Environ. Sci. Technol. **2010,** *44,* 9405–9412

Measurement of Dissolved Organic Matter Fluorescence in Aquatic Environments: An Interlaboratory Comparison

KATHLEEN R. MURPHY,*,†,‡ KENNA D. BUTLER,§ ROBERT G. M. SPENCER,|| COLIN A. STEDMON,⊥ JENNIFER R. BOEHME,# AND GEORGE R. AIKEN*,§

School of Civil and Environmental Engineering, The University of New South Wales, Sydney, New South Wales 2052, Australia, Smithsonian Environmental Research Center, P.O. Box 28, Edgewater, Maryland 21037, United States, U.S. Geological Survey, 3215 Marine Street, Boulder, Colorado 80303, United States, Woods Hole Research Center, 149 Woods Hole Road, Falmouth, Massachusetts 02540, United States, Department of Marine Ecology, National Environmental Research Institute, Aarhus University, Frederiksborgvej 399, DK-4000 Roskilde, Denmark, and BIO Division of Biological Infrastructure, National Science Foundation, 4201 Wilson Boulevard, Arlington, Virginia 22230, United States

Received July 13, 2010. Revised manuscript received October 26, 2010. Accepted October 26, 2010.

The fluorescent properties of dissolved organic matter (DOM) are often studied in order to infer DOM characteristics in aquatic environments, including source, quantity, composition, and behavior. While a potentially powerful technique, a single widely implemented standard method for correcting and presenting fluorescence measurements is lacking, leading to difficulties when comparing data collected by different research groups. This paper reports on a large-scale interlaboratory comparison in which natural samples and well-characterized fluorophores were analyzed in 20 laboratories in the U.S., Europe, and Australia. Shortcomings were evident in several areas, including data quality-assurance, the accuracy of spectral correction factors used to correct EEMs, and the treatment of optically dense samples. Data corrected by participants according to individual laboratory procedures were more variable than when corrected under a standard protocol. Wavelength dependency in measurement precision and accuracy were observed within and between instruments, even in corrected data. In an effort to reduce future occurrences of similar problems, algorithms for correcting and calibrating EEMs are described in detail, and MATLAB scripts for implementing the study's protocol are provided. Combined with the recent expansion of spectral fluorescence standards, this approach will serve to increase the intercomparability of DOM fluorescence studies.

* Corresponding author phone: +61 2 9385 5227; fax: +61 2 9313 8624; e-mail: krm@unsw.edu.au (K.R.M.); phone: (303)541-3036; fax: (303)541-3084; e-mail: graiken@usgs.gov (G.R.A.).
† The University of New South Wales.
‡ Smithsonian Environmental Research Center.
§ U.S. Geological Survey.
|| Woods Hole Research Center.
⊥ Aarhus University.
National Science Foundation.

Introduction

Fluorescence is a popular tool for studying and monitoring the concentration and nature of dissolved organic matter (DOM) in aquatic systems. In all natural waters, a fraction of the total dissolved organic matter pool that absorbs light (CDOM) also fluoresces in the UV and visible regions of the spectrum. This fluorescence is often measured across a range of excitation and emission wavelengths using Excitation−Emission Matrix Spectroscopy (EEMs), producing a 3-dimensional fluorescence intensity landscape in which the presence of distinctive peaks can provide indications of sources, behavior, and biogeochemical cycling of DOM (*1*−*3*). Presently, a diverse user community including scientists and engineers use fluorescence techniques to investigate aquatic systems. Implementation more broadly by the water science community is hampered by a lack of awareness regarding the proper use of fluorescence technologies and the absence of community consensus on methodologies and applications.

Superficially, fluorescence is an attractive method because data collection is straightforward, it provides information about both the amount and composition of fluorescent DOM, it can be readily employed in detector systems for a variety of process-based studies and separation techniques to study DOM composition, and it can be employed in situ, allowing for the collection of environmental data in real time (*4, 5*). However, the ease of data collection and the potentially powerful applications of fluorescence spectroscopy to monitor compositional changes in DOM belie the inherent complexity of the method. Fluorescence measurements of CDOM are influenced by artifacts (bias and error) from a range of sources, including the sample constituents (concentration and matrix effects), sample acquisition, handling, and measurement.

Broadly speaking, analytical artifacts may be related to instrumental configuration and essentially static, or else related to its operation or the samples themselves, and are potentially dynamic over time. Static artifacts result from imperfections in the optical components or their alignment, or variations in the efficiency at which different wavelengths of light are transmitted through the monochromators, resulting in distortions of the excitation or emission spectra. Further, the output of all polychromatic light sources is a function of wavelength and subject to time-dependent changes in wavelength distribution. Typically with modern instruments, these changes are largely accounted for automatically or as an option (measurement in ratio mode), in which the fluorescence signal obtained from each excitation wavelength is normalized to a measure of the intensity of the exciting beam (i.e., corrected for reference intensity (*6*)).

Quantifying analytical bias associated with DOM fluorescence measurements requires stable and well-characterized standard reference materials that fluoresce in the UV region; such certified fluorescence materials have only recently become available (*7, 8*) and have yet to be widely implemented. The standard procedure for accounting for spectral bias in fluorescence measurements used within the CDOM community and within this paper is to multiply the raw EEM by instrument-specific correction factors supplied by an instrument's manufacturer or generated in house (*1*).

Fluorescence data are often presented in arbitrary units, or, alternatively, total lamp output may be accounted for by

10.1021/es102362t
Published on Web 11/11/2010

normalizing intensities to an external standard, producing data expressed in units of intensity that should be comparable between instruments. Standards typically used are the slope of a quinine sulfate dilution series (*1, 9*) or the area under a clean water Raman scatter peak (*10*). In optically dense samples, absorption of the excitation or emission radiation by molecules in the sample matrix leads to the underestimation of emitted fluorescence, a phenomenon known as the inner filter effect (IFE). Several methods have been proposed allowing posthoc adjustment for inner filtering that require companion measurements of some kind (*11*−*14*). This study employed a commonly implemented algorithm for correcting moderately absorbent samples measured in a typical fluorometer utilizing right-angled geometry, which uses absorbance spectra to adjust fluorescence intensities (*14, 15*).

