Scattering Orthogonalization of Near-Infrared Spectra for Analysis of Pharmaceutical Tablets

ARTICLE *in* ANALYTICAL CHEMISTRY · FEBRUARY 2009

Impact Factor: 5.64 · DOI: 10.1021/ac802105v · Source: PubMed

2 AUTHORS:

Zhenqi Shi
Eli Lilly
27 PUBLICATIONS **351** CITATIONS
SEE PROFILE

Carl A Anderson
Duquesne University
48 PUBLICATIONS **1,181** CITATIONS
SEE PROFILE

Article

Scattering Orthogonalization of Near-Infrared Spectra for Analysis of Pharmaceutical Tablets

Zhenqi Shi, and Carl A. Anderson

Anal. Chem., **2009**, 81 (4), 1389-1396• DOI: 10.1021/ac802105v • Publication Date (Web): 22 January 2009

Anal. Chem. **2009,** *81,* 1389–1396

Scattering Orthogonalization of Near-Infrared Spectra for Analysis of Pharmaceutical Tablets

Zhenqi Shi and Carl A. Anderson*

Graduate school of Pharmaceutical Sciences, Duquesne University, Pittsburgh, Pennsylvania 15282, and Duquesne University Center for Pharmaceutical Technology, Pittsburgh, Pennsylvania 15282

The paper explores scattering orthogonalization as a preprocessing technique to reduce physical interference and maintain chemical information in near-infrared (NIR) spectra of pharmaceutical tablets. Samples used in this study were tablets compressed at five compression forces; they were composed of theophylline, lactose, and microcrystalline cellulose (PH200). The NIR spectra were orthogonalized against the reduced scattering coefficients (representative of physical interference of scattering), and concentrations of all constituents were predicted. The robustness of predictions was compared to the widely employed standard normal variate (SNV) for the specificity of removing interference representative of physical parameter (such as tablet density). Group-wise cross-validation (groups were based upon similar chemical composition) and prediction demonstrated the enhanced robustness on prediction of chemical information via scattering orthogonalization in comparison to SNV. When compared to the SNV, scattering orthogonalization demonstrated an improved capacity to reduce physical interference while maintaining spectral variance attributable to chemical information. The improved capacity is expected to be useful for spectroscopy-based multivariate model calibration and continuous model update.

The number of near-infrared (NIR) spectroscopic applications in pharmaceutical analysis has grown significantly in the past decade.[1] Much of its appeal is due to the fact that little to no sample preparation is required, and non-invasive detection makes NIR spectroscopy a suitable tool for online process monitoring. NIR spectroscopy contains both chemical and physical information which are typically extracted using multivariate modeling techniques; these models can then be used to monitor process variations during routine pharmaceutical unit operations. In most cases, data pretreatments are performed to suppress physical interference prior to calculation of multivariate models for predicting chemical information. Potential sources of physical interference include particle size of a powder mixture and density of a pharmaceutical compact, and so forth. Pretreatments are typically employed to reduce error induced by physical interference and enhance signal-to-noise in multivariate models designed to predict chemical information.

Since physical interference is typically observed as a baseline shift or slope in an NIR absorbance spectrum, the most widely used data pretreatments are additive and multiplicative scattering corrections, for example, standard normal variate (SNV) and multiplicative scattering correction (MSC). It is understood that scattering is both wavelength and absorption (constituent) dependent, and thus, single slope or intercept correction is often insufficient to remove the effects of physical interference.[2−4] It was recently reported that traditional scattering correction, such as MSC, removed chemical information from NIR spectra as a result of reducing baseline shift (attributed to scattering caused by physical interference) when NIR transmittance data (log $1/T$) were calibrated against gluten concentration.[4] Thus, an improved data pretreatment method is desirable, ideally one which selectively removes from spectra information attributable only to physical interference (i.e., noise for a chemical calibration).

There are two fundamental events that occur as NIR light impinges upon a sample, absorption and scattering.[5,6] Absorption reduces the intensity of photons of specific energy because of the altered molecular dipole of a bond. The parameter used to describe absorption is the absorption coefficient, μ_a, which is defined as the probability of absorption per unit path-length. Scattering is caused by variations of refractive index at an interface; therefore, physical parameters such as sample density and porosity can have a dramatic effect on scattering events. Examples of these in pharmaceutical products (such as tablets) include air−solid and solid−solid interfaces. Since diffuse reflectance is an elastic scattering event, there is no energy change or photon intensity loss as a result of scattering. Using the diffusion approximation to the radiative transfer equation, scattering can be described by the reduced scattering coefficient, μ_s', which is the probability for scattering per unit path-length. Absorption and reduced scattering coefficients together are referred to as the optical properties or optical coefficients of a material. Although absorption and scattering can be understood as independent interactions between photons and

* To whom correspondence should be addressed. E-mail: andersonca@duq.edu. Phone: 412-396-1102. Fax: 412-396-4660.

(1) Ciurczak, E. W.; Drennen, J. K. In *Handbook of Near-Infrared Analysis*; Burns, D. A.; Ciurczak, E. W.; Eds.; Marcel Dekker: New York, 2001; pp 609.

(2) Burger, T.; Kuhn, J.; Caps, R.; Fricke, J. *Appl. Spectrosc.* **1997**, *51*, 309–317.

(3) Burger, T.; Ploss, H. J.; Ebel, S.; Fricke, J. *Appl. Spectrosc.* **1997**, *51*, 1323–1329.

(4) Martens, H.; Nielsen, J. P.; Engelsen, S. B. *Anal. Chem.* **2003**, *75*, 394–404.

