Standardizing Early Oil Spill Detection for Drinking Water with Absorbance-Transmittance Excitation Emission Matrix (A-TEEM) Spectroscopy

吸光・透過補正３次元励起・蛍光マトリクス法（A-TEEM）分光法による飲料水中への流出油，早期検出の標準化について

Adam M. GILMORE

Oil spills into fresh water sources used for drinking water treatment present serious potential damage to the treatment plant infrastructure, the environment and consumer health. While the major fraction of most fuel and oil spill components are insoluble in water, smaller component molecules including Benzene, Toluene, Ethylbenzene and Xylene (BTEX) and other Polycyclic Aromatic Hydrocarbons (PAHs) are both soluble and fluorescent. These dissolved components, which can diffuse rapidly in the water body and are detectable in the µg/L range, can serve as early warning sentinels to prevent spill uptake using HORIBA's patented A-TEEM technology. Importantly, the A-TEEM facilitates spectral identification and linear quantification of these compounds in the presence of mg/L levels of natural and manmade Dissolved Organic Matter (DOM) by virtue of Inner-Filter Effect (IFE) correction of the fluorescence data. The A-TEEM detection limits are significant as exemplified by the carcinogen Benzene which is regulated in finished drinking water at 10 and 5 µg/L, respectively, by the World Health Organization (WHO) and United States Environmental Protection Agency (USEPA). Here we summarize the background and methodology associated with the recently published standard test method, D8431-22[1], with the American Society of Testing Materials (ASTM).

Keywords
BTEX, Extreme Gradient Boosting, Inner-Filter Effect, Naphthalene,
Parallel Factor Analysis, Polycyclic Aromatic Hydrocarbons

飲料水処理に利用される淡水源への油流出は，処理施設や環境，そして消費者の健康に深刻な影響を与える可能性がある。多くの燃料や油流出成分の大部分は水には溶けないが，ベンゼン，トルエン，エチルベンゼン，キシレン(BTEX)や多環芳香族炭化水素(PAHs)の中でも一部の低分子量成分は，水に溶けて素早く広がる。これらの成分は蛍光を発する性質を持っており，HORIBAの特許技術であるA-TEEM技術を用いることで，µg/Lレベルの検出が可能となり，油拡散が起こる前に警告を発し早期に対処することができる。水中に天然や人工の有機物(DOM)がmg/Lレベルで存在していても，A-TEEM測定で試料に吸収された蛍光に内部フィルター効果(IFE)補正を行うことにより，これらの化学物質を高い精度で識別し正確に測定することができる。A-TEEMでの検出限界性能は，発がん性物質であるベンゼンが，世界保健機関およびアメリカの環境保護庁によって，それぞれ10µg/Lおよび5µg/L以下に規制されている例からもわかるように有益である。本稿では、アメリカ材料試験協会(ASTM)が新たに公開した標準試験方法D8431-22[1]に関する背景と分析法についてまとめている。

キーワード
BTEX，勾配ブースティング，内部フィルター効果，ナフタレン，
並列因子分析，多環芳香族炭化水素

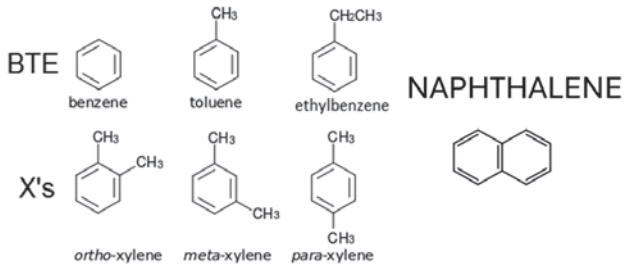

Figure 1 Chemical structures for Benzene, Toluene, Ethylbenzene (BTE), ortho-, meta- and para-Xylene (X's) and Naphthalene.

Introduction

The global demand of fuel oils conflicts directly with safe drinking water treatment with respect to the fact that both shipping and storage of oil can expose drinking water sources to spilled and leaked materials. Oil spills are dangerous to consumer health primarily as sources of carcinogens, including Benzene, that are regulated by the WHO and USEPA in the low µg/L concentration ranges. Oil spills can also significantly damage the environment, including natural flora and fauna, as well as damage the infrastructure of many types of drinking water treatment facilities. Thus, it is of direct benefit to be able to detect oil spills prior to uptake into a treatment facility as a primary means to protect the plant infrastructure and, most importantly, consumer health.

