

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/7363133>

Improving the Linearity of Infrared Diffuse Reflection Spectroscopy Data for Quantitative Analysis: An Application in Quantifying Organophosphorus Contamination in Soil

ARTICLE *in* ANALYTICAL CHEMISTRY · FEBRUARY 2006

Impact Factor: 5.64 · DOI: 10.1021/ac0509859 · Source: PubMed

CITATIONS

12

READS

26

6 AUTHORS, INCLUDING:



Avishai Ben-David

Edgewood Chemical Biological Center, Aber...

90 PUBLICATIONS 566 CITATIONS

SEE PROFILE

Improving the Linearity of Infrared Diffuse Reflection Spectroscopy Data for Quantitative Analysis: An Application in Quantifying Organophosphorus Contamination in Soil

Alan C. Samuels,^{*,†} Changjiang Zhu,[‡] Barry R. Williams,[§] Avishai Ben-David,[†] Ronald W. Miles, Jr.,[†] and Melissa Hulet[§]

Edgewood Chemical Biological Center, Aberdeen Proving Ground, Maryland 21010, Battelle Eastern Science and Technology Center, 1204 Technology Drive, Aberdeen, Maryland 21001, and EAI Corporation, 1308 Continental Drive, Abingdon, Maryland 21009

Diffuse reflection data are presented for ethyl methylphosphonate in a fine Utah dirt sample as a model system for organophosphate-contaminated soil. The data revealed a chemometric artifact when the spectra were represented in Kubelka–Munk units that manifests as a linear dependence of spectral peak height on variations in the observed baseline position (i.e., the position of the observed transmission intensity where no absorption features occur in the sample spectrum). We believe that this artifact is the result of the mathematical process by which the raw data are converted into Kubelka–Munk units, and we developed a numerical strategy for compensating for the observed effect and restoring chemometric precision to the diffuse reflection data for quantitative analysis while retaining the benefits of linear calibration afforded by the Kubelka–Munk approach. We validated our Kubelka–Munk correction strategy by repeating the experiment using a simpler system—pure caffeine in potassium bromide. The numerical preprocessing includes conventional multiplicative scatter correction coupled with a baseline offset correction that facilitates the use of quantitative diffuse reflection data in the Kubelka–Munk formalism for the quantitation of contaminants in a complex soil matrix, but is also applicable to more fundamental diffuse reflection quantitative analysis experiments.

Diffuse reflection (DR) spectroscopy (also referred to as diffuse reflectance infrared Fourier transform spectroscopy) has realized a broad range of applications in analytical chemistry. Both qualitative and quantitative work using this infrared absorption measurement method have been reported, but in the case of mid-infrared DR spectroscopy, the majority of the applications are qualitative in nature, often turning to multivariate analysis to ascertain various properties of materials under study. One reason for the paucity of studies that apply the DR method to quantitative analysis problems stems from the difficulty with which linearity in observed peak height with analyte concentration is realized with

accuracy, particularly in mid-infrared studies. The most common approach for analyzing DR data is to apply the Kubelka–Munk (KM) formalism to the absorption data.¹ The KM formalism provides a means for representing the absorption data in DR experiments such that in the limit of certain assumptions, including an infinite sample depth and isotropic scattering, reasonable linearity is achieved as a function of sample concentration. However, while the KM approach is attractive by virtue of its simplicity, it has certain known deficiencies. Band intensities as evaluated under the KM method are known to vary with baseline position error, often giving rise to unacceptable uncertainty levels in quantitative experiments. To avoid such errors, the DR data are often expressed in $\log(1/R)$ units, where R is the reflectance measurements instead of KM units.² Indeed, the $\log(1/R)$ band intensities do not vary with baseline offset error; however, the linearity in $\log(1/R)$ units is much poorer than that achievable when using the KM formalism. This is to be expected since the intrinsic radiative-transfer processes are known to be nonlinear. Linearity with analyte concentration is merely a convenience for quantitative analysis, but one which makes data reduction and interpretation much more facile. We have arrived at an alternative approach for quantitative analysis of DR data that takes advantage of an apparent relationship between the actual baseline position observed in the DR experiment and the magnitude of the observed band intensities. This approach allows the DR data to be analyzed quantitatively in KM units while at least partially rectifying the uncertainties that result as a function of baseline offset error in the DR experiment.

The application of DR spectroscopy for quantitative analysis spans a wide variety of applications, including the assessment of stearate coatings on magnesium hydroxide and calcium carbonate fillers,³ phosphate coatings on talc,⁴ surface-derivatized silica powders,⁵ adsorption densities of alcohols onto silica powders,⁶ siloxane deposition on keratin fiber surfaces,⁷ and pharmaceutical

* To whom correspondence should be addressed. E-mail: alan.samuels@us.army.mil.

[†] Edgewood Chemical Biological Center.

[‡] Battelle Eastern Science and Technology Center.

[§] EAI Corp.

(1) Kubelka, P.; Munk, F. Z. *Tech. Phys.* **1931**, *12*, 593–601.

(2) Griffiths, P. J. J. *Near Infrared Spectrosc.* **1995**, *3*, 60–62.

(3) Gilbert, M.; Petiraksakul, P.; Mathieson, I. *Mater. Sci. Technol.* **2001**, *17* (11), 1472–1478.

(4) Liu, Z.; Gilbert, M. J. *Appl. Polym. Sci.* **1996**, *59* (7), 1087–1098.

(5) Boroumand, F.; van den Bergh, H.; Moser, J. M. *Anal. Chem.* **1994**, *66* (14) 2260–2266.

