Nondestructive Total Excitation-Emission Fluorescence Microscopy Combined with Multi-Way Chemometric Analysis for Visually Indistinguishable Single Fiber Discrimination

6 authors, including:

Arsenio Muñoz de la Peña
Universidad de Extremadura
201 PUBLICATIONS   **4,688** CITATIONS

SEE PROFILE

Nirvani Mujumdar
University of Central Florida
10 PUBLICATIONS   **21** CITATIONS

SEE PROFILE

Emily C. Heider
University of Central Florida
27 PUBLICATIONS   **163** CITATIONS

SEE PROFILE

Héctor C Goicoechea
Universidad Nacional del Litoral
233 PUBLICATIONS   **5,318** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project   Strategies of multiway calibration for the enhancement of analytical methodologies in the bioanalytical, food and environmental fields View project

Project   Photoluminescence in BCN films View project

Nondestructive Total Excitation−Emission Fluorescence Microscopy Combined with Multi-Way Chemometric Analysis for Visually Indistinguishable Single Fiber Discrimination

Arsenio Muñoz de la Peña,[†,‖] Nirvani Mujumdar,[†] Emily C. Heider,[†] Hector C. Goicoechea,[‡] David Muñoz de la Peña,[§] and Andres D. Campiglia*,[†,⊥]

[†]Department of Chemistry, University of Central Florida, 4111 Libra Drive, P.O. Box 25000, Orlando, Florida 32816-2366, United States

[‡]Laboratorio de Desarrollo Analítico y Quimiometría (LADAQ), Cátedra de Química Analítica I, Facultad de Bioquímica y Ciencias Biológicas, Universidad Nacional de Litoral, Santa Fe S3000ZAA, Argentina

[§]Department of Automation and Systems Engineering, University of Seville, Seville, 41092, Spain

[‖]Department of Analytical Chemistry and IACYS, University of Extremadura, Badajoz, 06006, Spain

[⊥]National Center for Forensic Science, University of Central Florida, 12354 Research Parkway, Suite 225, Orlando, Florida 32826, United States

S Supporting Information

ABSTRACT: The potential of total excitation−emission fluorescence microscopy combined with multiway chemometric analysis was investigated for the nondestructive forensic analysis of textile fibers. The four pairs of visually indistinguishable fibers consisted of nylon 361 dyed with acid yellow 17 and acid yellow 23, acetate satin 105B dyed with disperse blue 3 and disperse blue 14, polyester 777 dyed with disperse red 1 and disperse red 19, and acrylic 864 dyed with basic green 1 and basic green 4. Excitation emission matrices were recorded with the aid of an inverted microscope and a commercial spectrofluorimeter. The full information content of excitation−emission matrices was processed with the aid of unsupervised parallel factor analysis (PARAFAC), PARAFAC supervised by linear discriminant analysis (LDA), and discriminant unfolded partial least-squares (DU-PLS). The ability of the latter algorithm to classify the four pairs of fibers demonstrates the advantage of using the multidimensionality of fluorescence data formats for the nondestructive analysis of forensic fiber evidence.
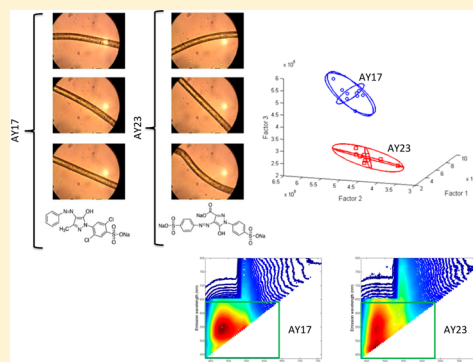
■ INTRODUCTION

The analysis of fibers from clothes is of paramount importance in forensic science, when investigating a crime scene. Trace fiber evidence has been probative in cases ranging from the 1963 JFK assassination,[1] to the Atlanta Child murders[2] of the early 1980s, or the 2002 Washington, DC, sniper case.[3] The fiber examiner typically performs a series of comparisons of the questioned fiber to a known fiber in an attempt to exclude the possibility that a questioned fiber and known fiber could have originated from a common source. If the two fibers are considered to be substantially different, then the hypothesis that the two fibers originated from a common source can be disregarded. The extension to which fibers of different origins can be discriminated is related to the analytical method used for their analysis.[4] A challenging aspect of forensic fiber examinations involves the comparison of fibers colored with visually indistinguishable dyestuffs. This is not an uncommon situation, as there are numerous indistinguishable fibers predyed with commercial dyes of virtually identical colors.

Minimal chemical structural variations are actually encouraged by the dye patent process and commercial competition.

Microscopy-based techniques currently used in forensic science laboratories include polarized light microscopy, infrared microscopy,[5] microspectrophotometry,[6−9] fluorescence microscopy,[10,11] and scanning electron microscopy coupled to energy dispersive spectrometry.[12] Differences in cross-sectional shape, type of fiber material (natural or synthetic), weave, and color often make possible to rule out a common source for the two samples. The main advantage of these techniques is their nondestructive nature, which preserves the physical integrity of the fibers for further court examination. Beyond microscopy— but still under the category of minimally destructive techniques—is pyrolysis coupled to gas chromatography. This tool is capable to compare the polymeric nature of synthetic and natural fibers at expenses of partial sample consumption.[13]

When fibers cannot be discriminated by nondestructive tests, the next step is to extract the questioned and the known fiber for further dye analysis. Thin-layer chromatography (TLC),[14,15] high-performance liquid-chromatography (HPLC)[16] and capillary electrophoresis (CE)[17,18] have been used to separate and identify colored dyes in fiber extracts. The ultimate selectivity belongs to mass spectrometry (MS) coupled to either HPLC or CE. HPLC-MS and CE-MS are able to differentiate textile dyes with similar molecular structures that provide similar elution times and optical spectra.[19−24] Unfortunately, MS techniques destroy the fiber just like all the other methods that provide chemical information based on previous dye extraction.

Research reports on the nondestructive analysis of fibers have proposed the use of diffuse reflectance infrared Fourier transform spectroscopy (DRIFTS) and Raman spectroscopy. When coupled to chemometric methods for spectral interpretation such as principal component analysis (PCA) and soft independent modeling of class analogies (SIMCA), DRIFTS was able to discriminate both dye color and reactive dye state on cotton fabrics.[25−27] Raman spectroscopy was able to characterize dyes in both natural and synthetic fibers via a combination of Fourier transform-Raman spectra and PCA analysis.[28−31]

Our group has focused on room-temperature fluorescence (RTF) spectroscopy. Although fluorescence microscopy is currently used in forensic laboratories for single fiber examination,[10,11] measurements are made with the aid of band-pass filters that provide very limited information on the spectral profiles of fibers. Our approach takes fluorescence microscopy to a higher level of selectivity with the collection of excitation emission matrices (EEMs).[32,33] EEMs—which refer to a series of emission spectra recorded at various excitation wavelengths—were recorded with the aid of a microscope coupled to a spectrofluorimeter. The subtraction of EEMs from visually indistinguishable fibers provided the best excitation wavelength for recording two-dimensional fluorescence spectra (first order data). The comparison of fluorescence spectra via PCA resulted in the accurate identification of fibers with no false positives.[32] The same approach was later applied to investigate laundering effects on textile fibers. The spectral fingerprints of brighteners and other detergent components adsorbed on the fibers improved fiber discrimination via RTF-EEMs-PCA.[33]

