Supporting Information

Monitoring the Behavior of Emerging Contaminants in Wastewater-Impacted Rivers Based on the Use of Fluorescence Excitation Emission Matrixes (EEM)

Massimiliano Sgroi[1], Paolo Roccaro[1], Gregory V. Korshin[2], Federico G.A. Vagliasindi[1,]*

[1] Department of Civil Engineering and Architecture, University of Catania, Viale A. Doria 6, 95125, Catania, Italy

[2] Department Civil and Environmental Engineering, University of Washington, Box 352700, Seattle, WA 98195-2700, USA

[*]Corresponding author. Phone: +39 095 738 2704; fax: +39 095 738 2748, e-mail address: fvaglias@dica.unict.it

Text (4)

Table (18)

Figure (14)

Text S1. Materials. All standards and reagents used during the study were of the highest purity commercially available (≥97% for all compounds). All native standards were procured from Sigma-Aldrich (St. Louis, MO). Labeled standards carbamazepine-d_{10} and N,N-diethyl-meta-toluamide-d_6 (DEET-d_6) were purchased from Cambridge Isotope Laboratories (Andover, MA), $^{13}C_3$-caffeine, $^{13}C_6$-diclofenac, ibuprofen-d_3, $^{13}C_6$-sulfamethoxazole were obtained by Sigma-Aldrich (St. Louis, MO), whereas trimethoprim-d_9 and sucralose-d_3 were purchased from Toronto Research Chemicals (Ontario, Canada).

HPLC grade formic acid and ammonium acetate were obtained from Sigma-Aldrich (St. Louis, MO). All solvents were of the highest purity available and suitable for LC-MS analysis. Methanol (LC/MS grade and HPLC grade), methyl tertiary butyl ether (MTBE) (HPLC grade), formic acid (LC/MS grade) and water (LC/MS grade) were obtained from Sigma-Aldrich (St. Louis, MO).

Text S2. Investigated rivers. The Simeto River with a length of 116 km and a basin catchment extension of 4029 km^2 is one of the main rivers in Sicily. Its main tributaries listed in the source to the mouth sequence are the Troina, Salso, Dittaino and Gornalunga Rivers. The Simeto River originates in the Nebrodi Mountains, in the northern part of the catchment basin where the soil geology is characterized by volcanic rocks. After the point of confluence with the Salso River, the Simeto River flows in an alluvial plain which extends to its mouth in the Mediterranean Sea. The river bed consists of rocks and coarse gravel sediments with little in-stream vegetation in the mountain part of the river, whereas the stretch of the river located in the alluvial plain is more densely vegetated and has finer sediments.

The San Leonardo is a much smaller river with a length of 49 km and a basin catchment surface of 461 km^2. The geology of the area exhibits volcanic and calcareous formations characterized by high permeability. The river bed is densely vegetated with a fine sandy and clayey sediment texture.

Both rivers flow through agricultural areas characterized by the prevalence of the cultivation of citruses. Due to the dry weather and the drought conditions typical of Sicily, the flow rates are strongly influenced by precipitation, especially in the case of the San Leonardo River characterized by a much smaller basin catchment area. Average, maximal and minimal flow rates for the last stretch of the rivers are compiled in Table S2.

Text S3. Analytical methods. Emerging organic contaminants (EOCs) concentrations in wastewater samples were determined using solid phase extraction (SPE) followed by HPLC-MS/MS[1]. All samples were spiked with isotopically labeled surrogate standards at concentrations varying from 100 to 200 ng/L depending on analytical sensitivity. Samples were then filtered at 0.7 µm by Whatman glass microfiber filter (Clifton, NJ) and extracted using 200 mg hydrophilic-lipophilic balance (HLB) cartridges (Water Corporation; Millford, MA). Cartridges were first preconditioned with 5 mL of MTBE, followed by 5 mL of methanol and 5 mL of ultrapure water. Samples were then loaded at 10 mL/min onto the cartridges which were subsequently rinsed with ultrapure water and dried under air flow. Extraction volumes were 1000 mL for all the samples. The analytes were then eluted with 5 mL of methanol followed by 5 mL of 10/90 (v/v) methanol/MTBE solution. The extract was evaporated to 1 mL using gentle nitrogen flow. Final extracts were transferred into 2-mL vials and stored in darkness at -20 °C until HPLC-MS/MS analysis.

The chromatographic separation was carried out with an Agilent 1200 HPLC (Santa Clara, California). The Ascentis Express reverse phase C18 column (4.6x50 mm) with a packing size of 2.7 µm was used to separate compounds analyzed in positive and negative electro spray ionization (ESI) mode. The mobile phase for ESI positive used two solvents comprising (A) ultrapure water with 0.1% formic acid and (B) methanol. With a constant flow rate of 0.3 mL/min, solvent B linearly increased from 0% to 100% in 25 min and it was held at this percentage till 35 min. Then, solvent B was lowered back to 0% in 15 min and held at this percentage for further 5 min. This resulted in a total run-time of 55 min. The mobile phase for ESI negative used a dual eluent system comprising (A) 1 mM ammonium acetate in ultrapure water buffered at pH 7.00 by adding ammonia and (B) methanol. The same flow rate, gradient and run-time of the ESI positive analysis was used for the analysis in negative ESI mode. The injection volume was 10 µL.

Mass spectrometry was performed using an Agilent 6410 triple quadrupole mass spectrometer (Santa Clara, California) equipped with a multi-mode ionization source operating in ESI mode. The mass spectrometer was operated in the multiple reaction monitoring (MRM) mode. Nitrogen was used as nebulizing and collision gas. In Table S5 are reported the optimized source dependent parameters of the mass spectrometer. MRM transactions used for quantification and the optimized fragmentor and collision energy voltages are reported in Table S6 for each compound.

Text S4. Fluorescence data acquisition and correction. Fluorescence EEM data were acquired for the range of excitation wavelengths from 220 nm to 450 nm at an interval of 5 nm and emission wavelengths from 250 nm to 580 nm at the interval of 1 nm. Excitation and emission slit widths were both set at 5 nm. Spectroscopic measurements have always been accomplished after 0.7 µm filtration (Whatman glass microfiber filter, Clifton, NJ). The Raman scatter effect was minimized by subtracting EEMs of pure deionized (DI) water from the sample EEMs; any

negative intensity values produced by this subtraction were converted to zero values[2]. Then, the emission intensity data were normalized to the Raman peak area of a emission wavelengths scan of DI water samples collected at the interval of 1 nm and related to an excitation wavelength of 350 nm to produce fluorescence intensities in Raman unit (RU)[2]. All the EEMs were subjected to inner filter effect correction according to the methodology proposed by MacDonald et al.[3]. Non-trilinear data related to the Rayleigh scattering were eliminated and interpolated from either side of the scatter band following the tutorial published by Murphy et al.[4].

Table S1. Typical uses and selected characteristics of EOCs examined in this study.

Compound	Use	Biodegradability	Log K_{ow}	Note	Reference
Ibuprofen	Analgesic	High	3.97		Kunkel and Radke[5]; Buser et al.[6]; Jekel et al.[7]; Trenholm et al.[8]
Caffeine	Stimulant	High	-0.07		Buerge et al.[9]; Jekel et al.[7]; Trenholm et al.[8]
Diclofenac	anti-inflammatory	moderate	4.51	high photosensitivity	Kunkel and Radke[5]; Tixier et al.[10]; Petrie et al.[11]
DEET	Insect repellant	moderate	2.18		Weeks et al.[12]; Trenholm et al.[8]
Trimethoprim	Antibiotic	moderate	0.91		Perez et al.[13]; Trenholm et al.[8]
Sulfamethoxazole	Antibiotic	moderate	0.89		Radke et al.[14]; Grunheid et al.[15]; Perez et al.[13];Benotti and Brownawell[16]; Trenholm et al.[8]
Carbamazepine	Anti-seizure	Low	2.45		Jekel et al.[7]; Clara et al.[17]; Petrie et al.[11]
Sucralose	Sweetener	Low	-0.4		Jekel et al.[7]; Soh et al.[18]; Anumol et al.[19]

Table S2. Average, minimal and maximal annual flow rates in the Simeto and the San Leonardo Rivers. Results reported by two hydrometric stations located in the last stretch of the rivers[20]

River	Hydrometric station	Geographical coordinates		Flow rate (m^3/s)		
		Latitude	*Longitude*	*Average*	*Min*	*Max*
Simeto	Passo Fico	37.41982°	14.99195°	2.91	0.88	7.60
San Leonardo	Villaggio Biviere	37.30357°	14.97766°	0.84	0.05	3.13

Table S3. Maximal temperature (T_{max}), minimal temperature (T_{min}) and monthly precipitation reported by the forecast service "Meteo Sicilia" in September 2013, April 2015 and October 2015 in the city of Catania

Pluviometric Station	Parameter	September 2013	April 2015	October 2015
Catania - San Giovanni Galermo	T_{min} (°C)	19.2	10.5	15.5
	T_{max} (°C)	29.8	20.3	23.7
	Monthly precipitation (mm)	44.6	5.0	338.4

Table S4. Water quality parameters measured along the Simeto River during the sampling campaign of September 2013.

Sampling Point	Conductivity (µS/cm)	pH	DOC (mg/L)	TN (mg/L)
Simeto River - September 2013				
R1	428	8.30	2.6	3.4
C1	722	8.12	2.7	0.9
R4	675	8.34	3.2	5.5
R5	786	8.50	2.8	3.0
R6	1230	8.36	2.7	1.9
R8	1407	8.24	5.1	4.3
R11	1555	8.43	5.2	3.0
C2	2200	8.23	5.9	2.1
R13	2010	8.36	4.4	2.7
R14	2450	7.99	4.8	3.3

Table S5. Water quality parameters measured along the Simeto and San Leonardo Rivers in April and October 2015.