This paper presents the results of a large scale interlaboratory study designed to assess and provide recommendations on EEM data comparability issues. Twenty laboratories participating in the blind study each analyzed a duplicated set of samples that included natural fluorophores spanning a range of DOM types likely to be encountered in natural aquatic systems together with some well-characterized fluorophores. Corrected EEMs submitted by participants were compared with EEMs corrected from raw data files under a consistent protocol, to assess how diversity in protocols affects between-laboratory comparability. Scans of fluorescence standards were examined for bias remaining in corrected spectra, and the wavelength dependency of precision in corrected EEMs examined in relation to whether intensities were calibrated to a quinine sulfate dilution series or to Raman area. It is intended that the results of this study will provide the basis for fruitful discussion of standard approaches to improve data collection and comparison among the increasing numbers of scientists and engineers using fluorescence spectroscopy in aquatic systems.

Methods

Samples. Participating laboratories were shipped a package that included 14 blind samples labeled A-N, together with a working solution of 1000 ppb quinine sulfate dihydrate (NIST Standard Reference Material 936a) dissolved in 0.1 N H_2SO_4. Samples consisted of blind duplicates of seven fluorescent solutions selected to encompass a range of aquatic environments, fluorophore concentration and compositions, and absorbance characteristics (see the Supporting Information). Two capsule-filtered (0.45 μm) fresh waters were included after dilution to absorbance in a 1-cm cell at $\lambda = 254$ nm (A_{254}) below 0.2: humic rich Penobscot River at Eddington, Maine (samples D, H) and anthropogenically affected water from Boulder Creek, Colorado, collected downstream from the wastewater treatment plant (samples E, K). Three well-characterized fulvic acid isolates were diluted in nanopure water to $A_{254\ nm} < 0.12$: Pony Lake (Antarctica) (IHSS Standard, 1/2006; samples C, N), Suwannee River (IHSS Standard, 1/1983; samples B, M), and Pacific Ocean (100 m depth, 2/1986 (*16*); samples A, L), respectively, representing a saline and hypereutrophic coastal pond, a black water river dominated by allochthonous sources, and marine DOM. A seawater Consensus Reference Material from the Florida Strait (700 m depth, lot 12/2007; samples F, I) was purchased from the University of Miami RSMAS in acidified sealed ampules and provided unlabeled to participants. Quinine sulfate dihydrate (QS) was diluted in 0.1 N sulfuric acid to an approximate concentration of 15 ppb for samples G and J. The six unacidified solutions were transferred to 40 mL amber glass vials with Teflon lined caps without headspace and refrigerated prior to shipment.

Samples were shipped to participants between April 10−15 and analyzed between April 19−June 18, 2008, with approximately two-thirds of participants complying with instructions specifying analysis within 2 weeks. In refrigerated control samples analyzed 4 times between April 17-May 30, no significant effects of storage on fluorescence intensities were observed (see the Supporting Information).

In addition, two sets of commercial polymer fluorescence reference materials (6BF, Starna Scientific Ltd.) were circulated between laboratories. Only a partial data set could be analyzed due to some cells becoming damaged before completion of measurements. Multiple scans were obtained for SR3: p-terphenyl, SR5: compound 610 and SR6: Rhodamine B.

Measurements. Participants performed a series of measurements on each unknown sample. Measurements included (1) UV absorbance (200−700 nm in 1-nm intervals) using individual laboratory specifications for blanking or baseline correction; (2) DOC concentrations determined according to the laboratory's standard procedures; (3) raw EEMs collected at specified excitation wavelengths (λ_{ex}) of 240−450 nm in 5-nm intervals and emission wavelengths (λ_{em}) of 300−600 nm in 2-nm intervals, (4) blank EEMs (as above) and Raman scans ($\lambda_{ex} = 350$ nm, $\lambda_{em} = 365$−450 nm in 0.5 nm intervals) of Milli-Q water, collected daily, (5) emission scan ($\lambda_{ex} = 350$ nm) of a QS dilution series (0, 20, 50, 100 ppb) developed from the supplied 1000 ppb working solution, and (6) single excitation or emission scans of 6BF cells. Laboratories submitted these raw data files along with spectral correction factors (determined in-house or supplied by the instruments manufacturers) and EEMs corrected according to the lab's usual protocols.

Protocols requested for the study included (1) warm up fluorometer lamps for 1 h before beginning measurements, (2) do not dilute samples, (3) analysis at 20 °C where possible, and (4) 5 nm slit widths and 0.25 s integration times for all scans. An electronic workbook noting the procedures implemented and reporting corrected data at requested wavelength positions was returned by each laboratory along with data files.

Data Analysis. *Absorbance, $SUVA_{254}$, and DOC.* Laboratories submitted absorbance scans in the form of either napierian absorption coefficients (α_λ, m^{-1}) or decadic absorbance (A_λ, no units); the two are related according to eq 1

$$\alpha_\lambda = 2.303 A_\lambda / L \quad (1)$$

where A_λ is the measured absorbance, λ is the wavelength (nm), and L is the optical path length in meters (*17*). All scans were converted to the decadic form of UV absorbance in a 1 cm cell. After comparing scans and discarding data that were below detection limits, with excessive noise, or that deviated substantially from the bulk of scans, a single best-average absorbance scan was determined for each sample from averaging the scans from the four laboratories that agreed most closely with the median of the data set. This best-average scan was used to generate inner filter correction factors (step 2 below) and calculate specific ultraviolet absorbance ($SUVA_{254}$). $SUVA_{254}$ is defined as A_{254} measured in inverse meters divided by the best-average DOC concentration (in $mgCL^{-1}$) (*18*).

EEM Correction Procedures. The first step implemented in the data analysis was to reproduce the lab-corrected EEMs using the raw EEMs and associated files supplied by each lab. No adjustments were made for differences in sample temperatures. These reproduced data are referred to hereafter as 'study-corrected EEMs'. Figure 1 illustrates a 5-step correction procedure that along with the equations below documents the process used to convert raw EEMs provided in arbitrary units (AU) to study-corrected EEMs expressed in Raman units (RU) and parts-per-billion quinine sulfate equivalent units (QSE). In the equations, element-wise

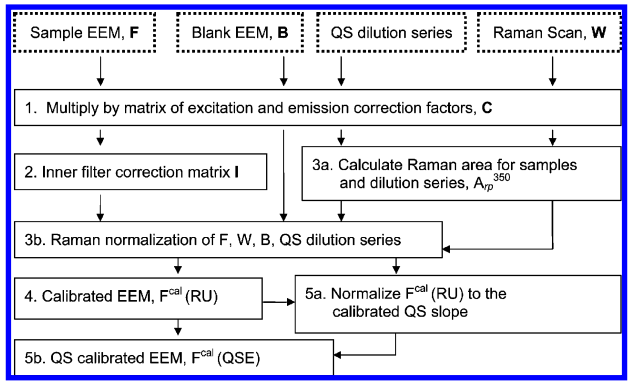

FIGURE 1. Flow diagram of EEM correction procedures implemented in this study.

multiplication is denoted by the symbol "•" to distinguish it from classical matrix multiplication.

