Label -free Approach to Simultaneous Determination of Three Endocrine Disruptors in River Water: Angular Sweep Total Fluorescence Spectroscopy Coupled to Multi-way Calibration

Anupama Vijayan[a] and John Prakash[a]*

[a] Department of Chemistry, School of Basic and Applied Sciences, Central University of Tamil Nadu, Thiruvarur, 610 005, India

* Corresponding Author.
Tel: +91 9497592682
Address: Department of Chemistry, School of Basic and Applied Sciences, Central University of Tamil Nadu, Neelakudy, Thiruvarur, 610 005, India
E-mail address: johnprakash@cutn.ac.in (John Prakash)

Electronic Supplementary Information

1. Structural Analysis of ASTFS Data: Model Selection

The data structure of the ASTFS data matrix (angle × emission) obtained for each sample does not conform to a low-rank bilinear (after unfolding) or trilinear model; it must be treated using multi-way regression strategies such as unfolded partial least squares (U-PLS) or N-PLS. Departures from low-rank behaviour typically arise from pronounced spectral overlap, inter-analyte interactions, or non-linear signal contributions that distort the multilinear structure of fluorescence data. In complex multi-component systems such as the CBZ–BPA–CBR mixture, these effects manifest as highly correlated spectral responses that preclude accurate modelling by conventional decomposition techniques. In such instances, careful examination of the low-rank subspace is essential to select a suitable regression model that can adequately represent the underlying analytical variance. Therefore, rank estimation was performed using Principal Component Analysis (PCA) and Singular Value Decomposition (SVD), allowing the

identification of the effective dimensionality, and ensuring the chosen regression model aligns with the true structure of the dataset.

PCA of augmented data matrix across both modes (emission and angle) **lacked a discernible elbow or clear, consistent stabilization plateau** across the instrumental modes. This confirms the data's severe **rank-ambiguity. The corresponding explained variance and residual plot are given in figure S1(A) and (B).**

Using SVD on the mode-1, mode-2, and mode-3 unfoldings, we observed a dominant first singular value (~75% variance explained) followed by a sharp drop to near-constant minor contributions (~1 or 2%) from higher-order components. A clear elbow at Rank 2, shows that only two principal components account for majority of the variance. Variance decay profiles for all three modes were nearly identical, the scree plot(Explained variance plot) for all three modes is given in figure S1(C). This result, when contrasted with the known chemical rank is a case of **rank-deficiency** and when contrasted with rank estimation via PCA, is a **rank inconsistency.** This behavior is commonly observed in systems with severe spectral overlap, such as in the present study, where the proximity of spectral features induces significant spectral collinearity.

Given this non-ideal structure, the more robust U-PLS (by eliminating the rank constraint via unfolding the data into two-way matrix) and N-PLS (by employing a more flexible tensor regression) were adopted. The final 4-Latent Variable model was chosen based on the system's chemical reality, a decision permitted by the flexibility of these PLS models, ensuring accurate quantification. . This approach potentially aided by RBL provides a reliable framework for modeling complex mixtures and overlapping spectral features beyond the capabilities of purely trilinear models.

Additionally note that, PCA and SVD differ fundamentally in their underlying assumptions and data preprocessing steps; for example, PCA involves centering the data and analyzing the covariance matrix to identify directions of maximum variance, whereas SVD performs a direct factorization of the raw data matrix without necessarily centering. Despite these differences, the variance explained and residual analyses within each respective method show consistent behavior across multiple data modes, reflects a stable linear correlation structure predominantly in two modes, supporting bilinearity.

