternative approach to partial least squares regression. *Chemom Intell Lab Syst.* 1993;18:251–263.

26. ICH. ICH harmonised tripartite guideline: validation of analytical procedures: text and methodology. *Fed Regist.* 1997;62:27463–27467.

27. Haaland DM, Thomas EV. Partial least-squares methods for spectral analyses, 1: relation to other quantitative calibration methods and the extraction of qualitative information. *Anal Chem.* 1988;60: 1193–1202.

28. Long GL, Winefordner JD. Limit of detection: a closer look at the IUPAC definition. *Anal Chem.* 1983;55:712A–724A.

29. Savitzky A, Golay MJE. Smoothing and differentiation of data by simplified least squares procedures. *Anal Chem.* 1964;36: 1627–1639.

30. Marbach R. On Wiener filtering and the physics behind statistical modeling. *J Biomed Opt.* 2002;7:130–147.