The Supporting Information includes a program for MATLAB ("FDOMcorrect.m") that allows flexible implementation of the algorithms described below, including optional inner filter correction, blank subtraction, and QS calibration steps. The method in this study used independent water Raman scans at $\lambda_{ex} = 350$ nm and 0.5-nm emission intervals to calculate Raman peak areas instead of extracting Raman scans from Milli-Q EEMs (2-nm emission intervals), for better peak resolution.

Step 1: Spectral corrections

Variations in spectral responses were accounted for by the element-wise multiplication of each sample (**F**) and water blank EEM (**B**) by the EEM-sized, instrument-specific correction matrix (**C**). The matrix **C** was calculated by calculating the outer product of the vector of excitation correction factors, $\mathbf{f}_{\lambda ex}^{Excorr}$, and the transposed vector of emission correction factors, $(\mathbf{f}_{\lambda em}^{Emcorr})^{\mathrm{T}}$

$$\mathbf{C} = \mathbf{f}_{\lambda ex}^{Excorr} \otimes (\mathbf{f}_{\lambda em}^{Emcorr})^{\mathrm{T}} \tag{2}$$

$$\mathbf{F}^{corr}(AU) = \mathbf{C} \bullet \mathbf{F} \tag{3}$$

$$\mathbf{B}^{corr}(AU) = \mathbf{C} \bullet \mathbf{B} \tag{4}$$

The water Raman scan $\mathbf{W}_{350,\lambda em}$ is the emission vector corresponding to $\lambda_{ex} = 350$ nm for a water blank sample

$$\mathbf{W}_{350,\lambda em}^{corr}(AU) = \mathbf{C}_{350,\lambda em} \bullet \mathbf{W}_{350,\lambda em} \tag{5}$$

For the linear 0−50 ppb portion of the QS dilution series, the slope $m_{350/450}$ was calculated from linear regression and multiplied by the correction factor at $\lambda_{ex}/\lambda_{em} = 350/450$ nm

$$m_{350/450}^{corr}(QSE^{-1}) = \mathbf{C}_{350/450} \bullet m_{350/450} \tag{6}$$

Step 2: Inner filter correction

Inner filtering was accounted for by element-wise multiplication of each spectrally corrected EEM, $\mathbf{F}^{corr}$, by a correction matrix, **I**, calculated for each wavelength pair from absorbance A_λ of the same sample. Equation 7 was proposed by Parker and Barnes (*14*) and simplified (*19*) by assuming a 0.5 cm path length of excitation and emitted light in a typical 1 cm cuvette. To avoid introducing absorbance-related variation into the study-corrected data set, correction factors were generated from each sample-specific 'best-average' absorbance scan and applied to samples from all laboratories. Exceptions were 6 samples from one laboratory diluted to $A_{300} < 0.02$; these were multiplied by the dilution factor

$$\mathbf{I}_{\lambda_{ex},\lambda_{em}} = 10^{0.5(A_{\lambda_{ex}} + A_{\lambda_{em}})} \tag{7}$$

$$\mathbf{F}^{IFE} = \mathbf{I} \bullet \mathbf{F}^{corr} \tag{8}$$

Step 3: Raman normalization

The sample EEMs, blanks, and dilution series slopes were then normalized to the Raman peak area of a wavelength-corrected Raman scan collected on the same day, to produce corrected fluorescence in Raman Units

$$\mathbf{F}^{cal}(RU) = \frac{1}{\mathrm{A}_{rp}^{350}} \mathbf{F}^{IFE} \tag{9}$$

$$\mathbf{B}^{cal}(RU) = \frac{1}{\mathrm{A}_{rp}^{350}} \mathbf{B}^{corr} \tag{10}$$

$$m_{350/450}^{cal}(RUQSE^{-1}) = \frac{1}{\mathrm{A}_{rp}^{350}} m_{350/450}^{corr} \tag{11}$$

where the corrected Raman peak area (A_{rp}^{350}) is a scalar value specific to each sample, defined as follows

$$A_{rp}^{350} = \int_{381}^{426} \mathbf{W}_{350,\lambda em}^{corr} d\lambda_{em} \tag{12}$$

In eq 12, A_{rp}^{350} represents the baseline-corrected area under the curve $\mathbf{W}_{350,\lambda em}^{corr}$ from eq 5 that is bounded by the emission wavelengths 381 and 426 nm. It is important to note that Raman areas can vary over time due to fluctuating PMT signal strength, resulting in A_{rp}^{350} that varies among samples and dilution series in a data set.

Step 4: Blank subtraction

Much of the water scatter signal can be removed by subtracting from each sample EEM the Raman calibrated blank EEM $\mathbf{B}^{cal}$, obtained as close as possible in time to the sample. Although documented for completeness, blank subtraction is omitted from the study protocol since its role in improving fluorescence measurement accuracy is questionable, particularly for dilute samples (*20*)

$$\mathbf{F}^{cal}(RU) = \mathbf{F}^{cal}(RU) - \mathbf{B}^{cal}(RU) \tag{13}$$

Step 5: Quinine-sulfate calibration

The Raman normalized and spectrally corrected EEMs ($\mathbf{F}^{cal}$) were calibrated against the QS dilution series (0−100 ppb), producing EEMs in QSE (1 QSE = 1 ppb quinine bisulfate in 0.1 N H_2SO4, measured at $\lambda_{ex}/\lambda_{em} = 350/450$ nm). The Raman calibrated EEMs were divided by the calibrated slope (eq 11) as follows

$$\mathbf{F}^{cal}(QSE) = \frac{1}{m_{350/450}^{cal}} \mathbf{F}^{cal}(RU) \tag{14}$$

Note that inner filter correction of the QS series was not required because only the low-intensity linear range of the dilution series was used to determine *s* according to the equation $y = s\,x + int$. Similarly, blank subtraction of the dilution series is unnecessary because this step affects only the intercept, *int*.

Assessing Variability. Before comparing spectra between laboratories, a new data set of 7 EEMs was created for each lab by averaging the measurements of duplicate samples. Throughout this paper, the terminology "X&Y" denotes the average of replicate EEMs X and Y; for example, E&K refers to the EEM obtained by averaging duplicate samples E and K of Boulder Creek water.