(5) Groenhuis, R. A. J.; Ferwerda, H. A.; Ten Bosch, J. J. *Appl. Opt.* **1983**, *22*, 2456–2462.

(6) Keijzer, M.; Star, W. M.; Storchi, R. M. *Appl. Opt.* **1988**, *27*, 1820–1824.

10.1021/ac802105v CCC: $40.75
Published on Web 01/22/2009

sample, optical coefficients are affected by variation of both physical and chemical properties of a sample.[5,6] It is well established that the reduced scattering coefficient is wavelength-dependent and typically larger than the absorption coefficient.[2,3,7] Thus, it is expected that NIR spectra are affected more by scattering than absorption.

Since absorption and scattering are the basic interactions in NIR diffuse reflectance, separation of absorption and scattering followed by spectral pretreatment via scattering signals is proposed here as an alternative to perform scattering correction. Different instrumental techniques have been utilized to separate absorption and reduced scattering coefficient based on mathematical evaluation of the diffusion approximation to the radiative transfer equation, including spatially-,[8] frequency-[9] and time-resolved spectroscopy,[7] and integrating sphere based reflectance and transmittance measurements.[2,3] The extracted optical coefficients were further reported to correlate with physical and chemical information of pharmaceutical samples.[2,3,7,9] Time- and frequency-resolved spectroscopy are not commonly employed because of the complexity and high cost of optical setup.[7,9] Integrating sphere based optical measurements require samples to be thin enough for transmittance measurements, which is not always suitable for practical applications.[2,3] The work herein employed spatially resolved spectroscopy to detect radially diffused reflectance on pharmaceutical tablets via an NIR chemical imager. The method is a suitable approach to measure optical properties of pharmaceutical materials because of the relative simplicity of the optical setup and its applicability for a variety of samples. However, spatially resolved spectroscopy has been reported to be less accurate for the prediction of the optical coefficients when compared to time- and frequency-resolved spectroscopy.[10,11] The major reason is that spatially resolved spectroscopy does not provide direct measurements of photon path length as time- and frequency-resolved spectroscopy, and it only relies on detection of light intensity attenuation to describe the absorption and scattering behaviors. Despite any limitations, it is still a practical method of measuring the reduced scattering coefficient for scattering correction considering its purpose is to correct relative scattering variations at different wavelengths for different chemical constituents in pharmaceutical materials.

Scattering correction via the reduced scattering coefficient was reported previously for pharmaceutical samples; time-resolved spectroscopy was utilized to obtain reduced scattering coefficients followed by scattering correction on spectra of pharmaceutical compact samples.[7] The study only provided comparison between raw spectra and spectra after scattering correction. It would be more evident to show the comparison between spectra after scattering correction via reduced scattering coefficients and spectra after normal preprocessing routines (e.g., SNV). Additionally, no chemometric calculation procedure was clearly outlined regarding to scattering correction via the reduced scattering coefficient. Thus, the work herein utilized the well-known net analyte signal (NAS) equation to orthogonalize the spectra to a matrix spanned by the reduced scattering coefficient representative of physical interference. The detailed calculation procedure can be found in the section of Material and Methods.

The diffusion approximation to the radiative transfer equation was initially developed for spatially resolved spectroscopy to derive the optical coefficients as a means of quantitatively and qualitatively characterizing pathological conditions of human tissue.[8] The basic assumption of the diffusion approximation is the infinite (or semi-infinite) sample medium, which has infinitely wide boundaries or can be treated under the condition that it is much wider than the spatial extent of the photon distribution. When the discrete boundaries for pharmaceutical tablets are taken into account, the reported equation is not applicable. Thus, Monte Carlo simulation in combination with the diffusion approximation to the radiative transfer equation was utilized to generate radially diffused reflectance patterns after taking into account the effects of the boundaries, source distribution, and media geometry on the diffuse reflectance. The simulated radially diffused reflectance data were used to establish a multivariate calibration model with the simulated absorption and reduced scattering coefficients, which was further used to predict the optical coefficients of pharmaceutical tablets based on empirically measured radially diffused reflectance measurements.

This work employed a design of experiments based on tablets with both physical (different tablet density) and chemical (different chemical composition) variations. The reduced scattering coefficient was utilized as a vector representative of physical interference to be removed from single-point NIR absorbance spectra via orthogonalization. The orthogonalized absorbance spectra were used to calibrate for chemical composition. This method was compared to those treated with widely used scattering correction methods, such as SNV, for their ability to specifically remove physical interference. Thus, scattering orthogonalization was evaluated here as a suitable alternative for scattering correction of NIR absorbance spectra of pharmaceutical compacts, emphasizing the removal of physical interference while preserving the pertinent chemical information. The objectives of the study are as follows.

(1) Explore the potential capacities of a commercially available NIR chemical imaging system on measuring radially diffused reflectance on pharmaceutical samples.

(2) Establish a Monte Carlo simulation based partial least-squares (PLS) multivariate model to predict absorption and reduced scattering coefficients of pharmaceutical materials.

(3) Demonstrate the capacity of scattering orthogonalization as a preprocessing technique to remove physical interferences and maintain chemical information in NIR spectra collected from tablets.

MATERIAL AND METHODS

Tablet Production. A fully balanced, three-constituent mixture design composed of anhydrous theophylline (Lot No. 92577, Knoll AG, Ludwigshafen, Germany), Lactose 316 Fast Flo NF Monohydrate (Lot No. 8502113061, Hansen Laboratories, New Berlin, WI), and microcrystalline cellulose (MCC, Avicel PH 200, Lot No. M427C, FMC BioPolymer, Mechanicsburgh, PA) was generated.