(6) Horr, T. J.; Ralston, J.; St. Smart, R. C. *Colloids Surf.* **1992**, *64* (1), 67–85.

tablets.^{8,9} We also note that several careful experimental and theoretical studies of the application of DR in quantitative analysis have been reported. Mattoo and Nabar studied the application of quantitative DR to the analysis of soils,¹⁰ then extended the study to systematically characterize the effect of highly reflecting and highly absorbing analytes,¹¹ and reported on the additivity and linearity of the KM functions in these cases. Schober and Löhmansröben reported on the DR spectroscopy of soil-related particulate materials and some real soil samples and elucidated the relative contributions of scattering versus absorbance to the spectral reflectance in the ultraviolet through the near-infrared wavelength region.¹² They found that significant differences in the absorbance and scattering terms lead to strong variations in the light penetration depth. Reeves et al. have extensively studied soils using the DR method, typically employing multivariate analysis tools to estimate soil parameters such as carbon and nutrient content.^{13,14} Hecht compared data from some typical experiments involving soluble dyes in scattering media using several continuum models including the KM formalism, the Pitts two-parameter representation, and the Rozenberg function as well as exact solutions arrived at via radiative-transfer computations.¹⁵ The Pitts formula was determined to give the best fit to the experimental data, and the author concluded that the representation better compensates for scatter anisotropy. Carroll and Doyle reported on a variable KM correction and assessed its performance as well as the accuracy with which spectral subtraction was possible.⁹ Hamadeh et al. discussed several methods for the preparation of calibration standards for quantitative DR spectroscopy and noted that particularly for analytes that have a hardness very different from the potassium bromide diluent it is important to grind the analyte and diluent separately and then thoroughly mix the resulting powders.¹⁶ Mandelis and co-workers¹⁷ critically reviewed a 1963 statistical theory of the absolute diffuse reflectance of powders promulgated by Melamed¹⁸ and later developed a self-consistent discontinuum theory to determine optical absorption coefficients from both reflectance and transmittance measurement data.¹⁹ Hecht pointed out that the KM function should be regarded as a limiting case, applicable only to dilute solutions under the inherent (and in most cases, faulty) assumption that homogeneous scattering layers afford isotropic and monotonic scattering, and absorption and anomalous dispersion effects are absent from the measured data.¹⁵ Boroumand and co-workers described an adapta-

tion of the KM theory to account for nonideal scattering in undiluted samples.²⁰ Krivacsy and Hlavay compared four different calibration methods for quantitative DR infrared spectroscopy and recommended multiple calibration standards plus a reference reflectance measurement.²¹ Dahm and Dahm suggested that linear KM plots are obtained more readily when a volume percent metric is used in lieu of the more common mass percent metric.²² Griffiths responded to the latter communication with an argument in favor of the use of $\log(1/R)$ for more reproducibility in quantitative DR measurements because as noted in the introduction, unlike the KM function, the $\log(1/R)$ approach does not yield baseline position dependence in the spectral band intensities.² Pedersen et al. used a forward model where regressions of mean spectra onto the measured spectra are computed using least-squares analysis.²³ This regression attempts to reduce the spread (uncertainty) in the computed average spectra and force all individual measurements to a common reference spectrum. This approach, dubbed extended inverted signal correction, attempts to correct for the additive baseline offset effect and a multiplicative scaling effect. The additive effect can be tied to scattering correction and the multiplicative effect can be tied to absorbance correction. Mattia et al. recently reported on an unusual enhancement in the observed band intensities of analytes embedded in alkali halide powders on exposure to water vapor and concluded that as the powders become hydrated their surfaces become smooth and the scattering factor drops to yield an enhancement in the measured KM spectrum.²⁴

We use both real and simulated DR data in this work to describe the effect of baseline offset on the observed band intensities and arrive at a recommended mathematical correction for baseline offset errors that reduces the statistical uncertainty in quantitative results estimated with the DR method. We demonstrate the utility of our approach by analyzing a relatively complicated problem, the assessment of levels of contamination of an organophosphonate agent in a soil matrix.

We are motivated by a need to understand the optical properties of contaminated soils in a continuing effort to evaluate spectroscopic approaches for rapid surface contamination monitoring applications. The problem of identifying hazardous contaminants such as persistent chemical warfare agents on surfaces such as sand and soil is of particular concern in our research. Diffuse reflection provides an information-rich optical method for the rapid assessment of chemical contamination and further provides one means of identifying contaminants. We engaged in a study of the DR method applied to the analysis of organophosphate contamination in a soil matrix for the express purpose of assessing the capacity of the method to (1) detect the presence of a given contaminant, (2) identify the contaminant, and (3) determine the extent of contamination by quantifying the amount of contaminant present. The study employed ethyl methylphosphonic acid (EMPA) in a soil matrix as a model for a ground contaminated by a persistent organophosphorus nerve agent.²⁵ We noted in that study

