

Application of 3D-fluorescence/PARAFAC to monitor the performance of managed aquifer recharge facilities

Maximilian Stahlschmidt, Julia Regnery, Andy Campbell and
Jörg E. Drewes

ABSTRACT

3D-fluorescence spectroscopy was used as a monitoring tool to describe the fate and transport of dissolved organic matter (DOM) during groundwater recharge using recycled water, imported water, and stormwater at a managed aquifer recharge site in California. The study was supplemented by analysis of conservative wastewater-derived trace organic chemicals using liquid chromatography coupled with tandem mass spectrometry. Parallel factor analyses (PARAFAC) yielded six different independent fluorophoric components by mathematically decomposing the excitation emission spectra. The results revealed that this approach was successful in showing the decrease of chromophoric DOM in the subsurface over time and distance during recharge and detecting anthropogenic contaminations that were introduced into the recharge basins, most likely from weed and vector control applications. PARAFAC was able to extract at least one herbicide with chromophoric features from surface and groundwater excitation-emission matrices, suggesting that this approach could also be applied as a pollution control tool for hazardous events.

Key words | dissolved organic matter, managed aquifer recharge, recycled water, three dimensional (3D)-fluorescence spectroscopy

Maximilian Stahlschmidt
Jörg E. Drewes (corresponding author)
Chair of Urban Water Systems Engineering,
Technical University of Munich,
Garching,
Germany
E-mail: jdrewes@tum.de

Julia Regnery
Jörg E. Drewes
Advanced Water Technology Center (AQWATEC),
Department of Civil and Environmental
Engineering,
Colorado School of Mines,
Golden,
CO,
USA

Andy Campbell
Inland Empire Utilities Agency,
Chino,
CA,
USA

INTRODUCTION

Rapid population growth, unpredictable rainfall patterns, lack of conventional fresh water sources, and uncertainties due to climate change are increasing the pressure on water resources around the world (Drewes 2009). As a response to these phenomena, the development of municipal water recycling to achieve drinking water augmentation and recharge of local groundwater resources has gained widespread interest over the past decades (Henderson *et al.* 2009; Drewes & Khan 2011). Managed aquifer recharge (MAR) systems are widely used to augment groundwater supplies and are characterized by relatively low environmental impacts and low capital costs (Drewes & Khan 2011). In MAR systems, such as riverbank filtration, soil aquifer treatment, or aquifer recharge and recovery, water is purposefully added to a groundwater system via natural (i.e. lakes, rivers) and/or engineered structures (i.e. injection

wells, infiltration basins) (Drewes 2009; Drewes & Khan 2011; Parsekian *et al.* 2014). MAR systems that provide sufficient soil passage during infiltration act as a natural filter attenuating turbidity, bacteria, viruses, trace organic compounds (e.g. pesticides, industrial chemicals), nitrogen, algae toxins, and other inorganic and organic constituents through a combination of hydro-geochemical and biological processes such as precipitation, biotransformation, sorption, and ion exchange (Ray *et al.* 2008; Rauch-Williams *et al.* 2010; Hoppe-Jones *et al.* 2010). Further advantages of MAR systems are subsurface storage and the ability to recover recharged water at a later time.

To ensure reliability of MAR in terms of water quality improvement using recycled water sources, water quality monitoring tools are urgently required to maintain not only full protection of public health but also to protect the

environment from anthropogenic contaminants and meet regulatory requirements (Henderson *et al.* 2009). Three dimensional (3D)-fluorescence spectroscopy has received increasing interest as a monitoring tool for chromophoric dissolved organic matter (DOM) in a range of applications including the monitoring of drinking water quality (Stedmon *et al.* 2011), wastewater treatment (Hambly *et al.* 2010), surface water (Hudson *et al.* 2007; Henderson *et al.* 2009), and disinfection byproduct formation potentials in drinking water (Hua *et al.* 2007). 3D-fluorescence spectroscopy is: (1) capable of characterizing DOM with chromophoric features in different water types (Chen *et al.* 2003; Weishaar *et al.* 2003; Murphy *et al.* 2008; Fellman *et al.* 2009; Kowalczyk *et al.* 2009; Hao *et al.* 2012), (2) cheap compared to other advanced analytical techniques, (3) rapid (i.e. minimal sample preparation is required) (Henderson *et al.* 2009), and (4) sensitive (i.e. one to three orders of magnitude more sensitive than ultraviolet (UV)-visible spectroscopy) (Hambly *et al.* 2010). In short, 3D-fluorescence spectroscopy records the fluorescence signals of water samples at different emission and excitation wavelengths. For each excitation wavelength, the excitation spectrum is scanned at the same time as the fluorescence-emission spectrum is recorded and results are arranged in a 3D grid (excitation \times emission \times intensity) (Henderson *et al.* 2009; Nir 2013). Visual comparison ('peak-picking') of excitation-emission wavelength pairs in unknown mixtures with identified fluorescence regions of analyzed standards (i.e. humic acid) allows for a qualitative assessment of sample composition (Guo *et al.* 2010; Dahm *et al.* 2013). However, visual comparison can lead to misinterpretation of excitation-emission matrices (EEMs) since fluorescence spectra might have overlaying features (Stedmon *et al.* 2003; Guo *et al.* 2010). Multi-way statistical analysis such as parallel factor analysis (PARAFAC) has been shown to be a powerful tool, decomposing these EEMs into their underlying chemical components to allow a semi-quantitative assessment of individual components (Murphy *et al.* 2013).

In this study, 3D-fluorescence spectroscopy and PARAFAC were used as a tool to characterize and differentiate chromophoric DOM components of natural and anthropogenic origin, with the aim of monitoring the fate and transport of recycled water recharged to the subsurface at a full-scale MAR site in California. Furthermore, DOM

fate and transport was investigated for three different water types (imported surface water, recycled water, and stormwater) that were used for recharge during specified time periods. To our knowledge, such a tracing experiment has not been accomplished at a full-scale site before. We hypothesize that 3D-fluorescence spectroscopy, in combination with multi-way statistical analysis (i.e. PARAFAC), is capable of distinguishing between DOM originating from different sources to monitor the fate and transport of recycled water DOM in groundwater during recharge.