(7) Abrahamsson, C.; Lowgren, A.; Stromdahl, B.; Svesson, T.; Andersson-Engels, S.; Johansson, J.; Folestad, S. *Appl. Spectrosc.* **2005**, *59*, 1381–1387.

(8) Farrell, T. J.; Patterson, M. S. *Med. Phys.* **1992**, *19*, 879–888.

(9) Torrance, S. E.; Sun, Z.; Sevick-Muraca, E. M. *J. Pharm. Sci.* **2004**, *93*, 1879–1889.

(10) Chance, B.; Leigh, J. S.; Miyake, H.; Smiths, D. S.; Nioka, S.; Greeenfield, R.; Finander, M.; Kaufmann, K.; Levy, W.; Young, M.; Cohen, P.; Yoshioka, H.; Boretsky, R. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 4971–4975.

(11) Sevick, E. M.; Chance, B.; Leigh, J.; Nioka, S.; Maris, M. *Anal. Biochem.* **1991**, *195*, 330–351.

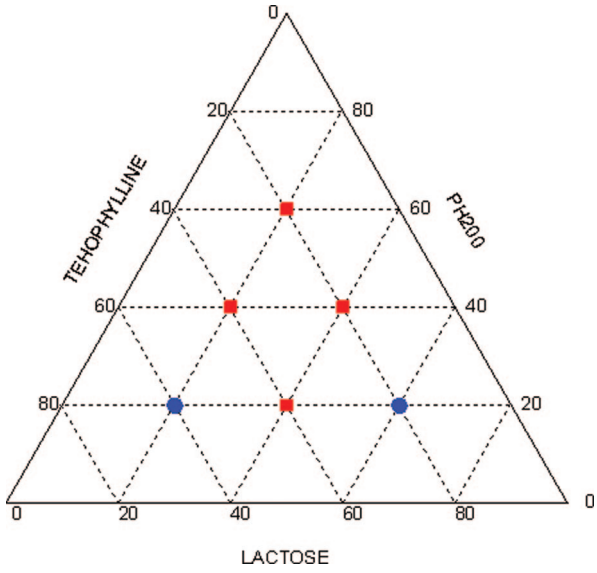

Figure 1. Tablet chemical composition. Squares were the calibration data points, and circles were the prediction data points.

A ternary diagram was utilized to generate six design points to incorporate different chemical compositions and remove any possibility of factor aliasing (Figure 1).

The tablet production process is described elsewhere.[12] Briefly, materials for each design point mixture were dispensed by weight according to the design matrix. The blending process was performed in 25 mL glass scintillation vials for 5 min cycle on a rotating Jar Mill (US Stoneware, East Palestine, OH, U.S.A.). The end-point was confirmed by an *ad hoc* PLS calibration model. Each blended mixture was compacted at 5 pressures (67.0, 117.3, 167.6, 217.8, 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 s was employed. In total, 30 compacts were produced with a target weight of 800 mg per tablet.

Single-Point NIR Reflectance Measurements. Single-point NIR reflectance measurements (expressed as $\log(1/R)$, or absorbance intensity) were acquired for all 30 tablets over the wavelength range of 1100−1620 nm at a 2 nm increments (FOSS NIR Systems 5000, FOSS NIRSystems, Inc., Laurel, MD). Thirty-two subscans were accumulated for each resultant sample spectrum. Among 30 absorbance spectra, 20 were used as calibration data set; the remaining 10 were used as prediction data set (Figure 1).

Radially Diffused Reflectance Measurements. Radially diffused reflectance images were collected under the optical settings described in Figure 2. Tablets were illuminated perpendicular to the light path of collection using an NIR chemical imaging system (MatrixNIR, Malvern Inc., MD) with 0.5× objective lens (field of view 1.72 cm × 2.15 cm). An integration time of 256 ms and a count of 16 co-adds were utilized throughout the image collection of all 30 tablets. The wavelength range was from 1050 to 1620 nm with a 5 nm interval. A tungsten halogen broadband light source (2.5 mW, Control Development Inc., IN) was attached to an optic fiber (400 μm, NA = 0.22) to transmit the light. Light emitted from the optic fiber was cast on the sample

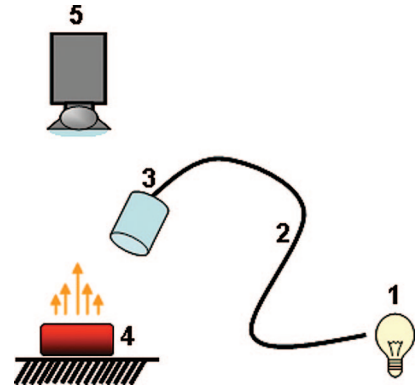

Figure 2. Optical setting of radially diffused reflectance measurements included (1) light source, (2) optic fiber, (3) focusing lens, (4) a sample tablet, and (5) imager.

through a focusing lens with an antireflective range (1050−1620 nm) to form a small spot of illumination (spot size = 2.4 mm). An incidence angle of 5−10° was used to avoid specular reflectance and focusing lens being situated inside the field of view of the imager. It has been reported that a small angle of incidence (5−10°) does not cause a significant difference in the spatially resolved reflectance.[13] The reflectance intensity inside the spot size within a range of radial distances from 0 mm to 1.2 mm (18 pixels under 0.5× objective) was used for data analysis because it provided radially diffused reflectance on tablet surfaces with an appropriate signal-to-noise ratio. Pixel reflectance intensities from the same radial ring area were summed and normalized by the corresponding radial ring area, which generated one radially diffused reflectance pattern per wavelength.