Sampling Point	Water Temperature (°C)	Conductivity (µS/cm)	pH	COD (mg/L O_2)	Ammonia (NH_4^+- N)	Nitrates (NO_3^- - N)	DO (mg/L)
Simeto River - October 2015							
R1	16.6	183	8.02	10	< 2	0.51	9
R2	19.3	289	7.94	9	< 2	1.63	7.93
C1	20	367	8.25	10	< 2	0.23	8.55
R3	20.9	306	7.84	6	< 2	1.34	9.29
R4	20.7	344	8.38	7	< 2	1.87	8.94
R5	21.7	362	8.27	8	< 2	1.81	8.67
R6	22.2	549	8.29	10	< 2	1.56	9.29
R7	25.5	955	8.43	22	< 2	5.11	9.6
R8	25.1	944	8.6	16	< 2	4.62	10.4
R9	26.1	952	8.62	16	< 2	4.66	11.78
R10	25.4	974	8.5	19	< 2	4.73	10.34
R12	22.9	796	8.23	16	< 2	1.02	8.49
C2	19.9	2220	7.56	23	< 2	1.78	8.03
R13	21.3	1269	8.27	25	< 2	3.81	7.32
R14	23	998	8.23	14	< 2	2.19	8.27
R15	23.2	1221	8.17	16	< 2	2.27	7.76
San Leonardo River - April 2015							
L1	19	957	7.91	2	< 2	2.21	8.6
L2	18.8	1244	7.71	17	6.11	1.56	6.91
L3	20.5	985	7.91	5	< 2	2.39	8.35
L4	21.8	1046	7.84	7	< 2	2.36	6.72
L5	21.5	1040	7.83	7	< 2	2.8	6.81
L6	21.7	1364	7.69	13	< 2	3.27	10.34
San Leonardo River - October 2015							
L1	23.4	939	7.85	13	< 2	1.48	6.55
L2	23	896	7.91	-	3.09	1.23	4.86
L3	23.3	888	7.79	16	< 2	2.72	5.26
L4	25.6	826	7.8	18	2.57	3.23	4.02
L5	23.8	818	7.81	18	< 2	3.75	3.64
L6	24.8	1201	7.8	30	< 2	3.11	3.9

Table S6. Optimized source dependent parameters of mass spectrometer

Parameter	ESI Positive	ESI Negative
Gas Temperature (°C)	350	350
Gas Flowrate (L/min)	10	10
Vaporizer Temperature (°C)	250	250
Nebulizer (psi)	60	60
Capillary (V)	2500	2000
Charging	2000	1500
Delta EMV (V)	0	0

Table S7. Multiple reaction monitoring (MRM) analysis parameters

Compounds	Precursor ion	Product ion	Fragmentor (V)	Collision energy (V)
ESI +				
Caffeine	195	138	110	15
$^{13}C_3$-caffeine	198	140	110	15
Carbamazepine	237	194	20	0
Carbamazepine-d_{10}	247	204	20	0
DEET	192	119	110	15
DEET-d_6	198	119	110	15
Sulfamethoxazole	254	156	80	10
$^{13}C_6$-sulfamethoxazole	260	98	80	10
Sucralose*	419	239	110	20
Sucralose-d_3*	425	243	110	20
Trimethoprim	291	230	110	22
Trimethoprim-d_9	300	234	110	22
ESI -				
Diclofenac	294	250	45	0
$^{13}C_6$-diclofenac	300	256	45	0
Ibuprofen	205	161	60	0
Ibuprofen-d_3	208	164	60	0

**Sucralose and Sucralose-d_3 were analyzed with the addition of a sodium adduct*

Table S8. Fluorescence features and their coordinates selected by peak-picking method.

Fluorescence peak	Description	Excitation wavelength (nm)	Emission wavelength (nm)
I_1	aromatic protein, tyrosine-like fluorescence	225	290
I_2	aromatic proteins, tryptophan-like fluorescence	225	340
I_3	fulvic-like and humic-like fluorescence	245	440
I_4	microbial byproducts, proteins, biopolymers an tryptophan-like fluorescence	275	345
I_5	humic-like fluorescence	345	440

Table S9. Concentrations (ng/L) of the investigated EOCs found in the Simeto River and San Leonardo River water during different sampling campaigns.

Sampling Point	Caffeine	Carbamazepine	DEET	Sucralose	Sulfamethoxazole	Trimethoprim	Diclofenac	Ibuprofen
Simeto River - October 2015								
R1	99	< MRL	16	< MRL	< MRL	< MRL	< MRL	< MRL
R2	82	16	12	32	2	3	< MRL	< MRL
C1	129	10	4	93	12	< MRL	< MRL	< MRL
R3	48	8	4	51	4	< MRL	< MRL	< MRL
R4	47	45	7	198	13	1	19	< MRL
R5	32	31	4	144	10	< MRL	5	< MRL
R6	24	31	< MRL	71	6	< MRL	< MRL	< MRL
R7	448	7	5	164	26	7	9	24
R9	291	9	16	181	47	6	16	30
R10	671	12	13	307	57	11	29	68
R12	49	4	4	< MRL	3	1	< MRL	< MRL
R13	279	28	25	381	32	2	13	< MRL
R14	175	17	18	270	15	2	< MRL	< MRL
R15	179	16	15	287	16	2	< MRL	< MRL
San Leonardo River - April 2015								
L1	78	1	9	< MRL	5	< MRL	< MRL	< MRL
L2	4263	Na	37	1032	81	21	147	290
L3	1137	17	31	312	28	na	56	58
L4	1133	45	35	548	52	14	121	59
L5	1092	49	34	711	57	15	130	70
L6	515	40	49	622	48	13	25	27
San Leonardo River - October 2015								
L1	111	< MRL	8	< MRL	< MRL	1	< MRL	< MRL
L2	4478	39	199	556	45	17	108	164
L3	891	30	140	417	51	5	81	106
L4	962	116	262	914	130	23	219	84
L5	573	96	180	669	97	11	165	< MRL

MRL = minimum reporting limit; na = non-available

Table S10. Fluorescing components determined by PARAFAC modeling and wavelengths corresponding to the positions of their excitation and emission maxima.

Component	Wavelengths of the maxima (nm)		Description	Previously reported by
	Excitation	Emission		
C1	235 & 310	410	Microbial humic-like fluorescence	Yu et al.[21]; Murphy et al.[22]; Ziegmann et al.[23]; Chen et al.[24].
C2	250 & 380	475	Humic-like fluorescence	Yu et al.[21]; Murphy et al.[22]; Yu et al.[25]; Chen et al.[24].
C3	225 & 275	340	Protein, tryptophan-like fluorescence	Yu et al.[21]; Murphy et al.[22]; Sanchez et al.[26].
C4	250 & 350	440	Wastewater/nutrient enrichment tracer, humic-like fluorescence	Yu et al.[21]; Murphy et al.[22]; Yu et al.[25].
C5	220	285	Protein, tyrosine-like fluorescence	Yu et al.[21]; Murphy et al.[22]; Sanchez et al.[26].

Table S11. Square correlation factor (R^2) values for the linear regressions between spectroscopic indexes and EOCs concentrations measured in the San Leonardo River in April 2015. Significant R^2 values are featured in a bold type

	Caffeine	Carbamazepine	DEET	Sucralose	Sulfamethoxazole	Trimethoprim	Diclofenac	Ibuprofen	Conductivity	COD
I_1	**0.868**	**0.976**	0.230	**0.913**	**0.890**	**0.993**	0.527	**0.883**	0.272	**0.725**
I_2	0.661	0.656	0.307	**0.702**	**0.835**	**0.787**	0.602	**0.692**	0.137	0.495
I_3	0.543	0.593	0.617	**0.750**	**0.803**	0.477	0.053	0.534	0.586	**0.858**
I_4	0.456	**0.822**	0.375	**0.905**	**0.856**	**0.868**	0.306	0.392	0.493	**0.741**
I_5	0.681	**0.913**	0.552	**0.944**	**0.975**	**0.913**	0.367	**0.789**	0.430	**0.816**
Φ_T	0.651	**0.923**	0.504	**0.993**	**0.967**	**0.856**	0.332	0.680	0.462	**0.829**
UV_{254}	0.586	**0.986**	0.531	**0.931**	**0.994**	**0.982**	0.500	0.605	0.375	**0.725**
C1	0.414	**0.852**	**0.761**	**0.876**	**0.900**	0.466	0.121	0.379	0.572	**0.791**
C2	0.660	0.691	0.533	**0.840**	**0.852**	0.611	0.103	0.655	0.640	**0.963**
C3	0.681	**0.913**	0.346	**0.980**	**0.942**	**0.907**	0.491	0.679	0.328	**0.716**
C4	**0.739**	**0.911**	0.471	**0.877**	**0.968**	**0.940**	0.469	**0.842**	0.351	**0.766**
C5	**0.931**	**0.961**	0.151	**0.845**	**0.795**	**0.980**	0.493	**0.937**	0.236	**0.695**
Conductivity	0.131	0.222	0.610	0.279	0.362	0.002	0.088	0.059	-	**0.776**
COD	0.569	0.486	0.529	**0.698**	**0.733**	0.404	0.026	0.511	**0.776**	-
n	*6*	*5*	*6*	*5*	*6*	*4*	*5*	*5*	*6*	*6*

Table S12. Square correlation factor (R^2) values for the linear regressions between spectroscopic indexes and EOCs concentrations measured in the San Leonardo River in October 2015. Significant R^2 values are featured in a bold type.

R^2	Caffeine	Carbamazepine	DEET	Sucralose	Sulfamethoxazole	Trimethoprim	Diclofenac	Ibuprofen	Conductiviy*	COD
I_1	**0.901**	0.460	0.279	0.172	0.525	0.235	0.297	**0.995**	0.000	0.409
I_2	0.126	0.682	0.155	0.562	0.472	0.018	0.625	0.004	0.007	0.098
I_3	0.095	**0.834**	0.526	0.543	**0.871**	0.349	0.692	**0.845**	0.945	**0.983**
I_4	0.221	0.500	**0.755**	0.190	0.340	0.436	0.314	0.378	0.405	**0.759**
I_5	0.034	**0.918**	0.633	**0.642**	**0.912**	0.453	**0.788**	**0.711**	0.985	**0.978**
Φ_T	0.069	0.429	**0.820**	0.362	0.633	0.480	0.416	**0.980**	0.626	**0.855**
UV_{254}	0.061	**0.762**	0.507	0.405	**0.723**	0.301	0.569	**0.854**	0.973	**0.973**
C1	0.077	**0.869**	0.532	0.562	**0.866**	0.365	**0.717**	**0.783**	0.970	**0.971**
C2	0.234	**0.795**	0.317	0.471	**0.810**	0.203	0.629	**0.874**	0.857	**0.927**
C3	0.270	**0.794**	0.274	0.561	0.603	0.086	0.674	0.123	0.022	0.190
C4	0.025	**0.934**	0.652	**0.664**	**0.917**	0.475	**0.807**	0.674	0.988	**0.976**
C5	**0.985**	0.327	0.207	0.098	0.418	0.223	0.195	**0.958**	0.004	0.528
Conductivity	0.015	**0.905**	0.633	0.596	**0.832**	0.450	0.749	0.610	-	-
COD	0.585	**0.953**	**0.899**	**0.757**	**0.824**	**0.680**	**0.847**	-	0.968	-
n	*5*	*4*	*5*	*4*	*4*	*5*	*4*	*3*	*5*	*4*

**R^2 values for conductivity are indicative of negative correlations, which is in disagreement with correlations related to previous collection events or other parameters, where correlations have always been positive.*

Table S13. Square correlation factor (R^2) values for the linear regressions between spectroscopic indexes and EOCs concentrations measured in the Simeto River, stretch A, in October 2015. Significant R^2 values are featured in a bold type.