Bootstrapping was used to calculate the between-laboratory variability in fluorescence measurements, because bootstrapped standard errors are not reliant on any distributional assumptions, such as normality (*21*). To determine bootstrapped means and standard errors, pseudosamples were created by randomly selecting 1 sample *n* times from

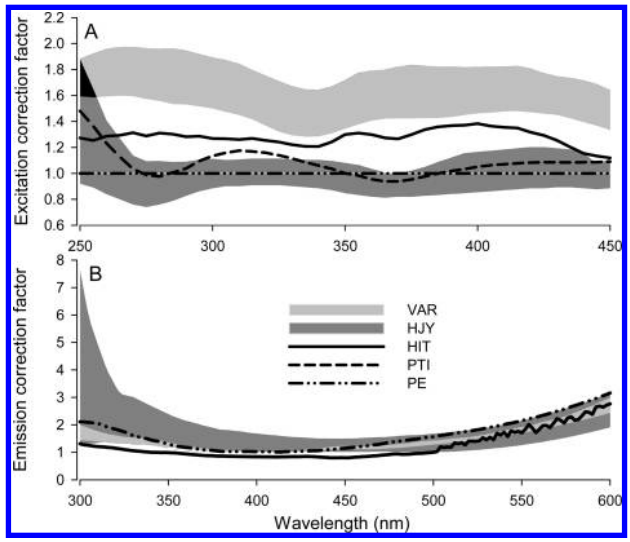

FIGURE 2. Ranges of submitted spectral correction factors according to instrument manufacturer for (A) excitation and (B) emission spectra. Instrument manufacturers were VAR: Varian, HJY: Horiba Jobin Yvon, HIT: Hitachi, PTI: Photon Technologies International, PE: Perkin-Elmer.

each data set of size n. This sample was replaced after each selection in order that some samples could be selected more than once and others not selected. When this operation is repeated many times, the distribution of the pseudosamples is similar to the underlying data distribution, and the standard deviation of the pseudosample means gives a bootstrap standard error of the estimate (*21*).

Due to the varying characteristics of data sets submitted by participants in this study, it was generally possible to include only a subset of laboratories in each analysis. The Supporting Information includes an allocation matrix indicating which data sets were included and the reasons for exclusions where applicable.

Results

Prevailing Methodologies. Nearly all of the spectrophotometers used in this study were from HORIBA Jobin Yvon (FluoroMax 2−4 and Fluorolog3, $n = 11$) or Varian (Cary Eclipse, $n = 6$), with only one representative from each of Photon Technologies International (QM4SE), Hitachi (F-4500), and Perkin-Elmer (LS-50). The HORIBA fluorometers use continuous Xenon light sources and can scan at speeds up to 160 nm/s. The other fluorometers use pulsed Xenon lamps and adjust scan rates automatically depending on user-specified averaging times. Samples were analyzed at temperatures between 16−27.5 °C.

Participating laboratories differed in whether or not they routinely implemented blank subtraction, adjusted data for IFE, applied spectral corrections, and normalized or calibrated the data − resulting in 9 unique combinations of data correction steps (see the Supporting Information). Of the 60% of laboratories that calibrated EEMs in this study, the majority submitted data in Raman units (RU), while the remainder submitted data in quinine sulfate equivalent units (QSE).

Figure 2 illustrates the range of spectral correction factors used by different research groups. Similarities between spectral correction factors for instruments produced by the same manufacturer are evident; FluoroMaxes had excitation correction factors near 1.0 for wavelengths between 280−450 nm with local minima at 280 and 370 nm, compared with factors near 1.7 for the Cary Eclipse and local minima at 340 and 450 nm. The FluoroMax also displayed a wider range of correction factors at $\lambda_{ex} < 300$ nm, with a 2-fold range (i.e., 1.0−1.8) reported at $\lambda_{ex} = 250$ nm. Emission correction factors also varied by instrument; for the FluoroMax, emission correction factors were particularly variable below 360 nm and above 480 nm. The Perkin-Elmer had similar spectral corrections to the FluoroMax, while the Hitachi was more similar to the Cary Eclipse. For the PTI, excitation correction factors were similar to the FluoroMax, but emission correction factors extracted from the PTI spectrophotometer were clearly erroneous, so are not reported.

Validation of Lab-Corrected EEMs. Most lab-corrected EEMs submitted by laboratories that implemented spectral and inner filter corrections could be reproduced closely using eqs 2−14. The greatest differences between lab-corrected and study-corrected EEMs were due to inner filter effects; seven laboratories did not dilute samples or apply inner filter corrections even though most samples were optically dense (correction factors >1.0, see the Supporting Information). All 12 laboratories that applied IFE corrections calculated correction factors from absorbance measurements, thus absorbance measurements were sometimes a significant source of variability between lab-corrected and study-corrected EEMs, particularly at low λ_{ex} and λ_{em} where absorbance is greatest but precision is typically lowest.

Almost half of the laboratories initially submitted one or more EEMs that were anomalous. These laboratories were invited to recheck their files and/or procedures, identify the cause of anomalies, and then resubmit the data. The reasons for the anomalous data were ostensibly (1) neglecting to adjust EEM intensities to account for dilution (data too low by a factor of 2−3); (2) neglecting to account for the effects of varying the PMT detector sensitivity/voltage between samples and Raman scans (data too high by orders of magnitude); (3) a malfunctioning macro associated with proprietary software (data too high by a factor of 2−3); (4) subtracting uncorrected blanks from corrected sample EEMs (data too low by a factor of 2); (5) mix-ups in which data for one or more samples were assigned to the wrong sample names (3 laboratories), or (6) submitting the wrong emission correction factors, and (7) negative numbers in raw EEMs (2 laboratories). The laboratory that submitted diluted EEMs had reported the results correctly in their workbook, whereas the remaining laboratories had reported incorrect results. Finally, although the emission correction factors submitted for the PTI spectrophotometer were erroneous, the instrument-corrected EEMs were not. Major issues were resolved in all data sets with two exceptions: (1) it was not possible to obtain correction factors for the PTI, and (2) for one lab, it was impossible to calibrate study-corrected EEMs. The analyses described in the remainder of this paper deal with this revised data set.

Spectral Comparisons. In Figures 3 and 4, study-corrected spectra are compared for common fluorescence reference materials. In Figure 3, the spectra for samples G&J (averaged) are compared with published technical spectra for quinine sulfate (*22, 23*). Deviations exceeding ±10% for the emission spectrum (Figure 3b) and ±15% on excitation (Figure 3d) are apparent. Figure 4 shows scans for duplicate sets (black and gray lines) of three 6BF reference cells. These deviated from median scans by more than 25% at 300 and 400 nm emission, despite small differences between duplicate blocks scanned on a single instrument (see the Supporting Information).