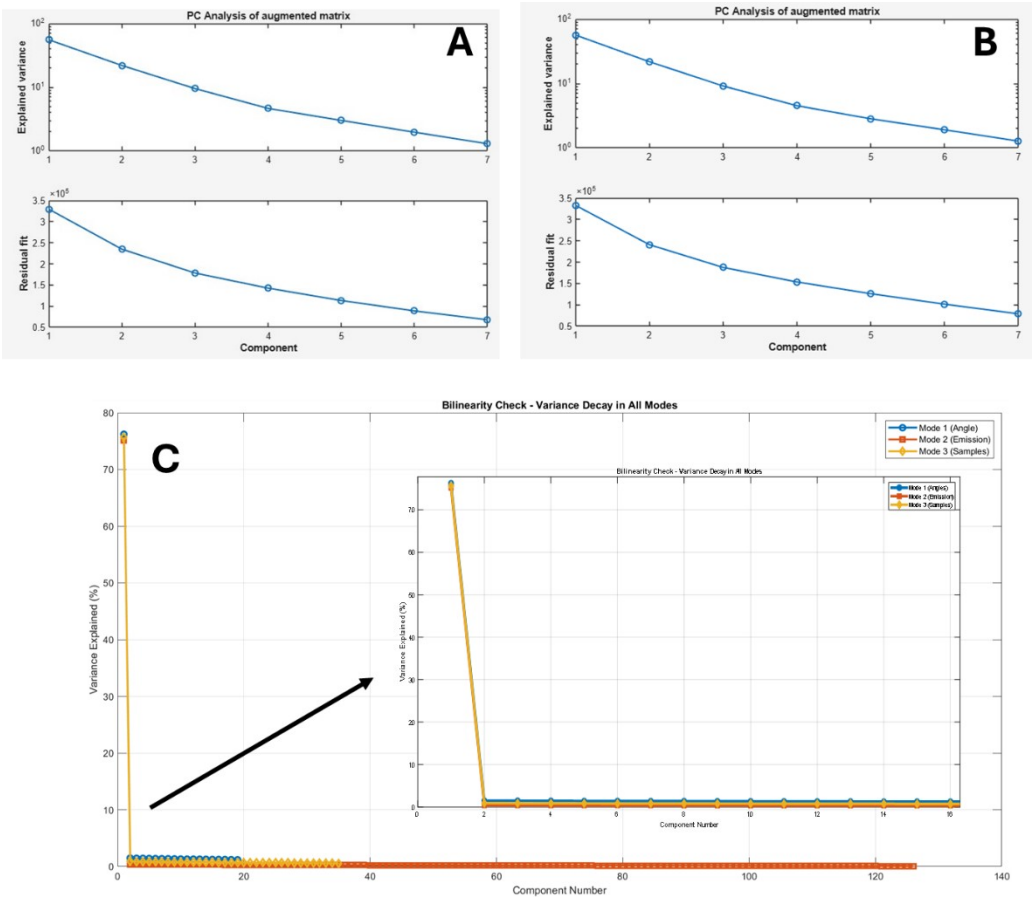

Figure S1: Principal Component Analysis (PCA) results showing explained variance versus number of components and residual fit versus number of components for (A) mode 1 (Emission) and (B) mode 2 (Angle), respectively. (C) SVD scree plot illustrating the variance distribution across components for all modes (Inset image: rescaled on x-axis)

2.C1 calibration sample data and validation set prediction results

Table S1: Concentrations of CBZ, BPA and CBR in calibration set C1

Sample	CBZ (µM)	BPA (µM)	CBR (µM)
1	0.5	0	0
2	0	1	0
3	0	0	0.25
4	1.5	1.4	0
5	3.5	0	0.4
6	0	0.8	0.1
7	5	1.2	0.05
8	3.5	1.6	0.1
9	4.5	2	0.2
10	3	1	0.25
11	2	0.8	0.3
12	1.5	0.6	0.35
13	0.5	0.2	0.5
14	0	0	0.05
15	0	0.2	0.1
16	0	0	0.2
17	1	0.4	0.3
18	2	0.8	0.5
19	2.5	0	0
20	2.5	2	0.3
21	3.5	1	0.25
22	4	1.4	0.35
23	5	1.8	0.45
24	1.5	1.4	0

Table S2: Actual concentrations versus N-PLS and unfolded-PLS predicted amounts of analytes in the samples of validation set V1

Sample	Analyte	Actual amount (µM)	N-PLS	U-PLS
1	CBZ	4.00	4.03	4.12
	BPA	1.80	1.32	1.33
	CBR	0.15	0.15	0.17
2	CBZ	1.50	1.41	1.15
	BPA	0.40	0.42	0.42
	CBR	0.45	0.44	0.45