Herein, we focus on the total fluorescence content of the EEMs. The entire data sets of fluorescence spectra recorded at various excitation wavelengths are compared with the aid of parallel factor analysis (PARAFAC). Since spectral profiles predicted with PARAFAC based methods provide chemical information on sample constituents, this second-order algorithm determines the number of fluorescence components that contribute to each EEM along with their individual excitation and emission profiles. The application of PARAFAC is carried out unsupervised and supervised by linear discrimination analysis (LDA).[34,35] Supervision refers to the provision of information about sample types to the model when studying a training set of samples, followed by a prediction step. LDA is based on the determination of linear discriminant functions. By maximizing the ratio of between-class variance and minimizing the ratio of within class variance, LDA supervised methods achieve maximum separation among classes and, therefore, superior classification performance than nonsupervised methods. PARAFAC and PARAFAC-LDA are then compared to

supervised discriminant unfolded partial least-squares (DU-PLS) for classification purposes.[36,37]

Theory. *PARAFAC.* This algorithm often achieves the decomposition of three-dimensional data arrays into two-dimensional spectral profiles for both qualitative and quantitative purposes.[38] If EEMs are arranged in a three-way array X of dimensions $I \times J \times K$, where I, J, and K are the number of samples, number of emission wavelengths, and number of excitation wavelengths, respectively, PARAFAC attempts to decompose it into three matrices A (scores), B, and C (loadings) with elements a_{in}, b_{jn}, c_{kn}, respectively, where n indicates the component number. An element of X is given by the following:

$$x_{ijk} = \sum_{n=1}^{N} a_{in} b_{jn} c_{kn} + e_{ijk} \tag{1}$$

where x_{ijk} is the fluorescence intensity for sample i at the emission wavelength j and excitation wavelength k and e_{ijk} indicates an element of the array E, which collects the variability not accounted by the model. For a given component n, the elements a_{in}, b_{jn}, and c_{kn} are arranged in the score vector $\mathbf{a}_n$ (whose elements are directly proportional to its concentration in each sample) and the loading vectors $\mathbf{b}_n$ and c_n, which estimate its emission and excitation profiles. The array of EEMs data is fitted to eq 1 by least-squares.

*LDA.* This algorithm calculates a surface separating sample groups by establishing a linear discriminant function that maximizes the ratio of the between-class and the within-class variances.[35,39] Categories are supposed to follow a multivariate normal distribution and be linearly separated. With the A score matrix of PARAFAC and the $I \times g$ dummy matrix $\mathbf{Y}$ of binary digits representing the group assignments (g is the number of categories), the best representation is obtained if the ratio of the between-class variance $\mathbf{Bc}$ matrix and the within-class variance $\mathbf{Wc}$ matrix is maximized. Suitable expressions for the matrices $\mathbf{Bc}$ and $\mathbf{Wc}$ are given by the following expressions:[40]

$$\mathbf{Bc} = (g - 1)^{-1} \mathbf{A}^T \mathbf{Y} (\mathbf{Y}^T \mathbf{Y})^{-1} \mathbf{Y}^T \mathbf{A} \tag{2}$$

$$\mathbf{Wc} = (I - g)^{-1} [\mathbf{A}^T \mathbf{A} - (g - 1) \mathbf{Bc}] \tag{3}$$

The canonical variate (CV) scores contain the successively maximized ratios between-groups variance/within-groups variance. They are obtained by PCA of the matrix ($\mathbf{Wc}^{-1} \mathbf{Bc}$) and projection of the data matrix A onto the first loadings. The samples are then plotted on a two- or three- dimensional space defined by the first CV scores for each sample.

DU-PLS. Although the mathematical foundations of U-PLS were originally developed for multivariate calibration purposes,[36] its application to the classification of samples has been reported extensively.[37,41,42] The main difference between U-PLS and discriminant U-PLS (DU-PLS) consists in the building of the dependent variable $\mathbf{y}$. For model calibration purposes, the variable $\mathbf{y}$ contains concentrations values. For discriminant analysis purposes, $\mathbf{y}$ contains a coding integer representing the class label of the samples. PLS regression is conducted between the instrumental response in $\mathbf{X}$ block (built with the unfolded original second-order matrix data) and the class label in $\mathbf{y}$ block using training samples, and the optimal number of latent variables is chosen based on the error rate by cross-validation.[43] The final model for A latent variables is used to predict the class label in the test set according to the following:
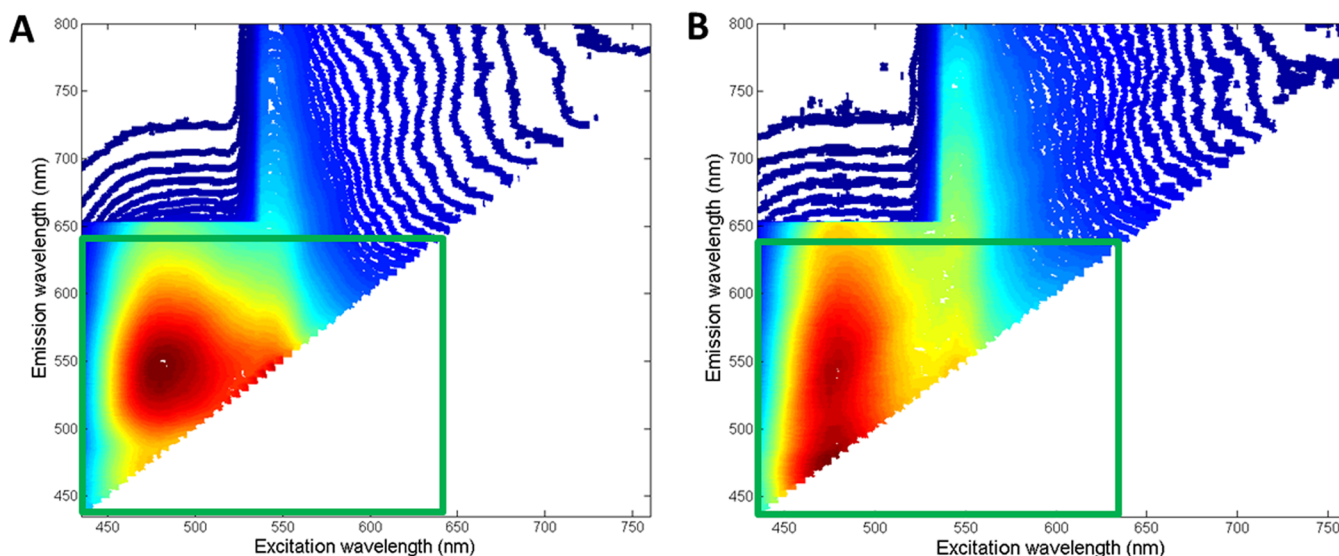
Figure 1. Contour plots of EEMs recorded from single AY17 (left) and AY23 (right) nylon 361 fibers. Each contour is the average of 10 EEMs recorded at different spots of each single fiber. Excitation range =435−760 nm. Emission range =435−800 nm. For the purpose of data analysis, the zones of the EEMs corresponding to Rayleigh dispersion were completed with NaN terms. The zones of the EEMs selected for PARAFAC are denoted in green (excitation range =435−630 nm; emission range =435−642 nm).