In most cases the majority of the mass of fuel oil spill components are insoluble in water and depending on their density may float or sink in fresh water. However, most fuel oils also contain a significant fraction of water soluble components in the form of BTEX and certain PAHs with relatively low molecular weight[2],[3]. Figure 1 shows the molecular structures for the BTEX and naphthalene compounds. These water-soluble components can diffuse and travel faster than the bulk of the insoluble oil in some streams and water sources. They are also highly fluorescent making them detectable as early warning spill indicators with HORIBA's patented A-TEEM technology [4],[5].

In this article we first describe the basic operation of the A-TEEM method with a special focus on how it facilitates rapid optical identification and quantification of BTEX and other components in the presence of an essentially ubiquitous background of naturally occurring DOM; DOM is usually present in the mg/L range in most fresh drinking water sources and finished water. Key aspects of the standard method development included evaluation of the method Ruggedness and Design of Experiment (DOE) which are discussed with a focus on the major potentially interfering signals from DOM and turbidity

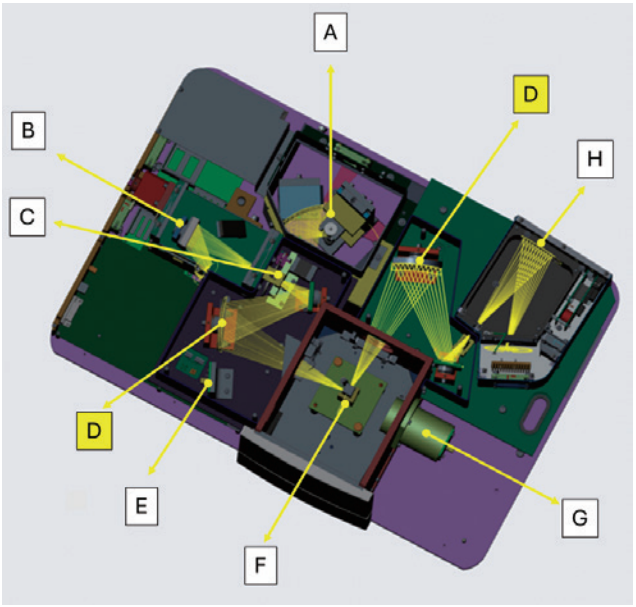

Figure 2 Patented Optical Bench Diagram for the HORIBA Aqualog® A-TEEM Spectrometer. (A). 150 W vertically mounted ozone-bearing exciting light source (200-1000 nm), (B). Double-subtractive monochromator (fixed 5 nm bandpass) with high-stray light rejecting holographic deep UV (250 nm) blazed gratings, (C). Order-sorted excitation optical path for absorbance and fluorescence excitation, (D). Geometrically matched all reflective excitation and emission sampling optics to eliminate color-dependent (chromatic) effects and ensure optical focus on the sample at all wavelengths, (E). A spectrally corrected reference diode detector (200-1000 nm) to account for changes in the light source intensity as a function of time and wavelength for signal stability, (F). A sample compartment with temperature, dry-gas, stirring and flow-cell compatibility, (G). A diode-based transmission detector capable of measuring from 200-1000 nm, and (H). A thermoelectrically cooled, aberration-corrected CCD-spectrograph (250-800 nm) with adjustable binning and gain to maximize the signal to noise with integration times ranging from 5 ms to 65 s.

(suspended particulates). The analytical methods including machine learning algorithms, and how these calibrations and validations are documented, are also discussed. The article concludes with a discussion of how this method can be applied with improved sensitivity to contemporary and future applications as well as a brief comparison to conventional, time consuming and expensive chromatographic methods.

Basic Theory and Operation of A-TEEM Spectroscopy

The patented A-TEEM technology[6], is exemplified in Figure 2 by the HORIBA Aqualog® (see caption for details). The optical bench consists of a powerful broadband (UV to NIR) white light source that is monochromated by a subtractive double monochromator followed by an automated order sorting filter wheel. The first order light is monitored immediately before the sample with a reference detector and then used for both the absorbance and fluorescence excitation source. The sample compartment uses all-reflective optics to geometrically and kinetically match the absorbance and the fluorescence excitation and emission paths. Fluorescence emission is measured

with a thermoelectrically cooled CCD-spectrograph. Operation involves scanning the absorbance and excitation wavelengths from red to ultraviolet (to minimize exposure to ionizing UV); at each abs/ex wavelength a complete emission spectrum is collected until a complete Excitation-Emission Matrix and matching absorbance spectrum is obtained. The signal processing is explained in detail elsewhere[6] noting the most significant feature of the A-TEEM analytical capacity is the coordinated correction of the primary and secondary inner-filter effects which would otherwise distort the EEM data with respect to Beer-Lambert linearity[6]-[9]. IFE correction is critical for the oil-spill detection method because the overlapping absorbance of the background DOM components, with concentrations that are both variable and usually much higher than BTEX, would lead to distorted, nonlinear estimates.