In total, 178 samples were collected at the Chino Basin in California from five different recharge basins over the period of one year (January 2014–January 2015) and analyzed by 3D-fluorescence spectroscopy. Subsequently, a site-specific PARAFAC model was developed to decompose EEM data into their underlying fluorescence spectra.

EXPERIMENTAL

Field site and sampling

In March 2014, the Metropolitan Water District of Southern California initiated a project to investigate the possibility of enhancing groundwater recharge using recycled water in the Chino Basin, California, using 3D-fluorescence spectroscopy as a monitoring tool for DOM in the aquifer. As part of the project, the Inland Empire Utilities Agency (IEUA) sampled recharge events with different source water types that were diverted to five selected recharge basins (Declez, Hickory, Turner 1, Turner 4, and RP3). Each of the chosen recharge basins at the Chino Basin comprises several cross-gradient and downstream monitoring wells as well as lysimeters at different depths (1.5, 3.1, 4.6, 7.6, and 10.7 m). Furthermore, several upstream municipal monitoring wells were monitored throughout the course of the study to determine the DOM baseline in background groundwater.

Between January 2014 and January 2015, a total of 178 samples were collected by IEUA field hydrologists during recharge events applying imported water to basins Declez and RP3 (February–March 2014), recycled water to basins Hickory, Turner 4, and RP3 (February–November 2014), and stormwater to basins Declez, RP3, and Turner 4 (January 2015). Recharge basin Turner 1 received a mix of

recycled water and stormwater (up to 90% recycled water) throughout the sampling campaign, and was omitted for the source water type specific assessment. In addition, source water grab samples were collected directly from delivery pipelines. A summary of all analyzed surface water and groundwater samples (including their respective ID) is provided in Table S1, Supplementary Information (available with the online version of this paper). Samples were collected and transported on ice for bulk parameter analyses (i.e. total organic carbon, total nitrogen, electrical conductivity) to the IEUA water quality laboratory in Ontario, California. Additionally, split samples for analysis of selected recycled water indicators (i.e. acesulfame-K, carbamazepine, dilantin, meprobamate, primidone, sucralose) were collected and shipped to Eaton Analytical Laboratory in Monrovia, California. Concentrations of these recycled water indicators are used to account for dilution and dispersion in the subsurface during recharge. Samples for 3D-fluorescence spectroscopy analyses were pre-filtered through a 0.45 μm filter at the IEUA laboratory prior to shipment to Colorado School of Mines (CSM) in Golden, Colorado. Samples were shipped on ice and were stored at 5 °C pending respective analyses.

During project start-up, the application of weed and vector control in the vicinity of the recharge basins was identified as a possible source of organic carbon that might influence analytical measurements. To investigate the contribution of anthropogenic chemicals on sample EEMs, three herbicides (diluted with water to concentrations applied at the recharge site) were provided by IEUA for further 3D-fluorescence spectroscopy analysis: DuPont Oust XP, Habitat Herbicide, and Monsanto Roundup Pro Concentrate. In addition, aqueous analytical standards of potential anthropogenic contaminants, the polycyclic aromatic hydrocarbon (PAH) phenanthrene (1.6 mg/L) and the pesticide carbofuran (50 mg/L), were prepared at CSM and analyzed by 3D-fluorescence spectroscopy.

Analytical methods

In general, the pre-filtered samples were analyzed within 72 hours after arrival at the CSM laboratory. Samples (10 mL volume) for dissolved organic carbon (DOC) analysis were acidified with phosphoric acid. DOC was

quantified using a Sievers 5310 TOC analyzer with auto-sampler (Ionics Instruments, Boulder, CO) according to *Standard Methods* (APHA 2012). Samples with DOC concentrations higher than 2 mg/L were diluted with ultrapure water to 2 mg/L for further spectroscopic analysis to minimize quenching effects on the fluorescence signal. Ultrapure water was obtained from an EMD Millipore Synergy UV-R system (Billerica, MA). 3D-fluorescence spectroscopy was carried out on a Horiba Jobin Yvon AquaLog spectrofluorometer (Edison, NJ) using a 1 cm quartz fluorospectrometer cell. UV absorbance at 254 nm ($UV_{254\text{nm}}$) was obtained from AquaLog absorbance scans at the respective wavelength. $UV_{254\text{nm}}$ data and DOC concentrations were used to calculate the specific UV absorbance (SUVA) for each sample. SUVA is defined as the ratio between $UV_{254\text{nm}}$ absorbance and DOC concentration, and is an indicator for the degree of aromaticity in the water sample (Weishaar et al. 2003).

The detailed analytical procedure for 3D-fluorescence measurements using the AquaLog spectrofluorometer is described elsewhere (Gilmore 2011). In brief, EEMs were obtained with the following settings: 1 s integration time, 230–599 nm excitation spectra at 3 nm steps, 156–933 nm emission spectra, 4 pixel (2.52 nm) emission increment, and medium CCD gain. By subtracting the measured blank EEM (ultrapure water) from the sample EEM, Raman scatter influences were eliminated. Compensation for the inner filter effect, which is the absorption of both excitation and emission light by the sample matrix (Stedmon et al. 2003; Nir 2013), as well as removal of Rayleigh scatter lines were applied to the scans by an algorithm provided by Aqualog software (version 3.6). EEMs were normalized to the daily measured integral of the Raman water peak area and corrected by their respective dilution factors to obtain the fluorescence intensity for the original undiluted samples. Thus, fluorescence intensity was reported in Raman Units (R.U.). Following this normalization procedure allowed for a comparison of observed spectra with EEMs measured in other studies at different instrument settings using various fluorescence spectrometers (Lawaetz & Stedmon 2009). Finally, corrected EEM contour plots were exported for excitation wavelengths 240–450 nm and emission wavelengths 250–600 nm, and the corrected EEM data were exported as

ASCII files and processed in Solo 7.9.2 (Eigenvector Research Inc., Wenatchee, WA) for PARAFAC analysis.