	Caffeine	Carbamazepine	DEET	Sucralose	Sulfamethoxazole	Trimethoprim	Diclofenac	Ibuprofen	Conductivity	COD
I_1	0.564	0.021	0.001	0.645	0.531	0.530	**0.737**	**0.856**	0.440	0.398
I_2	0.291	0.043	0.016	0.146	0.155	0.503	0.237	0.044	0.280	0.341
I_3	**0.732**	0.153	0.013	0.511	**0.706**	0.608	0.068	0.078	**0.936**	**0.880**
I_4	0.502	0.005	0.000	**0.726**	0.499	0.183	0.449	0.270	0.561	0.465
I_5	**0.817**	0.120	0.022	**0.651**	**0.788**	**0.691**	0.181	0.242	**0.913**	**0.860**
Φ_T	**0.734**	0.108	0.006	0.564	**0.690**	0.615	0.114	0.001	0.900	**0.851**
UV_{254}	**0.703**	0.145	0.039	0.510	**0.743**	0.572	0.077	0.535	**0.930**	**0.864**
C1	**0.737**	0.136	0.012	0.538	**0.716**	0.618	0.078	0.026	**0.940**	**0.871**
C2	**0.697**	0.157	0.009	0.469	0.662	0.572	0.041	0.213	**0.907**	**0.879**
C3	0.197	0.025	0.034	0.235	0.088	0.019	0.261	0.109	0.183	0.200
C4	**0.871**	0.100	0.032	**0.741**	**0.829**	**0.745**	0.308	0.591	**0.868**	**0.820**
C5	0.410	0.031	0.005	0.501	0.412	0.519	**0.796**	**0.973**	0.246	0.231
Conductivity	**0.714**	0.189	0.048	0.461	**0.784**	**0.707**	0.124	0.931	-	**0.818**
COD	**0.811**	0.266	0.062	0.357	0.603	**0.728**	0.044	0.020	**0.818**	-
n	*10*	*9*	*9*	*9*	*9*	*5*	*5*	*3*	*10*	*10*

Table S14. Square correlation factor (R^2) values for the linear regressions between spectroscopic indexes and EOCs concentrations measured in the Simeto River, stretch B, in October 2015. Significant R^2 values are featured in a bold type.

	Caffeine	Carbamazepine	DEET	Sucralose	Sulfamethoxazole	Trimethoprim	Diclofenac	Ibuprofen	Conductivity	COD
I_1	**0.915**	**0.945**	**0.947**	**0.819**	**0.945**	**0.805**	-	-	0.570	0.584
I_2	**0.791**	**0.819**	**0.890**	0.396	**0.755**	**0.724**	-	-	0.389	0.313
I_3	**0.992**	**0.997**	**0.967**	**0.990**	**0.989**	**0.930**	-	-	**0.798**	0.571
I_4	**0.944**	**0.969**	**0.961**	**0.875**	**0.968**	**0.844**	-	-	0.630	0.592
I_5	**0.983**	**0.993**	**0.963**	**0.981**	**0.994**	**0.906**	-	-	**0.765**	**0.603**
Φ_T	**0.982**	**0.994**	**0.983**	**0.931**	**0.979**	**0.909**	-	-	**0.721**	0.550
UV_{254}	**0.958**	**0.978**	**0.980**	**0.860**	**0.961**	**0.874**	-	-	0.650	0.541
C1	**0.990**	**0.998**	**0.974**	**0.976**	**0.988**	**0.923**	-	-	**0.772**	0.568
C2	**0.985**	**0.994**	**0.961**	**0.988**	**0.994**	**0.913**	-	-	**0.781**	0.599
C3	0.620	0.654	**0.748**	0.210	0.589	0.550	-	-	0.218	0.219
C4	**0.981**	**0.991**	**0.953**	**0.990**	**0.997**	**0.903**	-	-	**0.778**	**0.618**
C5	**0.823**	**0.864**	**0.816**	**0.936**	**0.935**	0.667	-	-	0.526	**0.807**
Conductivity	**0.824**	**0.784**	**0.717**	0.545	**0.757**	**0.874**	-	-	-	0.358
COD	0.483	0.523	0.422	**0.999**	0.673	0.323	-	-	0.358	-
n	*4*	*4*	*4*	*3*	*4*	*4*	*1*	*0*	*4*	*4*

Table S15. Square correlation factor (R^2) values for the linear regressions between spectroscopic indexes and water quality parameters measured in the Simeto River in September 2013. Significant R^2 values are featured in a bold type.

	Conductivity	DOC	TN
I_1	0.017	0.001	0.286
I_2	0.062	0.209	0.037
I_3	**0.741**	**0.963**	0.002
I_4	0.235	0.353	0.029
I_5	**0.741**	**0.975**	0.001
Φ_T	**0.706**	**0.853**	0.009
UV_{254}	**0.719**	**0.942**	0.019
C1	**0.736**	**0.889**	0.012
C2	**0.733**	**0.974**	0.000
C3	0.098	0.194	0.011
C4	**0.700**	**0.966**	0.000
C5	0.027	0.259	0.500
Conductivity	-	**0.665**	0.023
DOC	**0.665**	-	0.010
TN	0.023	0.010	-
n	*10*	*10*	*10*

Table S16. Summary of the regression parameters for correlations between fluorescence indexes and EOCs, conductivity, COD and DOC concentrations in the Simeto River. Reported parameters include slope, y intercept (y-int) and number of data points (n).

Compound	Index	Simeto A Oct 2015				Simeto B Oct 2015				Simeto Sept 2013			
		R^2	slope	y-int	n	R^2	slope	y-int	n	R^2	slope	y-int	n
Ibuprofen	C5	0.973	1051	-6.851	3	-	-	-	-	-	-	-	-
	I_1	0.856	752.73	-21.278	3	-	-	-	-	-	-	-	-
Caffeine	C5	-	-	-	-	0.823	8492.2	-188.05	4	-	-	-	-
	I_1	-	-	-	-	0.915	4959.4	-172.96	4	-	-	-	-
Sulfamethoxazole	C4	0.829	293.11	-23.487	9	0.997	163.87	-21.98	4	-	-	-	-
	I_5	0.788	162.58	-20.939	9	0.994	116.12	-33.298	4	-	-	-	-
Carbamazepine	C4	-	-	-	-	0.991	132.36	-14.795	4	-	-	-	-
	I_5	-	-	-	-	0.993	94.04	-24.043	4	-	-	-	-
Sucralose	C4	0.741	1208.2	-40.27	9	0.990	1056.6	36.098	3	-	-	-	-
	I_5	0.651	644.1	-23.251	9	0.981	756.24	-40.614	3	-	-	-	-
COD	C1	0.871	34.49	1.3835	10	0.568	37.643	-4.4876	4	-	-	-	-
	C2	0.879	50.958	-0.1016	10	0.599	44.717	-1.4835	4	-	-	-	-
	C4	0.820	80.88	0.1808	10	0.618	52.647	5.4647	4	-	-	-	-
	I_3	0.880	26.413	0.8785	10	0.571	26.645	-3.1536	4	-	-	-	-
	I_5	0.860	46.903	0.4106	10	0.603	36.884	2.0088	4	-	-	-	-
Conductivity	C1	0.940	2070.2	-91.14	10	0.772	1943.7	-77.239	4	0.736	2071.5	-112.42	10
	C2	0.907	2990.3	-164.45	10	0.781	2260.9	98.529	4	0.733	2469.2	-84.283	10
	C4	0.868	4805.9	-156.4	10	0.778	2614.5	460.88	4	0.700	4581.7	-323.76	10
	I_3	0.936	1573.3	-116.5	10	0.798	1394.9	-23.388	4	0.741	1505.4	-156.05	10
	I_5	0.913	2790.4	-143.57	10	0.765	1840.3	285.58	4	0.741	2674.1	-290.04	10
DOC	C1	-	-	-	-	-	-	-	-	0.889	4.105	1.0494	10
	C2	-	-	-	-	-	-	-	-	0.974	5.132	0.9668	10
	C4	-	-	-	-	-	-	-	-	0.966	9.7053	0.4024	10
	I_3	-	-	-	-	-	-	-	-	0.963	3.0966	0.8498	10
	I_5	-	-	-	-	-	-	-	-	0.975	5.5324	0.5547	10

Table S17. Summary of the regression parameters for correlations between fluorescence indexes and EOCs, conductivity and COD concentrations in the San Leonardo River. Reported parameters include slope, y intercept (y-int) and number of data points (n).

Compound	Index	San Leonardo Apr 2015				San Leonardo Oct 2015			
		R^2	slope	y-int	n	R^2	slope	y-int	n
Ibuprofen	C5	0.937	831.23	-34.113	5	0.958	157.2	56.699	3
	I_1	0.883	807.83	-57.752	5	0.995	177.5	26.316	3
Caffeine	C5	0.931	11387	-210.75	6	0.985	7401.5	-698.21	5
	I_1	0.868	10388	-361.89	6	0.901	6656.3	-1159.7	5
Sulfamethoxazole	C4	0.968	497.09	-64.868	6	0.917	483.73	-193.07	4
	I_5	0.975	354.08	-69.099	6	0.912	324.19	-203.17	4
Carbamazepine	C4	0.911	506.03	-73.827	5	0.934	509.8	-218.47	4
	I_5	0.913	343.25	-73.227	5	0.918	339.74	-227.44	4
Sucralose	C4	0.877	6465.2	-889.35	5	0.664	2148.2	-577.16	4
	I_5	0.944	4958.8	-1070.1	5	0.642	1419.5	-604.32	4
COD	C1	0.791	69.326	-13.587	6	0.971	13.917	3.4404	4
	C2	0.963	118.88	-21.83	6	0.927	26.446	-3.1432	4
	C4	0.766	93.177	-12.13	6	0.976	21.522	4.4594	4
	I_3	0.858	57.088	-16.618	6	0.983	12.082	0.6363	4
	I_5	0.816	68.216	-13.518	6	0.978	14.787	3.6629	4

Table S18. Summary of the regression parameters for correlations developed using normalized datasets. Reported parameters include slope, y intercept (y-int) and number of data points (n).