- (7) Klimisch, H. M.; Kohl, G. S.; Sabourin, J. M. *J. Soc. Cosmet. Chem.* **1987**, 38 (4), 247–262.
- (8) Yeboah, S. A.; Yang, W. J.; Griffiths, P. R. *Proc. SPIE—Int. Soc. Opt. Eng.* **1981**, 289, 118–120.
- (9) Carroll, J. E.; Doyle, W. M. *Proc. SPIE—Int. Soc. Opt. Eng.* **1981**, 289, 111–113.
- (10) Mattoo, B. N.; Nabar, B. S. *Indian J. Chem.* **1968**, 6 (1), 39–42.
- (11) Mattoo, B. N. *Indian J. Chem.* **1968**, 6 (7), 364–366.
- (12) Schober, L.; Löhmansröben, H.-G. *J. Environ. Monit.* **2000**, 2, 651–655.
- (13) Reeves, J. B., III; McCarty, G. W.; Reeves, V. B. *J. Agric. Food Chem.* **2001**, 49 (2), 766–772.
- (14) Reeves, J. B., III; Francis, B. A.; Hamilton, S. K. *Appl. Spectrosc.* **2005**, 59 (1), 39–46.
- (15) Hecht, H. G. *Anal. Chem.* **1976**, 48 (12), 1775–1779.
- (16) Hamadeh, I. M.; Yeboah, S. A.; Trumbull, K. A.; Griffiths, P. R. *Appl. Spectrosc.* **1984**, 38 (4), 486–491.
- (17) Mandelis, A.; Boroumand, F.; Vanden Bergh, H. *Appl. Opt.* **1990**, 29 (19), 2853–2860.
- (18) Melamed, N. T. *J. Appl. Phys.* **1963**, 34, 560–570.
- (19) Mandelis, A.; Boroumand, F.; Vanden Bergh, H. *Spectrochim. Acta, Part A* **1991**, 47A (7), 943–971.

- (20) Boroumand, F.; Moser, J. E.; Van den Bergh, H. *Appl. Spectrosc.* **1992**, 46 (12), 1874–1886.
- (21) Krivacsy, Z.; Hlavay, J. *Talanta* **1995**, 42 (4), 613–620.
- (22) Dahm, D. J.; Dahm, K. D. *J. Near Infrared Spectrosc.* **1995**, 3, 53–59.
- (23) Pedersen, D. K.; Martens, H.; Nielsen, J. P. *Appl. Spectrosc.* **2002**, 56, 1206–1214.
- (24) Mattia, J.; Painter, P. C.; Heaps, D. A.; Griffiths, P. R. *Appl. Spectrosc.* **2005**, 59 (1), 140–143.

that even though the absorption of the soil dominates the observed spectrum, soil particles tainted by liquid contaminants can exhibit a significant specular component due to the smooth surface and intrinsic index of refraction of the liquid contaminant on the soil particles. The specular component presented by liquids in an infrared reflection measurement produces anomalous dispersion effects on front surface reflection, which can give rise to the distortion, inversion, or complete disappearance of the expected absorption bands in the material.²⁶ On the other hand, chemicals adsorbed in a rough solid substrate such as sand or concrete may effectively create a composite, solidlike material that would be well suited to study using the DR technique by virtue of path length effects that enhance the observed DR signal of the analyte. Enhanced infrared absorption has been reported in the study of endolithic bacteria on the surface of rough rock.²⁷ Because many soils and clays are porous, powdered materials, they would be expected to produce volume-scattering phenomena, which may actually enhance the sensitivity of the optical spectroscopic measurement by virtue of effectively increasing the interaction path length with the incident radiation. Nevertheless, such materials present a complicated spectral matrix against which it may be difficult to recover the spectral features of the analyte.

Our previous results²⁵ demonstrated that within certain limits it is possible to partially recover the spectrum of the contaminant from that of the contaminated soil mixture by applying straightforward spectral subtraction techniques. We attributed the failure of the spectral subtraction method in recovering the entire neat EMPA spectrum to a number of factors, including anomalous dispersion and soil particle size-dependent scattering. However, we also noted during the systematic study that the absorption intensity for those bands that were recoverable was sensitive to measurement-to-measurement baseline position error in a nearly linear fashion. That is to say, the larger the baseline offset (from zero) observed, the larger the observed band intensity. This observation was a key motivation for the present study, and those data are revisited herein to assess this phenomenon and suggest a corrective measure with general application to other quantitative DR methods. To better understand the phenomenology, we studied a single-component system (caffeine powder) in addition to the multicomponent contaminated soil systems.

EXPERIMENTAL SECTION

Samples of "contaminated" soil were prepared by gravimetric addition of the ethyl methylphosphonate to the soil sample at a 1:10 w/w ratio. Diffuse reflection samples were prepared by weighing the analyte ("contaminated" soil) in a vial and then adding potassium bromide powder and again weighing the mixture to compute the mass-mass ratio of the analyte. All weighings were performed using a Sartorius model BP211D electronic analytical balance (Sartorius AG, Goettingen, Germany), which has a stated precision to the ± 0.02 mg level. We estimated the actual uncertainty in the balance measurement under our laboratory conditions of variable humidity to be ± 0.2 mg. Caffeine

(CAS No. 58-03-2, 99% purity) and ethyl methylphosphonate (EMPA, CAS No. 1832-53-7, 98% purity) were obtained from Aldrich (Sigma-Aldrich, St. Louis, MO) and used as received. Potassium bromide was obtained from Aldrich, ground in a mortar and pestle, and dried in a laboratory oven at 100 °C for at least 4 h prior to the sample preparation. A sample of fine topsoil was collected from a test site in the U.S. Army Dugway Proving Ground (DPG, UT) range. We characterized the particle size distribution of the DPG soil using an electron microscope and observed a roughly log-normal distribution in particle diameters with a mode around 2–4 μm and a broad tail indicating a relatively low (<10%) population of larger particle sizes of diameters around 24 μm .²⁵ Additional soil samples were obtained from the National Institutes of Standards and Technologies Standard Reference Material (NIST SRM) program (NIST, Gaithersburg, MD): NIST SRM 1944, New York–New Jersey Waterway Sediment. We mixed the powdered samples by shaking the sealed vial by hand or by transferring the powder mixtures to a Wig-L-Bug accessory (Thermo-Electron Corp., Madison, WI) and shaking them for a minimum of 2 min.