$$y_{test} = \mathbf{t}_{test}^{T}\mathbf{v} \tag{4}$$

where y_{test} is the label class predicted, $\mathbf{t}_{test}^{T}$ are the scores of test samples obtained by projection of $\mathbf{x}_{test}$ onto the training loadings, and $\mathbf{v}$ is the vector of the regression coefficients. In the ideal case scenario, the calculated values of y_{test}—for two classes of samples—are 1 or 2; in practice, y_{test} values are often close to 1 and 2. Therefore, in order to assign a test sample to a given class, it is necessary to establish thresholds for the y_{test} predicted values. The threshold is defined as the value that minimizes the number of false positives and false negatives.[44]

EXPERIMENTAL SECTION

Reagents and Materials. Acetate satin 105B, nylon 361, polyester 777 and acrylic 864 fabrics were acquired and dyed at Testfabrics, Inc. (West Pittston, PA). Textile dyes were purchased from Sigma-Aldrich (www.sigma-aldrich.com) at their highest available purity (% w/w). These included acid yellow 17 (AY17; 60%), acid yellow 23 (AY23; ≥ 85%), disperse blue 3 (DB3; 20%), disperse blue 14 (DB14; 97%), disperse red 1 (DR1; 95%), disperse red 19 (DR19; 97%), basic green 1 (BG1; ∼90%) and basic green 4 (BG4; > 95%). Nylon 361 fabrics were dyed with AY 17 and AY 23; acetate satin 105B with DB3 and DB14; polyester 777 with DR1 and DR19; and acrylic 864 with BG1 and BG4. All fabric samples were dyed separately, received in sealed packages and kept in the dark to avoid environmental exposure. Tweezers, blades, and scissors used to manipulate fabrics and isolate fibers were previously cleaned with methanol and visually examined under UV-light (254 nm) to avoid fluorescence contamination.

Fluorescence Microscopy. A spectrofluorimeter (Fluo-roMax-P; Horiba Jobin Yvon) connected to an epifluorescence microscope (BX-51; Olympus) via a bifurcated fiber-optic probe (Horiba Jobin-Yvon) were used to acquire EEMs. The spectrofluorimeter was equipped with a continuous 100 W pulsed xenon lamp with illumination from 200 to 1100 nm. Excitation and fluorescence emission spectra were recorded with two spectrometers holding the same reciprocal linear dispersion (4.25 nm mm^{-1}) and accuracy (±0.5 nm with 0.3 nm resolution). Both diffraction gratings had the same number of grooves per unit length (1200 grooves nm^{-1}) and were blazed at 330 nm (excitation) and 500 nm (emission). A photomultiplier tube (Hamamatsu, model R928) with spectral response from 185 to 850 nm was used for fluorescence detection operating at room temperature in the photon-counting mode. The acquisition of EEMs was computer-controlled with commercial software (Datamax).

Collection of excitation and emission radiation between the two instrumental units was facilitated with the two concave mirrors of a fiber-optic platform located in the sample compartment of the spectrofluorimeter. The microscope was equipped with two 50/50 beam splitters, one for the ultraviolet and the other for the visible spectral region. A 40X Visible (Olympus UPlanSApo 40X) objective lens was used for light collection in the visible (435−800 nm, 90% transmittance) spectral region. A rotating pinhole wheel, with various diameters varying from 0 to1000 μm, was located between the 50/50 beam splitter and the mirror that directed fluorescence either to the CCD camera (iDS UI-1450SE-C-HQ USB-camera) of the microscope or the emission fiber bundle of the spectrofluorimeter. Image acquisition was computer-controlled with commercial software (DataMax).

Recording EEMs. Otherwise noted, EEMs were recorded within the 435−800 nm excitation−emission range using 5 and 1 nm excitation and emission steps, respectively. Scatter interference from excitation radiation was avoided with the use of appropriate cutoff filters. EEMs from single fibers were recorded by placing the sample between two quartz glass slides. Each fiber was sampled 10 times by recording 10 EEMs from 10 spots randomly selected along the entire surface of the fiber.

Chemometric Analysis. All chemometric calculations were done using Mat Lab 8.0. Routines for PARAFAC were available in the Internet thanks to Bro.[38] A useful Mat Lab graphical interface provided a simple means of loading the EEM data matrices into the Mat Lab working space before running and analyzing data via PARAFAC and DU-PLS.[45,46] An in house Mat Lab routine was used for LDA calculations.[35]
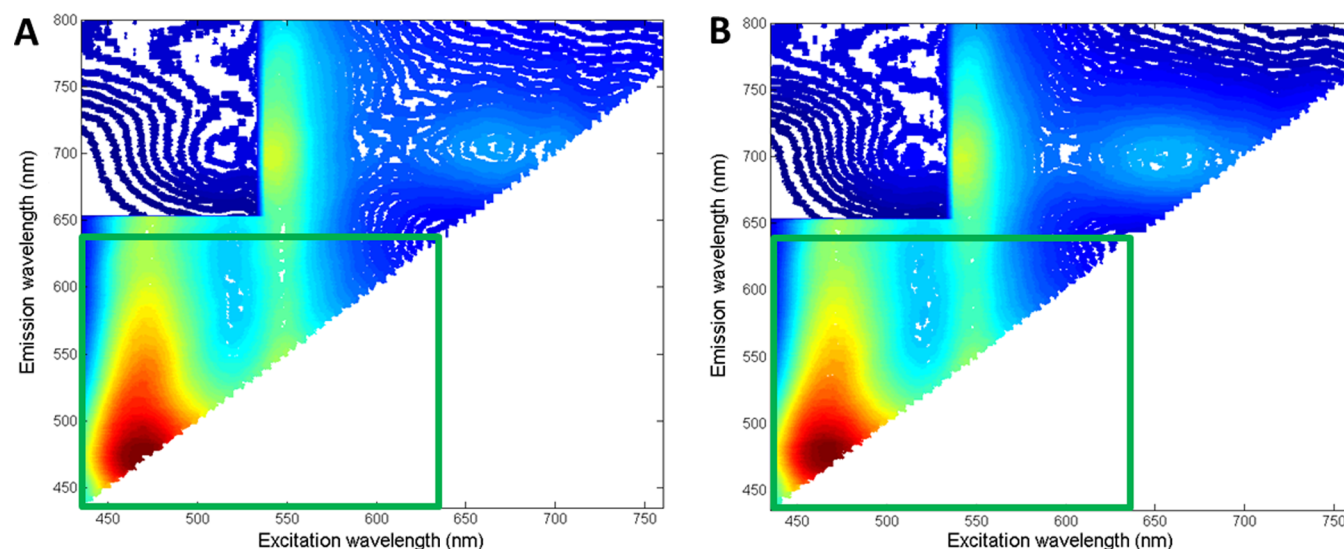
Figure 2. Contour plots of EEMs recorded from single BG1 (left) and BG4 (right) acrylic 864 fibers. Each contour is the average of 10 EEMs recorded at different spots of each single fiber. Excitation range =435−760 nm. Emission =435−800 nm. For the purpose of data analysis, the zones of the EEMs corresponding to Rayleigh dispersion were completed with NaN terms. The zones of the EEMs selected for PARAFAC are denoted in green (excitation range =435−630 nm; emission range =435−642 nm).
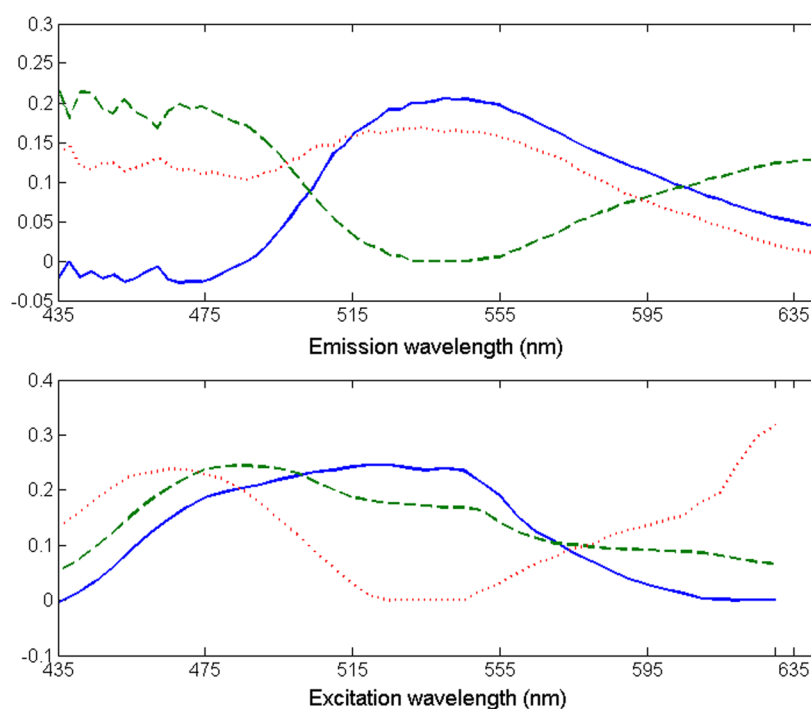