Compound	Index	Linear regression parameters				River stretches included in the combined correlation models
		R^2	slope	y-int	n	
Ibuprofen	C5	0.790	0.8287	-0.0954	11	Simeto A Oct 2015; San Leonardo Apr 2015; San Leonardo Oct 2015
	I_1	0.771	1.086	-0.4432	11	
Caffeine	C5	0.926	1.3117	-0.3117	15	Simeto B Oct 2015; San Leonardo Apr 2015; San Leonard Oct 2015
	I_1	0.879	1.55	-0.55	15	
Sulfamethoxazole	C4	0.864	2.2088	-1.2623	23	Simeto A Oct 2015; Simeto B Oct 2015; San Leonardo Apr 2015; San Leonardo Oct 2015
	I_5	0.822	2.1539	-1.2081	23	
Carbamazepine	C4	0.780	2.4439	-1.4208	13	Simeto B Oct 2015; San Leonardo Apr 2015; San Leonard Oct 2015
	I_5	0.845	2.8791	-1.8508	13	
Sucralose	C4	0.718	1.2885	-0.3674	21	Simeto A Oct 2015; Simeto B Oct 2015; San Leonardo Apr 2015; San Leonardo Oct 2015
	I_5	0.662	1.2136	-0.2838	21	
COD	C1	0.663	1.0511	-0.0567	24	Simeto A Oct 2015; Simeto B Oct 2015; San Leonardo Apr 2015; San Leonardo Oct 2015
	C2	0.683	1.2197	-0.2266	24	
	C4	0.652	1.1013	-0.1057	24	
	I_3	0.705	1.1222	-0.1277	24	
	I_5	0.694	1.131	-0.1356	24	
Conductivity	C1	0.848	1.134	-0.134	24	Simeto A Oct 2015; Simeto B Oct 2015; Simeto Sep 2013
	C2	0.819	1.1795	-0.1795	24	
	C4	0.770	1.2043	-0.2043	24	
	I_3	0.847	1.1719	-0.1719	24	
	I_5	0.828	1.22	-0.22	24	

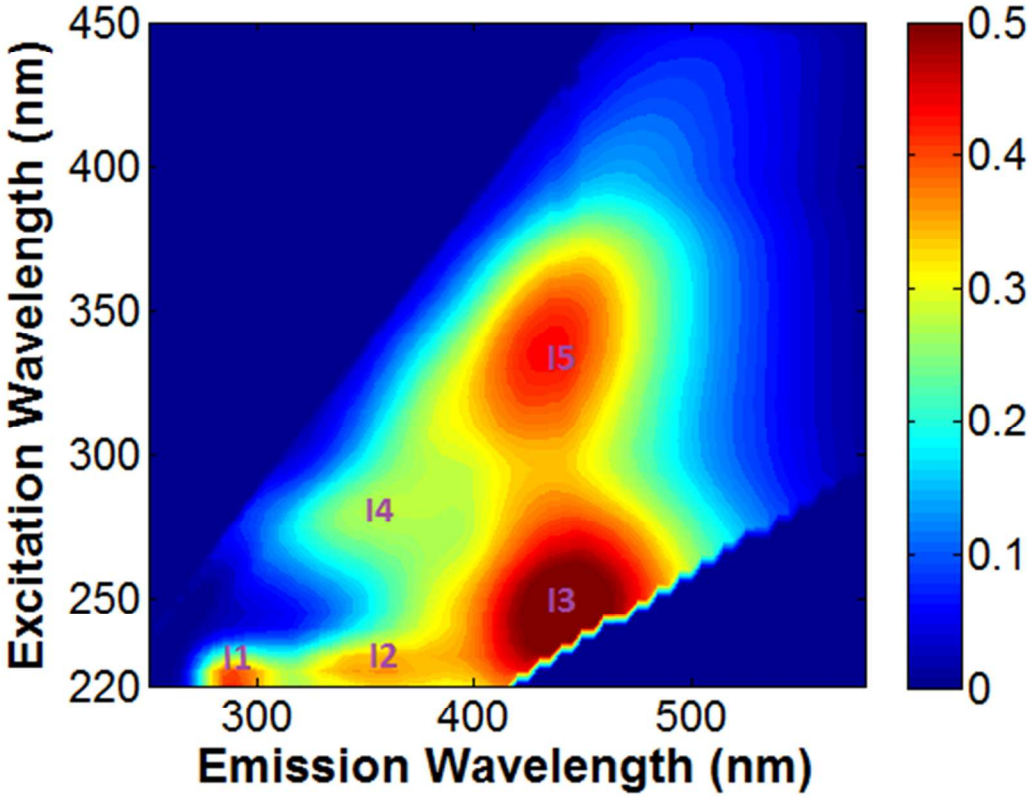


Figure S1. EEM of water sample collected at Point L2 in the San Leonardo River in April 2015 with shown locations of the five nominal assigned peak indexes.

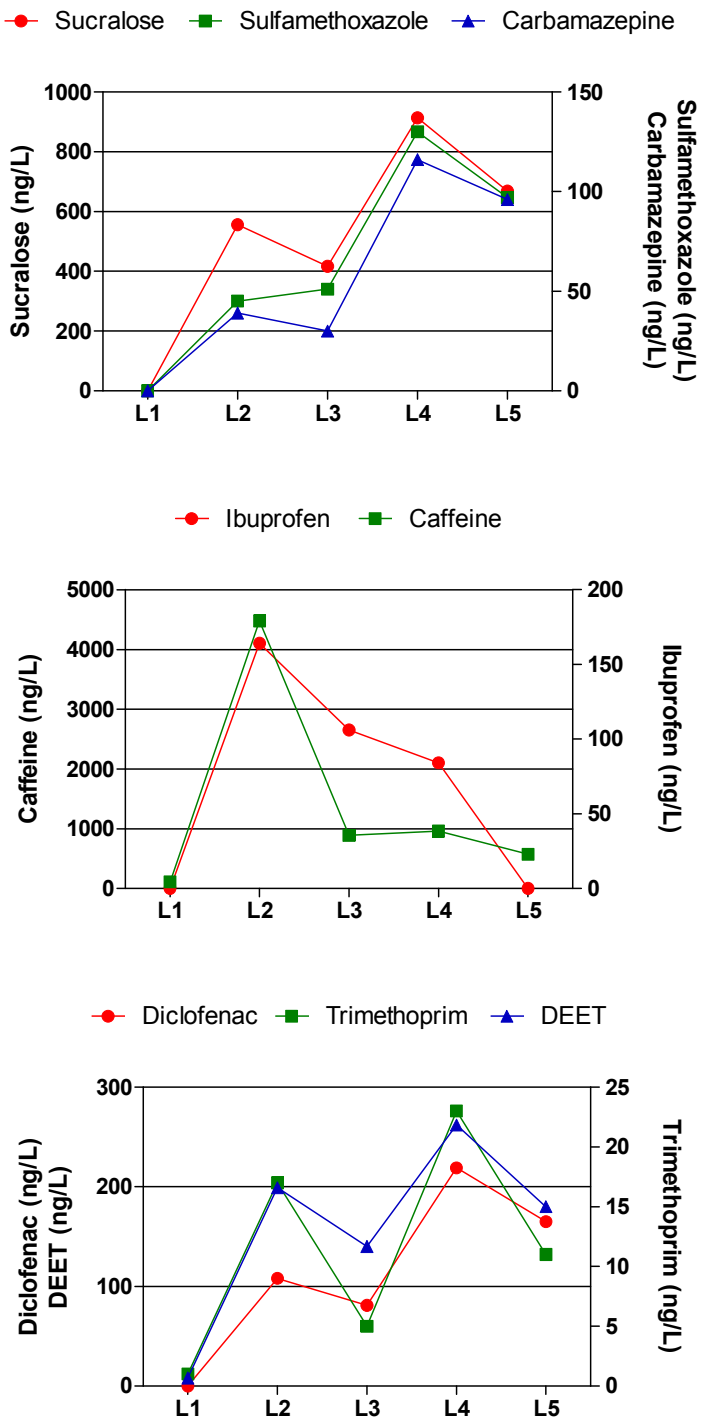
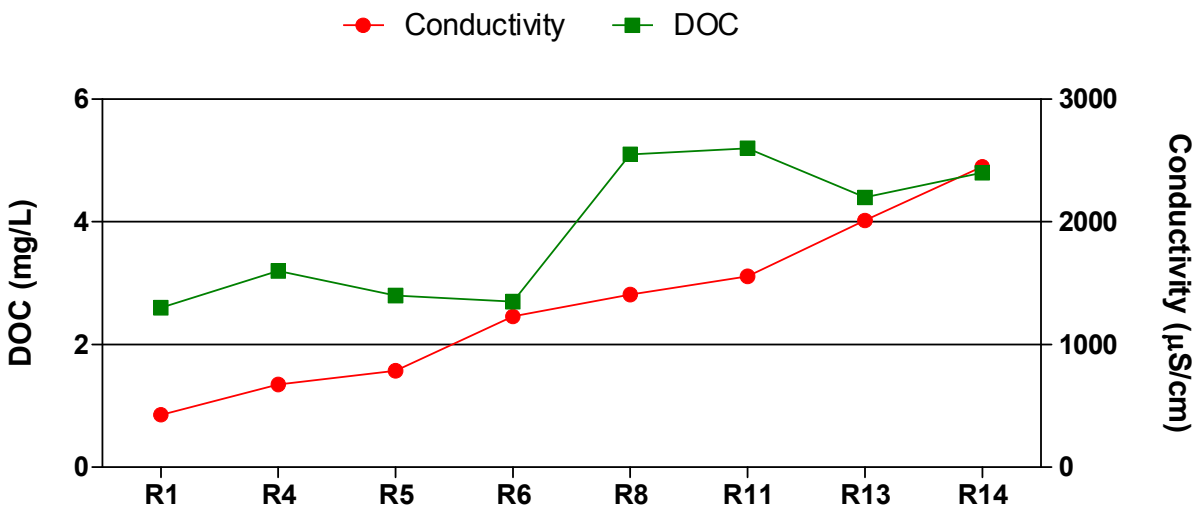


Figure S2. EOCs concentration profiles for the San Leonardo River. Sampling campaign of October 2015.

a)



b)

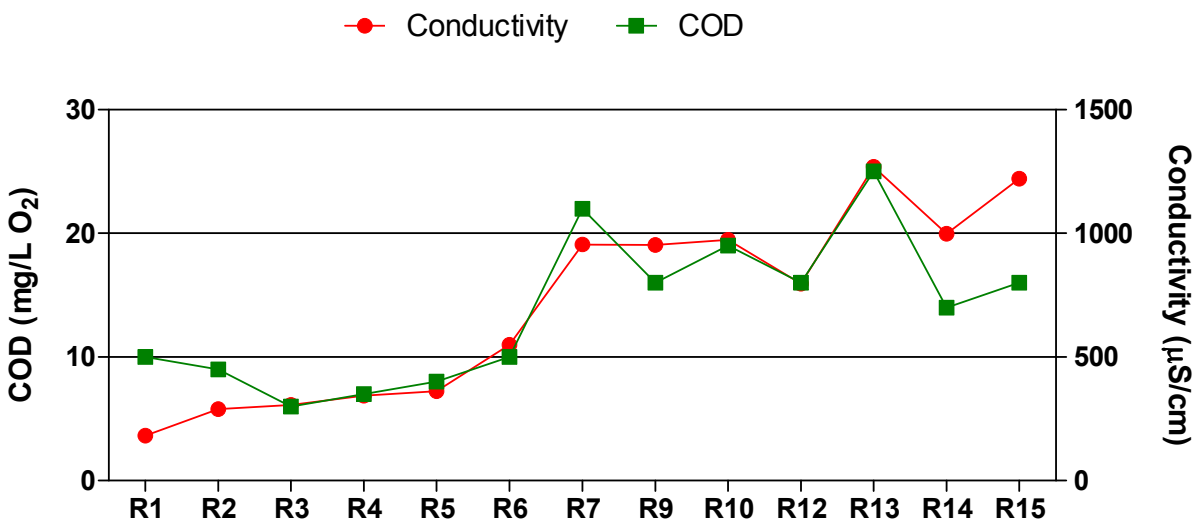
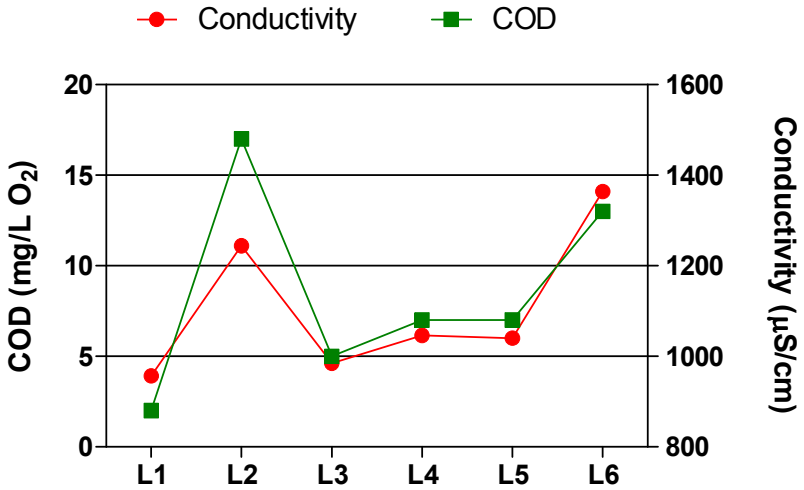
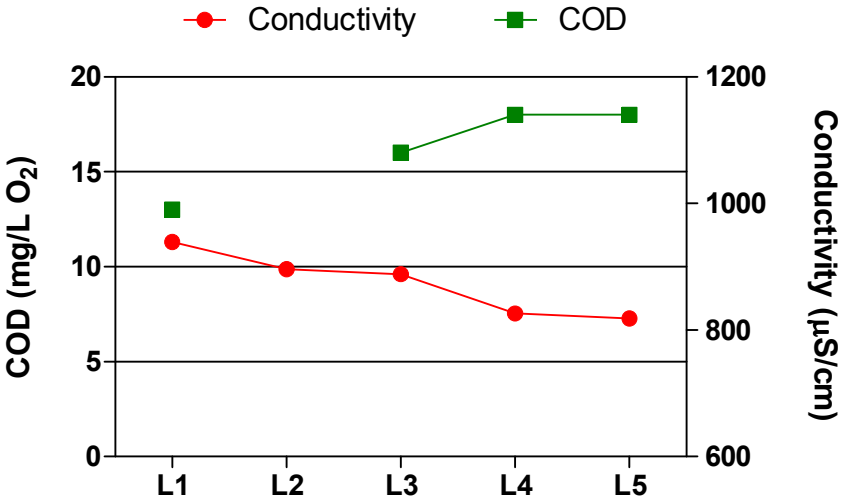