Statistical methods

PARAFAC is a three-way statistical method with its origin in psychometrics. The method can be applied to decompose fluorescence EEMs into a set of tri-linear terms and a residual array. The PARAFAC model is calculated by an alternating least square algorithm that minimizes the sum of squared residuals. That allows the separation of signals from a complex mixture of compounds such as DOM (Kowalczyk *et al.* 2009).

A PARAFAC model of a three-way array can be described by three loading matrices, **A**, **B**, and **C** with elements a_{if} , b_{jf} , and c_{kf} (Bro 1997):

$$x_{ijk} = \sum_{f=1}^F a_{if} b_{jf} c_{kf} + e_{ijk} \quad (1)$$

where $i = 1, \dots, I$; $j = 1, \dots, J$; $k = 1, \dots, K$

If Equation (1) is applied to EEMs, x_{ijk} is the intensity of fluorescence for the i th sample at emission wavelength j and excitation wavelength k ; e_{ijk} is the residual representing the variability not accounted for by the model. F defines the number of components in the model. Consequently, each f corresponds to a PARAFAC component. Each such component has I a -values, J b -values, and K c -values: one for each sample, one for each emission wavelength, and one for each excitation wavelength, respectively (Bro 1997; Stedmon *et al.* 2003; Bagthoth *et al.* 2011; Murphy *et al.* 2013).

Furthermore, the parameters a_{if} , b_{jf} , and c_{kf} have direct chemical interpretation in a valid model. The parameter a_{if} is directly proportional to the concentration of the f th analyte of the sample i and is defined as scores; b_{jf} and c_{kf} are related to the emission and excitation spectra, respectively, for the f th analyte and are defined as loadings (Bagthoth *et al.* 2011).

Since the loadings obtained with PARAFAC are normalized, all the quantitative information of the samples' fluorescence is given in the model scores. Maximum fluorescence intensity (F_{\max}) can be calculated for each component by multiplying each component's score with the

corresponding excitation and emission loadings at their λ_{\max} (Murphy *et al.* 2013; Prasad *et al.* 2014). For example, for the n th component F_{\max} is calculated as shown in Equation (2):

$$F_{\max_n} = \text{Score}_n \times Ex_n(\lambda_{\max}) \times Em_n(\lambda_{\max}) \quad (2)$$

F_{\max_n} is the fluorescence intensity at the maximum for the n th component, Score_n is the relative intensity of the n th component, $Ex_n(\lambda_{\max})$ is the maximum of the excitation loading of the n th component, and $Em_n(\lambda_{\max})$ is the maximum of the emission loading of the n th component derived from the model (Prasad *et al.* 2014).

The calculation of F_{\max} allows direct quantitative and qualitative comparison of the fluorescence signal of a given component, or the ratio of any two components within a sample set. However, if component A has a higher fluorescence signal than component B, it does not necessarily mean that component A has a higher concentration. Different fluorophores can have different efficiencies at absorbing and converting incident radiation to fluorescence (Murphy *et al.* 2013).

It is important to mention that components identified by PARAFAC in complex mixtures like surface water may represent overlapping spectra of fluorophores sharing the same or similar fluorescence properties (Andersen & Bro 2003; Murphy *et al.* 2008; Bagthoth *et al.* 2011). Since interferences and scatter can hinder the modeling process, successfully building a PARAFAC model is sensitive to specifying the correct number of components and the availability of a large number of representative samples (Murphy *et al.* 2008). While there are many ways to evaluate whether the correct number of components was chosen for a model, it is not always easy to assess the results as different diagnostic tools can give undetermined or even conflicting results (Murphy *et al.* 2008, 2013). Therefore a combination of several methods should be considered for real datasets.

Core consistency is a diagnostic tool, which evaluates the appropriateness of the model. A PARAFAC model can be represented as the relative difference between a so-called Tucker3-like core array, which is calculated from the data and the PARAFAC loadings, and a super diagonal core of ones (Kompany-Zareh 2012). The core consistency is usually expressed as a percentage, indicating how well the loadings represent variation in the data (Bro 1998; Andersen & Bro

2003). For lower numbers of components, the core consistency tends to start high, near 100%, and is decreasing to zero or even negative values if too many components are selected (Andersen & Bro 2003; Murphy *et al.* 2013). For real-world non-ideal datasets, Murphy *et al.* (2013) emphasized that core consistency is not always a reliable diagnostic tool to determine the correct number of PARAFAC components, because it does not provide information if the number of components is chosen correctly; however, it can be used as an indication that the model is not well specified (Bro 1998; Murphy *et al.* 2008, 2013). Another diagnostic tool is split-half analysis. First, the dataset is divided into two halves. Then, for each half a new PARAFAC model is generated. If the correct number of components was chosen, the excitation and emission spectral loadings of the two halves should be identical according to the uniqueness of PARAFAC. In other words, if the wrong number of components is chosen in the split-half experiment, it is very likely that the two models will not be equal due to the differences in the different samples (Andersen & Bro 2003; Murphy *et al.* 2013).

In this study, 82 samples were used to develop a site-specific six-component PARAFAC model. The number of components was determined by split-half analysis, core consistency test, and the identification and removal of outliers from the dataset. Models that could not pass the split-half analysis were rejected. A sample was considered as an outlier if it either contained some instrument error or artifact, or if it was properly measured but it was very different from other samples. Besides the above-mentioned tools, visual inspections of the spectral shapes as well as investigation of the excitation and emission spectra of each component were conducted to validate the correct number of components. The shape of the contour plots of the components should be clear without valleys; the emission spectra should further exhibit exactly one distinct maximum (Li *et al.* 2014).