Spectral variability between laboratories, and the effect of data corrections, are further exemplified by fluorescence ratios. In Table 1, the ratios of commonly reported humic-like fluorescence peaks A, C, M and protein-like fluorescence peaks T and B (*24*) are reported for selected samples in which these peaks were present. Applying corrections nearly always significantly changed the magnitude of intensity ratios. For example, the ratio of quinine sulfate emission at 450 nm when excited by 250 nm compared to 350 nm was 5.37 ±

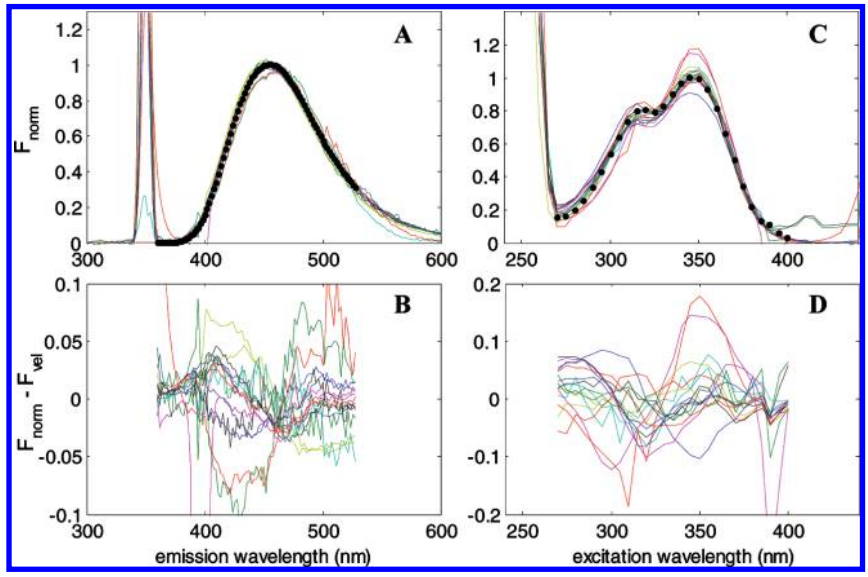

FIGURE 3. Comparison of study-corrected spectra for quinine sulfate with published spectra. Each curve represents the average intensity of replicates G&J from one laboratory (n = 17), normalized to the area under the curve. (A) Emission spectrum at Ex = 350 nm, normalized to the area under each curve between 410−530 nm and plotted against Velapoldi's (2004) technical emission spectrum; B) deviation of the emission spectra from the technical spectrum, (C) excitation spectrum at λ_{em} = 450 nm, normalized to the area under each curve between 270−390 nm and plotted against Velapoldi's technical excitation spectrum; and D) deviation of the excitation spectra from the technical spectrum.

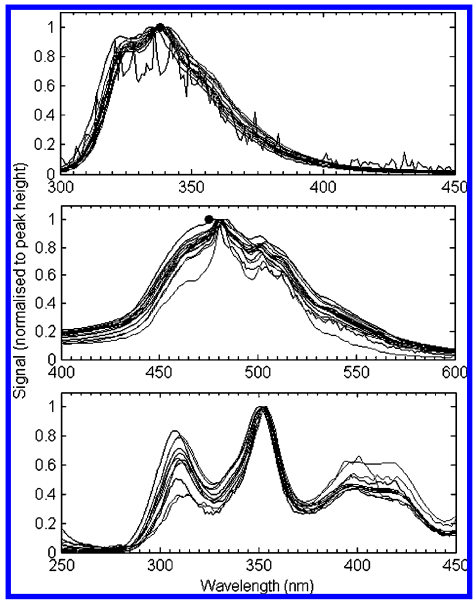

FIGURE 4. Study-corrected scans of Starna 6BF reference cells normalized to maximum peak height. Black and gray lines indicate two different sets of cells; black circles indicate positions of the main peak according to the manufacturer. A) Emission scan of SR3: p-terphenyl (λ_{ex} = 265 nm, n = 16); B) emission scan of SR5: compound 610 (λ_{ex} = 350 nm, n = 16); and C) excitation scan of SR6: Rhodamine B (λ_{em} = 562 nm, n = 13).

1.34 in lab-corrected EEMs, compared with a study-corrected EEM ratio of 5.61 ± 1.18 and an uncorrected EEM ratio of 4.69 ± 0.83. For the Boulder Creek sample, applying spectral corrections to the "uncorrected fluorescence index" (FI) used to distinguish 'terrestrial DOM' (low FI) from 'microbial DOM' (high FI) (*15*), decreased its magnitude by nearly 25%. Table 1 also shows the spectrally corrected (*25*) fluorescence index FI_c, for a range of samples. While lab-corrected and study-corrected ratios in Table 1 were not statistically different, the study-corrected ratios were usually more precise. Thus, the overall variability in ratios between laboratories as represented by coefficients of variation (CV = SD/mean) was typically lowest for study-corrected EEMs and highest for lab-corrected EEMs, with uncorrected EEMs slightly more variable than study-corrected EEMs. For example, FI_c for the Pony Lake sample had CVs of 10%, 13%, and 8% for uncorrected, lab-corrected, and study-corrected EEMs, respectively.

Fluorescence Intensity. Fluorescence intensities of peaks A, C, M, B and T in study-corrected EEMs are provided in the Supporting Information. Interlaboratory measurement precision varied between samples and between wavelengths and, as discussed below, between Raman versus QS normalized data. Typically, CVs were in the range of 10−20% for humic-like peaks A, C, and M and between 15−35% for protein-like peaks T and B (Table S7), while for replicate samples analyzed within the same laboratory, CVs were typically below 5% (Table S8), with peak T intensities reproduced least precisely. This was probably due to sporadic contamination, since peak T precision was similar for samples analyzed within 2 weeks compared with the entire data set and no biological production occurred in stored control samples (Table S9).

The wavelength dependency of measurement precision across the entire data set is illustrated for the Boulder Creek (E&K) sample in Figure 5, for which contours show changes in the relative standard error (RSE = SE/mean) as a function of λ_{ex} and λ_{em}. Among QS-calibrated EEMs, between laboratory precision was greatest in the neighborhood of 350/450 nm and decreased in the direction of both lower and higher emission wavelengths, reaching minima at the most extreme emission wavelengths. Among EEMs represented in Raman units, precision was greatest at emission wavelengths near 350−400 nm and also decreased in the direction of both lower and higher emission wavelengths, with the least precise measurements observed at high emission wavelengths (>520 nm). This dependence of precision upon the normalization method was also apparent in other samples, especially the samples from Penobscot River, Pony Lake, Suwannee River, and the central Pacific.