A-TEEM Signals for Naturally Occurring DOM and BTEX Components

As mentioned above naturally occurring DOM is normally present in all fresh drinking water sources and even in most finished drinking water since complete removal is not normally achieved with conventional treatments. The natural molecular composition of fluorescent DOM includes three major classes of compounds including humic acid-like, fulvic acid-like and aromatic amino acid like compounds. Combined these normally add up to at least 1 mg/L of total dissolved organic carbon but can vary widely to >10 mg/L or more in certain sources depending on weather and other conditions[6],[7]. Importantly, all three DOM component classes absorb light (excite) in the UV range from 240 nm. Their PARAFAC loadings also each exhibit distinct emission peak wavelengths with the humic acid like being the most redshifted (peak >450 nm) followed by the fulvic acid-like (peak around 415 nm) then aromatic amino acid-like (peak around 330-350 nm), itself being deepest DOM component in the UV range. Conventional wisdom explains that the relative red-shift for the higher molecular weight fluorophores is structurally associated with the higher extent of aromaticity (ring conjugation) with the single ring aromatic amino acid fluorophores showing the deepest UV emission for natural DOM.

The spectral signatures of the natural DOM A-TEEM signals described above contrast significantly with those from the BTEX and Naphthalene components as shown in Figure 3. Figure 3 compares raw surface water A-TEEM contour plots before (A: Control) and after (B: Spiked) spiking with 100 µg/L each of naphthalene and total BTEX. Naphthalene, which exhibits a very high fluorescence yield is most prominent with the excitation peaking around 275 nm and emitting around 325 nm. While this overlaps with the aromatic amino acid contours the absorbance (not shown) and emission contours are significantly different. BTEX in Panel 3B excites at a lower wavelength peak <270 nm and emits deeper in the UV peaking around 285 nm. Overall, it is clear both of these additions yield well distinguished spectral contours at the 100 µg/L concentration range. This is further exemplified in Figure 4 (adapted from reference[4]) which shows the PARAFAC loadings for raw fresh water samples from the same spiking experiments as Figure 3. The model yields four distinct components for the selected excitation-emission wavelength range; noting the humic acid like region is excluded (masked) for clarity. The 4 components in Figure 4 represent the A (Naphthalene), B (BTEX), C (Fulvic acid-like) and D (Aromatic amino acid-like) components. The component numbers (top of each panel) were assigned based on the score contribution(s) of each component to the overall model; the split-half validation matching for the model was 94.8%. Clearly, the most important distinguishing factor centers on the peak emission wavelengths for the BTEX compounds being considerably below 300 nm (peak for BTEX mixture spike is around 285 nm). This serves well to facilitate BTEX resolution since their emission is below that of any of the naturally occurring DOM components including the amino acid-like components. It is however important to consider that for most drinking water sources and finished water samples the amino acid-like signal intensity is the lowest

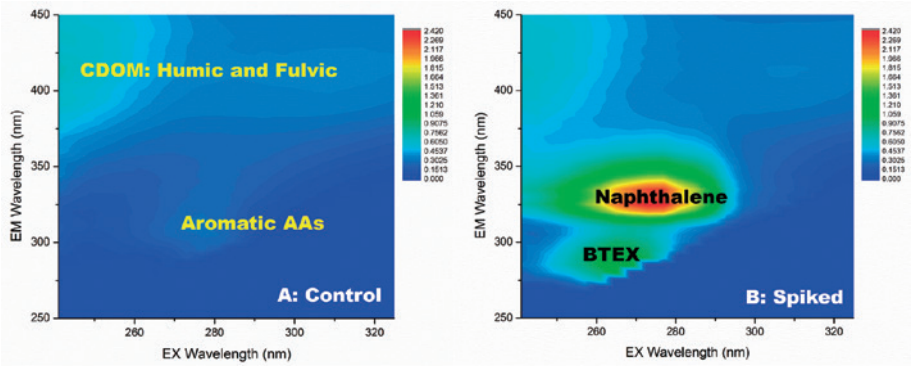

Figure 3 A-TEEM contour plots for a raw surface water control sample (A) with 1.86 mg/L dissolved organic carbon and a matching sample (B) spiked with 100 µg/L of Naphthalene and 100 µg/L of BTEX (25 µg/L of each compound). Signal contour areas associated with humic acid-like, fulvic acid-like and aromatic amino acid-like DOM components are labeled in Panel A. The signal contour areas associated with Naphthalene and BTEX are labeled in Panel B.