RESULTS AND DISCUSSION

Travel times and dilution

Concentrations of selected wastewater indicators were used to account for mixing effects of water originating from

different recharge events as well as dilution effects with native groundwater during recharge (data not shown). Travel time estimates were provided by IEUA based on long-term conductivity readings. Table 1 provides a summary of estimated dilution and travel times of water in the subsurface for lysimeters at recharge basins RP3, Declez, Hickory, Turner 4, and Turner 1. Based on trace organic chemical data, dilution can be neglected for lysimeters at all basins except Hickory basin, where more data is needed to draw sufficient scientific conclusions. Looking closer at the monitoring wells, the following conclusions can be drawn: at basins Declez and Turner 1, comparable concentrations of indicator chemicals to recharge surface water were found, which means that dilution can be neglected for monitoring wells at these two basins; at basins RP3 and Turner 4, dilution effects could not be determined since travel time of the recharged water has been estimated in excess of one year; at Hickory basin, high background concentrations of artificial sweeteners like acesulfame-K and sucralose were detected in the monitoring wells.

Fate of bulk parameters

To assess the attenuation of DOM in the subsurface, results were averaged for each recharged water type, independent of recharge basin location. Information provided by IEUA indicated similar performance across recharge basins Declez, RP3, Hickory, and Turner 4 during soil aquifer treatment. Table 2 summarizes the average DOC, UV_{254nm} , and SUVA values and standard deviations of water samples collected from recharge basins and lysimeters at 1.5, 3.1, 4.6, 7.6, and 10.7 m depth for recharge events applying imported water, recycled water, and stormwater, respectively. Recharge basins exhibited the following DOC concentrations on average: 8.1 ± 4.7 mg/L for imported water, 8.3 ± 4.4 mg/L for recycled water, and 7.9 ± 2.2 mg/L for stormwater. Comparing DOC values from a recharge basin to the values measured at a lysimeter at 7.6 m depth, the DOC is reduced in the subsurface by $71.1 \pm 10.7\%$ to 2.3 ± 0.9 mg/L for imported water, $77.3 \pm 7.1\%$ to 1.9 ± 0.6 mg/L for recycled water, and $78.0 \pm 4.9\%$ to 1.7 ± 0.4 mg/L for stormwater. Average UV_{254nm} -absorption values were determined for recharge basins to be 18.5 ± 13.3 m⁻¹ (imported water), 15.4 ± 10.7 m⁻¹ (recycled

Table 1 | Dilution and travel time of water for lysimeters at different depths at the basins RP3, Declez, Turner 4, Turner 1, and Hickory

Basin	Depth (m)	Travel time ^a (d)	Dilution ^b
RP3	1.5	3–5	Negligible
	3.1	9–15	Negligible
	4.6	25–30	Negligible
	7.6	35–50	Negligible
	10.7	75–90	Negligible
	73.8–79.9	390	High background concentration, but dilution non-determinable since travel time > 1 year (no sufficient data)
Declez	1.5	11	Negligible
	3.1	18	Negligible
	4.6	12–40	Negligible
	7.6	20	Negligible
	10.7	Not observed	Negligible
	47.2–53.3	25–40 ^c	Negligible
Turner 4	1.5	12	Negligible
	3.1	28–56	Negligible
	4.6	28–56	Negligible
	7.6	No useful time data	Negligible
	10.7	120	Negligible
	119.5–125.6	360–450	High background concentration, but dilution non-determinable since travel time > 1 year (no sufficient data)
Turner 1	1.5	7–14	Negligible
	3.1	14–28	Negligible
	4.6	21–35	Negligible
	7.6	40–50	Negligible
	10.7	No useful time data	Negligible
	115.8–121.9	135	Concentration comparable with recharged surface water (travel time > 4 months)
Hickory	3.1	Data not available	Data not available
	4.6	Data not available	Data not available
	7.6	Data not available	Data not available
	132.6–144.8	Data not available	High background concentration of artificial sweeteners

^aBased on long-term electrical conductivity readings by IEUA.

^bBased on trace organic chemical concentrations obtained during this study (data not shown).

^cWater level changes, electrical conductivity change may take longer.

water), and $16.8 \pm 4.4 \text{ m}^{-1}$ (stormwater). Likewise DOC concentration, $UV_{254\text{nm}}$ absorbance is reduced in the subsurface; samples from a lysimeter at 7.6 m depth exhibited the following $UV_{254\text{nm}}$ values: $4.8 \pm 2.0 \text{ m}^{-1}$ for imported water (74.1 ± 10.5% reduction), $3.4 \pm 1.2 \text{ m}^{-1}$ for recycled water (77.7 ± 7.6% reduction), and $3.6 \pm 1.2 \text{ m}^{-1}$ (78.8 ± 4.6% reduction) for stormwater. The SUVA values for the recharge basins exhibited distinctly different values for the three different recharge events: $2.5 \pm 1.9 \text{ L/mg m}$ for imported water, $1.8 \pm 0.4 \text{ L/mg m}$ for recycled water, and $2.2 \pm 0.1 \text{ L/mg m}$ for stormwater. At the lysimeter at 10.7 m depth, the following SUVA values were determined: $1.7 \pm 0.2 \text{ L/mg m}$ (imported water), $1.7 \pm 0.3 \text{ L/mg m}$ (recycled water), and $2.1 \pm 0.1 \text{ L/mg m}$ (stormwater),

indicating that SUVA (expressing DOM aromaticity) did decrease for the imported water, but change very little for recycled water and stormwater during subsurface travel.

