ACS OMEGA

Open Access

http://pubs.acs.org/journal/acsodf

Article

PCA and PLS Analysis of Lanthanides Using Absorbance and Single-Beam Visible Spectra

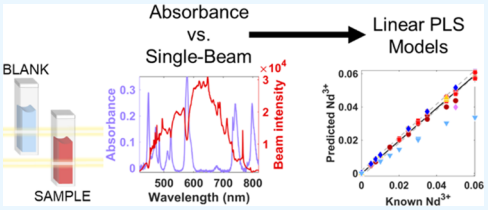
Hope E. Lackey,* Gilbert L. Nelson, Heather M. Felmy, Xiaofeng Guo, Samuel A. Bryan,* and Amanda M. Lines*

Cite This: *ACS Omega* 2024, 9, 33662−33670

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: During process monitoring applications, referenced optical spectroscopy, such as absorbance spectroscopy, can suffer from environmental and instrumental fluctuations that alter the intensity of irradiance reaching the spectrometer's detector at each detected frequency. Temperature, vibration, light source aging, instrument damage, detector aging, detector registry shifts, sampling cell degradation, and similar perturbations create situations in which a previously collected reference spectrum may no longer be valid for the current state of the system. This can lead to the calculation of poor-quality absorbance spectra that are unsuitable for qualitative or quantitative analysis based on prior calibration models. The use of single-beam spectra in the creation of multivariate calibration models circumvents the need for collecting and maintaining a stable reference spectrum throughout an ongoing chemical process. However, unlike absorbance spectra, which typically have a zero baseline, single-beam spectra contain a high background signal relative to an analyte signal, and they may also contain intense peaks from the light source. Here, multivariate principal component analysis (PCA) and partial least squares (PLS) regression models are built using single-beam and absorbance spectra to compare the efficacy of both types of spectra for qualitative and quantitative analyses of lanthanide solutions. A multileg fiber optic UV−visible spectrometer is utilized to collect samples under three distinct wavelength registries in three unique sampling cells and under lighting conditions spanning 0.2 to 2.0 relative transmittance. Under these conditions, single-beam spectral PCA models produced enhanced discrimination between sampling conditions, allowing spectra to be grouped by the instrumental conditions under which they were collected. Absorbance and single-beam PLS models produced equivalent quantitations of the lanthanide concentrations.



INTRODUCTION

Process monitoring, whether in-line or online, can provide vital information for maintaining the control, efficiency, and safety of chemical processes.[1,2] Process analytical technology (PAT) utilizing optical spectroscopy as a means of measurement has been applied in numerous fields, frequently being applied in food production and safety,[1,3] nuclear safeguards and reprocessing,[2,4] environmental monitoring,[5] and pharmaceutical[6,7] processes.

Optical spectroscopic techniques allow for the identification and quantification of a wide range of chemical species. Optical spectroscopy can utilize robust probes in a variety of commercially available designs, allowing in-line or online data collection even under radioactive,[4,8] thermally intensive,[9] and acidic or alkaline conditions.[10] Although optical spectroscopy does not always achieve the same sensitivity as other techniques, it is inexpensive, well-developed, easily implemented,[11] and adept at rapid analysis. These features allow it to be substituted in place of more expensive and time-consuming methods in applications where the required sensitivity falls within the range of the optical method.[4] Optical methods are also often nondestructive,[12] allowing them to be utilized before or currently with other techniques, such as in spectroelectrochemistry applications.[13] Many sensor configurations have been adapted to fit the analysis of an ongoing process on a variety of process scales, and the type of spectroscopy applied to each scenario can be chosen to fit the required analytical need.

Multivariate chemometric analysis is often applied to derive qualitative and quantitative data from complex optical spectra. Chemometrics may be used as an effective mathematical tool for online monitoring to detect when internal chemical changes occur in process streams. Several issues related to the successful analysis of a processing stream, such as the quantification of a species of interest in the presence of chemical interferents or shifts in environmental conditions, have been mitigated using chemometric modeling.[14] Multivariate signal processing methods have also been developed to

Received: March 6, 2024
Revised: May 24, 2024
Accepted: June 14, 2024
Published: July 23, 2024

combat multiplicative effects and baseline shifts in optical spectra,[15] which interfere with basic univariate analysis.

UV−visible (UV−vis) absorbance spectroscopy is useful for detecting molecular and atomic species.[16] However, absorbance spectroscopy requires the ability to acquire a stable reference spectrum,[17] which is utilized to calculate the absorbance spectrum of a sample. In an ongoing chemical process such as that encountered in PAT applications, system instabilities may occur for a variety of reasons and may originate at many points within the system. For instance, light sources may age or fluctuate in output, fiber optics can suffer damage, and process cells may become damaged or contaminated.[1] Another challenge to reference stability is the shift of frequency registries in spectrometers as detectors, gratings, and other instrument components degrade over time or are damaged.[18] Such changes cause a collected reference to no longer be valid for the continued calculation of absorbance spectra.[17] Furthermore, such changes to sampling conditions introduce uncalibrated variance in the data set of a chemometric model.

Sampling condition variation can invalidate calibration models unless a new reference can be collected under the new lighting conditions or unless the models can be rebuilt using data collected under new conditions.[18] Calibration transfer methods may be used to overcome changes in sampling conditions, but calibration transfer often requires remeasuring standards under new conditions.[19,20] Resampling can cost considerable time and money depending on the process, and it may be impossible depending on the accessibility of the system and the required samples. Robust calibration techniques must remain viable despite system perturbations and thus not require expensive and time-consuming model updating.[20]

This work seeks to create robust calibration models that maintain adequate performance despite significant changes in instrumental conditions between spectral measurements. The motivation for this study derives from the desire to maintain a reliable online process for monitoring data streams during unexpected system changes. Additionally, the ability to diagnose the instrumental cause of the spectral changes is desirable. Thus, this study focuses on overcoming practical limitations in the application of UV−vis absorbance spectroscopy and other referenced techniques to online monitoring situations. Specifically, due to the challenges of maintaining a stable reference spectrum in complex, online monitoring applications, this study re-examines the common practice of utilizing referenced absorbance spectra[21] compared to unreferenced, single-beam spectra.[17,22]

Here, multivariate spectral data are utilized to build chemometric models for the prediction of lanthanide ion concentrations in aqueous solutions. This work focuses on the measurement of two rare earth elements, namely, neodymium and praseodymium. Although more sensitive detection methods exist for rare earth elements, such as total reflection fluorescence,[23,24] laser-induced breakdown spectroscopy,[25] and inductively coupled plasma mass spectrometry,[23] visible spectroscopy was chosen here due to its aforementioned merits as a nondestructive online monitoring method.[7] The detection of neodymium and praseodymium is used to demonstrate a chemometric treatment of visible spectra, but the technique is applicable to many visibly absorbing species. Traditional absorbance spectra are utilized in chemometric models, along with single-beam, unreferenced spectra. These spectra are collected after various alterations are made to sampling cells, fiber optics, and incident light beam intensity. These changes result in a transmittance range of 0.20 to 2.0 relative to control conditions. Chemometric models are built from absorbance and single-beam spectra, and the qualitative and quantitative results are examined. Models built from absorbance and single-beam spectra are compared quantitatively using the errors in the predicted concentration of two analytes, neodymium (Nd) and praseodymium (Pr) nitrate.