Prepared sample mixtures were loosely packed into the sample well in the carriage of a Thermo-Nicolet SmartCollector compound parabolic concentrator (CPC) diffuse reflectance accessory (Thermo-Electron Corp.), while pure KBr was added to the reference well, and the powders were leveled using the flat edge of a spatula in accordance with the manufacturer's recommended method. Systematic studies performed by TeVrucht and Griffiths describe the effect of sample packing for diffuse reflectance measurement;²⁸ we chose to adhere to the routine sample preparation method described here as it represents the most common mode of sample preparation for diffuse reflection measurements. The SmartCollector was mounted in a Thermo-Nicolet model 670 Nexus Fourier transform infrared spectrometer (Thermo-Electron Corp.). All experimental data were collected on this instrument by taking the ratio of the single-beam sample spectrum to the pure KBr reference well. Interferograms were acquired with a mercury–cadmium–telluride wide-range (MCT/B) detector using the following instrumental parameters: 4-cm⁻¹ resolution, 128 scans, 4 \times gain, and 2 \times oversampling. Single-sided interferograms were phase corrected, apodized, and Fourier transformed using the instrument's Omnic software. A Mertz phase correction and a Happ-Genzel apodization function were applied with no zero-filling to yield the single-beam spectra.

A simple model to interpret our measured results describes the observed peak height in KM units as a function of the observed baseline offset as follows. We start by defining the KM function as the reflectance spectrum that results when the ratio of single-beam spectrum of an infinitely thick layer of the sample dispersed in nonabsorbing matrix (e.g., powdered KBr) to that of a pure nonabsorbing matrix is calculated:

$$f(R_\infty) = \frac{(1 - R_\infty)^2}{2R_\infty} \quad (1)$$

where R_∞ is the reflection spectrum of the infinitely thick layer of sample under the KM formalism and $f(R_\infty)$ is proportional to the

(25) Samuels, A. C.; Williams, B. R.; Miles, R. W., Jr.; Zhu, C.; Ben-David, A. *Proc. SPIE* **2003**, 5085, 46–51.

(26) Coleman, P. B., Ed. *Practical Sampling Techniques for Infrared Analysis*; CRC Press: Boca Raton, FL, 1993.

(27) Holman, H.-Y. N.; Perry, D. L.; Hunter-Cevera, J. J. *Microbiol. Methods* **1998**, 34, 59–71.

(28) TeVrucht, M. L.; Griffiths, P. R. *Appl. Spectrosc.* **1989**, 43 (8), 1492–1494.

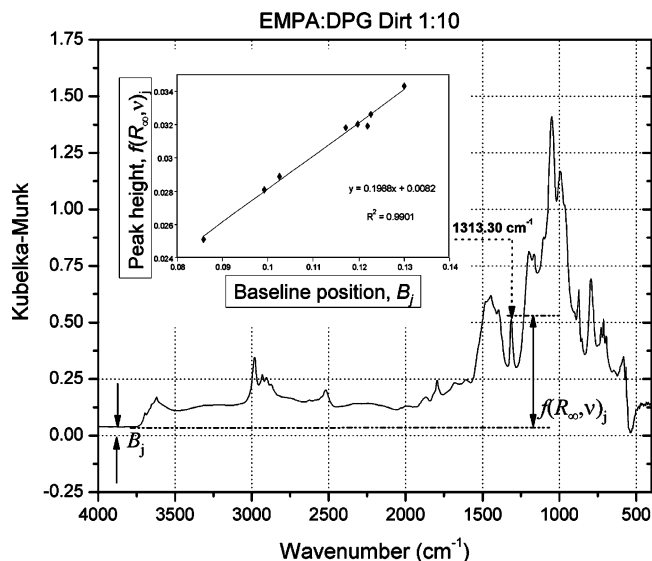


Figure 1. Kubelka–Munk spectrum of EMPA/DPG soil (1:10 w/w) 4.5% in KBr with the indications of baseline offset (B_j), observed peak intensity $f_{\text{obs}}(R_{\infty}, \nu)_j$ for 1313.3 cm^{-1} and its left baseline. B_j is averaged between 3980 and 3990 cm^{-1} . Inset is the linear plot between peak intensity and the observed position of the left baseline of 1313.3 cm^{-1} peak.

ratio of an absorption component to a scattering component. When we perform a set of j measurements of a sample or samples, we observe a range of peak heights $f_{\text{obs}}(R_{\infty}, \nu)_j$ in the KM spectrum as a function of spectral frequency ν . When we observe a nonzero baseline offset (B_j) in some of the measurements, we are tempted to use a simple baseline translation to yield a corrected value for $f(R_{\infty}, \nu)_j$:

$$f(R_{\infty}, \nu)_j = f_{\text{obs}}(R_{\infty}, \nu)_j - B_j \quad (2)$$

Instead, we observed a linear relationship (Figure 1 inset) between the KM intensity and the baseline position:

$$f(R_{\infty}, \nu)_j = f_0(R_{\infty}, \nu)_j + A(\nu)B_j \quad (3)$$

where $f_0(R_{\infty}, \nu)_j$ is hypothesized here to be the “true” peak height value for the spectral band and $A(\nu)$ is the slope of the line observed at each wavelength as a function of multiple measurements presenting varying levels of nonzero baseline offset B_j and hence varying levels of peak magnitude versus peak local minimum magnitude (i.e., the slope of the line in the inset of Figure 1). If a simple baseline offset correction is applied in the conventional sense along with multiplicative scatter correction (MSC),²⁹ the resulting plots in $\log(1/R)$ units present difficulties for quantitative analysis as seen in Figure 2. Note that the MSC method is a variance scaling method that reduces the spread in the data. The raw and MSC corrected data present nonlinear behavior and a significant nonzero y-intercept. The baseline offset correction (eq 2) improves the y-intercept at the expense of the linearity of the data on this scale. Using eq 3, we derive the