Monte Carlo Simulation Based PLS Calibration Model to Predict Optical Coefficients. A modified version of Monte Carlo simulation in multilayered tissues (MCML)[14] was used to account for the tablet geometry and mimic the pseudo-random source distribution used in the radially diffused reflectance measurement. A refractive index of 1.58 and an anisotropy factor of 0.9 were used for all simulation data (values used in the simulation are discussed in the section of Results and Discussion). The required photon number was calculated based on an empirical relationship.[13] A grid system of the same size as the measured radially diffused reflectance patterns was used to store the simulated radially diffused reflectance. Each simulated radially diffused reflectance was recorded in a row vector containing 18 columns as only radial distances from 0 mm to 1.2 mm (18 pixels under 0.5× objective) were used for empirical measurements. Simulated pixel reflectance intensities from the same radial ring area were processed the same as the measured pixel reflectance intensities.

A center component design was used to generate a design matrix containing different combinations of absorption and reduced scattering coefficients (Figure 3). Within the designed samples, 13 pairs (including 5 replicates at the center pair) of optical coefficients were used for calibration and 9 pairs were used for validation. Each pair of optical coefficients (both absorption and reduced scattering coefficient) was used as the input for Monte Carlo simulation, which generated a corresponding radially

(12) Short, S. M.; Cogdill, R. P.; Anderson, C. A. *AAPS Pharm. Sci. Tech.* **2007**, *8*, 96.

(13) Kienle, A.; Lilge, L.; Patterson, M. S.; Hibst, R.; Steiner, R.; Wilson, B. C. *Appl. Opt.* **1996**, *35*, 2304–2314.

(14) Wang, L. H.; Jacques, S. L.; Zheng, L. Q. *Comput. Methods Programs Biomed.* **1995**, *47*, 131–146.

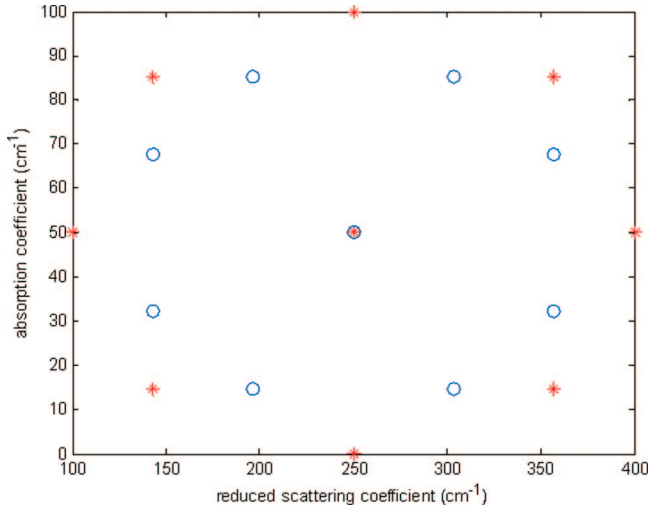

Figure 3. Center component design of absorption and reduced scattering coefficients. Asterisks were calibration data points. Circles were validation data points.

diffused reflectance pattern as the output of the simulation. Thus, the final calibration data was X (13 ×18) and Y (13 × 2). The final validation data was X (9 × 18) and Y (9 × 2). X here was simulated radially diffused reflectance; Y here contained simulated absorption and reduced scattering coefficients in the center component design.

Data pretreatments were performed prior to PLS calibration between simulated radially diffused reflectance and simulated optical coefficients. First, a reference standard (Intralipid 20%, Sigma-Aldrich Inc., MO) was used to address the difference of absolute reflectance intensity between simulated and empirical measurements. The radially diffused reflectance images of Intralipid were collected using the experimental setup shown in Figure 2. The known optical coefficients of Intralipid at 1064 nm were used as the input to perform Monte Carlo simulation.[15] The ratio between measured and simulated radially diffused reflectance of Intralipid at 1064 nm was further used to supplement the magnitude difference of simulated radially diffused reflectance in the center component design. Second, logarithmic conversion was performed on simulated radially diffused reflectance to facilitate PLS calibration by capturing the approximately linear relationship between the logarithmic radially diffused reflectance patterns and the optical coefficients.[16] Calibration and validation of the simulated patterns, and prediction of the measured patterns, were performed by the same data pretreatments as mentioned above.

Scattering Orthogonalization. After the Monte Carlo simulation based PLS calibration model was used to predict the absorption and reduced scattering coefficients of all 30 tablets from the measured radially diffused reflectance pattern at each wavelength, the mean reduced scattering coefficient of 20 tablets in the calibration data set was calculated. Because of differences in wavelength range and wavelength intervals between single-point NIR spectra and radially diffused reflectance, a cubic spline interpolation was used to supplement the mean reduced scattering coefficient for 1100−1620 nm at 2 nm intervals. Subsequently, the supplemented reduced scattering coefficient was utilized as a factor to be orthogonalized with respect to the single-point NIR absorbance spectra in the calibration and validation data set separately to remove the physical interference (i.e., tablet density). The following equations were utilized:

$$\mathbf{V} = \mathbf{X}_c[\mathbf{M}/\mathrm{norm}(\mathbf{M})] \quad (1)$$

where $\mathbf{X}_c$ is the mean centered single-point NIR spectra, $\mathbf{M}$ is the supplemented reduced scattering coefficient, and norm was used to calculate the length of $\mathbf{M}$ to normalize $\mathbf{M}$ within the range between 0 and 1.

$$\mathbf{P} = \mathbf{X}_c^{+}\mathbf{V} \quad (2)$$

$\mathbf{X}_c^{+}$ is the pseudo-inverse of the mean-centered single-point NIR spectra matrix

$$\mathbf{X}_{c\_orth} = \mathbf{X}_c(\mathbf{I} - \mathbf{P}\mathbf{P}^{T}) \quad (3)$$

$\mathbf{X}_{c\_orth}$ is the spectra after orthogonalizing to the reduced scattering coefficient, $\mathbf{I}$ is the identity matrix, and the superscript T indicates a transpose in the matrix calculation.