Figure 3. Extracted emission (top) and excitation (bottom) PARAFAC profiles taken from an AY17 and an AY23 nylon 361 fiber. Spectral profiles are based on ten EEM replicates per fiber.

■ RESULTS AND DISCUSSION

The focus of our studies was then placed on fibers containing structurally similar dyes that are not easily differentiated by visual inspection under the microscope. The similarity of the investigated fibers—nylon 361 fibers dyed with AY17 and AY23; acetate satin 105B fibers dyed with DB3 and DB14; polyester 777 fibers dyed with DR1 and DR 19; and acrylic 864 fibers dyed with BG1 and BG4—is shown in Figures S1−S4. In addition to the contribution of the textile dye to the fluorescence spectrum of the fiber, our approach considers the contribution of fluorescence impurities—i.e., impurities

imbedded into the fibers during fabrication of garments—as a reproducible source of fiber comparison.[32,33] Since the purity of all the reagent dyes is lower than 100%, the presence of unidentified fluorophores in the investigated fibers is possible.

Figures 1 and 2 show EEMs recorded from AY17 and AY23 nylon fibers and BG1 and BG4 acrylic fibers. The similarity of the EEMs certainly challenges the pairwise discrimination of AY17/AY23 and BG1/BG4 fibers on the sole bases of visual comparison. Comparison of EEMs for purposes of fiber identification requires a statistical figure of merit that could
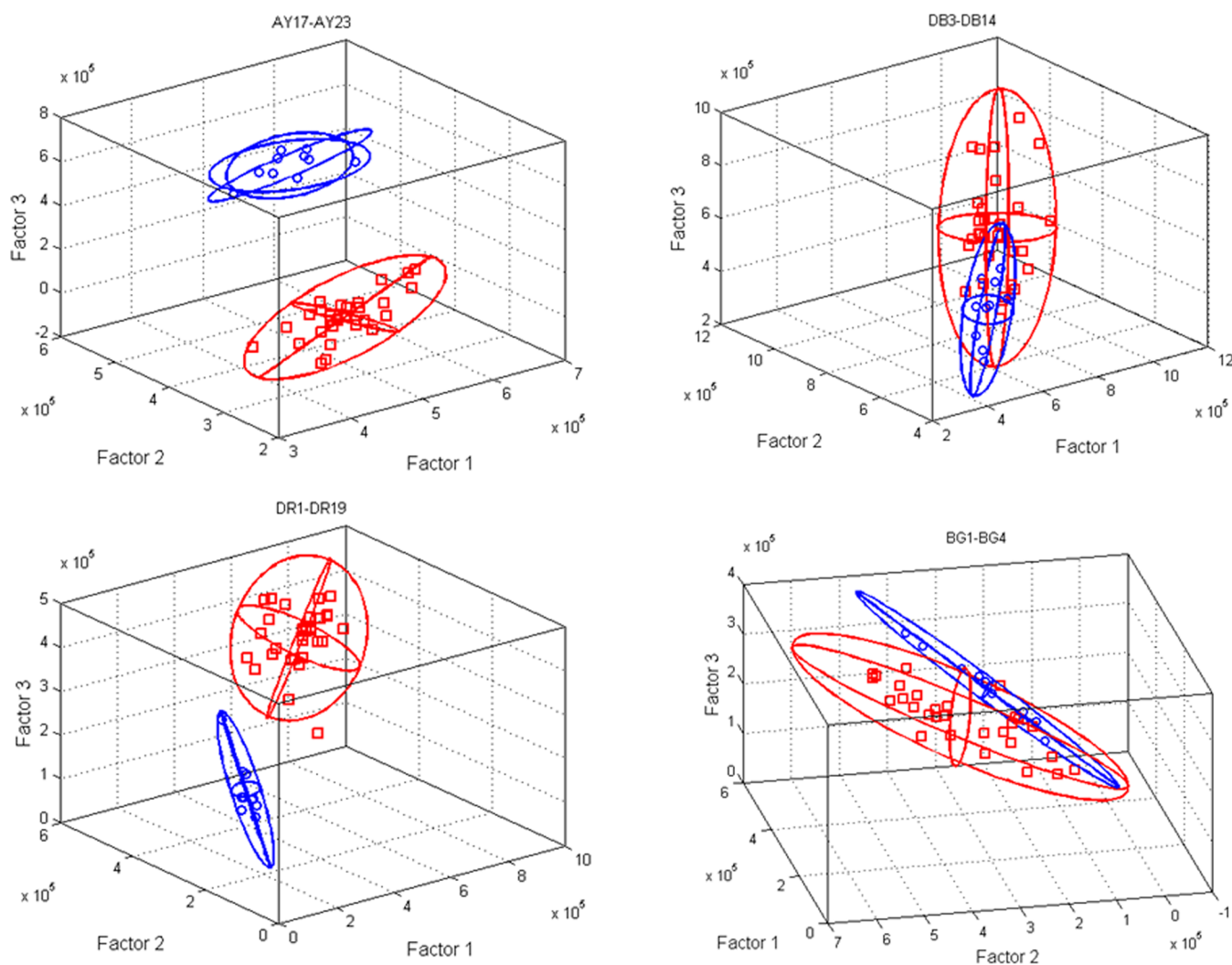
Figure 4. PARAFAC scores (3 components model) for 40 samples of (top left) AY17 (1 fiber, 10 replicates; blue circles) and AY23 (3 fibers, 10 replicates each; red squares) nylon 361 fibers; (top right) DB3 (1 fiber, 10 replicates; blue circles) and DB14 (3 fibers, 10 replicates each; red squares) acetate satin 105B fibers; (bottom left) DR 1 (1 fiber, 10 replicates; blue circles) and DR19 (3 fibers, 10 replicates each; red squares) polyester 777 fibers; and (bottom right) BG1 (1 fiber, 10 replicates; blue circles) and BG 4 (3 fibers, 10 replicates each; red squares) acrylic 864 fibers. The three-dimensional projection of the 95% confidence ellipse of the data collected from each type of fiber is included to facilitate visualization of the obtained results.