Grab samples collected directly from the delivery pipes revealed DOC, $UV_{254\text{nm}}$, and SUVA values of 2.3 mg/L, 4.1 m^{-1} , and 1.8 L/mg m for imported surface water and $4.7 \pm 0.4 \text{ mg/L}$, $6.4 \pm 0.6 \text{ m}^{-1}$, and $1.4 \pm 0.01 \text{ L/mg m}$ for recycled water, respectively. DOC, $UV_{254\text{nm}}$, and SUVA values in stormwater samples collected directly from the stormwater channel were in the range of $8.8 \pm 3.2 \text{ mg/L}$, $19.1 \pm 5.9 \text{ m}^{-1}$, and $2.2 \pm 0.1 \text{ L/mg m}$, respectively. Oddly, unexpectedly high DOC concentrations were detected in some of the recharge basins at times, irrespective of the quality of water applied. In general, samples with higher DOC

Table 2 | Mean and standard deviation of DOC, UV_{254nm}, and SUVA for imported water, recycled water, and stormwater recharge events determined at recharge basins (RB) and corresponding lysimeters (LYS) at different depths

Source water	Depth (m)	DOC (mg/L)	UV _{254nm} (m ⁻¹)	SUVA (L/mg m)
Imported water	0 (RB)	8.1 ± 4.7	18.5 ± 13.3	2.5 ± 1.9
	1.5 (LYS)	3.1 ± 1.1	7.5 ± 3.3	2.4 ± 0.3
	3.1 (LYS)	3.0 ± 1.0	7.0 ± 3.1	2.3 ± 0.4
	4.6 (LYS)	2.4 ± 0.6	5.1 ± 1.4	2.1 ± 0.3
	7.6 (LYS)	2.3 ± 0.9	4.8 ± 2.0	2.0 ± 0.3
Recycled water	10.7 (LYS)	1.9 ± 0.5	3.1 ± 0.5	1.7 ± 0.2
	0 (RB)	8.3 ± 4.4	15.4 ± 10.7	1.8 ± 0.4
	1.5 (LYS)	4.4 ± 2.2	11.3 ± 8.1	2.5 ± 0.6
	3.1 (LYS)	3.1 ± 0.9	5.9 ± 1.7	2.0 ± 0.3
	4.6 (LYS)	2.6 ± 0.9	5.4 ± 1.6	2.4 ± 1.7
Stormwater	7.6 (LYS)	1.9 ± 0.6	3.4 ± 1.2	1.8 ± 0.3
	10.7 (LYS)	1.8 ± 0.4	3.0 ± 1.1	1.7 ± 0.3
	0 (RB)	7.9 ± 2.2	16.8 ± 4.4	2.2 ± 0.1
	1.5 (LYS)	2.4 ± 0.7	5.2 ± 0.9	2.3 ± 0.8
	3.1 (LYS)	2.2 ± 0.4	5.2 ± 0.5	2.4 ± 0.5
	4.6 (LYS)	2.2 ± 0.7	5.0 ± 1.3	2.3 ± 0.3
	7.6 (LYS)	1.7 ± 0.4	3.6 ± 0.8	2.1 ± 0.3
	10.7 (LYS)	1.7 ± 0.6	3.5 ± 0.9	2.1 ± 0.1

concentrations also revealed increased values of UV_{254nm} absorbance. Correlation of DOC and UV_{254nm} for all water types at all recharge basins is depicted in Figure 1, and indicates that several recharge basins received additional sporadic input of organic carbon from non-point sources throughout the course of this study.

Application of herbicides for weed control in the vicinity of the recharge facility was identified as one of these

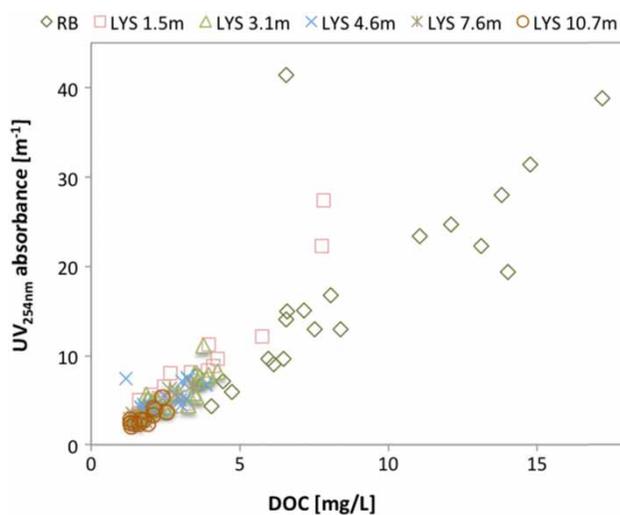


Figure 1 | UV_{254nm} absorbance plotted over DOC for all recharge basin (RB) and lysimeter (LYS) samples (all source water types).

non-point sources for elevated DOC. As summarized in Table 3, additional organic carbon was introduced to the recharge basins via wet and dry deposition, even taking dilution with recharged water into account. All three herbicides are composed of a carbon structure and although biodegradable, the active ingredient in DuPont's Oust XP, sulfometuron methyl (C₁₅H₁₆N₄O₅S, CAS 74222-97-2), for instance, exhibits a reported half-life in water and soil of around 30 days.

EEMs of potential anthropogenic contaminants

While the herbicide Monsanto Roundup Pro Concentrate revealed only a weak fluorescence signal, the fluorescence signals of the herbicides DuPont Oust XP and Habitat Herbicide, the PAH phenanthrene, and the pesticide carbofuran exhibited distinct and well-defined peaks as illustrated in Figure 2. In addition, the Habitat Herbicide exhibited fluorophoric features similar to humic- and fulvic-like derived fluorophores; humic-like material fluoresce in the region λ_{ex/em}: 250–470/380–580 nm, fulvic-like material have fluorophoric properties in the wavelength region λ_{ex/em}: 220–250/380–580 nm (Dahm et al. 2013).

Phenanthrene is generally formed during incomplete combustion of organic matter via natural processes or human activities (Wang et al. 2010). The pesticide carbofuran is marked under the trade name Furadan and is widely used to combat agricultural pests (Alves et al. 2002).

PARAFAC model

The PARAFAC model development was initiated with a series of two to seven components using an EEM dataset of 82 samples. Outliers were removed from the dataset.