The results in Figure 5 are interesting considering that the data used in both figures were identical except that in Figure B, the data from Figure A have been multiplied by the Raman-normalized slope of the QS dilution series ($m_{350/450}^{cal}$)

TABLE 1. Intensity Ratios (Mean ± CV) for the Main Aquatic Organic Matter Fluorescence Peaks (A, C, M, B, T) and Indices (Uncorrected FI (*15*) and Corrected FI_c (*25*)) in Uncorrected (uncor.), Lab Corrected (lab-cor.), and Study Corrected (study-cor.) EEMs

					ratio		
ratio	sample	source	numerator $\lambda_{ex}/\lambda_{em}$	denominator $\lambda_{ex}/\lambda_{em}$	uncor.	lab-cor.	study-cor.
FI	E&K	Boulder Creek	370/450	370/500	1.90 ± 9%	1.60 ± 11%	1.52 ± 7%
FI_c	E&K	Boulder Creek	370/470	370/520	2.45 ± 10%	1.96 ± 12%	1.90 ± 8%
FI_c	C&N	Pony Lake	370/470	370/520	1.86 ± 10%	1.49 ± 13%	1.45 ± 8%
FI_c	I&F	Marine standard	370/470	370/520	2.52 ± 112%	1.58 ± 27%	1.51 ± 26%
FI_c	A&L	Pacific Ocean	370/470	370/520	1.84 ± 9%	1.50 ± 12%	1.44 ± 6%
FI_c	D&H	Penobscot Rr.	370/470	370/520	1.69 ± 9%	1.36 ± 13%	1.32 ± 8%
M/A	C&N	Pony Lake	320/410	250/450	0.65 ± 11%	0.54 ± 17%	0.49 ± 12%
M/A	I&F	Marine standard	320/410	250/450	0.68 ± 19%	0.56 ± 29%	0.53 ± 17%
M/C	D&H	Penobscot Rr.	320/410	350/450	0.93 ± 6%	0.90 ± 8%	0.91 ± 4%
A/C	G&J	Quinine sulfate	250/450	350/450	4.69 ± 18%	5.37 ± 25%	5.61 ± 21%
B/M	A&L	Pacific Ocean	270/304	300/410	1.10 ± 39%	1.65 ± 39%	1.91 ± 24%
T/C	E&K	Boulder Creek	280/340	350/450	0.66 ± 17%	0.90 ± 21%	0.93 ± 14%
T/A	E&K	Boulder Creek	280/340	250/450	0.52 ± 17%	0.58 ± 19%	0.58 ± 16%

for the corresponding laboratory. Each dilution series passed through the origin and had a slope (QSE/RU) in the range of 62.3−113.6, indicating that multipliers for converting between QSE and RU were instrument-specific.

Discussion

There is little doubt that the absence of a widely agreed standard protocol for treating raw fluorescence EEMs in the aquatic sciences is impeding scientific progress by fractionating research output into numerous data sets that are difficult or impossible to compare quantitatively. In this study, participating laboratories used any of nine unique routines implemented with in-house programs to produce corrected EEMs. This diversity of data correction procedures reduces transparency, decreases the efficiency of knowledge transfer, is a source of confusion among new researchers in the field, and probably contributed to the high rate of quality assurance problems observed during this study.

Failures in quality assurance typically resulted in over- or under-reporting of fluorescence intensities by a factor of approximately 2−3. Formalization and implementation of sound laboratory procedures for tracking samples within all laboratories would improve data quality and comparability. There is an urgent need for user-friendly software to facilitate data organization and implementation of standard correction and normalization routines; toward this goal, the Supporting Information includes the MATLAB program 'FDOMcorrect.m', which flexibly implements the study's EEM correction and calibration protocol. Reported $SUVA_{254}$ values differed by a factor of approximately 2.3 depending upon whether calculated from napierian (A_e) or decadic (A) absorbance. Clarifying the derivation of absorbance-related data is essential to reporting data unambiguously.

Diversity in correction procedures among participants affected between-laboratory precision, as indicated by consistently higher variability in lab-corrected fluorescence ratios compared with (study-corrected) ratios obtained using a consistent protocol. In optically dense samples, inner-filter corrections were a large component of this variability. All laboratories that applied IFE correction used the absorbance method (eq 7), indicating that absorbance measurement error contributed to variability between laboratories. IFE correction procedures that do not require absorbance measurement have been proposed (*13, 26, 27*) but have yet to be fully evaluated. Sample dilution not only is a popular alternative to mathematical IFE correction but also has constraints and disadvantages (*26*).

Within-laboratory (inter-replicate) precision was typically high with CVs below 5% in at least 89% of replicate measurements of commonly reported fluorescence peaks (A, C, M, B). T-peak precision was lowest, possibly due to more frequent contamination. Between-laboratory variability varied with wavelength, sometimes exceeding ±25%, although relative uncertainties below 10% are achievable even in scatter-affected EEM regions (*28*). As observed previously (*6, 29*), fluorescence intensities and ratios in uncorrected data sets were significantly different than those in corrected data sets. Some laboratories choose not to correct spectra

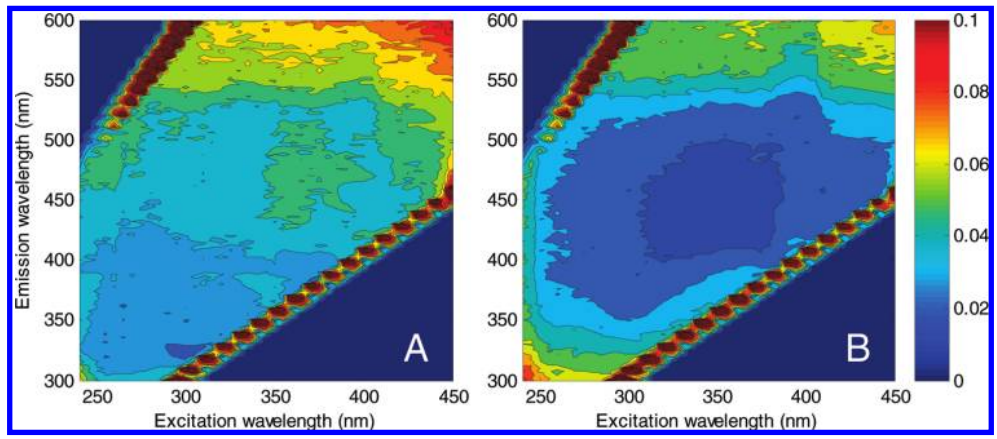

FIGURE 5. Bootstrapped interlaboratory relative standard errors for the study corrected and averaged Boulder Creek EEM in A) Raman units and B) QSE. The EEMs ($n = 14$) used in both figures were identical except that in part B, they were multiplied by the Raman-normalized slopes of QS dilution series.

on the basis that within-laboratory measurements on the same instrument are internally consistent; however, this has not prevented other researchers from attempting to draw spectral comparisons with many such studies. Further, as shown by Holbrook et al. (*6*), elimination of instrument bias can improve the accuracy of fluorophore detection, particularly at low λ_{ex} and λ_{em}. For these reasons, spectral correction is considered a critical component of producing publication-ready EEMs in most applications.