EXPERIMENTAL SECTION

Materials. Nitric acid (concentrated, 70%) was acquired from Sigma-Aldrich (St. Louis, MO), as were iron(III) nitrate nonahydrate (>99.999%, trace metal basis) and copper(II) nitrate trihydrate (99−104%). Neodymium nitrate hexahydrate (99.9%), praseodymium nitrate hexahydrate (99.9%), and holmium nitrate pentahydrate (99.9%) were purchased from Strem Chemicals (Newburyport, MA). Ruthenium chloride (45−55% Ru content) was acquired from Sigma-Aldrich (St. Louis, MO). Molar absorptivity spectra for each chemical are shown in Figure S2. Deionized (DI) water of 18 MΩ·cm resistivity was used in sample preparation.

A sealed standard solution of didymium perchlorate (RM-DL) was purchased from Starna Scientific. Didymium is a mixture of neodymium and praseodymium.

Spectrometer. The UV−vis spectrometer is a thermoelectrically cooled (TE) charge-coupled device (CCD)-based detector capable of transmitting and receiving light through fiber optic cables purchased from Spectra Solutions Inc. (Norwood, MA). Figure 1 (top) shows each major component of the system.

Light is directed via a fiber optic from an external light source into a port on the spectrometer's face, indicated in

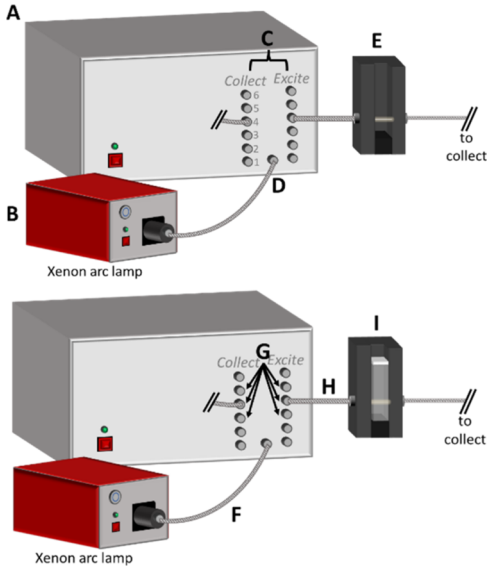


Figure 1. Diagram of the multitrack UV−visible spectrometer utilized in this study and the points at which alterations were made. (A) Spectrometer housing containing CCD arrays and internal fibers for each track. (B) External light source. (C) Excitation and collection ports for each internal track. (D) Port through which external light is fed into the internal beam splitter. (E) Stainless steel holder for sampling cells with fiber optic attachments. (F) Lamp fiber. (G) Three instrument tracks utilized in this study. (H) Excitation fiber. (I) Cell holder with sampling cell in place.

https://doi.org/10.1021/acsomega.4c02202
ACS Omega 2024, 9, 33662−33670

Table 1. Descriptions of the Alterations Made to Sampling Conditions, as well as the Transmittance of a Water Spectrum Collected under Each Condition Relative to a Water Spectrum Collected under the Control Condition

title	description	relative transmittance	change point in Figure 1
control	Single fiber to and from sample cell holder.	1.0	n/a
γ-irradiated	γ-irradiated fiber to sample cell holder, and unirradiated fiber from sample cell holder to spectrometer.	1.6	H
alternate track 1	Single fiber to and from sample cell holder, on a secondary track.	1.2	G
alternate track 2	Single fiber to and from sample cell holder, on a tertiary track.	0.7	G
coupled fiber	Three coupled fibers to sample cell holder, and single fiber from sample cell holder.	1.1	H
cell filter	Fiber to cell holder containing filter, fiber from filter holder to sample cell holder, single fiber from sample cell holder to spectrometer.	0.25 to 1.6	H
lamp aperture	Single fiber to and from sample cell holder, an aperture attenuator in the light path between the lamp and beam splitter.	0.32 to 2.0	F
lamp filter	Cell holder containing filter in the light path between the lamp and beam splitter, single fiber to and from sample cell holder.	0.27 to 1.8	F
plastic cuvette	Single fiber to and from sample cell holder, with samples in plastic 1 cm cuvettes.	1.9	I
scintillation vial	Single fiber to and from sample cell holder, with samples in half dram glass vials.	0.20	I

Figure 1D. Excitation light was provided by a SLS205 Xenon arc lamp purchased from ThorLabs Inc. (Newton, NJ).

Incoming light is split into six fibers, which can then be directed out of six excitation ports through fiber optic cables. There are six corresponding collection ports that direct light onto six different regions of the CCD detector. These ports are shown in Figure 1C. Each pair of excitation and collection locations is referred to as a "track," resulting in six tracks in the instrument. The CCD detector collects a signal for a length of time, called the integration time, specified in the SpectraSoftAbsorb software from Spectra Solutions Inc. Each track is collected for the same amount of time.

The detected range for the spectrometer is 395−874 nm, but the included range for analysis is limited to 417−820 nm due to low signal below 417 nm and sharp, intense source lines above 820 nm.

For each sample, 100 UV−vis spectra were taken. To reduce the noise and data processing time, the spectra were averaged into a single spectrum per sample.

Data Processing Software. Spectral extraction, averaging, and plotting were conducted using Matlab2022b (Mathworks, Natick, MA). Spectral preprocessing and chemometric analyses were performed using PLS Toolbox 9.1 (Eigenvector Research Incorporated, Manson, WA).

Chemometric Methods and Measures of Error. Principal component analysis (PCA), principal component regression (PCR), and partial least squares (PLS) are well-defined elsewhere.[26−28] Briefly, PCA is a decomposition method, and PCR and PLS are regression methods. These algorithms produce model components called principal components in PCA and PCR and latent variables in PLS. Model components capture variance within the matrices used to calibrate the models. These model components are mathematically significant but do not necessarily capture the spectroscopic signatures of pure chemical constituents.

Here, PCA is used to classify spectra based upon the system alterations made during spectral collection and to examine trends in the variance present in absorbance versus single-beam spectra. PCR is used to add quantitative detail in the PCA analysis. PLS analysis is used to regress the concentrations of neodymium and praseodymium.