Figure S3. DOC, COD and conductivity concentration profiles for the Simeto River. Sampling campaign of a) September 2013 b) October 2015.



a)



b)

Figure S4. Conductivity and COD concentration profiles for the San Leonardo River. Sampling campaign of a) April 2015 b) October 2015.

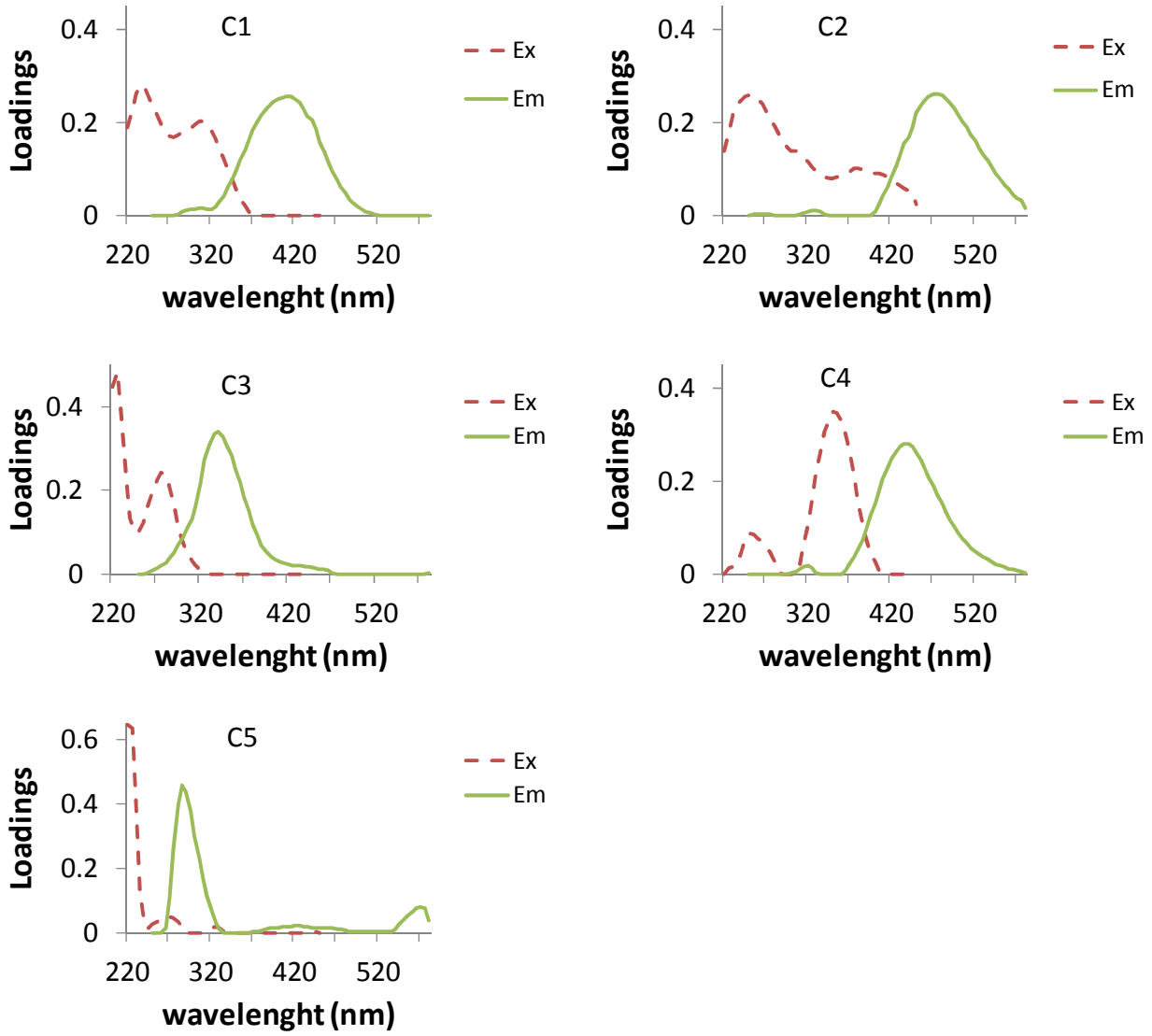


Figure S5. Excitation and emission loadings of fluorescing components C1 - C5.

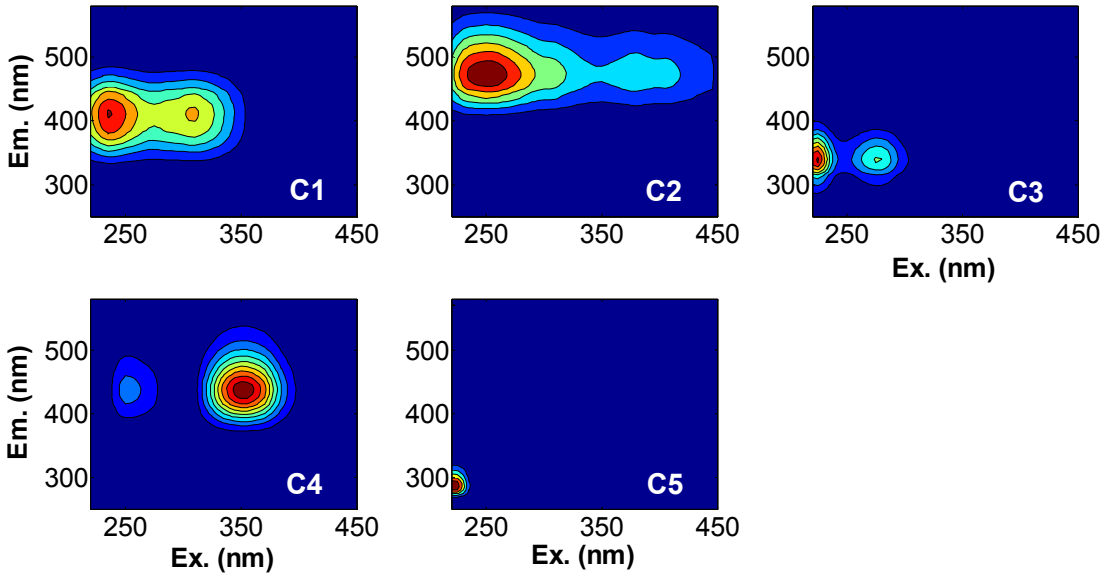


Figure S6. Spectral fingerprints of fluorescing components C1 - C5.

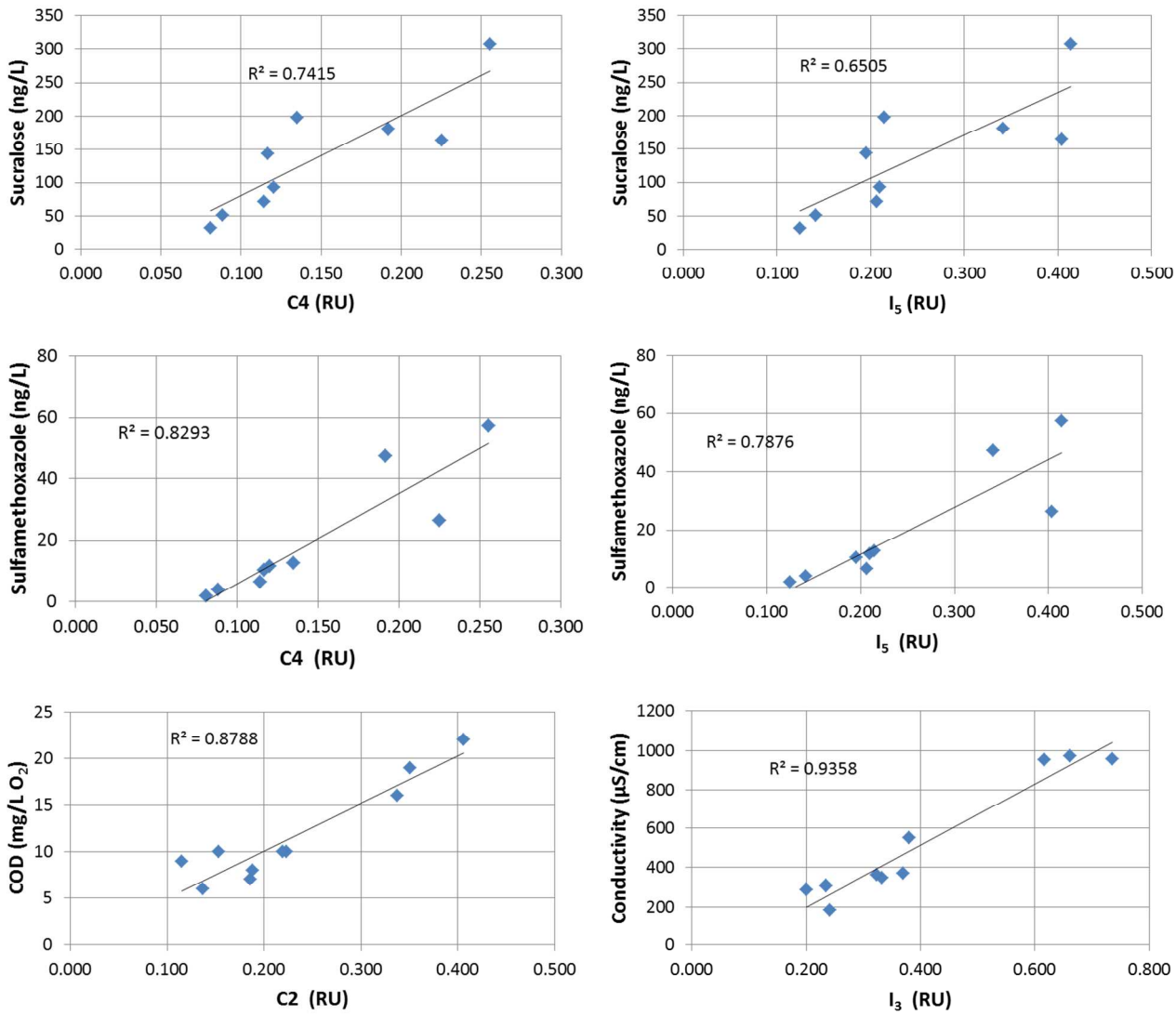


Figure S7. Correlations between EOCs, conductivity, COD and fluorescence indexes for samples from the Simeto River, stretch A, collected in October 2015 (n=9-10).