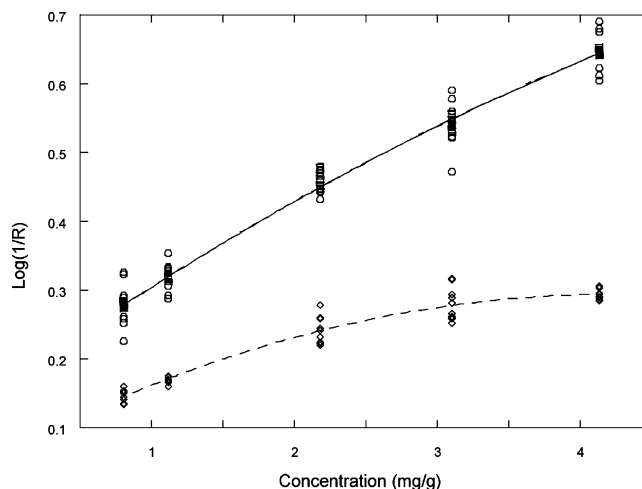


Figure 2. $\log(1/R)$ as a function of concentration of EMPA-tainted soil in KBr for 10 replicate measurements at each of five concentrations of the analyte in KBr. A polynomial fit is presented to underscore the nonlinear nature of the observed absorption in the case of the raw data (circles) and multiplicative scatter corrected (squares) and baseline correction plus multiplicative scatter correction (diamonds) treatments. Note that multiplicative scatter correction reduces measurement-to-measurement scatter, but the data present a severe nonzero intercept. The baseline peak correction plus multiplicative scatter correction improve the intercept somewhat at the expense of linearity.

following proposed correction $f'_0(\nu)_j$ to the observed spectral band intensity:

$$f'_0(R_{\infty}, \nu)_j = f(R_{\infty}, \nu)_j - A(\nu)B_j \quad (4)$$

where the difference between the derived ($f'_0(R_{\infty}, \nu)_j$) versus true ($f_0(R_{\infty}, \nu)_j$) spectral band intensity is the measurement uncertainty not accounted for by the baseline offset effect on the KM spectrum. We will refer to eq 4 as the baseline peak correction (BPC). The derived spectral intensity therefore uses the observed baseline offset effect to apply a more accurate correction to the KM spectrum than would be realized by a simple translation of the spectrum back to a zero baseline ($f'_0(R_{\infty}, \nu)_j = f_0(R_{\infty}, \nu)_j - B_j$) as would be done in routine baseline correction methods. We now compute the mean square error as a function of ν for n measurements:

$$\frac{1}{n} \sum_{j=1}^n [f(R_{\infty}, \nu)_j - \overline{f(R_{\infty}, \nu)}]^2 \quad (5)$$

and for convenience we further define a scalar metric that provides a qualitative sense of the residual mean square error over all m measured frequencies:

$$\frac{1}{nm} \sum_{\nu=1}^m \sum_{j=1}^n [f(R_{\infty}, \nu)_j - \overline{f(R_{\infty}, \nu)}]^2 \quad (6)$$

If our linear relationship in eq 3 holds true, then we should realize a reduction in the statistical uncertainty in the residual error in KM values for replicate measurements of DR samples upon applying our proposed correction, eq 4. We apply two levels of

(29) Barnes, R. J.; Dhanoa, M. S.; Lister, S. J. *Appl. Spectrosc.* **1989**, *43* (5), 772–777.

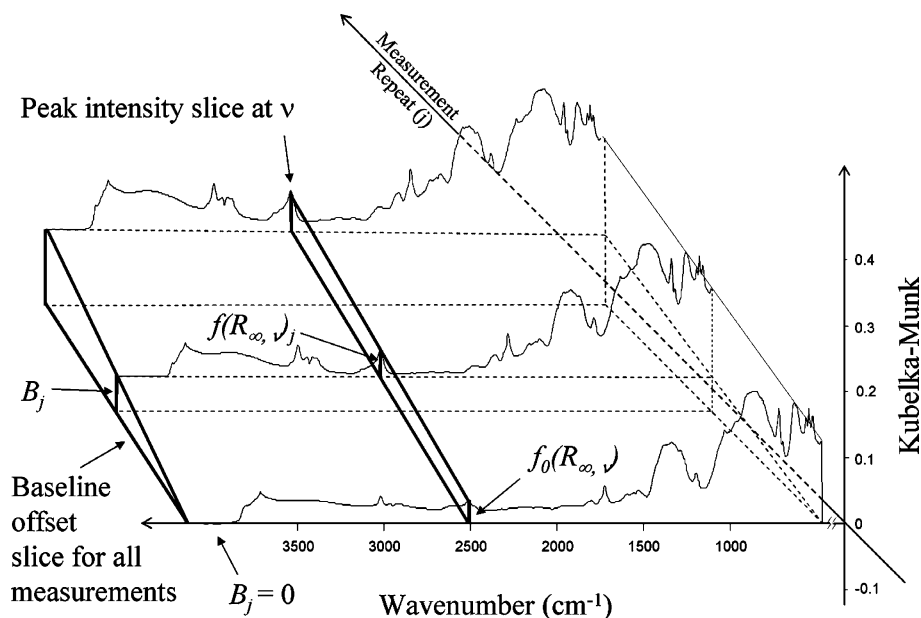


Figure 3. Schematic description of acquiring baseline offset B_j and peak intensity $f(R_{\infty}, \nu)_j$ at certain ν for one of multidiffuse reflection measurements of the same sample in Kubelka–Munk scale. In this plot, $f(R_{\infty}, \nu)_j$ approaches its hypothesized “true” value $f_0(R_{\infty}, \nu)_j$ when $B_j = 0$. Our algorithm achieves BPC on the Kubelka–Munk scale by deriving “ $f_0(R_{\infty}, \nu)_j$ ” or $f'_0(R_{\infty}, \nu)_j$ from $f(R_{\infty}, \nu)_j$ for each ν by utilizing the linear relationship between baseline offset slice and peak intensity slice at ν . We then perform the KMC operation as described in the text by applying MSC followed by BPC.