be obtained from the chemometric algorithms proposed in the present work.

Forensic comparisons involve fibers found in the crime scene (questioned fibers) and fibers collected from a known, suspect source (known fibers). All the comparisons made in this study assumed that only one fiber was found in the crime scene. To account for potential batch-to-batch variations in textile dyeing process, each "questioned fiber" was compared to three visually indistinguishable fibers collected from three garments that were dyed separately.

PARAFAC. Ten EEMs were recorded from each one of the investigated fibers. Each comparison of visually indistinguishable fibers—AY17 versus AY23, DB3 versus DB 14, DR1 versus DR19 and BG1 versus BG4—was based on 40 EEMs; i.e., Ten EEMs recorded from each one of the "questioned fiber" (AY17, DB3, DR1, or BG1) and 30 EEMs recoded from three "known fibers" (AY23, DB14, DR19, or BG4). Excitation was made within 435 and 800 at 5 nm increments. Fluorescence was recorded between 345 and 800 at 1 nm increments. Under these parameters, each EEM resulted in a matrix with 66 × 366

data points (excitation × emission). As shown in Figures 1 and 2, those rather large data matrixes included wavelength regions with either residual background fluorescence or no fluorescence data (Rayleigh scattering). These regions were excluded from calculations to optimize computational time. No significant loss was observed when data analysis time was further reduced by computing one fluorescence data point every 3 nm intervals. The resulting EEMs consisted of matrixes with 66 × 122 data points (excitation × emission). The wavelength regions selected for PARAFAC analysis are denoted as green rectangles in Figures 1 and 2. They included data points 1–40 in the excitation mode (435–650 nm) and data points 1–70 in the emission mode (432–642 nm). Missing data points due to Rayleigh scattering were completed with NaNs terms and handled by expectation maximization.[38]

Spectral deconvolution of EEMs via PARAFAC was carried out with the four pairs of investigated fibers. PARAFAC was first applied without supervision. The model was estimated using certain constraints such as non-negativity on all three modes.[47−49] The test provided the best fit for a three
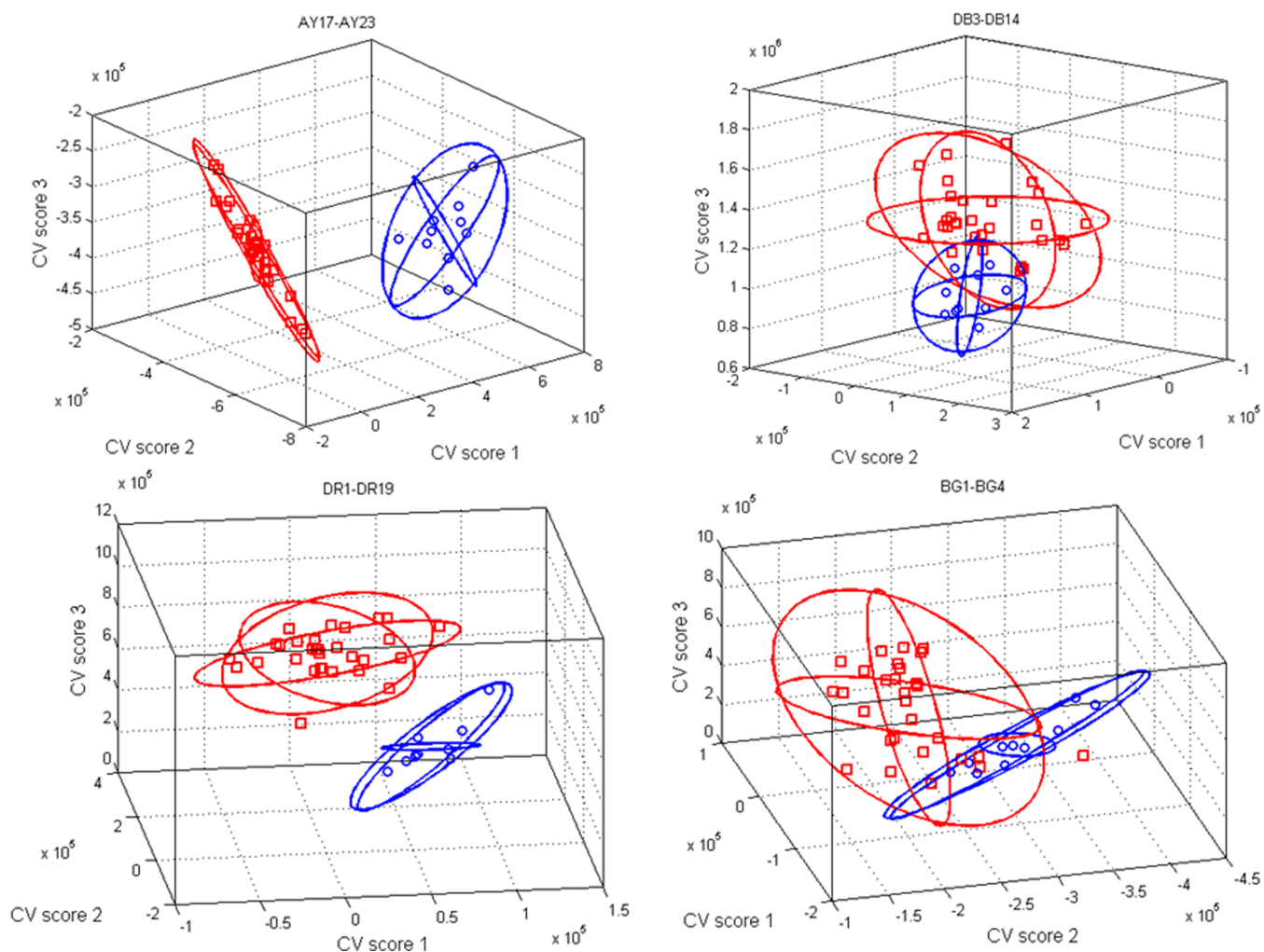
Figure 5. LDA CV scores (3 components model) for 40 samples of (top left) AY17 (1 fiber, 10 replicates; blue circles) and AY23 (3 fibers, 10 replicates each; red squares) nylon 361 fibers; (top right) DB3 (1 fiber, 10 replicates; blue circles) and DB14 (3 fibers, 10 replicates each; red squares) acetate satin 105B fibers; DR 1 (1 fiber, 10 replicates; blue circles) and DR19 (3 fibers, 10 replicates each; red squares) polyester 777 fibers; and (bottom right) BG1 (1 fiber, 10 replicates; blue circles) and BG 4 (3 fibers, 10 replicates each; red squares) acrylic 864 fibers. The three-dimensional projection of the 95% confidence ellipse of the data collected from each type of fiber is included to facilitate visualization of the obtained results.