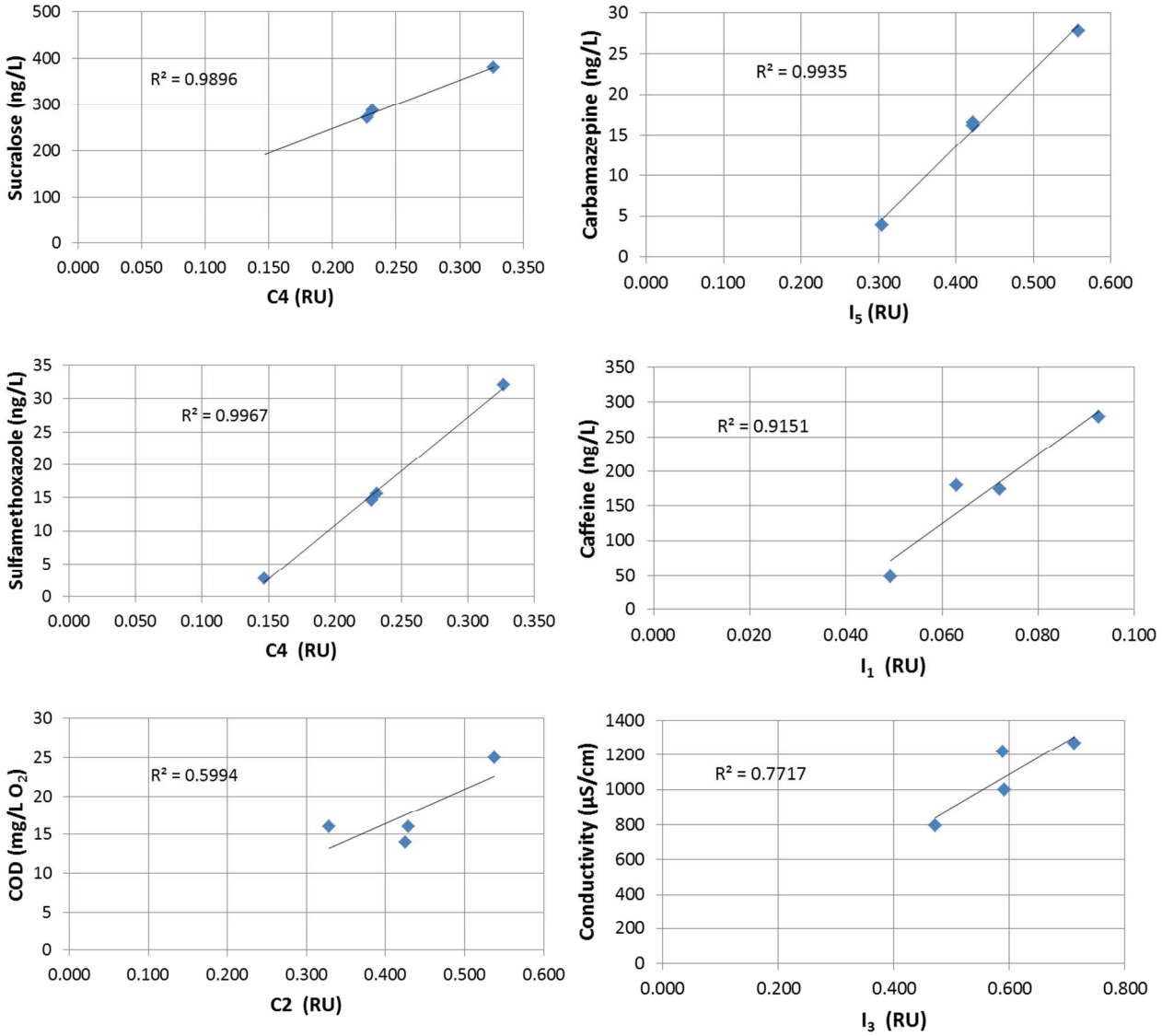


Figure S8. Correlations between EOCs, conductivity, COD and fluorescence indexes for samples from the Simeto River, stretch B, collected in October 2015 (n=3-4).

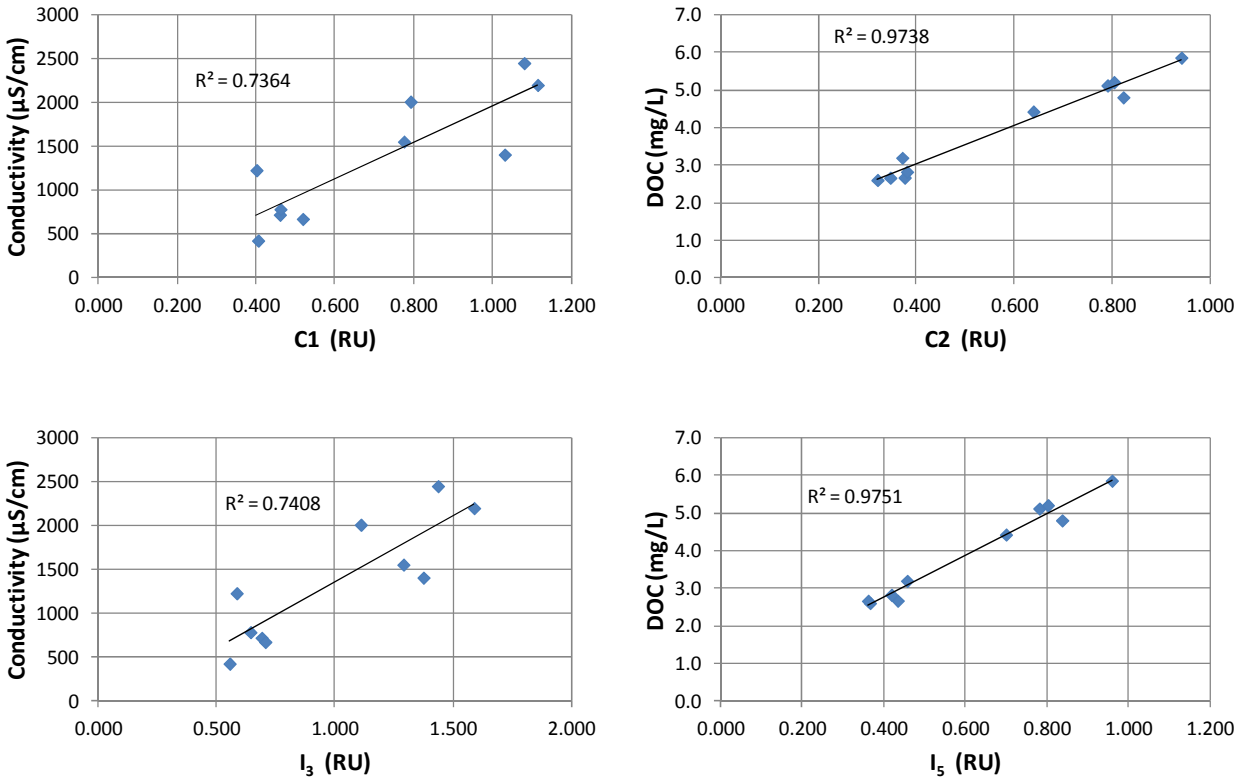


Figure S9. Correlations between conductivity, DOC and fluorescence indexes for samples from the Simeto River collected in September 2013 (n=10).