Comparison of Calibration Models to Predict Chemical Information of Tablets. Single point NIR absorbance spectra were used to establish a PLS-II calibration model with nominal concentration (%, w/w) for all three components in the design matrix. Scattering orthogonalization was used as a preprocessing technique and was compared to conventional data pretreatment of SNV to explore the capacity of removing physical interference while maintaining chemical variation. Each design point was considered as a group of tablets containing identical chemical composition but different physical interference (i.e., tablet density). Group-wise cross validation was utilized to compare the robustness of predictions based on different preprocessing techniques. Prediction was subsequently performed according to the design matrix (Figure 1).

RESULTS AND DISCUSSION

The reported radially diffused reflectance detected by spatially resolved spectroscopy was originally based on measurements via optic fiber. It typically requires either a powerful light source or a sensitive detector to achieve a reasonable signal-to-noise ratio.[8,17] However, the availability of the NIR chemical imager provides an easy access to measure spatial distribution of radially diffused reflectance. Compared to frequency- and time-resolved spectroscopy, the easy optical setup in the current experiment demonstrates the potential of using a chemical imaging system to capture radially diffused reflectance on a variety of pharmaceutical samples.

Only radially diffused reflectance inside the point illumination was used in the study because of the better signal-to-noise ratio inside the illumination spot than that outside the point illumination. This approach was used for the following reasons. First, incidence angle of 5−10° avoided the specular reflectance, thereby allowing only diffuse reflectance to be detected by the imager. Second,

(15) van Staveren, H. J.; Moes, C. J. M.; van Marle, J.; Prahl, S. A.; van Gemert, M. J. C. *Appl. Opt.* **1991**, *30*, 4507–4514.
(16) Pham, T. H.; Eker, C.; Durkin, A.; Tromberg, B. J.; Andersson-Engels, S. *Appl. Spectrosc.* **2001**, *55*, 1035–1045.
(17) Nichols, M. G.; Hull, E. L.; Foster, T. H. *Appl. Opt.* **1997**, *36*, 93–104.

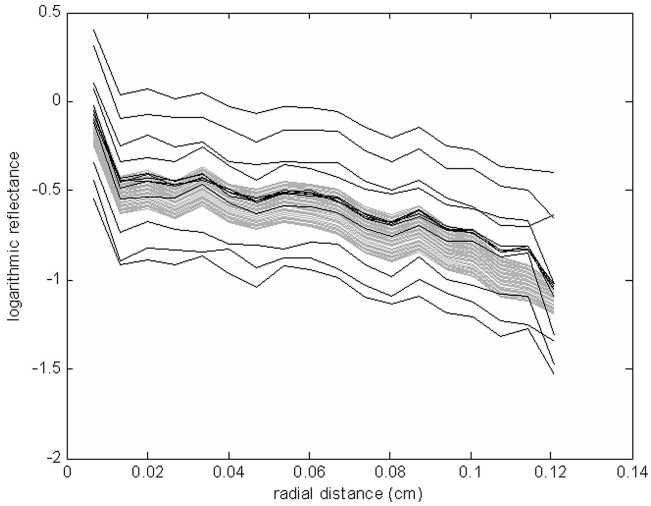

Figure 4. Example of radially diffused reflectance. Black: simulated radially diffused reflectance after magnitude correction by reference standard. Gray: measured radially diffused reflectance for one single tablet.

the key to successfully calibrate and predict optical coefficients was the comparability between simulated and measured radially diffused reflectance. On the basis of pseudorandom distribution of light illumination, simulated radially diffused reflectance after correction by reference standard was able to appropriately mimic the shape and magnitude of measured radially diffused reflectance inside the point illumination (Figure 4). Thus, radially diffused reflectance from inside the point illumination was still considered appropriate for separation of absorption and scattering.

Monte Carlo simulation was used to generate simulated radially diffused reflectance data. Monte Carlo simulation is a stochastic model used to capture the average behavior of random variables to describe the desired physical quantities using the random variables. There are several important considerations for the present system. First, a large number of independent replicates are required for the average behaviors to approximate the true physical quantities. Thus, an empirical equation was used to calculate the required photon number to generate the similar noise characteristics across the space of center component design.[13] Second, because of the unavailability of reported refractive indices for most pharmaceutical materials, a refractive index of 1.58 was used in the simulation. It was the only reported value for lactose, a commonly used excipient in the pharmaceutical industry.[18] Third, an anisotropy factor has yet to be reported for a pharmaceutically relevant material. However, because pharmaceutical materials are clearly not isotropic, the anisotropy factor must be larger than 0 and smaller than 1 to be characterized by the forward scattering.[19] To avoid the potential effects of anisotropy factor on simulated radially diffused reflectance, a constant anisotropy factor of 0.9 was selected.

Partial least-squares (PLS) was used to calibrate the simulated radially diffused reflectance with simulated absorption and reduced scattering coefficients. PLS is a mathematical method that describes the covariance between multivariate data (simulated radially diffused reflectance) and response variables (simulated optical coefficients) by means of a small number of orthogonal variables or principal components. After mean-centering, a multivariate PLS calibration model using two principal components captured 99.70% of the X variance and 94.40% of the Y variance, with RMSEC = 13.7934 cm^{-1} and RMSEV = 9.6123 cm^{-1} (Figure 5A and 5B). Before the PLS model was used to predict optical coefficients from measured radially diffused reflectance, validation by Intralipid 20% at 1064 nm was also performed. The predicted optical coefficients can be found in Table. 1. The closeness between predicted and reported optical coefficients indicate the current PLS model was reasonably calibrated and tested.