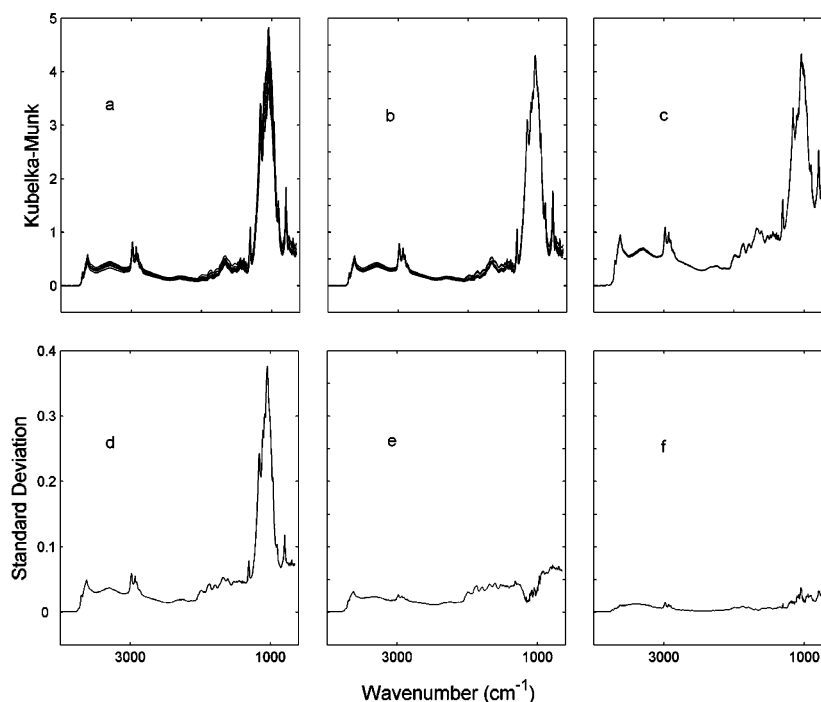


Figure 4. Overlay and corresponding standard deviation of 10 EMPA-tainted soil measurements at concentration 4.13 mg/g analyte in KBr under raw (baseline subtraction performed), MSC, and KMC. (a–c) Overlay plots of raw, MSC, and KMC; (d–f) standard deviation plots corresponding to plots a–c.

corrective processing to our DR data to evaluate the proposed relationship. First, we perform routine multiplicative scatter correction²⁹ to eliminate measurement-to-measurement variance attributable to normal measurement error (i.e., variance scaling the data). We used the MSC function in the Thermo-Electron Grams software package to perform this part of the analysis. Second, we apply our linear correction to the DR data, eq 4, to rescale the DR spectra and at least partially rectify the effect of

variance in the observed baseline position on the KM value. We define the combination of MSC with BPC as an overall KM correction (KMC) to the DR data. Figure 3 presents the conceptual basis for our derived correction to the KM spectrum.

To verify that the observed phenomena (effect of nonzero baseline on peak intensity) were indeed general to the diffuse reflection experiment rather than a peculiarity of the CPC accessory, a set of caffeine data in potassium bromide were also

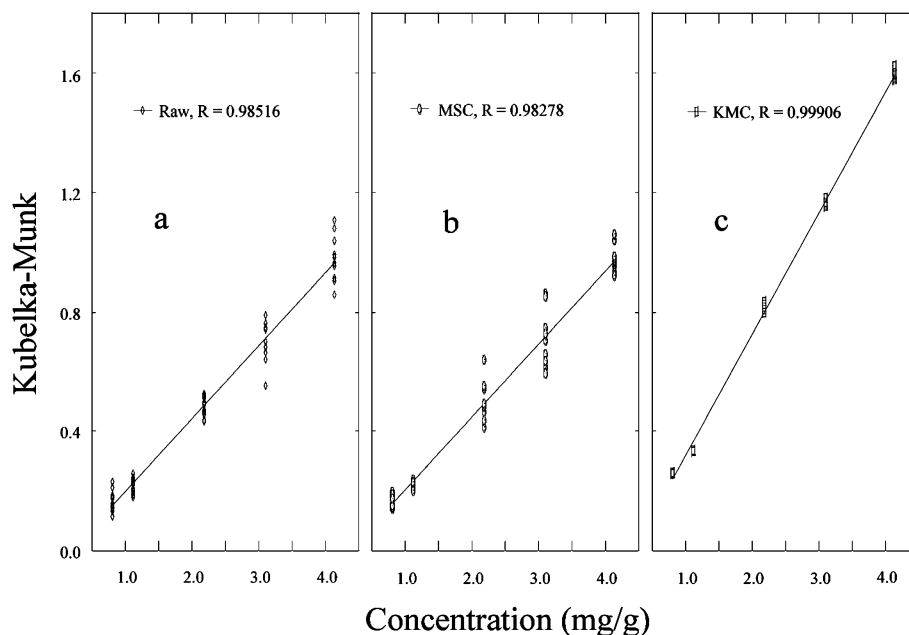


Figure 5. Kubelka–Munk vs concentration and linear fitting plots for the EMPA-tainted soil 1313.3- cm^{-1} peak at concentrations of 0.804, 1.12, 2.18, 3.10, and 4.13 mg/g under raw, MSC, and KMC, measured by CPC and 10 measurements for each sample. (a–c) correspond to raw, MSC, and KMC, respectively.