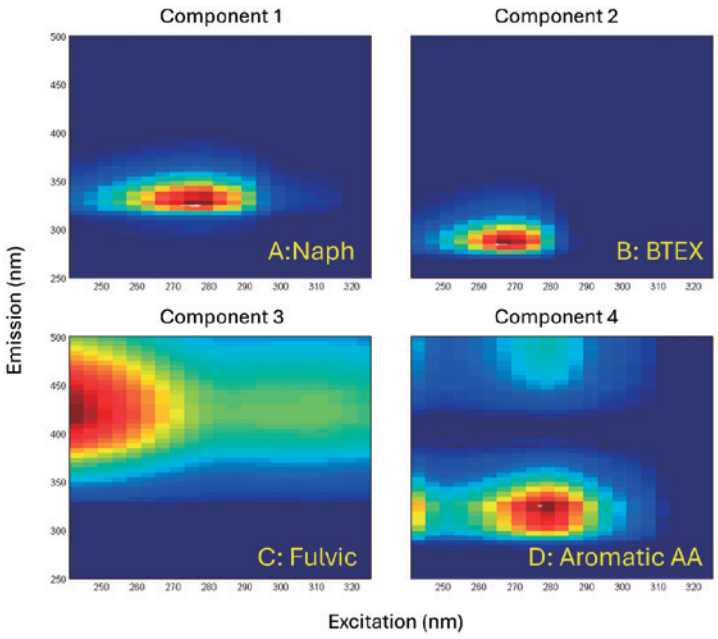

Figure 4 PARAFAC component loading plots for a 4 component model of the Raw water sample shown in Figure 3 spiked with varying levels from 0-100 µg/L of Naphthalene and 0-100 µg/L of BTEX. The component numbers 1 and 2 correspond to the assignments made based on linear regression correlations of the component scores for Naphthalene and BTEX (see ref.[4]); components 3 and 4 for fulvic acid-like and aromatic amino acid-like compounds were assigned based on earlier models from measured samples of Raw and Finished water from the same treatment plant.

of the three major DOM components. It however is more important to consider the BTEX absorbance/excitation and emission spectra overlap strongly with the UV absorbance for all three DOM components dictating that IFE correction is imperative for accurate BTEX and naphthalene quantification.

Experimental Design, Calibration and Validation for ASTM D8431-22

Having established the basis for BTEX and naphthalene detection above here we explain how this information was used to develop and publish ASTM D8431-22 for the purpose of early warning detection of these oil spill components. The method development began with a Design of Experiment (DOE) protocol for a Ruggedness test (see ASTM Practice E1169[10]) where all essential variables were varied to gauge their influence on the method calibration. Key variables tested included filtration (0.45 µm), turbidity from 0 to 20 Nephelometric Turbidity units, DOM concentration (0-15 mg/L), temperature and stirring. The use of 0.45 µm filters both mitigated the effects of turbidity and was needed to obtain the 'dissolved organic fraction' of which the water soluble BTEX and naphthalene are key constituents. It was determined the method was stable up to 15 mg/L DOM whereas beyond this nonlinear IFE effects due to Beer-Lambert linear deviation were observed; dilution could be applied to compensate for DOM concentrations >15 mg/L. Temperature was determined to be best stabilized at 20°C and stirring was insignificant for the dissolved components. Importantly, there was no significant evidence of loss by volatilization for the BTEX or naphthalene components using the method noting best performance includes rapid sampling and analysis and use of Teflon stoppered 4 ml quartz fluorescence cuvettes.