An example of the predicted absorption and reduced scattering coefficients for a single tablet are illustrated in Figure 6A and 6B. This figure illustrates the basis for the following observations. First, the reduced scattering coefficient was inversely correlated to the absorption coefficient, in which a valley in the reduced scattering coefficient corresponded to a peak in the absorption coefficient. This observation supports the argument that scattering correction should be minimized (down-weighed) in spectral regions where dominating chemical constituents absorb very strongly.[4] Second, as described by the Mie scattering theory, the scattering coefficient is (in general) inversely proportional to wavelength while the absorption coefficient increases with wavelength. Thus, the difference between the reduced scattering and absorption coefficients is smaller at longer wavelengths compared to shorter wavelengths, further confirming that scattering is wavelength-dependently larger than absorption.[2,3,7] It also indicates that the absorption becomes more dominant in the spectra at longer wavelengths as compared to shorter wavelengths.

After orthogonalization of the reduced scattering coefficient, it was observed in Figure 7A that the variance reduction seemed to have a wavelength dependency. Original spectral variance decreased at wavelength regions (1350−1600nm), while spectral variance still maintained at other wavelength regions (1100−1350 nm). This can be attributed to the wavelength- and absorption-dependence of the reduced scattering coefficient. Comparatively, since SNV removed baseline and slope simultaneously for each single spectrum, the spectral variance was reduced at all the wavelengths (Figure 7B).

A ternary diagram was used to generate six design points. Since there were three levels for each chemical component, the three outermost design points were not chosen for calibration development because of their lack of leverage for each component during group-wise cross-validation. Thus, one outermost point combined with three inner points was used as the calibration data set while the remaining two outermost points were reserved for validation.

Group-wise cross-validation and prediction were used to compare the robustness of prediction on independent data set between scattering orthogonalization and SNV as spectral pretreatments before PLS-II calibration. It was suggested that orthogonalization required one less principal component to achieve the lowest RMSECV and RMSEP (Figure 8). It was shown in Table. 2 that RMSECV and RMSEP were better when employing scattering orthogonalization as preprocessing technique (4.09% and 4.02%) compared to SNV (4.27% and 6.23%). The preprocessed spectra after SNV and orthogonalization were also used for PLS-I

(18) Pan, T. S.; Sevick-Muraca, E. M. *J. Pharm. Sci.* **2006**, *95*, 530–541.
(19) Cogdill, R. P.; Drennen, J. K. In *Spectroscopy of Pharmaceutical Solids*; Brittain, H. G., Ed.; Taylor & Francis: New York, 2006; pp 313−412.

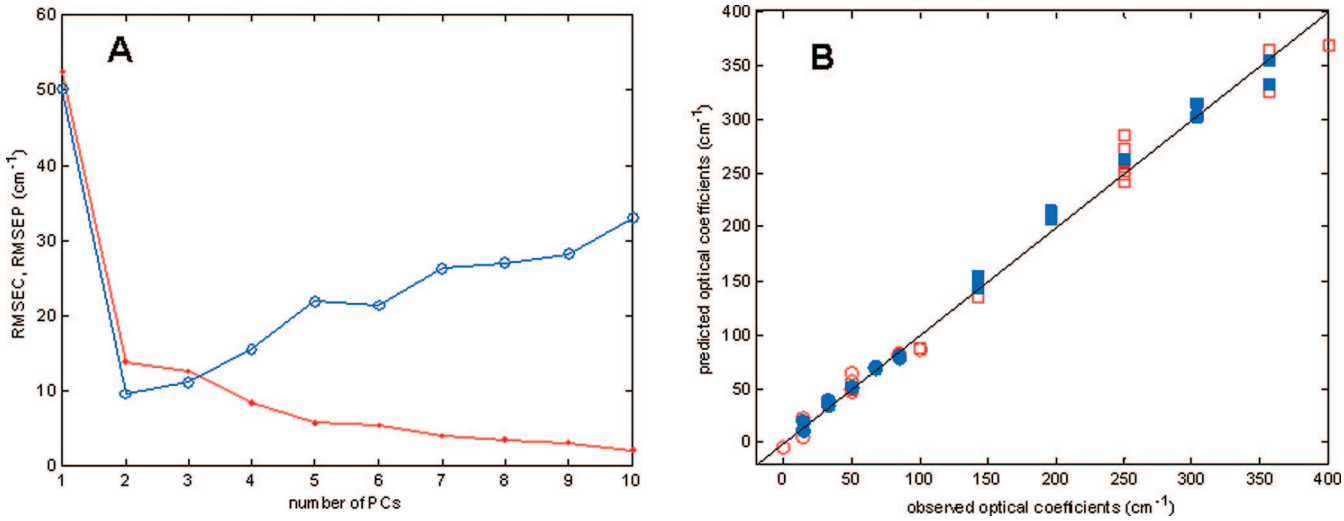

Figure 5. Scree plot (A) and correlation plot (B) of Monte Carlo simulation based PLS model. A: dotted line represented RMSEC, circled line represented RMSEV. B: open circle was calibrated absorption coefficients, open square was calibrated reduced scattering coefficients, closed circle was validated absorption coefficients, and closed square was validated reduced scattering coefficients.

Table 1. Validation of PLS Model by Intralipid at 1064 nm

	reported value	predicted values[a]
μ_a (cm^{-1})	10	9.9 ± 0.2
μ_s' (cm^{-1})	262	270.7 ± 0.5

[a] Predicted values were based on three replicate measurements.