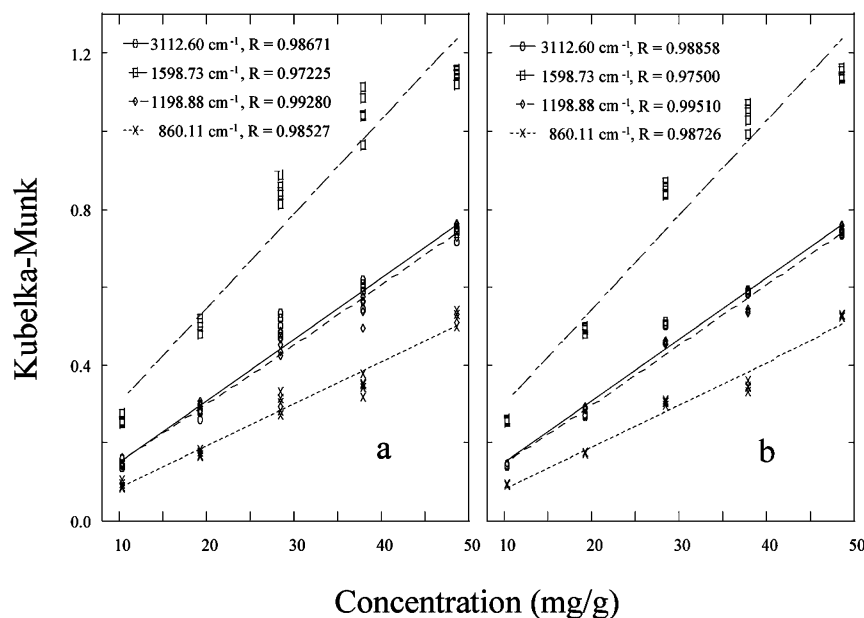


Figure 6. Kubelka–Munk vs concentration and linear fitting plots for caffeine KBr; four peaks at concentration 10.38, 19.26, 28.45, 37.96, and 48.65 mg/g under MSC and KMC, measured by CPC and five measurements for each sample. In this figure, legend circle, square, diamond, and cross stand for 3112.6, 1598.7, 1189.9, and 860.1 cm^{-1} correspondingly. (a) and (b) correspond to MSC and KMC, respectively.

acquired in Griffiths' laboratory (Moscow, ID) using a Harrick Cricket (Harrick Scientific, Ossining, NY) diffuse reflection accessory mounted in a Thermo-Nicolet Magna 760 FT-IR spectrometer (Thermo-Electron Corp) equipped with a deuterated triglycine sulfate detector.

RESULTS AND DISCUSSION

A representative DR spectrum of EMPA-tainted DPG soil is presented in Figure 1 along with some key observations regarding the effect of observed departures in the apparent baseline on the spectral intensity of the spectral bands in KM units. For a given measurement j , we observed a linear relationship between the

magnitude of the local peak minimum and the peak height $f(R_{\infty}, \nu)_j$ of the observed absorption bands. A plot of the observed intensity of the 1313.30- cm^{-1} band relative to the baseline offset over several replicate measurements of a given sample is presented as an inset in Figure 1 along with the equation of the line to demonstrate the linear relationship that we observed. The slope of the line that results from this plot gives us $A(\nu)$ in our BPC algorithm (eq 4).

The result of our two-step analysis applied to the DR data from the EMPA–DPG soil mixtures is presented in Figure 4. Note that the spread in the raw data is reduced by the MSC algorithm

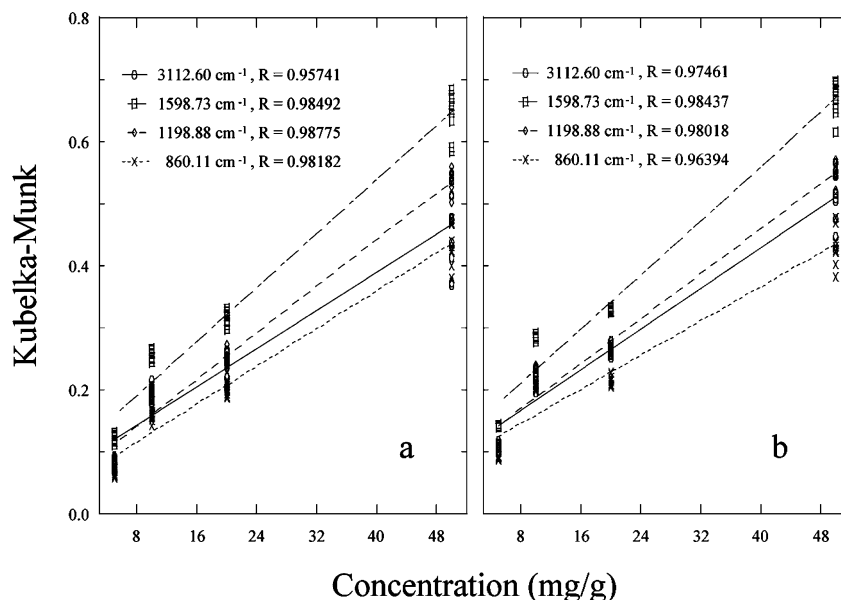


Figure 7. Kubelka–Munk vs concentration and linear fitting plots for caffeine in KCl; four peaks at concentration 5.00, 10.00, 20.00, and 50.00 mg/g under MSC and KMC, measured by Cricket and 10 measurements for each sample. In this figure, legend circle, square, diamond, and cross stand for 3112.6, 1598.7, 1189.9, and 860.1 cm^{-1} correspondingly. (a) and (b) correspond to MSC and KMC, respectively.