The Ruggedness test set the stage for the final method conditions for the sample preparation and instrument settings, see D8431-22[1] for details, noting accessing this controlled document requires a fee to ASTM. The basic outline of the method includes first standardizing the instrument response as a function of the integration time and CCD camera settings including the gain and binning. This is accomplished using a sealed Type I water Raman Scattering Unit (RSU) sample. The acquired RSU value acts as an external reference standard to account for changes in sample concentration independent of the aforementioned instrument settings. Samples are simply filtered (0.45 µm) through a nylon or glass fiber filter to avoid extractable compounds that may absorb in the UV range as an interference. The scan range was adjusted to include the BTEXN components with the excitation range from 240-325 nm with a 4 nm increment and the emission range from 250-800 nm with a 5 nm binning/interpolation. The default integration time was 1 s and Medium Gain with the sample compartment at 20°C. A blank file was acquired using these conditions with a sample of Type I water and a clean cuvette. The method cites ASTM guidelines for cuvette and glassware cleaning instructions for water-borne oil samples. Subsequent samples were evaluated against the blank sample and the time-date-stamped data files exported for multivariate model calibration and validation/application.

Table 1 Limits of Detection and Quantification* (low and high range) for *BTEXN* based on the Single Lab Study.

Analyte	Low Range	Detection Limit (µg/L)	Quantification Limit (µg/L)	High Range	Detection Limit (µg/L)	Quantification Limit (µg/L)
Benzene		2.52	7.62		6.45	19.55
Toluene		0.45	1.35		5.69	17.23
Ethylbenzene		0.45	1.35		5.69	17.23
Xylenes		0.45	1.35		5.69	17.23
Naphthalene		0.7	2.12		3.98	12.05

*Based on ASTM Practice E2617[11], the Limits of Detection and Quantification were calculated using the respective formulae: *LOD* = 3.3 *s*/*S* and *LOQ* = 10 *s*/*S*, where s is the standard deviation of the Y-intercept and S is the slope of the linear relationship between the regression model predicted and the target concentration values.

Calibration of the Extreme Gradient Boost (XGB) machine learning algorithms for regression (quantification) and discrimination (qualification model for Pass/Fail contaminant threshold testing) were made using the Eigenvector Inc. Solo software. Implementation of the regression and discrimination models used the HORIBA Multi-Model Predictor (HMMP) tool which is a commercial add-in application for Eigenvector Inc.'s Solo and Partial Least Squares toolboxes; the HMMP tool facilitates concurrent batch analysis of samples with either multiple regression models (possibly including multiple algorithms) simultaneously or a single discrimination model including multiple class assignments.

The machine learning algorithms for the method were calibrated and validated according to the ASTM guidelines defined in (E2617[11] and E2691[12]) which respectively deal with multivariate model calibration for empirical models and specifics for pharmaceutical and manufacturing applications. A specific experimental calibration design was applied for the standard preparations of the BTEX and naphthalene compounds to avoid to minimize issues with collinearity effects of the standards. Calibration curves were measured to determine the repeatability, recovery and the limits of detection and quantification for each individual compound and the sum of BTEXN. It was found that use of both high- (>50 µg/L) and low-range (<50 µg/L) calibrations for BTEX and (>20 and <20 µg/L) for naphthalene was useful to insure accurate detection across the wide working range of the method; noting naphthalene exhibited a significantly higher fluorescence yield per weight than any of the BTEX or natural DOM components. The single-lab calibration published in D8431-22 was all measured by spiking the raw source water before filtration to yield detection and quantification limits for the BTEXN compounds.

For the single lab calibration, the coefficient of determination (R^2) for the predicted values was >0.998 for all compound models in both the low and high range calibrations with the exception of Benzene which was 0.988 for the low range calibration. The recovery (%) values for all compounds in both the high and low range calibrations ranged from 99.4 % to 102.7 % except for Benzene in the low range calibration which yielded 21 % at 5 µg/L, which was the lowest concentration tested. Table 1 summarizes the limits of detection for all compounds including the sum of BTEXN for both the low and high range calibrations. Benzene yielded an LOD of 2.52 µg/L in the low range and 6.45 µg/L in the high range calibration. Benzene also yielded the highest detection and quantification limits for both the high and low range calibrations. This indicates Benzene, compared to the other compounds, has the lowest fluorescence yield and or it is associated with higher relative levels of background interference. This subject will be addressed in the scheduled full interlaboratory study. Possible improvements could include increasing the integration time to maximize signals in the Benzene region when signals fall below a threshold level; this would require adjusting the RSU factor to account for the concentration.