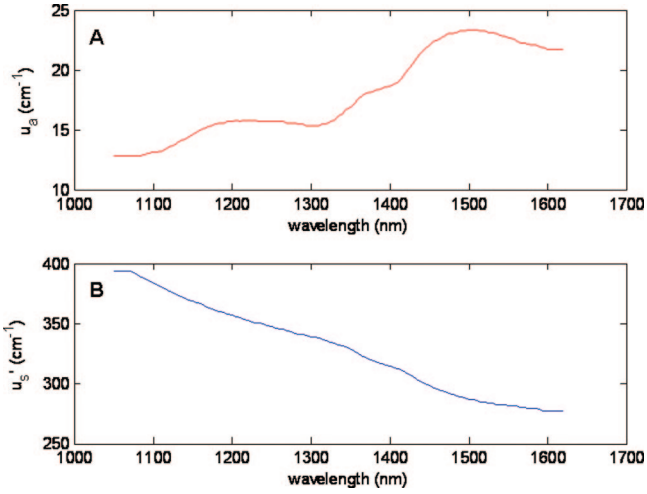

Figure 6. Example spectra of absorption coefficient (A) and reduced scattering coefficient (B).

model calibration of individual components. The results were found to be similar to what was observed in PLS-II (data not shown). Thus, cross-validation and prediction results demonstrated the benefits of using scattering orthogonalization for scattering correction to predict chemical information (concentrations in this case).

According to the available literature,[20] orthogonalization is designed to remove the portion of signal related to **P** in eq 3, and as such, the orthogonalized spectral matrix is related to everything but **P**, which was the factor of the reduced scattering coefficient.

(20) Lorber, A. *Anal. Chem.* **1986**, *58*, 1167–1172.

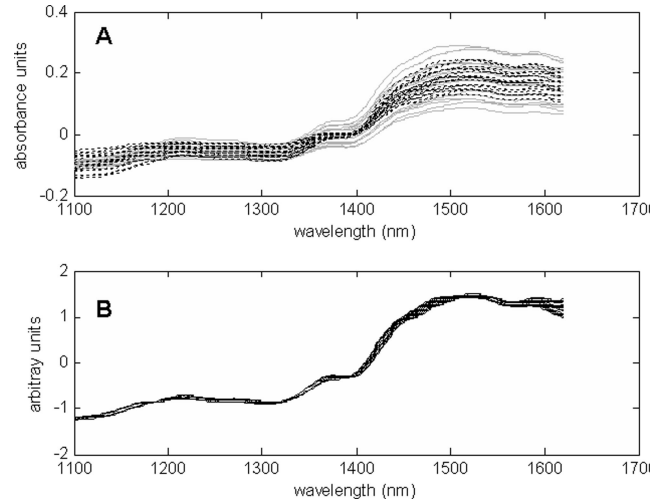

Figure 7. Absorbance spectra comparison after different data pretreatments. A: gray represented raw absorbance spectra of three-component tablets; black dot represented the spectra after orthogonalization of reduced scattering coefficient. B: absorbance spectra after SNV.

The average reduced scattering coefficient within the calibration data set was used as the vector to perform orthogonalization to represent overall wavelength- and absorption- dependent scattering properties. After scattering orthogonalization, the regression vectors corresponding to the three constituents were observed to be similar to pure-component spectra (Figure 9). Since scattering orthogonalization inhibited the interference of reduced scattering coefficients by down-weighing the absorption signals during scattering correction in this region, adequate spectral variations remained in the spectra to be correlated with chemical information. On the other hand, SNV, which removed the spectral variance at all the wavelengths, intermixed the chemical absorption and physical interference at certain regions. For example, regression vectors of each component after SNV showed anomalous features as indicated by arrows, which can not be attributed to any known signal from the chemical constituents (verified by comparison to pure components). Thus, it was demonstrated that

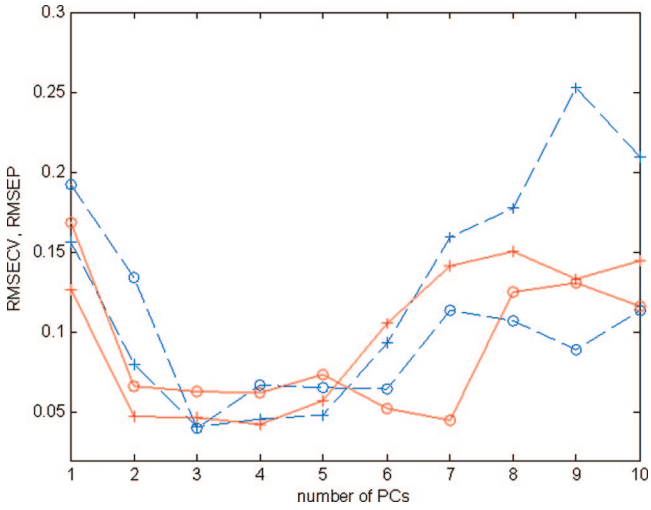

Figure 8. Comparison between SNV and orthogonalization as data pretreatment. Plus sign symbols represented RMSECV, circles represented RMSEP; solid lines were the results after SNV, dashed lines were the results after orthogonalization.

Table 2. Cross-Validation and Prediction Results after Using Different Data Pretreatments

	PCs	RMSEC (%)	RMSECV (%)	RMSEP (%)
SNV	4	1.97	4.27	6.23
Orth[a]	3	2.99	4.09	4.02

[a] The results were based on scattering orthogonalization.