An important development in the EEM correction protocol presented in this paper is that it combines the principles of two prevailing but currently unrelated methods for calibrating fluorescence in Raman- (*30*) and QSE- (*1*) units in a single algorithmic procedure, which as a biproduct calculates an instrument-specific factor for converting between the two units of measurement ($m_{350/450}{}^{cal}$). Currently, QS-calibrations are typically performed by normalizing sample intensities to Raman intensities at either fixed (*1*) or varying (*9*) λ_{ex}. While QS-calibrated measurements computed according to either of these methods are comparable, since Raman heights are instrument dependent (*10*), data reported in Raman height units cannot be compared between instruments. Intensities reported in Raman (area) units are also often incomparable, due to varying λ_{ex} used to calculate Raman areas (*10, 30*). The program "FDOMcorrect.m" corrects fluorescence EEMs collected in S/R mode, calibrates them to Raman and QSE units, and calculates the between-unit conversion factor ($m_{350/450}{}^{cal}$). The user chooses the excitation wavelength for Raman normalization (default λ_{ex} =350 nm) and whether to use height or area (default is area). In applications where the absolute magnitude of fluorescence intensities is relevant - for example, the diagnosis of polluted water based on fluorescence intensity thresholds (*31, 32*), the widespread reporting of $1/m_{350/450}{}^{cal}$ (QSE/RU) at specified λ_{ex} used would allow for simple conversion between RU and QSE.

Due to $m_{350/450}{}^{cal}$ varying between laboratories in this study, EEMs normalized to the slope of the QS dilution curve were most precise in the region of $\lambda_{ex}/\lambda_{em} = 350/450$ nm, whereas EEMs normalized to the area under the water Raman peak at $\lambda_{ex} = 350$ nm were most precise at lower excitation and emission wavelengths. This illustrates that both QS and Raman area behave as would be expected of a fluorescence standard, by maximizing precision in the region of the wavelengths selected for measuring the standard. While operator-related uncertainty in dilution series slopes reduces the accuracy of QS calibrated EEMs, this appeared less important in this study than were instrument-specific differences affecting Raman-normalized intensities. The degree to which variation in $m_{350/450}{}^{cal}$ between instruments may be reduced by improving spectral corrections is important to clarify in future studies.

Scans of solid and liquid reference materials corrected according to the study's protocol indicated substantial variability due in part to inadequate spectral corrections. Manufacturer-supplied instrument correction factors can become decreasingly reliable due to aging optical and electro-optical components and general wear and tear. Whereas QS has long been recommended as a spectral standard for the humic peak region (*22*), new traceable short (<400 nm) and long (>550 nm) wavelength standards, together with easily implemented procedures for checking the accuracy of spectral correction factors, should be extremely valuable in this regard (*28, 33*). Both liquid (*7, 8*) and solid glass (*34*−*36*) fluorescence standards encompassing a broad spectral range are available.

Acknowledgments

We thank all of the participating laboratories for providing the data described in this paper. The unpublished manuscript benefited greatly from comments by A. Baker, L. Larson, K. Merriman and anonymous reviewers. This study was undertaken for the AGU Chapman Conference on Organic Matter Fluorescence convened by A. Baker and P. Coble and funded by the US Geological Survey and NASA Grant NNH04AA62I. Use of brand names in this paper is for identification purposes only and does not imply endorsement by the U.S. Geological Survey.

Supporting Information Available

Eleven appendices are provided summarizing sample characteristics, instrumentation, and procedures at participating laboratories, further methods details, and additional analyses. Also included are two scripts for implementing EEM corrections in MATLAB. This material is available free of charge via the Internet at http://pubs.acs.org.

Literature Cited

(1) Coble, P. G.; Mopper, K.; Schultz, C. S. Fluorescence contouring analysis of DOC intercalibration experiment samples: A comparison of techniques. *Mar. Chem.* **1993**, *41*, 173–178.

(2) Mopper, K.; Schultz, C. A. Fluorescence as a possible tool for studying the nature and water column distribution of DOC components. *Mar. Chem.* **1993**, *41* (1−3), 229–238.

(3) Coble, P. G.; Green, S. A.; Blough, N. V.; Gagosian, R. B. Characterization of dissolved organic matter in the Black Sea by fluorescence spectroscopy. *Nature* **1990**, *348*, 432–435.

(4) Chen, R. F. In situ fluorescence measurements in coastal waters. *Org. Geochem.* **1999**, *30* (6), 397–409.

(5) Boehme, J.; Wells, M. Fluorescence variability of marine and terrestrial colloids: Examining size fractions of chromophoric dissolved organic matter in the Damariscotta River estuary. *Mar. Chem.* **2006**, *101* (1−2), 95–103.

(6) Holbrook, R. D.; DeRose, P. C.; Leigh, S. D.; Rukhin, A. L.; Heckert, N. A. Excitation-emission matrix fluorescence spectroscopy for natural organic matter characterization: A quantitative evaluation of calibration and spectral correction procedures. *Appl. Spectrosc.* **2006**, *60* (7), 791–799.

(7) Pfeifer, D.; Hoffmann, K.; Hoffmann, A.; Monte, C.; Resch-Genger, U. The calibration kit spectral fluorescence standards - A simple and certified tool for the standardization of the spectral characteristics of fluorescence instruments. *J. Fluoresc.* **2006**, *16* (4), 581–587.

(8) Tsvirko, M. P.; Svetashev, A. G. A calibration kit of fluorescent standards for the UV spectral range of 230−410 nm. *Opt. Mater.* **2009**, *31* (12), 1842–1844.

(9) Hoge, F. E.; Vodacek, A.; Blough, N. V. Inherent optical properties of the ocean: Retrieval of the absorption coefficient of chromophoric dissolved organic matter from fluorescence measurements. *Limnol. Oceanogr.* **1993**, *38* (7), 1394–1402.

(10) Lawaetz, A. J.; Stedmon, C. A. Fluorescence Intensity Calibration Using the Raman Scatter Peak of Water. *Appl. Spectrosc.* **2009**, *63* (8), 936–940.