fluorescence components model in all cases. The three fluorescence components improve the robustness of the modeling with the prior knowledge of the inexistence of negative fluorescence intensities. In all cases, PARAFAC was initialized with the loadings giving the best fit after a small number of trial runs, selected from the comparison of the results provided by generalized rank annihilation and several orthogonal random loadings.[50] Convergence was achieved when the relative change in fit was 1×10^{-6}.

The number of PARAFAC components was analyzed by two different procedures: the statistical test of the core consistency diagnostic (CORCONDIA)[51] and the analysis of residuals.[38] The core consistency analysis consists in studying the structural model based on the data and the estimated parameters of gradually augmented models. Under this prospective, the model is considered appropriate when the addition of other combinations of components does not provide a considerable improvement of the fitting. An additional indication for the correct number of components is the dropping of the core consistency to value below 50.[38] However, the analysis of residuals considers the residual fit of the PARAFAC model as a

function of increasing number of factors. The appropriate model is the one which is not statistically different from the model leading to the minimum residual fit. Additional indicators for reaching the correct number of components include the visual inspection of the retrieved profiles. Forcing PARAFAC to extract more components than the number of real constituents produces two different types of extra profiles, profiles composed of random noise and/or profiles that appear to repeat. Analysis of our data yielded three factors for each one of the analyzed pairs of fibers. Figure 3 shows the emission and excitation loadings of the three-component PARAFAC model for AY17 and AY23 nylon fibers. The two first components are interpreted as the contributions of each of the component dyes, as the excitation and emission maxima of the dyes closely match those of the excitation and emission profiles obtained by PARAFAC. The third component may be attributed to the presence of other fluorescent components or fluorescent impurities present in the analyzed textiles.

Figure 4 shows the tridimensional plots of PARAFAC scores 1, 2, and 3 for each one of the investigated pairs. To facilitate the visualization of pairwise comparison, each plot includes the
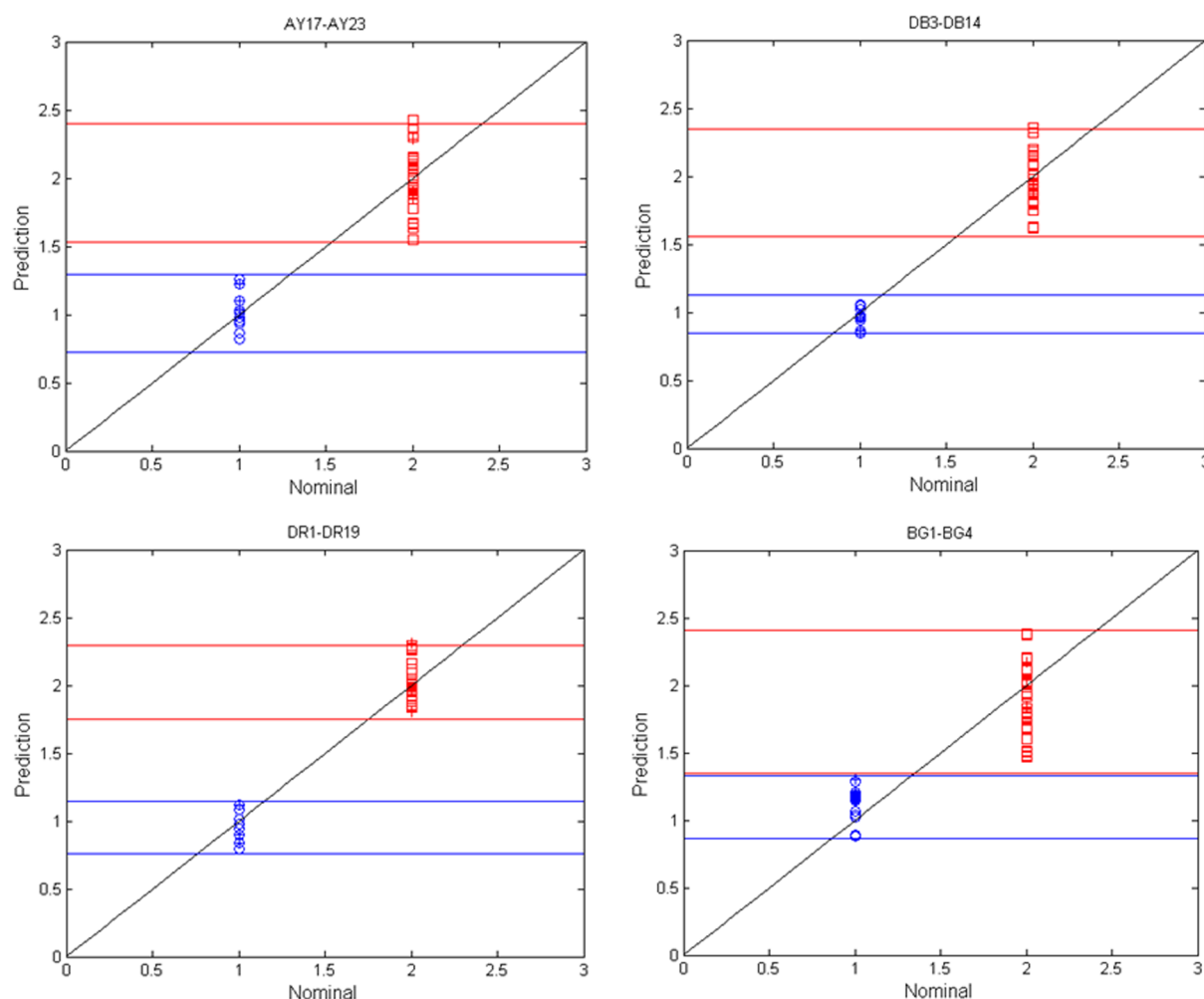
Figure 6. Plot of the DU-PLS (3 components model) predicted vs nominal coded values for 40 samples of (top left) AY17 (7 calibration samples = blue circles; 3 validation samples = blue crosses) and AY 23 (27 calibration samples = red squares; 3 validation samples = red crosses) nylon 361 fibers; (top right) DB3 (7 calibration samples = blue squares; 3 validation samples = blue crosses) and DB14 (27 calibration samples = red squares; 3 validation samples = red crosses) acetate satin 105B fibers; (bottom left) DR1 (7 calibration samples = blue circles; 3 validation samples = blue crosses) and DR19 (27 calibration samples = red squares; 3 validation samples = red crosses) polyester 777 fibers; and (bottom right) BG1 (7 calibration samples = blue squares; 3 validation samples = blue crosses) and BG4 (27 calibration samples = red squares; 3 validation samples = red crosses) acrylic 864 fibers.