3	CBZ	2.00	2.08	2.00
	BPA	0.60	0.95	0.95
	CBR	0.40	0.42	0.37
4	CBZ	1.00	0.74	0.97
	BPA	0.20	0.72	0.72
	CBR	0.05	0.06	0.02
5	CBZ	1.50	1.15	1.27
	BPA	0.40	0.75	0.75
	CBR	0.10	0.09	0.09
6	CBZ	2.00	1.99	2.07
	BPA	0.60	1.01	1.01
	CBR	0.15	0.11	0.14
7	CBZ	4.00	3.90	3.78
	BPA	1.60	1.45	1.45
	CBR	0.40	0.40	0.41

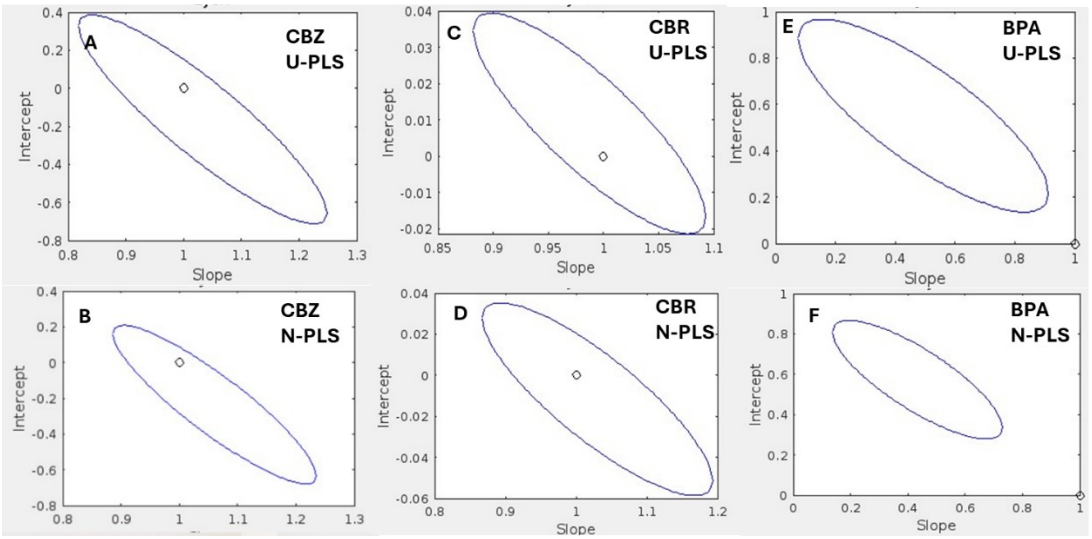

Figure S2: EJCR plots comparing U-PLS and N-PLS models for three analytes in aqueous medium: (A, B) CBZ, (C, D) CBR, and (E, F) BPA

3. Theoretical determination of molecular count per micelle using Poisson's ratio

Molecule to micelle ratio is calculated using following equation:

$$x = \frac{Analyte\ concentration}{Micelle\ concentration} \quad \ldots\ldots\ldots\ldots\ldots (s1)$$

Where,

$$Micelle\ concentration = \frac{Concentration\ of\ surfantant - CMC}{Aggregation\ number} \quad \ldots\ldots\ldots\ldots\ldots (s2)$$

Therefore,

$$x = Analyte\ concentration \times \left[\frac{Aggregation\ number}{Concentration\ of\ surfantant - CMC} \right] \quad \ldots\ldots\ldots\ldots\ldots (s3)$$

Table S3: P(1) and P(2) values calculated for CTAC concentrations ranging from 2 mM to 50 mM at a fixed analyte concentration of 7 μM.

CTAC Concentration (mM)	Micelle Concentration (mM)	Analyte to Micelle Ratio (x)	Probability P(1) (%)	Probability P(2) (%)
2	0.008	0.875	36.475	5.621
3	0.017	0.412	27.287	1.637
5	0.035	0.200	16.375	0.355
10	0.080	0.088	8.059	0.153
15	0.123	0.057	5.384	0.085
20	0.167	0.042	4.027	0.053
25	0.212	0.033	3.192	0.035
30	0.256	0.027	2.628	0.020
40	0.344	0.020	1.960	0.013
50	0.433	0.016	1.575	0.000