The predictive ability of the regression models for neodymium and praseodymium concentrations was measured by root mean square errors of calibration (RMSEC), cross-validation (RMSECV), and prediction (RMSEP).[7,28] The RMSEC is calculated by using the difference between the known solution composition and the composition predicted from the calibration set spectra by the PLS model. The RMSEP is calculated using the difference between the known and predicted compositions of the independent validation set spectra.

The coefficient of determination (R^2) is utilized as a measure of the fit between the gravimetrically known concentrations of samples and their concentrations, as measured by PLS models. An ideal fit would have an R^2 of 1, in which the PLS predictions do not vary from the gravimetric measurements.[7] Plots of the known versus chemometrically measured values are termed parity plots.

The Hotelling T-squared (T^2) measures the distance of a data point from the center of the data cloud within a space defined by the model's latent variables.[29] A sample with high T^2 scored very high or low on one or more model components, indicating that it is near or beyond the edge of the calibrated signal variation. The Q-residuals are a measure of unexplained variance, that is, variance uncaptured by the model.[30]

Instrumental Conditions. Spectra were collected under a variety of instrumental conditions, in which a portion of the spectrometer setup (Figure 1) was modified from the control conditions. The conditions are summarized in Table 1. The conditions are shown in diagrams in Figure S1 and described in greater detail in the Supporting Information.

The locus of change for each condition is noted in Table 1, referencing the lower portion of Figure 1. Each instrumental condition is designed to mimic a potential disturbance that can occur in a spectrometric system during an online monitoring process. Briefly, conditions that mimic damage or alteration to a fiber optic leading to the sampling cell are the γ-Irradiated, Alternate track 1, Alternate track 2, Coupled fiber, and Cell filter conditions. The Lamp filter and Lamp aperture conditions mimic light source degradation. Alternate track 1 and Alternate track 2 mimic wavelength registry shift and detector sensitivity changes over time. The Plastic cuvette and Scintillation vial represent a change in the sample cell itself.

The Scintillation vial condition is the most drastic instrumental change, as it involves the use of a round sampling cell with a shorter path length than the 1 cm cuvettes used in all other studies. This condition is regarded as a condition of critical fault in the system.

https://doi.org/10.1021/acsomega.4c02202
ACS Omega 2024, 9, 33662−33670

RESULTS AND DISCUSSION

PCA, PCR, and PLS are linear methods. Absorbance is known to exhibit a linear relationship with chromophore concentration under certain conditions according to the Beer–Lambert law. The particulars of the Beer–Lambert law are available from multiple sources,[31,32] and the derivation[32,33] demonstrates that a near linear relationship of path length, concentration, and absorbance of a solution containing a chromophore is established alongside a logarithmic relationship. The emergence of a linear relationship in both absorbance and single-beam spectra is discussed in the Supporting Information.

The following section discusses the collected spectra and compares linear models of the absorbance and single-beam spectra.

Spectral Types. Single-beam and absorbance spectra are modeled in this study. Single-beam spectra are defined here as the digitized signal produced after photons pass through a sample and contact the CCD detector. The absorbance spectra used for modeling are calculated using references that are recollected approximately every five samples, and absorbance is calculated as in eq 1 in the Supporting Information.

Figure 2 shows absorbance (A) and single-beam (B) spectra of a sample containing 60 mM neodymium nitrate, collected under each instrumental condition and attenuation level. Figure 2B of the single-beam spectra reveals that the intensity of the background signal, that is, unabsorbed photons, varies in each condition, as does the depth of the absorbance bands.

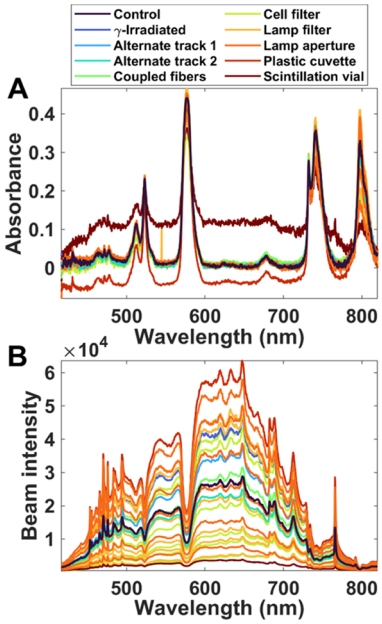


Figure 2. (A) Absorbance spectra of 60.0 mM neodymium nitrate under varying instrumental conditions. (B) Single-beam spectra corresponding to the spectra shown in (A).

Supporting Information Figure S7 and Figure S8 show single-beam and absorbance spectra of a series of neodymium nitrate with samples collected under each instrument condition displayed in a separate plot.

Data Set Definitions. A total of 1121 averaged spectra were collected and assigned to data sets for PCA or PLS. Each data set contains a calibration set and a validation set, with each set having a spectral matrix and a corresponding concentration matrix. Calibration samples are used to build the chemometric models' regression vectors, and the validation samples are utilized to test the error of the models when measuring new spectra.[27,34]

Though only neodymium and praseodymium are regressed using PLS, additional chromophores of copper(II), ruthenium(III), iron(III), and holmium are included in the sample sets as interfering absorbers. Data sets are further described in the Supporting Information.

Spectral Processing. Spectral processing is often required in online monitoring applications due to unstable or unexpected conditions that arise and perturb spectra.[14]

Absorbance and single-beam spectra require preprocessing due to differences in the incident light, baseline fluctuations, lack of linearity, or wavelength misalignment. The preprocessing steps used in this study's chemometric models are listed in Table 2.

Table 2. Preprocessing Steps Applied to Spectral Data before the Calibration of Multivariate Chemometric Models

model	spectra	preprocessing
PCA, PCR, Instrumental condition	Single-beam	$\log_{10}$, Savitsky–Golay 1st derivative,[a] 2-norm, GLSW[b]
PCA, PCR, Instrumental condition	Absorbance	Savitsky–Golay 1st derivative,[a] GLSW[b]
PLS, Nd regression	Single-beam	$\log_{10}$, 1-norm, Savitsky–Golay 1st derivative,[a] GLSW,[b] mean center
PLS, Nd regression	Absorbance	Savitsky–Golay 1st derivative,[a] GLSW,[b] mean center
PLS, Pr regression	Single-beam	$\log_{10}$, EMSC,[c] Savitsky–Golay 1st derivative,[a] mean center
PLS, Pr regression	Absorbance	Savitsky–Golay 1st derivative,[a] mean center

[a]Savitsky–Golay derivatives used a 2nd-order polynomial, 29 point window width. [b]Generalized least squares weighting (GLSW) filter used a downweighting parameter of 0.02. [c]Extended multiplicative scatter correction (EMSC) utilized a classical least squares algorithm and 2nd-order polynomial.