Table 3 | Bulk organic parameters of herbicides (in aqueous solution) deployed for weed control

Herbicide	Proportion in solution (%)	DOC (mg/L)	UV _{254nm} (m ⁻¹)	SUVA (L/mg m)
Habitat Herbicide	0.09	427.4	9,684	22.7
Monsanto Roundup Pro	1	338.0	50	0.2
DuPont Oust XP	0.02	243.4	1,399	5.8

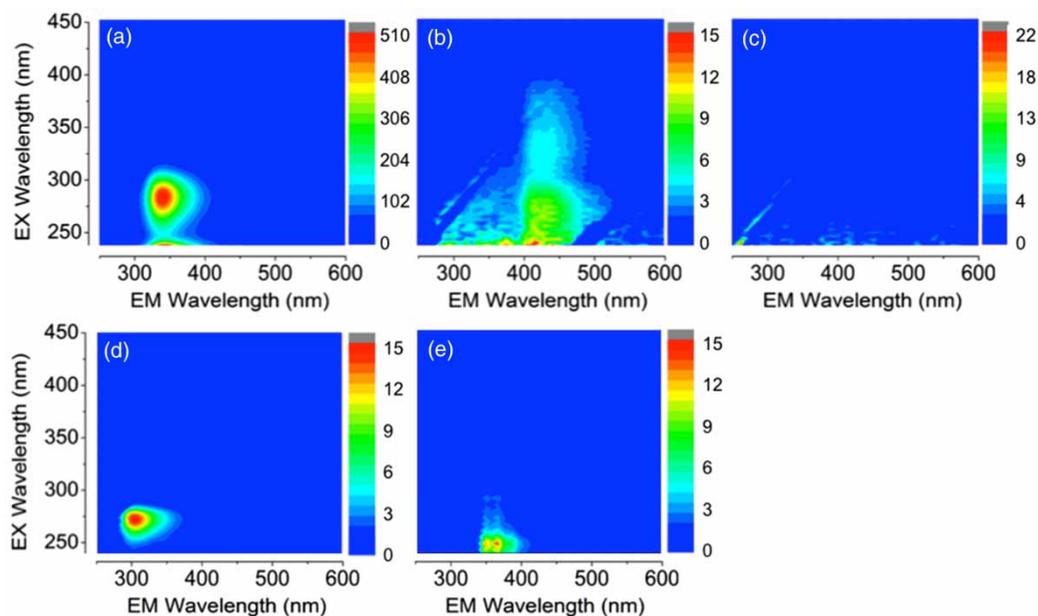


Figure 2 | EEMs of the herbicides DuPont Oust XP (a), Habitat Herbicide (b), Monsanto Roundup Pro Concentrate (c), the PAH phenanthrene (d), and the pesticide carbofuran (e).

From all the models that were established, the six-component PARAFAC model was assessed to be the most appropriate as the model exhibited the best agreement in the split-half validation (i.e. 84% agreement between excitation and emission loading modeled for the two parts of the combined data set and the whole data set), core consistency test (68%), and visual inspections of the spectral shapes. Models that included EEM data of the five individual anthropogenic contaminants could not be validated. Figure 3 summarizes the contour plots, the excitation and emission spectra of the identified individual components, as well as examples of matching components identified by other studies.

Component 1 had an emission maximum at 407 nm, an excitation spectrum with a maximum at 246 nm, and a shoulder at 324 nm. The fluorescence signal in this region results from the presence of both carbon-carbon double bonds and aromatic carbon bonds and is referred to as fulvic-like organic matter (Henderson *et al.* 2009). Component 2 revealed an emission maximum at 483 nm, an excitation spectrum with a maximum below 240 nm, and a secondary excitation band at 369 nm. Fluorescence phenomena in this area represent humic-like organic matter (Murphy *et al.* 2008; Singh *et al.* 2013). Component 3 had an emission maximum at 354 nm, an excitation

maximum below 240 nm, and a shoulder at 294 nm. While this overlaps with the region of amino acids, free or bound in proteins (Murphy *et al.* 2008; Kowalczyk *et al.* 2009), and tryptophan-like components (Henderson *et al.* 2009), it was suspected that component 3 might represent the PAH phenanthrene (Wang *et al.* 2010). Phenanthrene fluoresced at the same wavelength region $\lambda_{\text{ex/em}}$: <250/348–384 nm (Wang *et al.* 2010 and Figure 2), although the shape of component 3 was rather unique. Component 4 revealed an emission maximum at 368 nm, an excitation maximum below 240 nm, and a secondary excitation band at 279 nm. Henderson *et al.* (2009) referred to it as tryptophan-like components. Component 5 had an emission maximum at 327 nm, an excitation maximum at 279 nm, and a shoulder below 240 nm. This peak was also found in other studies, and overlaps with the region of amino acids, free or bound in proteins (Murphy *et al.* 2008; Kowalczyk *et al.* 2009) and tryptophan-like fluorescence (Fellman *et al.* 2009; Singh *et al.* 2013). In addition, Jiji *et al.* (1999) reported two peaks that belong to the pesticides carbaryl ($\lambda_{\text{ex/em}}$: 270–320 nm) and 1-naphthol ($\lambda_{\text{ex/em}}$: 282–335 nm).

Although component 5 may relate to a combination of compounds, it was concluded that component 5 represents predominantly the herbicide DuPont Oust XP, as component 5 revealed two clear excitation and emission