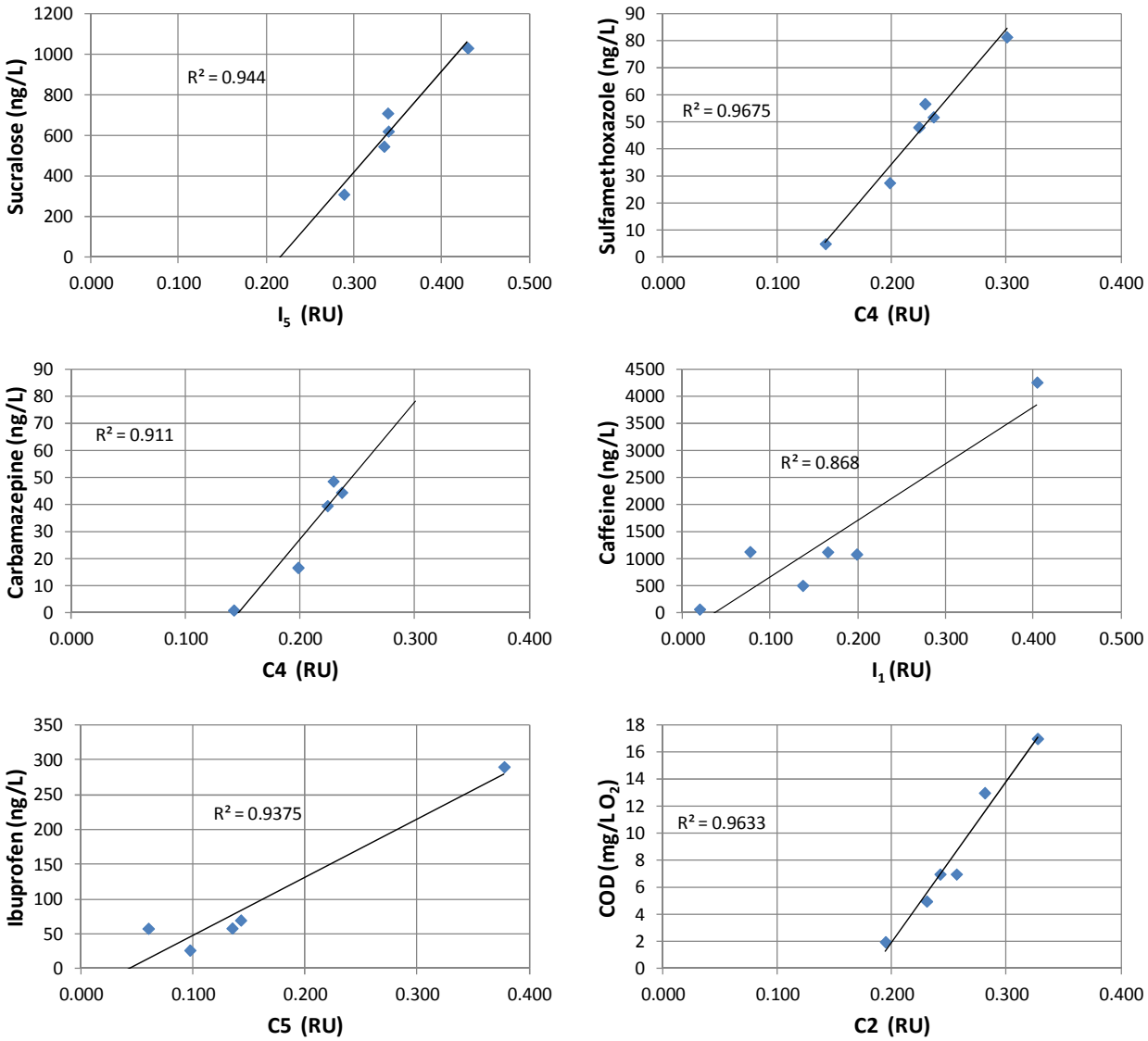


Figure S10. Correlations between EOCs, COD and fluorescence indexes for samples from the San Leonardo River collected in April 2015 (n=5-6).

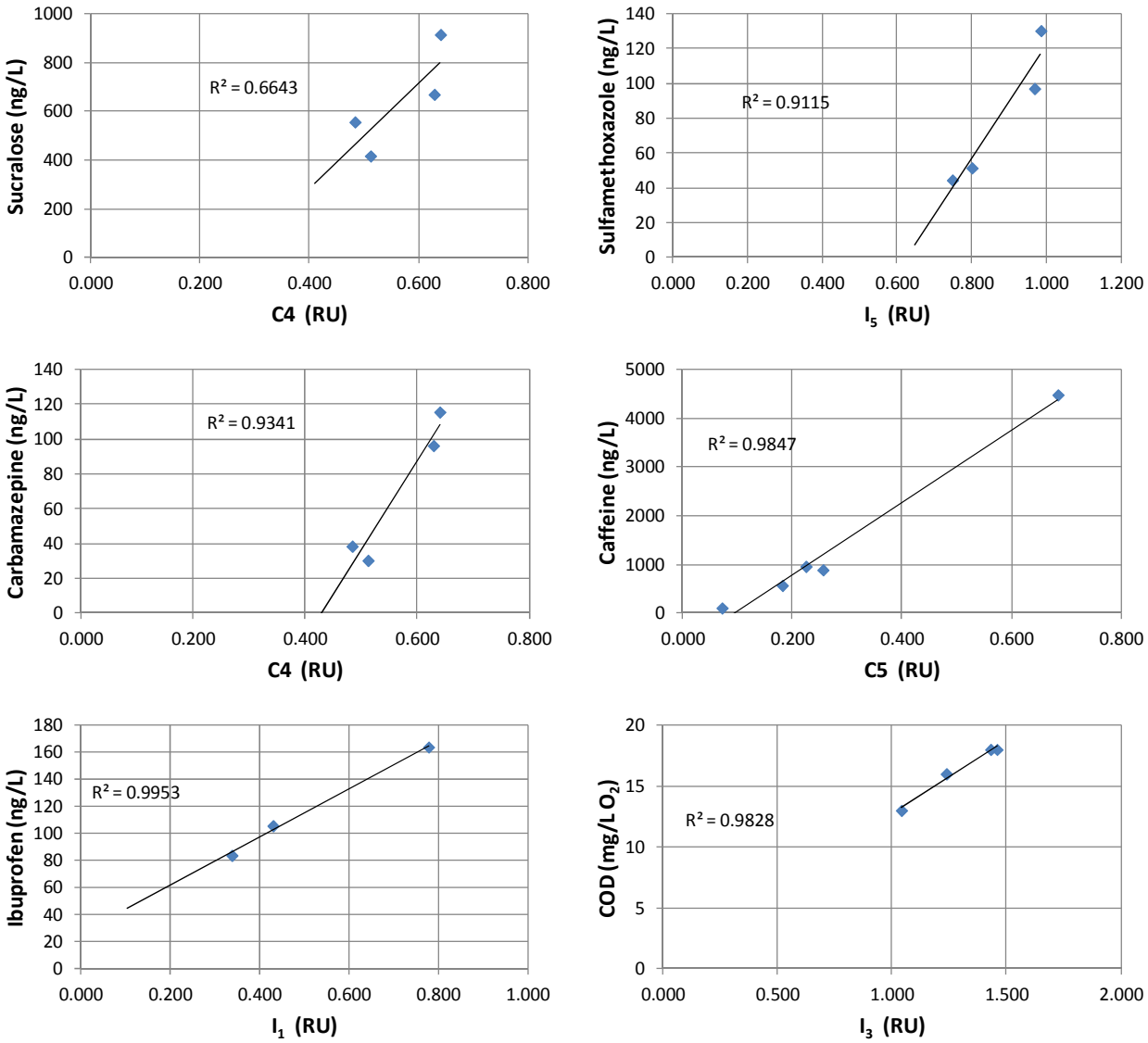


Figure S11. Correlations between EOCs, COD and fluorescence indexes in samples from the San Leonardo River collected in October 2015 (n=3-5).

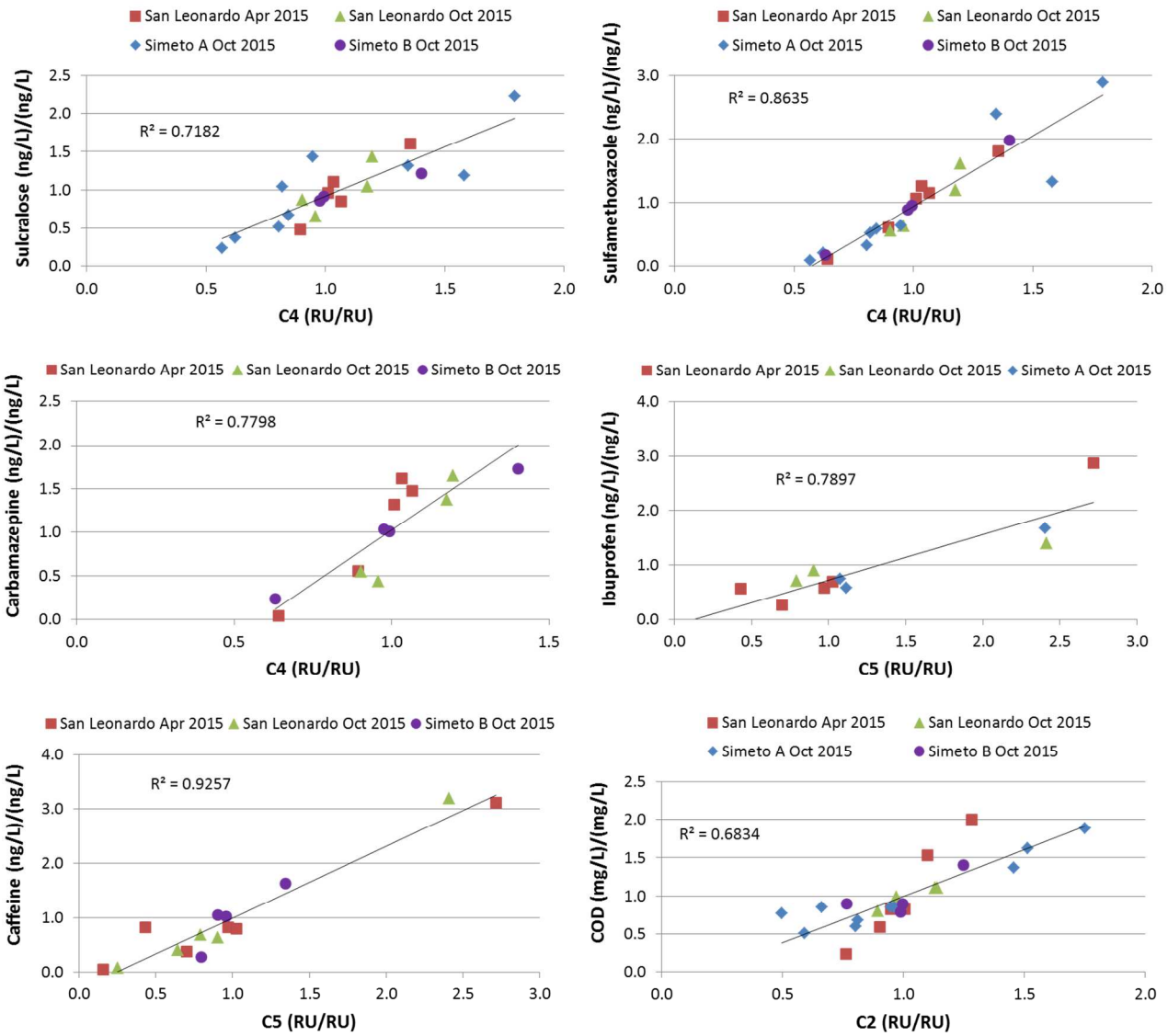


Figure S12. Correlations between EOCs, COD and PARAFAC components normalized by their corresponding average values. The shown datasets include the data for all sampling events reported in this study (n=11-24).

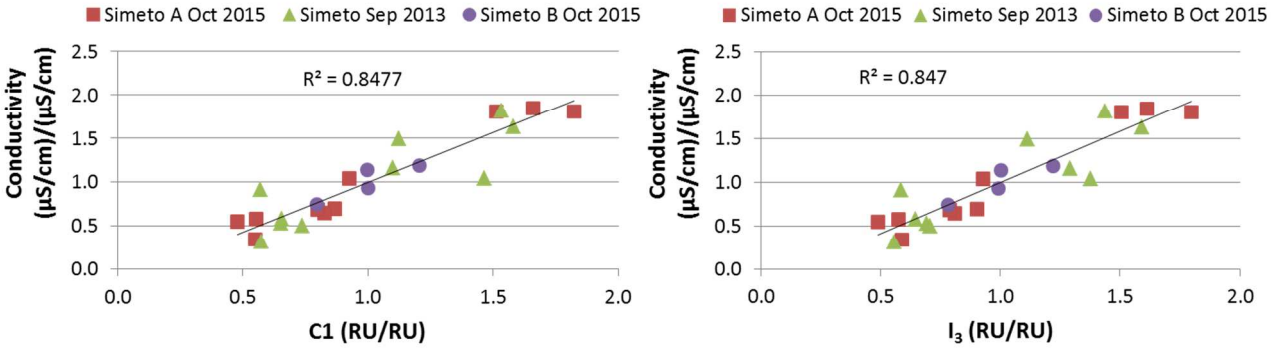


Figure S13. Correlations between conductivity and fluorescence indexes normalized by their corresponding average values. The shown datasets include the data for all sampling events reported in this study (n=24).