Before any other signal processing occurred, wavelength registries for the three tracks of the instrument were aligned via a first- or second-order polynomial transformation. This is discussed further in the Supporting Information Axial Alignment section.

For PCA and PCR analysis, the following preprocessing schemes were used to distinguish the spectral variation unique to each instrumental alteration rather than the spectral variation caused by the presence of chromophores. Single-beam spectra were preprocessed by applying a base-10 logarithm followed by a Savitsky–Golay first derivative.[35] Next, normalization of the area under the curve was applied. Finally, a generalized least squares weighting (GLSW) clutter filter was applied. The GLSW filter creates a downweighting matrix using singular value decomposition of the preprocessed spectral matrix, which is applied to spectra to reduce spectral variance that does not correlate with the assigned system alteration, in this case, the instrumental conditions.[36] Additional α (downweighting) values were tested in the PLS Toolbox for the GLSW filter with only moderate improvement to model error at smaller α values. To avoid overfitting, 0.02 was chosen.

https://doi.org/10.1021/acsomega.4c02202
ACS Omega 2024, 9, 33662–33670

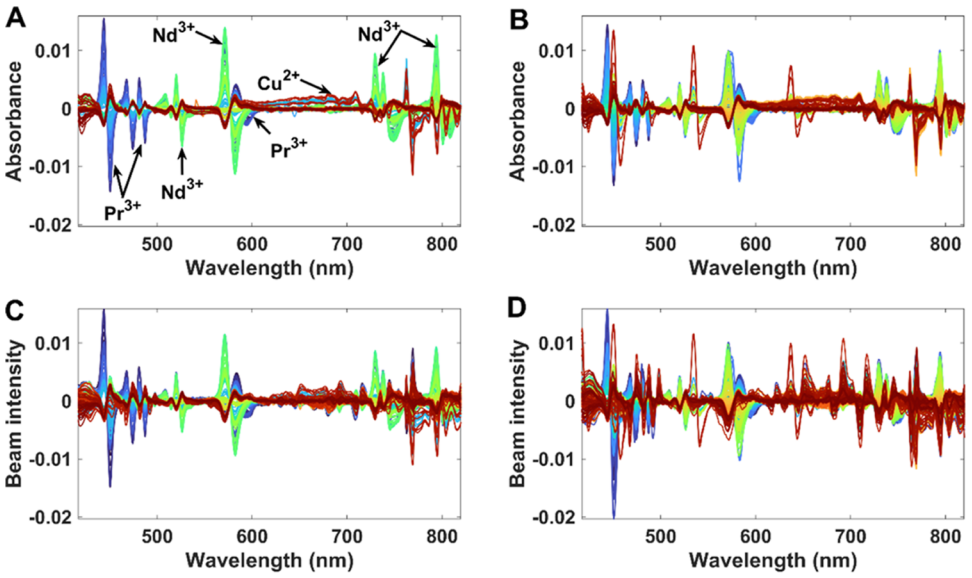


Figure 3. (A) Calibration absorbance spectra after preprocessing and (B) validation absorbance spectra after preprocessing via a Savitsky−Golay 1st derivative and mean centering. (C) Inverted calibration single-beam spectra after preprocessing and (D) inverted validation single-beam spectra after preprocessing via application of $\log_{10}$, extended multiplicative scatter correction, Savitsky−Golay 1st derivative, and mean centering.

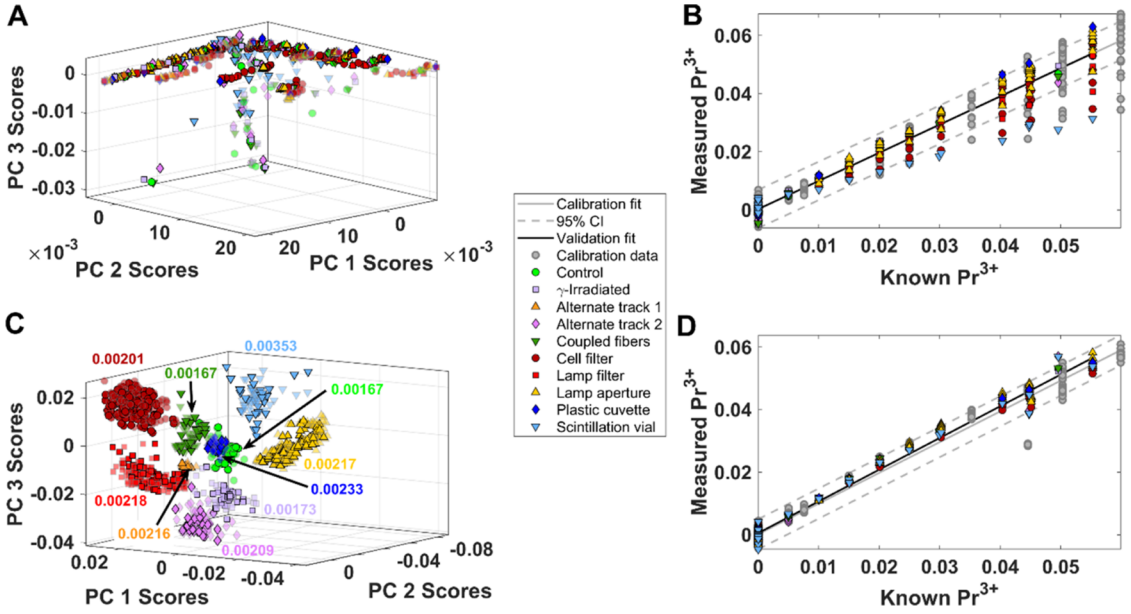


Figure 4. (A) Scores for sample spectra on the first through third components of a PCA model of absorbance. (B) Gravimetrically measured versus PCR-measured praseodymium concentration from a PCR model of absorbance spectra shown in part (A). (C) Scores for sample spectra on the first through third components of a PCA model of single-beam spectra. The overlaid, colored numbers are RMSEP for the data of each instrumental condition: 0.00167 for Control, 0.00173 for γ-Irradiated, 0.00216 for Alternate track 1, 0.00209 for Alternate Track 2, 0.00167 for Coupled fibers, 0.00201 for Cell filter, 0.00218 for Lamp filter, 0.00217 for Lamp aperture, 0.00233 for Plastic cuvette, and 0.00353 for Scintillation vial conditions. Partially transparent markers indicate calibration samples, and opaque markers indicate validation samples. (D) Gravimetrically measured versus PCR-measured praseodymium concentration from a PCR model of single-beam spectra shown in part (C).

Absorbance spectra were preprocessed using a Savitsky−Golay first derivative and a GLSW clutter filter. When a 2-norm was applied, as in the single-beam spectra, all correlative information was lost, with sample scores becoming normally distributed across all PC axes.