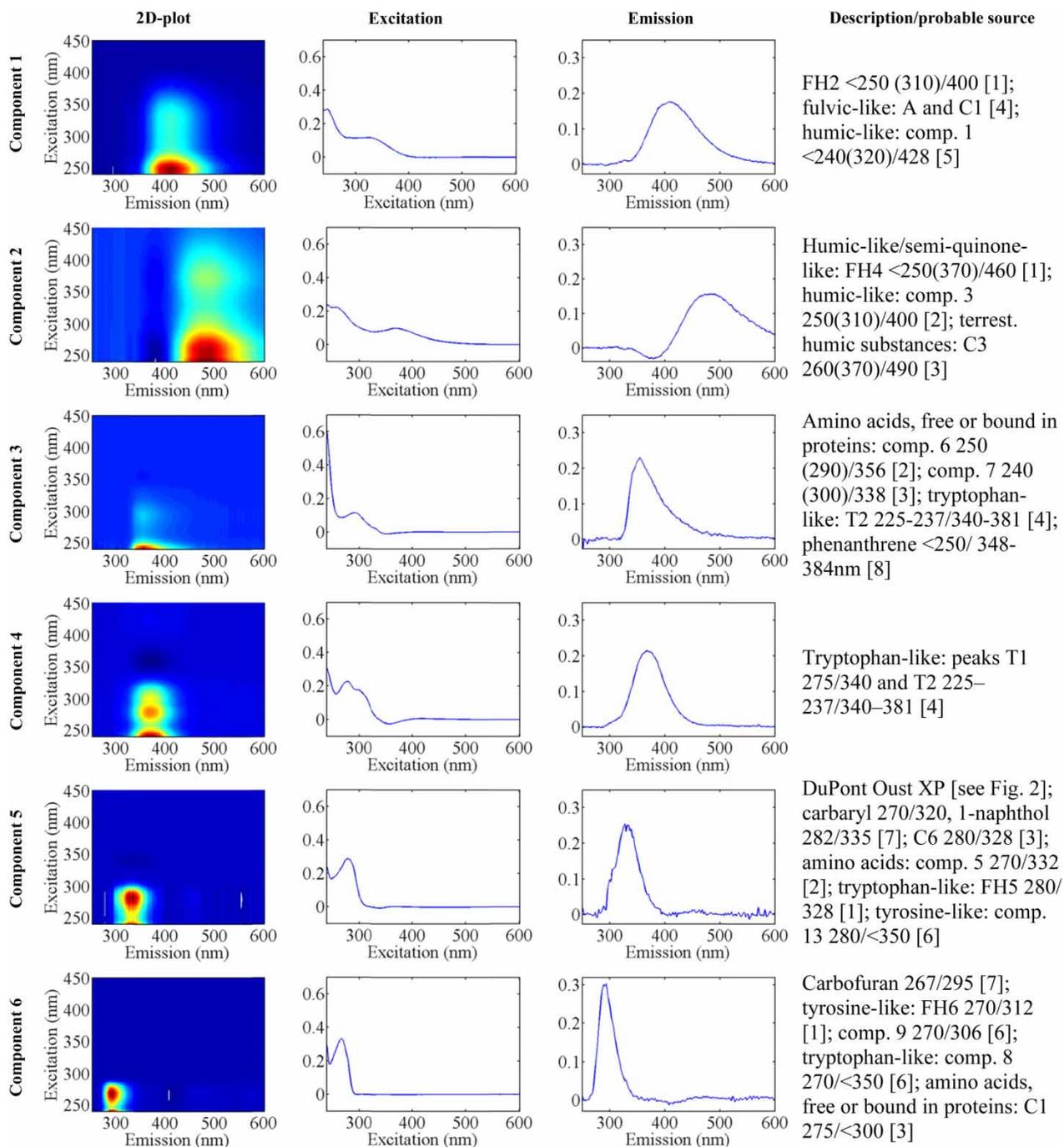


Figure 3 | Output of the PARAFAC analysis showing contour plots, excitation and emission spectra, and probable sources of six components identified in the data set ($n = 82$); 99.34% of variation was captured by the model. Core consistency was determined to be 68%. Split half analysis was successfully developed (84%). References: (Jiji et al. 1999 [7]; Murphy et al. 2008 [3]; Fellman et al. 2009 [6]; Henderson et al. 2009 [4]; Kowalczyk et al. 2009 [2]; Wang et al. 2010 [8]; Stedmon et al. 2011 [5]; Singh et al. 2013) [1].

maxima (Figure 3). DuPont Oust XP exhibited exactly the same maxima with the same spectral shape (see Figure 2) while the protein peaks and the pesticide peaks of carbaryl and 1-naphthol revealed only one maximum. Component 6

had an emission maximum at 293 nm, an excitation maximum at 267 nm, and a shoulder below 240 nm. Fluorescence in this region was found in a broad range of environments and is thought to represent tyrosine-like

components (Fellman *et al.* 2009; Singh *et al.* 2013) and tryptophan-like components (Fellman *et al.* 2009). In addition, Jiji *et al.* (1999) reported that the pesticide carbofuran fluoresced in that region ($\lambda_{\text{ex/em}}$: 267–295 nm). The spectral shape of carbofuran (Figure 2) only revealed one maximum, but clearly overlapped with the spectral shape of component 6.

Attenuation of chromophoric DOM

After validation of the six-component PARAFAC model, the fate of the components during recharge events with different source water types was tracked using their calculated F_{max} values. F_{max} values of all components present in imported source water, recycled source water, stormwater, and up-gradient background wells are depicted in Figure 4.

Summarized by recharged water type, the distribution of each component in the subsurface at different depths at the recharge basins is illustrated in Figure 5.

Comparing the F_{max} values for each component from recharge basin to the monitoring wells, it can be summarized that F_{max} of each component was attenuated during water recharge (Figure 5). F_{max} can be reduced either by biodegradation of DOM or dilution with native groundwater (Hoppe-Jones *et al.* 2010). Since dilution is negligible for

most of the lysimeters (Table 1), attenuation of the fluorescence signal is mainly related to microbial metabolism and/or sorption during travel through the subsurface. In general, F_{max} was higher for fulvic-like component 1 than any of the other components for each recharge event (Figure 5 and Table S1, Supplementary Information). While these findings suggest that all the samples were predominated by fulvic-like components, it has to be considered that fluorescence intensity is not only proportional to concentration but also to quantum yield (Baghoth *et al.* 2011). The differences in the relative intensities of the components are not only related to differences in concentration but to a combination of concentration and/or quantum efficiencies of the individual fluorophores.