(Figure 4b) but that the concomitant spectral band intensity error remains. After the application of our KMC, the residual error in spectral band intensities is noticeably reduced. We present the KM plot of the 1313.30- cm^{-1} band of EMPA in our EMPA–soil mixture in Figure 5 as raw, MSC corrected, and KMC corrected. The performance of the KMC processing is more lucid in the KM plot of the DR data. Note that the MSC correction does not affect the slope of the line whereas the KMC correction results in a significant change in the slope of the KM plot in addition to the dramatic reduction in the measurement uncertainty.

We acknowledged two key challenges with the generalization of our KMC algorithm to quantitative DR experiments. First, the EMPA–soil mixture is a poor model for conventional quantitative DR analysis in that it is an exceedingly complex system consisting of a broad particle size distribution and a heterogeneous mixture of solid and liquid analytes. Second, all our measurements were performed on a particular and somewhat unique apparatus, the CPC optics in the Thermo-Nicolet SmartCollector accessory. We wanted to make certain that the artifacts that we observed in our data were indeed general and not some peculiar result of studying the difficult heterogeneous system or some unique feature of the CPC accessory. We therefore engaged in a study of caffeine using the CPC experiment and asked an associate to perform a similar set of measurements on caffeine using a completely different DR accessory. The results of these experiments are summarized in the KM plots (Figures 6 and 7) of the data that were generated by both laboratories, and we found that indeed the KMC algorithm improved the precision of both sets of data acquired on caffeine. The magnitude of the precision enhancement is more readily realized by inspection of the mean square error metric that we defined in eq 6. The mean square error for four wavelengths of the data analyzed by normal MSC methods and by the KMC algorithm are presented in Table 1. The data clearly show the marked decrease in the spread in the data afforded by the KMC treatment.

Table 1. Mean Square Error Metric Defined in Eq 6 for Four Wavelengths of the EMPA-Tainted Dugway Soil Data Analyzed by Normal MSC Methods and by the KMC Algorithm

mass EMPA (mg)/g of soil	peak (cm^{-1})	raw	MSC	KMC
0.804	2983.4	0.004 936 9	0.001 310 0	0.000 116 1
	1795.4	0.003 367 2	0.001 536 3	0.000 073 6
	1313.3	0.004 936 9	0.001 310 0	0.000 116 1
	873.6	0.004 102 9	0.001 570 0	0.000 083 3
	spectral av	0.005 667 8	0.000 330 9	0.000 030 6
1.12	2983.4	0.003 169 1	0.001 448 0	0.000 141 6
	1795.4	0.001 896 4	0.000 947 1	0.000 256 6
	1313.3	0.003 169 1	0.001 448 0	0.000 141 6
	873.6	0.002 637 6	0.001 309 4	0.000 241 7
	spectral av	0.001 278 4	0.000 147 7	0.000 033 5
2.18 mg/g	2983.4	0.003 886 2	0.024 064 9	0.001 758 1
	1795.4	0.002 426 2	0.017 320 8	0.001 719 9
	1313.3	0.004 738 6	0.023 467 0	0.002 025 9
	873.6	0.004 148 9	0.022 644 2	0.001 667 7
	spectral av	0.014 273 5	0.003 731 8	0.000 749 0
3.10 mg/g	2983.4	0.014 002 3	0.038 088 3	0.000 271 8
	1795.4	0.010 591 5	0.025 826 0	0.000 332 7
	1313.3	0.013 144 3	0.035 305 6	0.000 317 6
	873.6	0.012 798 6	0.033 450 4	0.000 214 4
	spectral av	0.018 787 6	0.003 964 6	0.000 096 3

We also investigated the commutation rules associated with the KMC approach. Specifically, we were interested in the degree to which the mathematical treatments could be applied in a different order. We found that the optimal improvement results from the application of BPC after detrending the data with the MSC algorithm, suggesting that the BPC approach is effective on the condition that the main reason for the departure in the baseline position is the KM effect itself rather than simple scatter due to measurement-to-measurement error. We also found that the BPC algorithm offers no benefit to data that are reduced in $\log(1/R)$ units; the baseline-dependent offset correction is understandably peculiar to the linearity errors associated with the KM theory.

Simulated data are presented in Supporting Information that demonstrate the phenomenological effect of offsetting the baseline on the data on the observed peak height in $\log(1/R)$ scale vis-à-vis KM units, and the successful improvements in the data realized by the application of the KMC algorithm.

CONCLUSIONS

Quantitative DR analysis under KM theory affords better linearity than that reduced in $\log(1/R)$ units, facilitating calibrations in an intuitive manner and lessening the risk of overfitting data to accommodate nonlinearities. However, baseline position variance in the KM data space results in observed peak intensity changes that are proportional to the observed offset. We offer a simple empirical correction to KM data, the KMC, that compensates for the phenomenological relationship observed between the measured measurement-to-measurement baseline drift and the spectral band intensities in KM units.

ACKNOWLEDGMENT

We acknowledge the contribution of Prof. Peter Griffiths and his students Lacey Averett and David Heaps in the preparation of a set of caffeine DR data to verify that the algorithms we developed were equally applicable to data acquired using different instrumentation and operators. We also thank Prof. Griffiths for many useful discussions on the topics presented here. C.Z. thanks Mr. Darren Emge, Edgewood Chemical Biological Center, for his help with the development of the Matlab (The MathWorks, Inc., Natick, MA) code used in this analysis. This effort was funded by the Department of Defense's Joint Science and Technology Office.

SUPPORTING INFORMATION AVAILABLE

Additional information as noted in text. This material is available free of charge via the Internet at <http://pubs.acs.org>.

Received for review June 3, 2005. Accepted November 10, 2005.

AC0509859