The preprocessed calibration and validation sets utilized in PLS regression of praseodymium concentration are shown in Figure 3(A,B) for absorbance and (C,D) for single-beam spectra. The single-beam spectra have been inverted for ease of comparison to the absorbance spectra. Single-beam spectra for praseodymium regression were preprocessed as shown in Figure S12 in the Supporting Information, with the preprocessed spectra shown in Figure 3C for the calibration set. The raw single-beam spectra are subjected to a base-10 logarithm, a 2nd-order extended multiplicative scatter correction (EMSC) algorithm,[37] a Savitsky−Golay first derivative, and mean centering. Mean centering here serves to reduce constant baseline effects and center the quantification models at an origin of zero.[28] The EMSC algorithm is capable of isolating chemical and physical sources of variance

https://doi.org/10.1021/acsomega.4c02202
ACS Omega 2024, 9, 33662−33670

Table 3. Partial Least Squares Root Mean Square Errors of Calibration (RMSEC), Cross-Validation (RMSECV), and Prediction (RMSEP) of Neodymium Nitrate and Praseodymium Nitrate Concentrations in Molarity from PLS Models of UV−Vis Spectra

	absorbance			single-beam		
analyte	RMSEC	RMSECV	RMSEP	RMSEC	RMSECV	RMSEP
Nd	1.02×10^{-3}	1.10×10^{-3}	2.25×10^{-3}	9.50×10^{-4}	9.68×10^{-4}	2.12×10^{-3}
Pr	2.52×10^{-3}	2.65×10^{-3}	3.66×10^{-3}	2.07×10^{-3}	2.11×10^{-3}	2.97×10^{-3}

within a data set, including multiplicative effects, to bring variant spectra closer to a common baseline.[37]

Absorbance spectra for praseodymium regression were preprocessed as shown in Figure S13 in the Supporting Information, through the application of a Savitsky−Golay first derivative and mean centering.

Figure 3 shows that while the preprocessed absorbance and single-beam spectra vary in intensity, the peak profiles are very similar. The single-beam spectra contain additional non-chemical artifacts compared to the absorbance spectra due to sharp and intense peaks in the xenon arc lamp's irradiance spectrum, which can be seen particularly in the region above 650 nm in parts (C,D).

For PLS analysis, differing preprocessing schemes were utilized for the regression of praseodymium and neodymium. This is discussed in the Supporting Information, along with a description of the neodymium model preprocessing.

In all of the above cases, the broad bands of ruthenium and copper ions are largely removed by the first derivative. Therefore, these preprocessing schemes are suitable only for quantification from narrow bands, and new preprocessing schemes would need to be developed to quantify analytes with broad spectral bands.

Overall, these preprocessing schemes reduce nonlinearities in the spectral matrices, establishing linear ranges within both single-beam and absorbance spectra such that linear models, such as PCA, PCR, and PLS, can be applied to the spectra.

Principal Component Analysis and Regression. PCA models can reveal underlying structures in spectral data sets and identify the profiles of the key sources of variance within the data set. Here, PCA is utilized to identify the instrumental conditions under which spectra were collected. It is also utilized to elucidate the relative importance of variance sources within the data set, as discussed in the Supporting Information. PCR, an extension of PCA, is utilized to compare the prediction errors introduced by each instrumental condition.

In this study, spectra were collected under ten instrumental conditions. In each condition, part of the system was altered, creating a unique profile of irradiance that was transmitted through the sample and collected as a single-beam spectrum. Not every system can be easily altered in such a way. However, if a system can be subjected to known alterations and calibration spectra collected when those alterations are in effect, then it may be possible to classify, via chemometric modeling, the unique change in spectral profile caused by that alteration. In doing so, new spectra can be fed into the classification model, and the type of alteration can be identified. This holds value for processing systems with known failure points, such that failure points can be identified via collected optical spectra, and the failure point can be repaired or adjusted to restore the system to its proper configuration.

The results of the PCA models of absorbance and single-beam spectra are shown in Figure 4, with part (A) displaying the scores of each sample on the first through third principal components for the absorbance spectral model and part (C) displaying the same scores for the single-beam spectral model.

In the single-beam model, most samples are grouped based on the instrumental condition under which they were measured. In Figure S11(B), the PCs that are closest to the PC space are replotted, and a greater degree of separation is revealed in the fourth PC space between these conditions.

In Figure 4C, the RMSEP values for neodymium predictions from a PCR single-beam model of each condition are listed next to the corresponding group of scores. The RMSEP was calculated by calibrating a PCR model using the Control spectra and the concentration of neodymium nitrate and then predicting the concentration of neodymium nitrate from the spectra taken under each system condition. The largest RMSEPs are seen in the conditions which are farthest from the Control condition in the primary PCs, PC 1 and 2. These high RMSEP conditions are the Lamp aperture, Alternate track 2, Cell filter, and Scintillation vial sets. The scintillation vial spectra have a great degree of separation in PC 2, separating them from the other conditions. This is to be expected due to the difference in path length, cell shape, and cell material utilized in the collection of these samples. The position of additional conditions is discussed in the Supporting Information.

The PCA scores from the model of absorbance spectra are shown in Figure 4A, where no discrimination between sampling conditions is seen. Discrimination of instrumental conditions also does not occur in the fourth PC space, as shown in Figure S11 (A). Instead, the primary variance in the absorbance spectra is caused by the concentration of the neodymium and praseodymium cation chromophores. This lack of grouping is due to the referencing that occurs in the calculation of absorbance. Here, the unique background signals of each condition are present in the single-beam reference and single-beam sample spectra. The conditional signals are eliminated in nonabsorbing regions during the calculation of transmittance, as discussed in the Supporting Information. Thus, referencing by this method removed the ability of chemometric models to distinguish and classify the conditions based upon spectral signatures, as the absorbance model PCs captured chemical variation instead of the more characteristic background variation capturable in the single-beam spectra.

Partial Least Squares Regression. PLS models were constructed for absorbance and single-beam spectra for the prediction of neodymium and praseodymium molar concentration. Three latent variables were selected due to the presence of multiple absorbing species within the calibration set.