4. Micellar-assistance for enhanced fluorescence visibility- fluorescence spectra of individual analytes

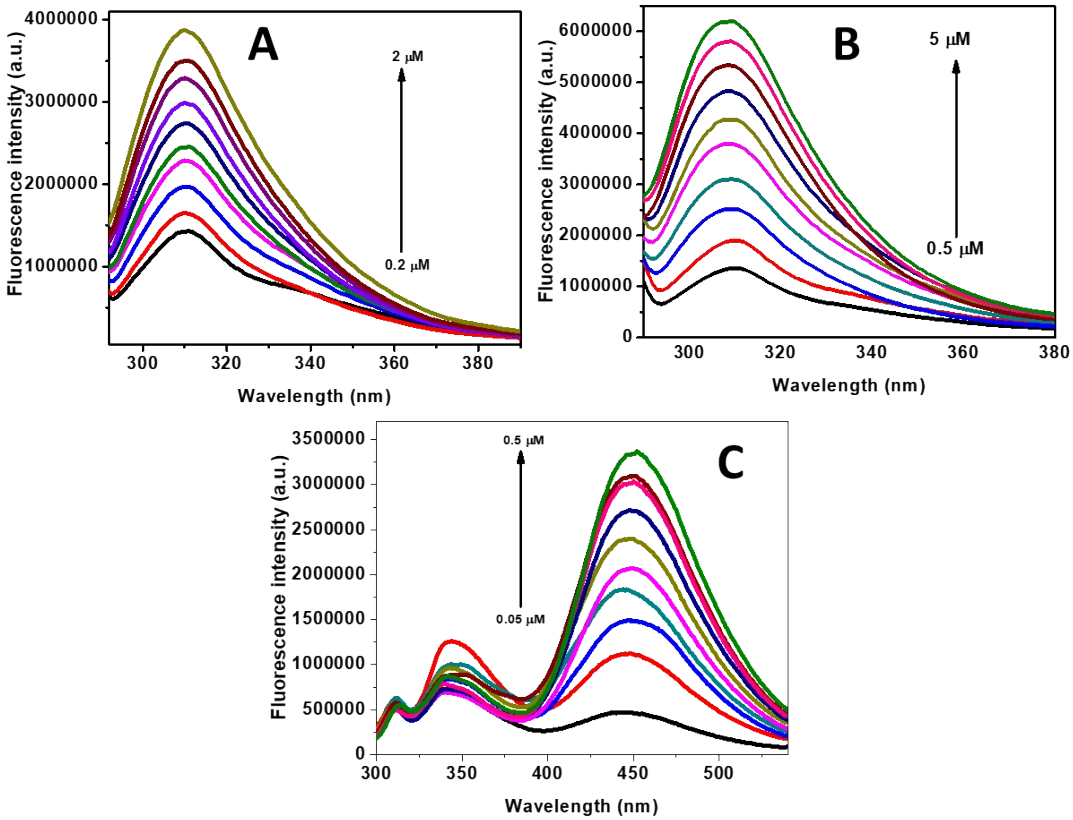

Figure S3: Fluorescence spectra of (A) BPA (0.2-2 µM) (B) CBZ (0.5-5 µM), and (C) CBR (0.05 -0.5 µM) in 15 mM CTAC, excitation at 280 nm. (incubation time: 12 hours)

5. C2 Calibration sample data

Table S4: Concentrations of CBZ, BPA and CBR in calibration set C2

Sample	CBZ (µM)	BPA (µM)	CBR (µM)
1	2	0	0
2	0	1	0
3	0	0	0.25
4	3.5	0	0.4
5	5	1.2	0.15
6	4.5	1.6	0.2
7	4	1.8	0.25
8	3.5	2	0.3
9	3	1	0.35
10	2	0.8	0.4
11	1.5	0.6	0.45
12	0.5	0.2	0.05
13	5	1	0.1
14	4	0.6	0.15
15	3.5	0.2	0.25

16	2	1.8	0.45
17	1.5	1.6	0.4
18	0.5	1.2	0.3
19	0	1.2	0.25
20	0	0.4	0.1
21	1	0.6	0.2
22	4	0	0
23	3.5	0.8	0
24	1.5	1.2	0.35
25	0	0	0.5
26	2.5	1.6	0.15