(11) Gu, Q.; Kenny, J. E. Improvement of Inner Filter Effect Correction Based on Determination of Effective Geometric Parameters Using a Conventional Fluorimeter. *Anal. Chem.* **2009**, *81* (1), 420–426.

(12) Kubista, M.; Sjoback, R.; Eriksson, S.; Albinsson, B. Experimental correction for the inner-filter effect in fluorescence-spectra. *Analyst* **1994**, *119* (3), 417–419.

(13) Luciani, X.; Mounier, S.; Redon, R.; Bois, A. A simple correction method of inner filter effects affecting FEEM and its application to the PARAFAC decomposition. *Chemom. Intell. Lab.* **2009**, *96* (2), 227–238.

(14) Parker, C. A.; Barnes, W. J. Some experiments with spectrofluorimeters and filter fluorimeters. *Analyst* **1957**, *82*, 606–618.

(15) McKnight, D. M.; Boyer, E. W.; Westerhoff, P. K.; Doran, P. T.; Kulbe, T.; Andersen, D. T. Spectrofluorometric characterization of dissolved organic matter for indication of precursor organic material and aromaticity. *Limnol. Oceanogr.* **2001**, *46* (1), 38–48.

(16) Malcolm, R. The uniqueness of humic substances in each of soil, stream and marine environments. *Anal. Chim. Acta* **1990**, *232*, 19–30.

(17) Verhoeven, J. W. Glossary of terms used in photochemistry (IUPAC Recommendations 1996). *Pure Appl. Chem.* **1996**, *68* (12), 2223–2286.
(18) Weishaar, J. L.; Aiken, G. R.; Bergamaschi, B. A.; Fram, M. S.; Fujii, R.; Mopper, K. Evaluation of Specific Ultraviolet Absorbance as an Indicator of the Chemical Composition and Reactivity of Dissolved Organic Carbon. *Environ. Sci. Technol.* **2003**, *37* (20), 4702–4708.
(19) Lakowicz, J. R. *Principles of fluorescence spectroscopy*, 3rd ed.; Plenum Press: New York, 2006.
(20) Zepp, R. G.; Sheldon, W. M.; Moran, M. A. Dissolved organic fluorophores in southeastern US coastal waters: correction method for eliminating Rayleigh and Raman scattering peaks in excitation-emission matrices. *Mar. Chem.* **2004**, *89* (1−4), 15–36.
(21) Efron, B. Nonparametric estimates of standard error - the jackknife, the bootstrap and other methods. *Biometrika* **1981**, *68* (3), 589–599.
(22) Velapoldi, R. A.; Mielenz, K. D. *Standard reference materials: A fluorescence standard reference material- quinine sulfate dihydrate; NBS Special Publication 260-64*; National Bureau of Standards: Washington, DC, 1980.
(23) Velapoldi, R. A.; Tønnesen, H. H. Corrected emission spectra and quantum yields for a series of fluorescent compounds in the visible spectral region. *J. Fluoresc.* **2004**, *14* (4), 465–472.
(24) Coble, P. G. Characterization of marine and terrestrial DOM in seawater using excitation-emission matrix spectroscopy. *Mar. Chem.* **1996**, *51*, 325–346.
(25) Maie, N.; Parish, K. J.; Watanabe, A.; Knicker, H.; Benner, R.; Abe, T.; Kaiser, K.; Jaffe, R. Chemical characteristics of dissolved organic nitrogen in an oligotrophic subtropical coastal ecosystem. *Geochim. Cosmochim. Acta* **2006**, *70* (17), 4491–4506.
(26) Larsson, T.; Wedborg, M.; Turner, D. Correction of inner-filter effect in fluorescence excitation-emission matrix spectrometry using Raman scatter. *Anal. Chim. Acta* **2007**, *583* (2), 357–363.
(27) Lutz, H. P.; Luisi, P. L. Correction for inner filter effects in fluorescence spectroscopy. *Helv. Chim. Acta* **1983**, *66* (7), 1929–1935.
(28) Resch-Genger, U.; Hoffmann, K.; Hoffmann, A. Standardization of fluorescence measurements - Criteria for the choice of suitable standards and approaches to fit-for-purpose calibration tools. In *Fluorescence Methods and Applications: Spectroscopy, Imaging, and Probes*; Wolfbeis, O. S., Ed.; Blackwell Publishing: Oxford, 2008.
(29) Hofstraat, J. W.; Latuhihin, M. J. Correction of fluorescence-spectra. *Appl. Spectrosc.* **1994**, *48* (4), 436–447.
(30) Determann, S.; Reuter, R.; Wagner, P.; Willkomm, R. Fluorescent matter in the early Atlantic-ocean. 1. Method of measurement and near-surface distribution. *Deep-Sea Res., Part I* **1994**, *41* (4), 659–675.
(31) Henderson, R. K.; Baker, A.; Murphy, K. R.; Hambly, A.; Stuetz, R. M.; Khan, S. J. Fluorescence as a potential monitoring tool for recycled water systems: A review. *Water Res.* **2009**, *43* (4), 863–881.
(32) Murphy, K. R.; Ruiz, G. M.; Dunsmuir, W. T. M.; Waite, T. D. Optimized parameters for fluorescence-based verification of ballast water exchange by ships. *Environ. Sci. Technol.* **2006**, *40* (7), 2357–2362.
(33) DeRose, P. C.; Wang, L.; Gaigalas, A. K.; Kramer, G. W.; Resch-Genger, U.; Panne, U. Need for and metrological approaches towards standardization of fluorescence measurements from the view of national metrology institutes. In *Standardization and Quality Assurance in Fluorescence Measurements I: Techniques*; Resch-Genger, U., Ed.; Springer-Verlag: Berlin, 2008.
(34) DeRose, P. C.; Early, E. A.; Kramer, G. W. Qualification of a fluorescence spectrometer for measuring true fluorescence spectra. *Rev. Sci. Instrum.* **2007**, *78* (3), .
(35) DeRose, P. C.; Smith, M. V.; Mielenz, K. D.; Blackburn, D. H.; Kramer, G. W. Characterization of standard reference material 2941, uranyl-ion-doped glass, spectral correction standard for fluorescence. *J. Lumin.* **2008**, *128* (2), 257–266.
(36) DeRose, P. C.; Smith, M. V.; Mielenz, K. D.; Blackburn, D. H.; Kramer, G. W. Characterization of Standard Reference Material 2940, Mn-ion-doped glass, spectral correction standard for fluorescence. *J. Lumin.* **2009**, *129* (4), 349–355.

ES102362T