The RMSE results for neodymium and praseodymium are shown in Table 3. RMSEs represent the uncertainty of the calibration (RMSEC) and the measurement of new samples (RMSEP).[34,38] Praseodymium, with its primary bands occurring in a region of low detector sensitivity, is more

https://doi.org/10.1021/acsomega.4c02202
ACS Omega 2024, 9, 33662−33670

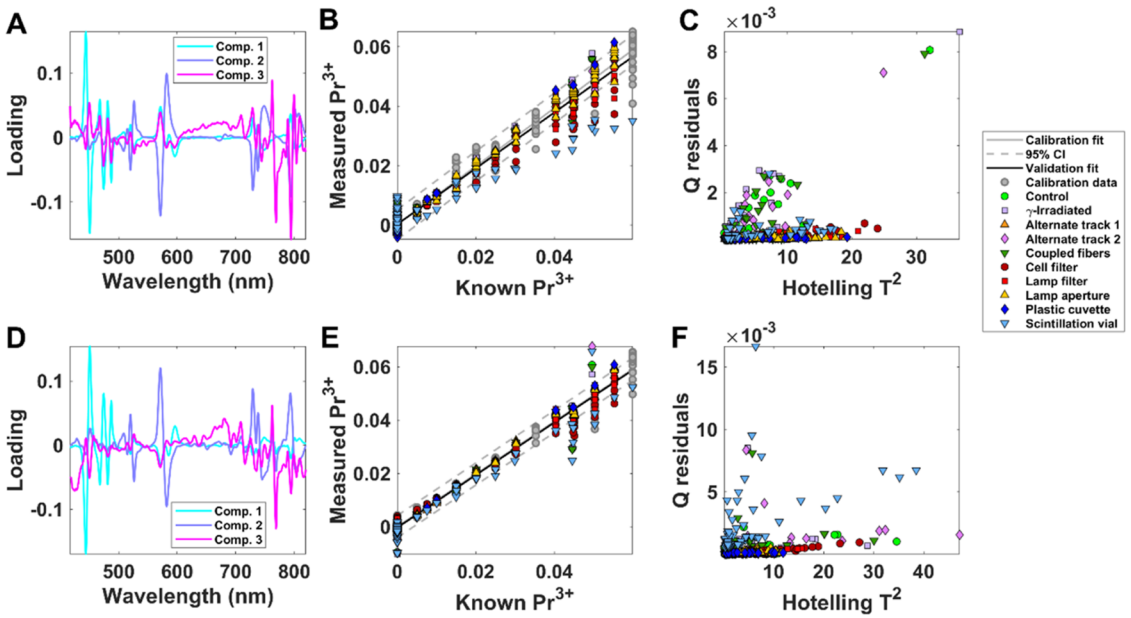


Figure 5. Results of the PLS model of the absorbance data set: (A) latent variables, (B) parity plot of praseodymium concentration as known and as measured by the model, and (C) Q-residuals versus Hotelling T-squared. Results of the PLS model of the single-beam data set: (D) latent variables, (E) parity plot of praseodymium concentration as known and as measured by the model, and (F) Q-residuals versus Hotelling T-squared. The rightmost legend applies to plots (B), (C), (E), and (F).

prone to the multiplicative effects of detector nonlinearities. Due to this, neodymium regression models tended to have lower RMSE values than praseodymium models.

Due to the preprocessing methods utilized to remove baseline variation in absorbance spectra and to remove background signal in single-beam spectra, the absorbance and single-beam models produce nearly equivalent errors in all models for RMSEC, RMSECV, and RMSEP. The results of the predictions can be seen in the parity plots in Figure 5B for the absorbance model and in (E) for the single-beam model for praseodymium. An ideal fit in these plots would show the data points falling on a 1:1 line where the known and predicted values are identical. The R-squared value of the calibration fit was 0.98 or above in both models, and the R-squared value of the validation fit was 0.96 or above. Parity plots for both praseodymium and neodymium models can be found in Figure S14.

The plots for PLS models show the stacking of data points. This occurs where identical samples were measured at multiple attenuation levels. Due to the multiplicative effects present in both the absorbance and single-beam spectra, the predicted value from spectra at higher attenuations is lower than the true value due to the recorded peak magnitude being lower in those cases.

The latent variable profiles in Figure 5A,D appear similar in intensity and shape, though the single-beam model's latent variables are inverted. Despite light source artifacts in the single-beam data set, seen primarily above 700 nm, its predictive ability remains competitive with models of traditional absorbance spectra. A more detailed discussion of the latent variables is in the Supporting Information.

In the absorbance data set, the Q-residuals, shown in Figure 5C, display high Q values for a large number of Control and γ-Irradiated condition samples, which are not expected to have high variance, as both conditions are middling in light intensity compared to the rest of the data set. In the case of the single-beam models, the Q-residuals, shown in Figure 5F, are high for the scintillation vial samples, which were intentionally excluded from the calibration set and which were collected in cells with a path length and geometry different than the rest of the data set. Thus, these samples were intentionally deviant from the data set, and the single-beam model's ability to identify these samples as aberrant provides a useful tool for flagging abnormal and undesirable conditions in process monitoring applications.

To improve models and lower prediction error, two approaches are readily available. First, additional multiplicative filters could be tested for the data sets. While EMSC is shown here and elsewhere[39] to be effective for correcting nonlinear effects, other approaches have been designed for specific spectral challenges and for reducing computational requirements for large data sets, which are common in PAT applications.[15] Second, nonlinear models can be utilized. For instance, neural networks have been applied to absorbance data sets where referencing poses a challenge.[40] However, locally weighted regression, a simpler type of nonlinear model that creates local PCR or PLS models, may be more desirable for applications wherein each computational step must be clearly identified.

CONCLUSIONS

In this study, it was demonstrated that chemometric models of single-beam spectra, after signal processing, provide an equivalent quantitative value as do models of absorbance spectra. Single-beam spectral models contain additional information that allows for qualitative analysis of the instrumental conditions during spectral measurement.

Referenced spectroscopic methods are challenged by systems where references cannot be reliably collected, either due to the nonexistence of a suitable reference material or due to prohibitive cost or inconvenience of reference spectra collection. Single-beam spectra, the signal produced by a

https://doi.org/10.1021/acsomega.4c02202
ACS Omega 2024, 9, 33662−33670

detector as photons impact it, have been shown in this work to be usable in place of absorbance when signal referencing becomes impractical. Here, signal processing produces similar spectral signatures from single-beam spectra compared with traditional absorbance spectra. Through multivariate, linear chemometric modeling, comparable errors in the prediction of the chromophore concentration are achieved from models of single-beam and absorbance spectra. The prediction errors remain comparable even when uncalibrated absorbers and sampling conditions are included in the data sets. The prediction of the concentration of broad absorbers poses additional challenges in the signal processing step, and future studies will examine methods by which broad-range signals can be extracted from single-beam data without the need for the reference collection. PCA's ability to distinguish instrumental conditions from single-beam spectra indicates that classification models may be used in a hierarchical modeling structure to first assign a condition class to a spectrum and regress analyte concentrations from the spectrum based on a calibration model built for the specific class, providing reduced prediction errors; future work may explore such avenues.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.4c02202.