projections of the 95% confidence ellipses over the three planes defined by their corresponding axes. The prediction interval for the multivariate normal distribution yielded an ellipse consisting of **x** vectors satisfying the following equation:

$$(\mathbf{x} - \mu)^T \sum{}^{-1} (\mathbf{x} - \mu) \leq \chi_k^2(p) \quad (6)$$

where μ is the mean, $\sum$ is the covariance matrix and $\chi_k^2(p)$ is the quantile function for probability p of the χ^2 distribution with k degrees of freedom, where k is the dimension of the data. The axes are defined by the eigenvectors of the covariance matrix and the radius of each axis is equal to 2.796 times the square root of the corresponding eigenvalue. The value 2.796 is obtained from the square root of the χ^2 distribution with three degrees of freedom and 95% confidence interval.[52] As shown in Figure 4, pairwise discrimination was only obtained for AY17/AY23 and DR1/DR19 fibers.

No significant improvement was observed with a number of components higher than the one suggested by the analysis of residuals. The inability of PARAFAC to differentiate DB3 from DB14 fibers and BG1 from BG4 fibers could be attributed to its nonsupervised nature.

PARAFAC-LDA. An approach that often improves the screening capability of PARAFAC is to submit the resulting scores to supervised LDA.[47,48] With this approach, it was possible to discriminate AY17 from AY23 fibers and DB3 from DB14 fibers. Better pairwise discrimination than with PARAFAC was obtained for DR1/DR19 and BG1/BG4 fibers, even though the overlap of the clusters could result in misidentification of the dye on the fiber (see Figure 5). The inability of PARAFAC-LDA to completely differentiate between them led us to attempt their discrimination via DU-PLS.

DU-PLS. Since U-PLS is unable to process data files with NaN terms, the first step toward the application of DU-PLS was to select an appropriate range of sensors for each pair of fibers. Optimization of sensor ranges for each pair of fibers provided the following sensor data for the excitation and emission modes, respectively: 26 to 70 and 1 to 16 (AY17/AY23 fibers); 26 to 70 and 1 to 16 (DB3/DB14 fibers); 20 to

70 and 1 to 12 (DR1/DR19 fibers); and 10 to 70 and 1 to 5 (BG1/BG4 fibers).

The number of optimum latent variables (h) was estimated via the leave-one-sample-out cross-validation approach[52] using a 34-sample set per fiber pair; i.e., Seven samples from each type of "questioned fiber" (AY17, DB3, DR1, or BG1) and 27 from the "three known fibers" (AY23, DB14, DR19, or BG4). The optimum h value was estimated with the ratio $F(h) =$ PREES $(h < h^*)$/PRESS(h^*); where PRESS $= \sum(c_{i,act} - c_{i,pred})^2$, h was the trial number of factors and h^* corresponded to the minimum PRESS. By selecting an h value that led to a probability of less than 75% that $F > 1$, three factors were found for each pair of fibers.

The discriminant ability of DU-PLS was tested on 6 validation samples using the 34 samples as the calibration training set along with the coded values for each category of each pair of fibers. The predicted versus nominal code values are shown in Figure 6. The confidence interval for each category was estimated as the product of the calculated standard deviations of the results for the training samples— namely, 0.14 for AY17 and 0.21 for AY23; 0.07 for DB3 and 0.24 for DB14; 0.10 for DR1 and 0.15 for DR19; and 0.13 for BG1 and 0.28 for BG4 - and the Student t-value with $n-1$ degrees of freedom for each category. On the basis of Figure 6, it is safe to state that the all the investigated fibers were clearly predicted and classified. Since the mode of operation of PLS for discriminating purposes (DU-PLS) has not been completely understood yet, it is not possible to generalize that DU-PLS will always perform better than PARAFAC-LDA. However, it is known that DU-PLS works better than PARAFAC-LDA for a small number of groups,[53−55] as it is the case of the present work. Although the EEMs recorded from single fibers belonging to the same fiber pair were almost identical, the prediction ability of the DU-PLS algorithm was enough to perform a successful supervised classification of samples.

CONCLUSIONS

Nondestructive techniques that can either discriminate between similar fibers or match a known to a questioned fiber—and still preserve the physical integrity of the fibers for further court examination—are highly valuable in forensic science. The work presented here provides a valuable data format for the fluorescence differentiation of visually indistinguishable fibers. In the case of virtually identical EEMs, an additional data treatment step involving an appropriate chemometric algorithm is necessary. On the basis of the obtained results, it can be concluded that LDA-supervised PARAFAC shows better discriminating potential than unsupervised PARAFAC. However, LDA-supervised PARAFAC was unable to discriminate between acetate satin 105B fibers dyed with DB3 and DB14 and acrylic 864 fibers dyed with BG1 and BG4. The best discrimination was obtained with the supervised DU-PLS model, which allowed the pairwise differentiation of the four pairs of investigated fibers. The ability of DU-PLS to distinguish the questioned from the known fibers excludes the possibility that the visually indistinguishable fibers originated from the common source.

ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.analchem.6b00264.

Bright field images of visually indistinguishable fibers (PDF)

AUTHOR INFORMATION

Corresponding Author

*Tel.: 497-823-4162. Fax: 407-823-2252. E-mail: andres.campiglia@ucf.edu (A.D.C.).

Notes

The authors declare no competing financial interest.

REFERENCES

(1) *The Warren Commission's Report*; St. Martin's Press: New York, 1964.

(2) Deadman, H. Fiber Evidence and the Wayne Williams Trial, US Government Document J1.14/8a:F44, Federal Bureau of Investigation, US Department of Justice, FBI Law Enforcement Bulletin, March and May, 1984.

(3) Oien, C. T. *Case Management Issues From Crime Scene To Courtroom*, Clearwater, FL, 2007 [URL: http://nfstc.org/projects/trace/].

(4) Brettell, T. A.; Butler, J. M.; Saferstein, R. *Anal. Chem.* **2005**, 77, 3839−3860.

(5) Kirkbride, K.; Tungol, M. *Infrared Microspectroscopy of Fibers*, 2nd ed.; CRC: New York, 1999.

(6) Adolf, F.; Dunlop, J. *Microspectrophotometry/Color Measurement*, 2nd ed.; CRC: New York, 1999.

(7) Grieve, M. C.; Biermann, T. W.; Davignon, M. *Sci. Justice* **2001**, 41, 245−260.

(8) Grieve, M. C.; Biermann, T. W.; Davignon, M. *Sci. Justice* **2003**, 43, 5−22.

(9) Biermann, T. W. *Sci. Justice* **2007**, 47, 68−87.

(10) Abu-Rous, M.; Schuster, K. C.; Adlassnig, W.; Lichtscheidl, I. *Melliand Int.* **2007**, 13, 382−384.

(11) Mueller, M.; Murphy, B.; Burghammer, M.; Snigireva, I.; Riekel, C.; Gunneweg, J.; Pantos, E. *Appl. Phys. A: Mater. Sci. Process.* **2006**, 83, 183−188.