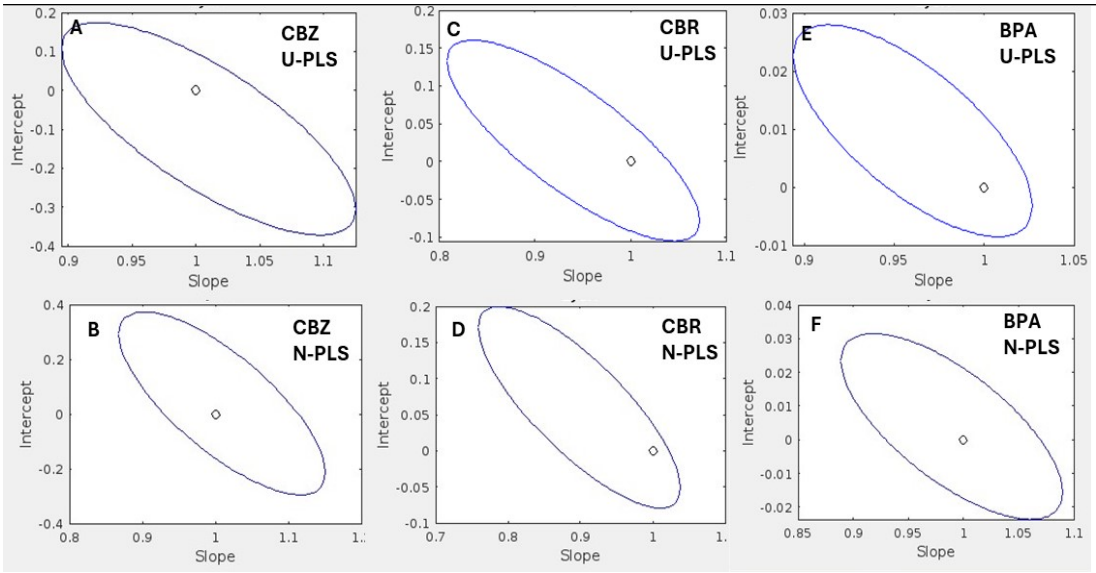

Figure S4: EJCR plots comparing U-PLS and N-PLS models for three analytes in CTAC medium: (A, B) CBZ, (C, D) CBR, and (E, F) BPA