The SI contains details about the instrumental conditions used to create unique irradiation spectra during the collection of visible-range spectra. It contains plots of pure component spectra of each chromophore within the detected range. It contains a discussion and several supporting figures regarding the univariate and linear analysis of absorbance and single-beam visible spectra, as well as plots of each type of spectra. It also contains further details on the signal processing methods used prior to the calculation of chemometric models. Further results of PCA and PLS models are presented (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Hope E. Lackey − *Pacific Northwest National Laboratory, Richland, Washington 99352, United States; Department of Chemistry, Washington State University, Pullman, Washington 99164, United States*; Phone: 1 (509) 372-6205; Email: hope.lackey@pnnl.gov

Samuel A. Bryan − *Pacific Northwest National Laboratory, Richland, Washington 99352, United States; Department of Chemistry, Washington State University, Pullman, Washington 99164, United States;* orcid.org/0000-0002-8826-0880; Phone: 1 (509) 375 5648; Email: sam.bryan@pnnl.gov

Amanda M. Lines − *Pacific Northwest National Laboratory, Richland, Washington 99352, United States; Department of Chemistry, Washington State University, Pullman, Washington 99164, United States*; Phone: 1 (509) 375 5689; Email: amanda.lines@pnnl.gov

Authors

Gilbert L. Nelson − *Department of Chemistry, The College of Idaho, Caldwell, Idaho 83605, United States*

Heather M. Felmy − *Pacific Northwest National Laboratory, Richland, Washington 99352, United States;* orcid.org/0000-0002-2548-7690

Xiaofeng Guo − *Department of Chemistry, Washington State University, Pullman, Washington 99164, United States;* orcid.org/0000-0003-3129-493X

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.4c02202

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Research was supported by the U.S. Department of Energy, Office of Nuclear Energy, and was performed at the Pacific Northwest National Laboratory, a multiprogram national laboratory operated by Battelle for the U.S. Department of Energy. Pacific Northwest National Laboratory is operated by Battelle Memorial Institute for the U.S. Department of Energy under contract DE-AC05-76RL01830.

■ REFERENCES

(1) Berg, T. H. A.; Ottosen, N.; van den Berg, F.; Ipsen, R. Inline UV-Vis spectroscopy to monitor and optimize cleaning-in-place (CIP) of whey filtration plants. *LWT* **2017**, *75*, 164−170.

(2) Burr, T.; Budlong-Sylvester, K.; Myers, K.; Demuth, S.; Bakel, A.; Krebs, J.; Bryan, S.; Orton, C.; Ehringer, M.; Garcia, H. Roles for process monitoring in nuclear safeguards at aqueous reprocessing plants. *JNMM* **2012**, *40* (2), 42.

(3) Wen, X. D.; Yang, Q. L.; Yan, Z. D.; Deng, Q. W. Determination of cadmium and copper in water and food samples by dispersive liquid-liquid microextraction combined with UV−vis spectrophotometry. *Microchem. J.* **2011**, *97* (2), 249−254.

(4) Cipiti, B.; McDaniel, M.; Bryan, S.; Pratt, S.. *Cost Effective Process Monitoring using UV−VIS−NIR Spectroscopy. IAEA-CN-−220*, International Atomic Energy Agency (IAEA), 2015.

(5) (a) Chen, J.; Wang, D. N.; Ramachandran, A.; Chandran, S.; Li, M. Z.; Varma, R. An open-path dual-beam laser spectrometer for path-integrated urban NO2 sensing. *Sensors Actuat A:Phys.* **2020**, *315*, No. 112208. (b) Dupont, M. F.; Elbourne, A.; Cozzolino, D.; Chapman, J.; Truong, V. K.; Crawford, R. J.; Latham, K. Chemometrics for environmental monitoring: a review. *Anal. Methods* **2020**, *12* (38), 4597−4620.

(6) (a) Chew, W.; Sharratt, P. Trends in process analytical technology. *Anal. Methods* **2010**, *2* (10), 1412−1438. (b) Chen, Z.-P.; Fevotte, G.; Caillet, A.; Littlejohn, D.; Morris, J. Advanced Calibration Strategy for in Situ Quantitative Monitoring of Phase Transition Processes in Suspensions Using FT-Raman Spectroscopy. *Anal. Chem.* **2008**, *80* (17), 6658−6665.

(7) Rolinger, L.; Rüdt, M.; Hubbuch, J. A critical review of recent trends, and a future perspective of optical spectroscopy as PAT in biopharmaceutical downstream processing. *Anal. Bioanal. Chem.* **2020**, *412* (9), 2047−2064.

(8) Tse, P.; Bryan, S. A.; Bessen, N. P.; Lines, A. M.; Shafer, J. C. Review of on-line and near real-time spectroscopic monitoring of processes relevant to nuclear material management. *Anal. Chim. Acta* **2020**, *1107*, 1−13.

(9) Schroll, C. A.; Lines, A. M.; Heineman, W. R.; Bryan, S. A. Absorption spectroscopy for the quantitative prediction of lanthanide concentrations in the 3LiCl−2CsCl eutectic at 723 K. *Anal. Methods* **2016**, *8* (43), 7731−7738.

(10) Lackey, H. E.; Nelson, G. L.; Lines, A. M.; Bryan, S. A. Reimagining pH Measurement: Utilizing Raman Spectroscopy for Enhanced Accuracy in Phosphoric Acid Systems. *Anal. Chem.* **2020**, *92* (8), 5882−5889.