(12) Grieve, M. C.; Dunlop, J.; Haddock, P. *J. Forensic Sci.* **1988**, 33, 1332−1344.

(13) Challinor, J. M. *Fiber Identification By Pyrolysis Techniques*, 2nd ed.; CRC: New York, 1999.

(14) Wiggins, K. G. *Thin Layer Chromatographic Analysis For Fiber Dyes*, 2nd ed.; CRC: New York, 1999.

(15) Wiggins, K. G.; Holness, J. A. *Sci. Justice* **2005**, 45, 93−96.

(16) Griffin, R, Speers, J. *Other Methods Of Colour Analysis: High Performance Liquid Chromatography*; CRC: Boca Raton, FL, 1999.

(17) Trojanowicz, M.; Wojcik, L.; Urbaniak-Walczak, K. *Chem. Anal.* **2003**, 48, 607−620.

(18) Xu, X.; Leijenhorst, H.; Van de Hoven, P.; de Koeijer, J.; Logtenberg, H. *Sci. Justice* **2001**, 41, 93−105.

(19) Tuinman, A. A.; Lewis, L. A.; Lewis, S. A. *Anal. Chem.* **2003**, 75, 2753−2760.

(20) Huang, M.; Yinon, J.; Sigman, M. E. *J. Forensic Sci.* **2004**, *49*, 238−249.

(21) Stefan, A. R.; Dockery, C. R.; Baguley, B. M.; Vann, B. C.; Nieuwland, A. A.; Hendrix, J. E.; Morgan, S. L. *Anal. Bioanal. Chem.* **2009**, *394*, 2087−2094.

(22) Petrick, L. M.; Wilson, T. A.; Fawcett, W. R. *J. Forensic Sci.* **2006**, *51*, 771−779.

(23) Morgan, S. L., Vann, B. C., Baguley, B. M., Stefan, A. R. *Advances In Discrimination Of Dyed Textile Fibers Using Capillary Electrophoresis/Mass Spectrometry*; Clearwater Beach, FL, 2007.

(24) Soltzberg, L. J.; Hagar, A.; Kridaratikorn, S.; Mattson, A.; Newman, R. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 2001−2006.

(25) Suzuki, E. *Forensic Applications Of Infrared Spectroscopy*; Regents/Prentice Hall: NJ, 1993.

(26) Gilbert, C.; Kokot, S. *Vib. Spectrosc.* **1995**, *9*, 161−167.

(27) Kokot, S.; Crawford, K.; Rintoul, L.; Meyer, U. *Vib. Spectrosc.* **1997**, *15*, 103−111.

(28) Jochem, G.; Lehnert, R. *J. Sci. Justice* **2002**, *42*, 215−222.

(29) Thomas, J.; Buzzini, P.; Massonnet, G.; Reedy, B.; Roux, C. *Forensic Sci. Int.* **2005**, *152*, 189−197.

(30) Massonnet, G.; Buzzini, P.; Jochem, G.; Stauber, M.; Coyle, T.; Roux, C.; Thomas, J.; Leijenhorst, H.; Van Zanten, Z.; Wiggins, K. G.; Russel, C.; Chabli, S.; Rosengarten, A. *J. Forensic Sci.* **2005**, *50*, 1028−1038.

(31) Casado, F.; Leona, M.; Lombardi, J. R.; Van Duyne, R. P. *Acc. Chem. Res.* **2010**, *43*, 782−791.

(32) Appalaneni, K.; Heider, E. C.; Moore, A. F. T.; Campiglia, A. D. *Anal. Chem.* **2014**, *86*, 6774−6780.

(33) Mujumdar, N.; Heider, E. C.; Campiglia, A. D. *Appl. Spectrosc.* **2015**, *69* (12), 1390−1396.

(34) Vandeginste, B. G. M., Massart, D. L., Buydens, L. M. C., De Jong, S., Lewi, P. J., Simeyers-Verbeke, J. *Handbook Of Chemometrics And Qualimetrics, Part B*; Elsevier: Amsterdam, 1998.

(35) Kemsley, E. K. *TrAC, Trends Anal. Chem.* **1998**, *17*, 24−34.

(36) Indahl, U. G. *J. Chemom.* **2014**, *28*, 168−180.

(37) Barker, M.; Rayens, W. *J. Chemom.* **2003**, *17*, 166−173.

(38) Bro, R. *Chemom. Intell. Lab. Syst.* **1997**, *38*, 149−171.

(39) Berrueta, L. A.; Alonso-Salces, R. M.; Heberger, K. *J. Chromatogr. A* **2007**, *1158*, 196−214.

(40) Arruda, A. F.; Goicoechea, H. C.; Santos, M.; Campiglia, A. D.; Olivieri, A. C. *Environ. Sci. Technol.* **2003**, *37*, 1385−1391.

(41) Ouertani, S. S.; Mazerolles, G.; Boccard, J.; Rudaz, S.; Hanafi, M. *Chemom. Intell. Lab. Syst.* **2014**, *133*, 25−32.

(42) Brereton, R. G.; Lloyd, G. R. *J. Chemom.* **2014**, *28*, 213−225.

(43) de Almeida, M. R.; Correa, D. N.; Rocha, W. F. C.; Scafi, F. J. O.; Poppi, R. J. *Microchem. J.* **2013**, *109*, 170−177.

(44) Ballabio, D.; Consonni, V. *Anal. Methods* **2013**, *5*, 3790−3798.

(45) Olivieri, A. C.; Wu, H. L.; Yu, R. Q. *Chemom. Intell. Lab. Syst.* **2009**, *96*, 246−251.

(46) Olivieri, A. C., Escandar, G. M. *Practical Three-Way Calibration*, 1st ed.; Elsevier: Amsterdam, 2014.

(47) Malinowski, E. R. *J. Chemom.* **1989**, *3*, 49−60.

(48) Christensen, J. H.; Hansen, A. B.; Mortensen, J.; Andersen, O. *Anal. Chem.* **2005**, *77*, 2210−2217.

(49) Guimet, F.; Boque, R.; Ferre, J. *Chemom. Intell. Lab. Syst.* **2006**, *81*, 94−106.

(50) Damiani, P. C.; Duran-Meras, I.; Garcia-Reiriz, A.; Jimenez Giron, A.; Muñoz de la Peña, A.; Olivieri, A. C. *Anal. Chem.* **2007**, *79*, 6949−6958.

(51) Bro, R.; Kiers, H. A. L. *J. Chemom.* **2003**, *17*, 274−286.

(52) Haaland, D. M.; Thomas, E. V. *Anal. Chem.* **1988**, *60*, 1193−1202.

(53) Berrueta, L. A.; Alonso-Salces, R. M.; Heberger, K. *J. Chromatog. A* **2007**, *1158*, 196−214.

(54) Guimet, F.; Boqué, R.; Ferré. *Chemom. Intell. Lab. Syst.* **2006**, *81*, 94−106.

(55) Arancibia, J. A.; Boschetti, C. E.; Olivieri, A. C.; Escandar, G. M. *Anal. Chem.* **2008**, *80*, 2789−2798.