6. River Sample Analysis

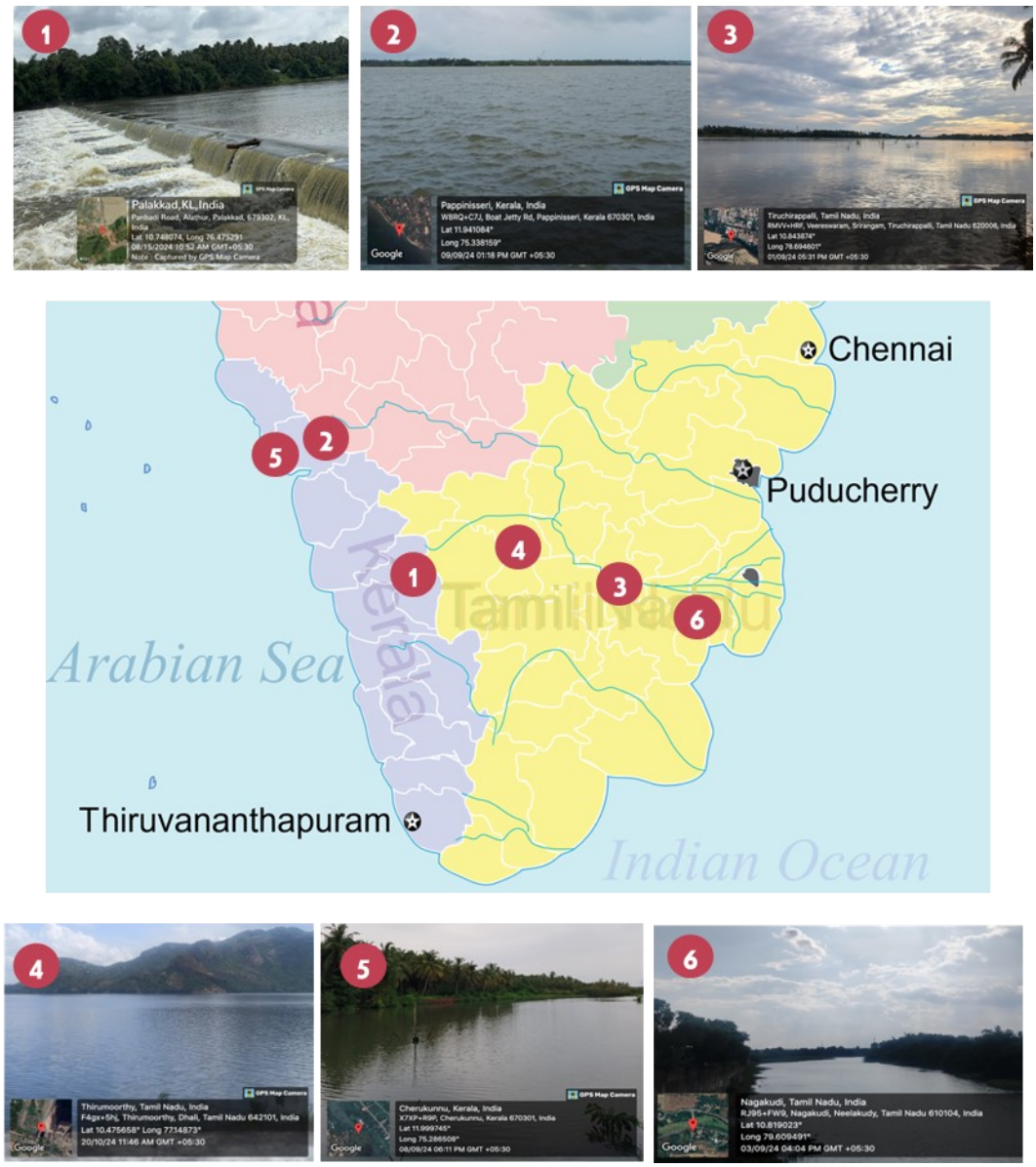

Figure S5: Depiction of the Geo-tagged locations of real sample collection on a map of South India

Table S5: Locations of sample collection

Sample	Location
1	Alathur, Palakkad, Kerala
2	Pappinisseri , Kannur, Kerala
3	Srirangam, Thiruchirapalli, Tamil Nadu
4	Thirumurthy, Udumalpet, Tamil Nadu
5	Cherukkunnu, Kannur, Kerala
6	Nagakudi, Thiruvarur, Tamil Nadu

Table S6: Actual concentrations (µM), predicted values from N-PLS and U-PLS models, and corresponding recovery percentages for CBZ, CBR and BPA

			CBZ		
Sample Number	**Actual (µM)**	**NPLS Predicted (µM)**	**Recovery (%)**	**UPLS Predicted (µM)**	**Recovery (%)**
1	3.5	3.7	107.2	2.5	72.5
2	2	2.6	128.4	1.9	94.1
3	3.5	3.7	107.1	3.5	99
4	1.5	1.3	83.4	1.5	104.8
5	2	2	98.5	2.1	103
6	1.5	1.3	87.1	2.0	134.9

			CBR		
Sample Number	**Actual (µM)**	**NPLS**	**Recovery (%)**	**UPLS**	**Recovery (%)**
1	0.4	0.41	103.2	0.39	99.4
2	0.3	0.27	90.6	0.27	91.1
3	0.15	0.14	94.8	0.15	103.5
4	0.15	0.16	107.4	0.17	112.7
5	0.3	0.31	105	0.36	120.4
6	0.4	0.39	97.2	0.39	98.7

			BPA		
Sample Number	**Actual (µM)**	**NPLS**	**Recovery (%)**	**UPLS**	**Recovery (%)**
1	1.8	1.8	100.2	2.2	120.97
2	0.8	0.7	83.2	0.8	101.1
3	0.7	0.8	113.1	0.5	72.3
4	2.4	2.5	103.3	2.5	104.9
5	2	1.9	94.1	2.1	103.3
6	1	0.8	81.2	0.8	82.2