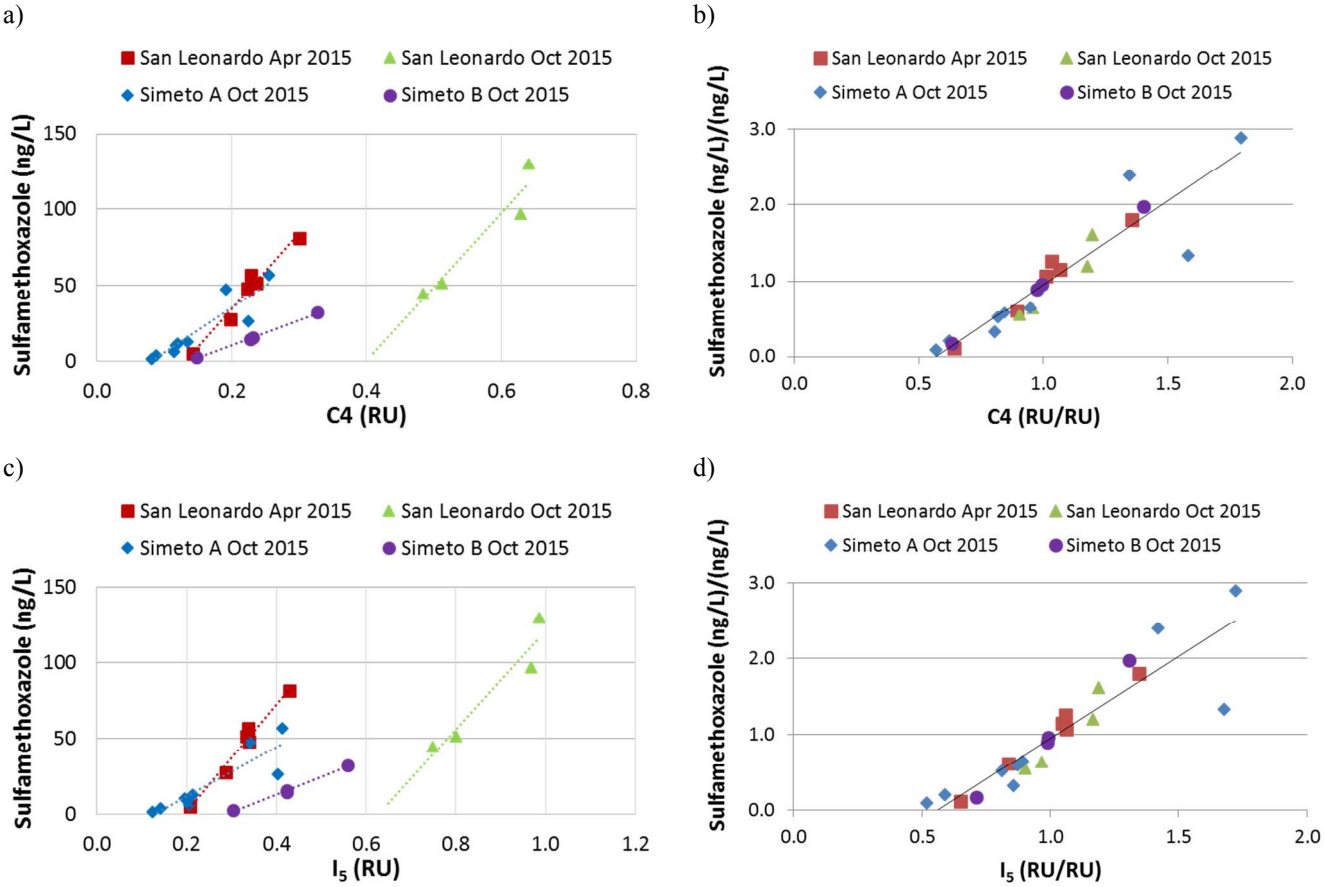


Figure S14. Comparison of correlations developed for sulfamethoxazole and obtained using normalized (b and d) and non-normalized (a and c) datasets (n=23).

Literature cited

1. Sgroi M, Roccaro P, Korshin G V., et al. Use of fluorescence EEM to monitor the removal of emerging contaminants in full scale wastewater treatment plants. *J Hazard Mater*. 2016. doi:10.1016/j.jhazmat.2016.05.035.
2. Murphy KR, Butler KD, Spencer RGM, Stedmon C a, Boehme JR, Aiken GR. Measurement of Dissolved Organic Matter Fluorescence in Aquatic Environments: An Interlaboratory Comparison RID B-8217-2009 RID B-5841-2008. *Environ Sci Technol*. 2010;44(24):9405-9412. doi:10.1021/es102362t.
3. MacDonald BC, Lvin SJ, Patterson H. Correction of fluorescence inner filter effects and the partitioning of pyrene to dissolved organic carbon. *Anal Chim Acta*. 1997;338(1-2):155-162. doi:10.1016/S0003-2670(96)00306-6.
4. Murphy KR, Stedmon CA, Graeber D, Bro R. Fluorescence spectroscopy and multi-way techniques. PARAFAC. *Anal Methods*. 2013;5(23):6557. doi:10.1039/c3ay41160e.
5. Kunkel U, Radke M. Biodegradation of acidic pharmaceuticals in bed sediments: Insight from a laboratory experiment. *Environ Sci Technol*. 2008;42(19):7273-7279. doi:10.1021/es801562j.
6. Buser HR, Poiger T, Markus D. Muller. Occurrence and Environmental Behavior of the Chiral Pharmaceutical Drug Ibuprofen in Surface Waters and in Wastewater. *Environ Sci Technol*. 1999;33:2529-2535. http://lib3.dss.go.th/fulltext/Journal/Environ Sci. Technology1998-2001/1999/no.15/15,1999 vol.33,no15,p.2529-2535.pdf.
7. Jekel M, Dott W, Bergmann A, et al. Selection of organic process and source indicator substances for the anthropogenically influenced water cycle. *Chemosphere*. 2015;125(January 2015):155-167. doi:10.1016/j.chemosphere.2014.12.025.

8. Trenholm RA, Vanderford BJ, Holady JC, Rexing DJ, Snyder SA. Broad range analysis of endocrine disruptors and pharmaceuticals using gas chromatography and liquid chromatography tandem mass spectrometry. *Chemosphere*. 2006;65(11):1990-1998. doi:10.1016/j.chemosphere.2006.07.004.

9. Buerge IJ, Poiger T, Müller MD, Buser HR. Combined sewer overflows to surface waters detected by the anthropogenic marker caffeine. *Environ Sci Technol*. 2006;40(13):4096-4102. doi:10.1021/es052553l.

10. Tixier, C.; Singer, H.P.; Oellers, S.; Müller SR. Occurrence and fate of carbamazepine, clofibric acid, diclofenac, ibuprofen, ketoprofen, and naproxen in surface waters. *Environ Sci Technol*. 2003;37(6):1061-1068. doi:Doi 10.1021/Es025834r.

11. Petrie B, Barden R, Kasprzyk-Hordern B. A review on emerging contaminants in wastewaters and the environment: Current knowledge, understudied areas and recommendations for future monitoring. *Water Res*. 2014;72(0):3-27. doi:10.1016/j.watres.2014.08.053.

12. Weeks JA, Guiney PD, Nikiforovz AI. Assessment of the environmental fate and ecotoxicity of N,N-diethyl-m-toluamide (DEET). *Integr Environ Assess Manag*. 2012;8(1):120-134. doi:10.1002/ieam.1246.

13. Pérez S, Eichhorn P, Aga DS. Evaluating the biodegradability of sulfamethazine, sulfamethoxazole, sulfathiazole, and trimethoprim at different stages of sewage treatment. *Environ Toxicol Chem*. 2005;24(6):1361-1367. doi:10.1897/04-211R.1.

14. Radke M, Lauwigi C, Heinkele G, Múrdter TE, Letzel M. Fate of the antibiotic sulfamethoxazole and its two major human metabolites in a water sediment test. *Environ Sci Technol*. 2009;43(9):3135-3141. doi:10.1021/es900300u.

15. Grünheid S, Amy G, Jekel M. Removal of bulk dissolved organic carbon (DOC) and trace organic compounds by bank filtration and artificial recharge. *Water Res*. 2005;39(14):3219-3228. doi:10.1016/j.watres.2005.05.030.

16. Benotti MJ, Brownawell BJ. Distributions of pharmaceuticals in an urban estuary during both dry- and wet-weather conditions. *Environ Sci Technol*. 2007;41(16):5795-5802. doi:10.1021/es0629965.

17. Clara M, Strenn B, Kreuzinger N. Carbamazepine as a possible anthropogenic marker in the aquatic environment: Investigations on the behaviour of Carbamazepine in wastewater treatment and during groundwater infiltration. *Water Res*. 2004;38(4):947-954. doi:10.1016/j.watres.2003.10.058.

18. Soh L, Connors K a, Brooks BW, Zimmerman J. Fate of sucralose through environmental and water treatment processes and impact on plant indicator species. *Environ Sci Technol*. 2011;45:1363-1369. doi:10.1021/es102719d.

19. Anumol T, Clarke BO, Merel S, Snyder SA. Point-of-use devices for attenuation of trace organic compounds in water. *J Am Water Works Assoc*. 2015;107(9):E474-E485. doi:10.5942/jawwa.2015.107.0129.

20. Regione Siciliana. Piano di tutela delle acque della Sicilia. 2007.

21. Yu H, Liang H, Qu F, et al. Impact of dataset diversity on accuracy and sensitivity of parallel factor analysis model of dissolved organic matter fluorescence excitation-emission matrix. *Sci Rep*. 2015;5(October 2014):10207. doi:10.1038/srep10207.

22. Murphy KR, Hambly A, Singh S, et al. Organic matter fluorescence in municipal water recycling schemes: Toward a unified PARAFAC model. *Environ Sci Technol*. 2011;45(7):2909-2916. doi:10.1021/es103015e.

23. Ziegmann M, Albert M, Muller M, Frimmel FH. Use of fluorescence fingerprints for the estimation of bloom formation and toxin production of mycrocystis aeruginosa. *Water Res*. 2011;45:5110-5122.

24. Chen ML, Price RM, Yamashita Y, Jaffe R. Comparative study of dissolved organic matter from groundwater and surface water in the Florida coastal Everglades using multi-dimensional spectrofluorometry combined with multivariate statistics. *Appl Geochemistry*. 2010;25(6):872-880.

25. Yu H, He PJ, Shao M. Novel insights into sludge dewaterability by fluorescence excitation-emission matrix combined with parallel factor analysis. *Water Res*. 2010;44(3):797-806.

26. Sanchez N, Skeriotis AT, Miller CM. Assessment of dissolved organic matter fluorescence PARAFAC components before and after coagulation-filtration in a full scale water treatment plant. *Water Res*. 2013;47:1679-1690.