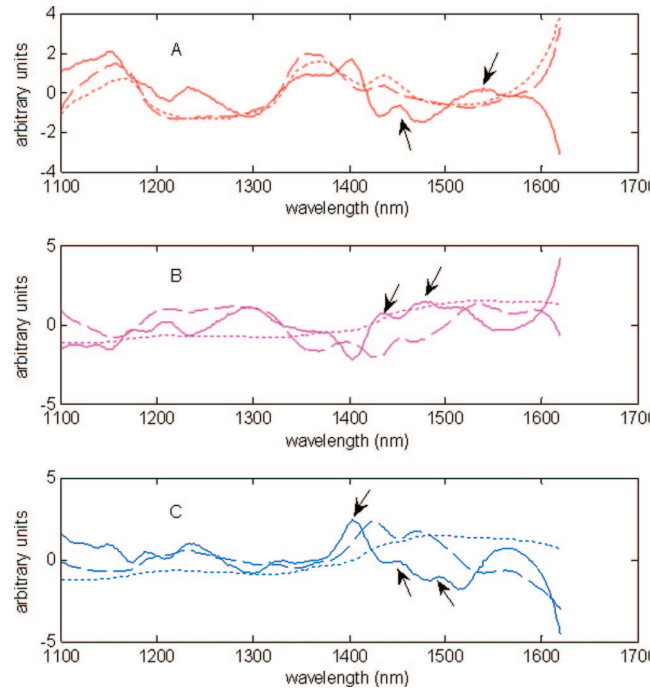

Figure 9. Comparison between pure component spectra and regression vectors after using SNV and orthogonalization as data pretreatments. A, B, and C represented the results for theophylline, lactose, and PH200, respectively. Solid lines represented the regression vector after SNV, dashed lines represented the regression vector after orthogonalization, dotted lines represented pure component spectra. Arrows indicate intermixed spectral features between chemical information and physical interference by SNV.

scattering orthogonalization improved the capacity to selectively reduce physical interference and maintain chemical information to establish a calibration model related only to chemical variation.

To our best knowledge, this is the first study applying spatially resolved spectroscopy to determine the optical properties of pharmaceutical materials. Although imaging sensor and radially diffused reflectance inside the point illumination were used under the current experimental setting, simple optic fiber detection plus comparable Monte Carlo simulation outside the spot size is also capable to perform prediction of reduced scattering coefficients and scattering orthogonalization. Thus, the enhanced scattering correction capacity of scattering orthogonalization demonstrated by this study and the comparatively simple optical setup of spatially resolved spectroscopy suggests a promising combination of spatially resolved spectroscopy with conventional NIR spectroscopy (single point) to perform scattering correction within the instrument. The resultant NIR absorbance spectrum is a scattering-orthogonalized spectrum, which can be directly used to correlate with the chemical makeup of the samples.

The similarity between regression vector after using scattering orthogonalization as data pretreatment and pure component spectra indicates the importance of applying reasonable preprocessing techniques based on solid understanding of scattering properties to reduce physical interference and maintain the signal of interest. Thus, scattering orthogonalization can not only be applied to remove physical interference when a calibration model is designed to predict chemical information but also be applied to remove potential variations in future samples related to scattering properties. For routine process monitoring via NIR on a pharmaceutical unit operation, it is advantageous to periodically collect sample spectra of both absorption and reduced scattering coefficients to trace the chemical and physical variations as well. Using the updated reduced scattering coefficient and necessary scattering orthogonalization can directly improves the robustness of the calibration model and reduce the potential error in the model prediction caused by physical variations, such as raw material variations because of different vendors (e.g., the same chemical composition with slight particle size variations) and scale-up issues (e.g., small changes due to optical interfaces, density of samples, and sample presentation).

The advantages of scattering orthogonalization demonstrated above are expected to be applicable for use in industries beyond pharmaceuticals, such as agriculture and food. First, the simple optic setup of spatially resolved spectroscopy allows measurements on a variety of sample forms, such as grain and cheese, and so forth. Second, physical interference caused by batch, production area, and year can be reduced via the measurement of the reduced scattering coefficient followed by scattering orthogonalization in a manner identical to that applied to pharmaceutical products. Thereby, the robustness of a spectroscopy-based multivariate model can be enhanced.

CONCLUSIONS

An NIR chemical imager is capable of performing radially diffused reflectance measurements on pharmaceutical samples leading to the determination of optical coefficients of the samples. A Monte Carlo simulation based PLS calibration model predicted absorption and reduced scattering coefficients for pharmaceutical tablets based on the radially diffused reflectance measurements

via the chemical imager. Orthogonalization of single-point NIR absorbance spectra utilizing the measured reduced scattering coefficients allowed both wavelength- and absorption- dependent scattering correction. The comparison of regression vectors demonstrates the improved ability of scattering-orthogonalization over conventional baseline correction (e.g., SNV) to reduce physical interference (i.e., tablet density) while maintaining chemical information. Further, orthogonalization of the NIR signal to the reduced scattering coefficient improved the calibration model statistics obtained for prediction of tablet constituents, relative to SNV. General applicability of the scattering-orthogonalization can be foreseen as (1) a scattering correction method integrated into NIR spectrometers; (2) a routine calibration model update tool to minimize changes to the calibration due to physical variations in the samples related to the reduced scattering coefficient. The proposed orthogonalization technique is expected to be applicable to any sample type that has a significant contribution to its spectrum from light scattering properties.

ACKNOWLEDGMENT

The authors would like to thank Steve Short for his efforts in preparing the compacts and collecting single-point NIR absorbance spectra, and Ryanne Palermo for her skillful scientific and grammatical editing.

Received for review October 6, 2008. Accepted December 27, 2008.

AC802105V