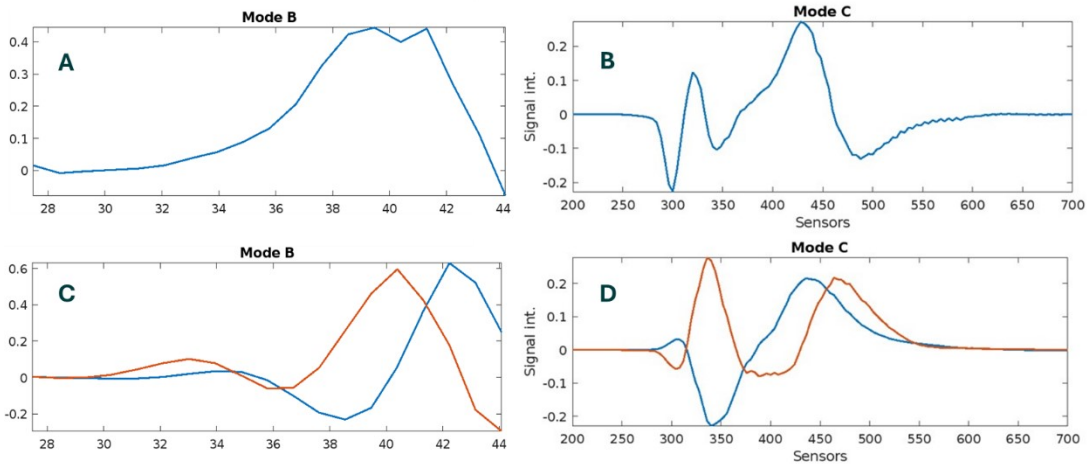

Figure S6: RBL profiles obtained from N-PLS/RBL and U-PLS/RBL analyses of BPA in real samples: (A) angle mode and (B) emission mode profiles from N-PLS/RBL; (C) angle mode and (D) emission mode profiles from U-PLS/RBL.

Table S7: Average Residuals before and after application of RBL for real samples

Model	Before	After
N-PLS	2.91 e5	2.2 e5
U-PLS	1.78 e5	5.92 e4

7. Effect of Ionic strength on fluorescence of BPA in CTAC medium

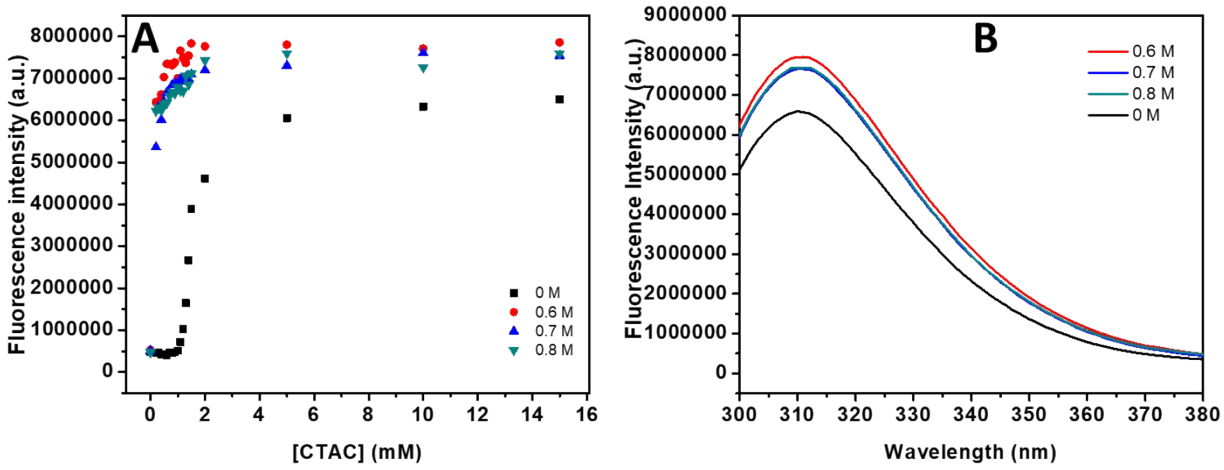

Figure S7: (A) Fluorescence intensity of 2 µM BPA in 0-15 mM CTAC medium at ionic strength = 0, 0.6, 0.7 and 0.8 M (B) Fluorescence spectra of 2 µM BPA in 15 mM CTAC at ionic strength = 0, 0.6, 0.7 and 0.8 M