(11) Havlik, I.; Beutel, S.; Scheper, T.; Reardon, K. F. On-Line Monitoring of Biological Parameters in Microalgal Bioprocesses Using Optical Methods. *Energies* **2022**, *15* (3), 875.
(12) Gredilla, A.; Fdez-Ortiz de Vallejuelo, S.; Elejoste, N.; de Diego, A.; Madariaga, J. M. Non-destructive Spectroscopy combined with chemometrics as a tool for Green Chemical Analysis of environmental samples: A review. *TrAC-Trend Anal. Chem.* **2016**, *76*, 30−39.
(13) Chatterjee, S.; Bryan, S. A.; Schroll, C. A.; Heineman, W. R. Method and Apparatus for Simultaneous Spectroelectrochemical Analysis. US Patent 8,585,880, 2013.
(14) Debus, B.; Kirsanov, D.; Ruckebusch, C.; Agafonova-Moroz, M.; Babain, V.; Lumpov, A.; Legin, A. Restoring important process information from complex optical spectra with MCR-ALS: Case study of actinide reduction in spent nuclear fuel reprocessing. *Chemometr. Intell. Lab. Syst.* **2015**, *146*, 241−249.
(15) Chen, Z. P.; Morris, J.; Martin, E. Extracting Chemical information from spectral data with multiplicative light scattering effects by optical path-length estimation and correction. *Anal. Chem.* **2006**, *78* (22), 7674−7681.
(16) Armenta, S.; Garrigues, S.; de la Guardia, M. Recent developments in flow-analysis vibrational spectroscopy. *TrAC-Trend Anal. Chem.* **2007**, *26* (8), 775−787.
(17) Lu, G.; Zhou, X.; Arnold, M. A.; Small, G. W. Multivariate calibration models based on the direct analysis of near-infrared single-beam spectra. *Appl. Spectrosc.* **1997**, *51* (9), 1330−1339.
(18) Defernez, M.; Wilson, R. H. Infrared spectroscopy: Instrumental factors affecting the long-term validity of chemometric models. *Anal. Chem.* **1997**, *69* (7), 1288−1294.
(19) Feudale, R. N.; Woody, N. A.; Tan, H.; Myles, A. J.; Brown, S. D.; Ferré, J. Transfer of multivariate calibration models: a review. *Chemometr. Intell. Lab.* **2002**, *64* (2), 181−192.
(20) Chen, Z. P.; Li, L. M.; Yu, R. Q.; Littlejohn, D.; Nordon, A.; Morris, J.; Dann, A. S.; Jeffkins, P. A.; Richardson, M. D.; Stimpson, S. L. Systematic prediction error correction: A novel strategy for maintaining the predictive abilities of multivariate calibration models. *Analyst* **2011**, *136* (1), 98−106.
(21) Mayerhöfer, T. G.; Mutschke, H.; Popp, J. Employing theories far beyond their limits—The case of the (Boguer-) Beer−Lambert Law. *ChemPhysChem* **2016**, *17* (13), 1948−1955.
(22) Ding, Q.; Small, G. W.; Arnold, M. A. Evaluation of Data Pretreatment and Model Building Methods for the Determination of Glucose from Near-Infrared Single-Beam Spectra. *Appl. Spectrosc.* **1999**, *53* (4), 402−414.
(23) Zawisza, B.; Pytlakowska, K.; Feist, B.; Polowniak, M.; Kita, A.; Sitko, R. Determination of rare earth elements by spectroscopic techniques: a review. *J. Anal. Atom. Spectrom.* **2011**, *26* (12), 2373−2390.
(24) Kirsanov, D.; Panchuk, V.; Goydenko, A.; Khaydukova, M.; Semenov, V.; Legin, A. Improving precision of X-ray fluorescence analysis of lanthanide mixtures using partial least squares regression. *Spectrochim Acta B* **2015**, *113*, 126−131.
(25) Akhmetzhanov, T. F.; Popov, A. M. Direct determination of lanthanides by LIBS in REE-rich ores: comparison between univariate and DoE based multivariate calibrations with respect to spectral resolution. *J. Anal. Atom. Spectrom.* **2022**, *37* (11), 2330−2339.
(26) Hao, S. L.; Shao, L. M. Determining the number of principal factors by eigenvector comparison of the original bi-linear data matrix and the one reconstructed from key variables. *Chemometr. Intell. Lab.* **2015**, *149*, 17−23.
(27) Brereton, R. G.; Jansen, J.; Lopes, J.; Marini, F.; Pomerantsev, A.; Rodionova, O.; Roger, J. M.; Walczak, B.; Tauler, R. Chemometrics in analytical chemistry—Part II: Modeling, validation, and applications. *Anal. Bioanal. Chem.* **2018**, *410* (26), 6691−6704.
(28) Lackey, H. E.; Sell, R. L.; Nelson, G. L.; Bryan, T. A.; Lines, A. M.; Bryan, S. A. Practical Guide to Chemometric Analysis of Optical Spectroscopic Data. *J. Chem. Educ.* **2023**, *100* (7), 2608−2626.
(29) Lourenço, N. D.; Lopes, J. A.; Almeida, C. F.; Sarraguca, M. C.; Pinheiro, H. M. Bioreactor monitoring with spectroscopy and chemometrics: a review. *Anal. Bioanal. Chem.* **2012**, *404* (4), 1211−1237.
(30) Wise, B. M.; Gallagher, N. B. The process chemometrics approach to process monitoring and fault detection. *J. Process Contr.* **1996**, *6* (6), 329−348.
(31) (a) Swinehart, D. F. The Beer-Lambert Law. *J. Chem. Educ.* **1962**, *39* (7), 333. (b) Coachsea. Derivation of beers law. YouTube, 2011.Skoog, D. A.; West, D. M. *Analytical Chemistry*; Holt, Rinehart and Winston, Inc., 1974.
(32) (a) Pinkerton, R. C. Beer's law without calculus. *J. Chem. Educ.* **1964**, *41* (7), 366. (b) Sánchez, H. R. Seven derivations of the Beer-Lambert law. *Spectrosc. Lett.* **2021**, *54* (2), 133−139.
(33) Mayerhöfer, T. G.; Popp, J. Beer's Law − Why absorbance depends (almost) linearly on concentration. *ChemPhysChem* **2019**, *20* (4), 511−515.
(34) Westad, F.; Marini, F. Validation of chemometric models − A tutorial. *Anal. Chim. Acta* **2015**, *893*, 14−24.
(35) Savitzky, A.; Golay, M. J. E. Smoothing + differentiation of data by simplified least squares procedures. *Anal. Chem.* **1964**, *36* (8), 1627.
(36) Martens, H.; Høy, M.; Wise, B. M.; Bro, R.; Brockhoff, P. B. Pre-whitening of data by covariance-weighted pre-processing. *J. Chemometr.* **2003**, *17* (3), 153−165.
(37) Afseth, N. K.; Kohler, A. Extended multiplicative signal correction in vibrational spectroscopy, a tutorial. *Chemometr. Intell. Lab* **2012**, *117*, 92−99.
(38) Næs, T.; Martens, H. Multivariate calibration. II. Chemometric methods. *TrAC-Trend Anal. Chem.* **1984**, *3* (10), 266−271.
(39) (a) Martens, H.; Nielsen, J. P.; Engelsen, S. B. Light scattering and light absorbance separated by extended multiplicative signal correction. application to near-infrared transmission analysis of powder mixtures. *Anal. Chem.* **2003**, *75* (3), 394−404. (b) Martens, H.; Stark, E. Extended multiplicative signal correction and spectral interference subtraction: New preprocessing methods for near infrared spectroscopy. *J. Pharmaceut. Biomed.* **1991**, *9* (8), 625−635.
(40) Magnussen, E. A.; Solheim, J. H.; Blazhko, U.; Tafintseva, V.; Tøndel, K.; Liland, K. H.; Dzurendova, S.; Shapaval, V.; Sandt, C.; Borondics, F.; Kohler, A. Deep convolutional neural network recovers pure absorbance spectra from highly scatter-distorted spectra of cells. *J. Biophotonics* **2020**, *13* (12), No. e202000204.

https://doi.org/10.1021/acsomega.4c02202
ACS Omega 2024, 9, 33662−33670