With respect to current WHO guidelines and USEPA maximum contamination limits (MCL) Table 2 lists the values for each BTEX compound. The LOD levels for all compounds in the low range calibration (Table 1) fall below both the guidelines and MCL values listed noting for Benzene the LOQ was above the USEPA MCL. It is important to note the WHO and USEPA MCL values represent the finished (distributed to customer) water values. This is especially relevant to the scope of D8431-22 which pertains only to raw untreated source water. Conventional water treatment processes may include treatment with activated carbon, polymers, coagulants for hydrophobic materials and filtration all of which would be expected to reduce soluble BTEX compound levels significantly via adsorption; noting volatilization would be another expected source of loss for BTEX during treatment and distribution.

Table2 Comparison of the WHO Guidelines and USEPA Maximum Contamination Limit (MCL) for the BTEX compounds in finished distributed to consumer) drinking water.

Compound (µg/L)	WHO Guidelines	USEPA MCL
Benzene	10	5
Toluene	700	1000
Ethylbenzene	300	700
Xylenes	500	10000

Conclusion

This article introduces and describes a new A-TEEM standard method for rapidly detecting water soluble oil spill components, primarily BTEX, in raw source water for early warning purposes to prevent uptake of potentially harmful materials by a drinking water treatment plant. The D8431-22 method was recently published by the ASTM D19.06 organics subcommittee on water quality based on a single-lab calibration study and is pending a full interlaboratory study due in 2027. The method is advantageous being rapid (<5 min per sample) and requiring minimal sample preparation (only filtration). The A-TEEM method is comparable in sensitivity requirements to the conventional Solid Phase Microextraction Gas-Chromatography Flame Ionization detection which can require up to 25 min run time among other time-consuming preparative steps. The D8431-22 method thus shows promise for rapid, accurate detection of BTEX in a wide variety of freshwater sources further noting it can be conceptually adapted to other types of water production as shown by Madhav and Gilmore[13] for hydroelectric dams and fuel oil spills during oil-assisted generator startup procedures.

* Editorial note: This content is based on HORIBA's investigation at the year of issue unless otherwise stated.

References

[1] *ASTM D8431-22.* Standard Test Method for Detection of Water-soluble Petroleum
Oils by A-TEEM *Optical Spectroscopy and Multivariate Analysis.*

[2] Bobra, M., "Solubility behaviour of petroleum oils in water," *Environment Canada Report,* (1992).

[3] Guard, H.E., NG, J. and Laughlin, R.B., "Characterization of gasolines, diesel fuels and their water soluble fractions," *Naval Biosciences Laboratory Report,* (1993).

[4] Adam M. Gilmore, Linxi Chen, "Water Soluble Fraction (WSF) contaminant detection using machine-learning Absorbance-Transmission Excitation Emission Matrix (A-TEEM) spectroscopy," *Proc. SPIE 11233,* Optical Fibers and Sensors for Medical Diagnostics and Treatment Applications XX, 112330H (20 February 2020); doi: 10.1117/12.2556434.

[5] Adam M. Gilmore, Linxi Chen, "Optical early warning detection of aromatic hydrocarbons in drinking water sources with absorbance, transmission and fluorescence excitation-emission mapping (A-TEEM) instrument technology," *Proc. SPIE 10983,* Next-Generation Spectroscopic Technologies XII, 109830E (13 May 2019); doi: 10.1117/12.2522498.

[6] Gilmore, A., Tong, M. System and Method for Fluorescence and Absorbance Analysis, *U.S. Patent 10,168,310 B2, 2019,* Issued January 1, 2019.

[7] Gilmore, A. Determination of water treatment parameters based on absorbance and fluorescence. *US 11874226B2.* Issued January 16, 2024.

[8] Gilmore, A., Tomioka, Y. Determination of water treatment parameters based on absorbance and fluorescence. *US 10184892.* Issued January 22, 2019.

[9] Gilmore AM. How to collect National institute of standards and Technology (NIST) traceable fluorescence excitation and emission spectra. *Methods in Molecular Biology.* 2014;1076:3-27.

[10] *ASTM E1169-21.* Standard Practice for Conducting Ruggedness Tests.

[11] *ASTM E2617–17R24.* Standard Practice for Validation of Empirically Derived Multivariate Calibrations.

[12] *ASTM E2891–20.* Standard Guide for Multivariate Data Analysis in Pharmaceutical Development and Manufacturing Applications.

[13] Madhav, H., Gilmore, A.M. (2024) Identification of oil contamination in process water using fluorescence excitation emission matrix (FEEM) and parallel factor analysis (PARAFAC). *Water Sci Technol* (2024) 90 (3): 908–919.

Adam M. GILMORE

Fluorescence Product Manager
Fluorescence Division
HORIBA Instruments Incorporated,
Ph.D.