Anthropogenic contribution to 3D-fluorescence during recharge

PARAFAC analysis was able to identify the samples that were influenced by anthropogenic chemicals due to increased F_{max} values of individual components. Figure 6 illustrates the EEMs of surface water samples collected at recharge basins RP3, Hickory, Turner 4, RP3, and Declez at different dates during recharge with recycled water and imported water, respectively. The influence of the Habitat

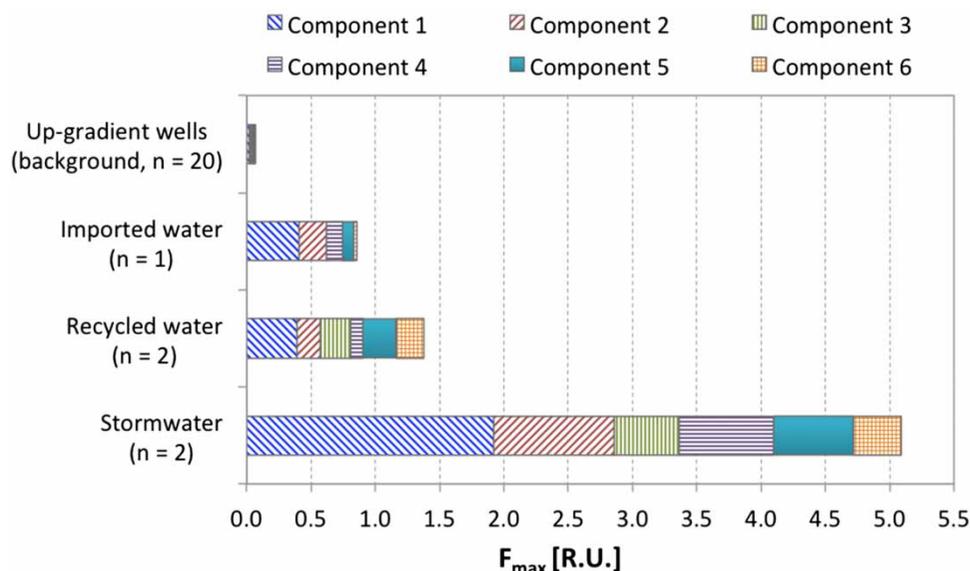


Figure 4 | F_{max} of six components identified by PARAFAC in source water grab samples and up-gradient monitoring well samples. Sample numbers are provided in parentheses.

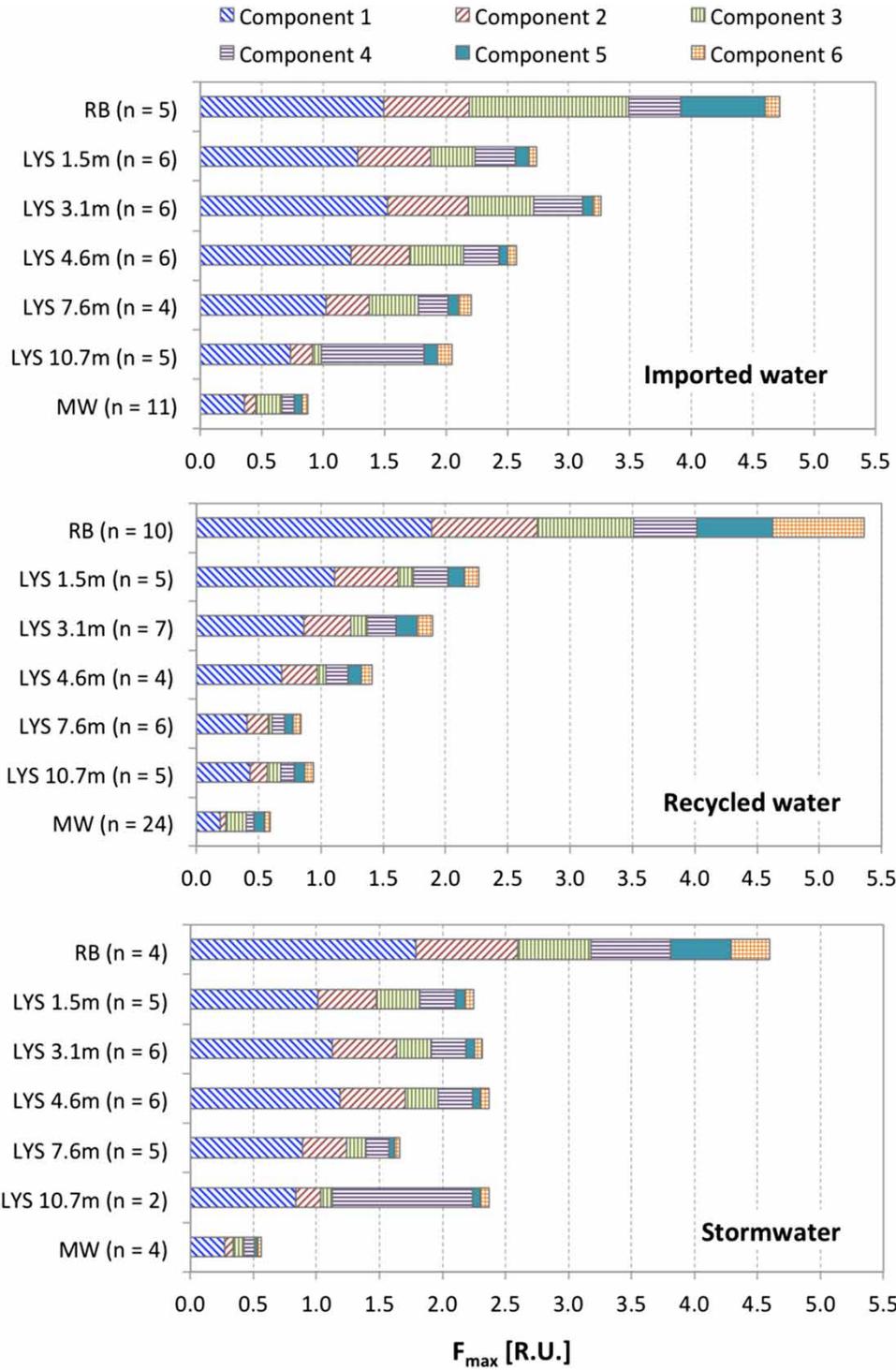


Figure 5 | Maximum fluorescence intensity F_{max} of six components identified by PARAFAC in samples collected from recharge basins (RB), lysimeters at different depths (LYS), and monitoring wells (MW) during imported water, recycled water, and stormwater recharge events, respectively. Sample numbers are provided in parentheses.

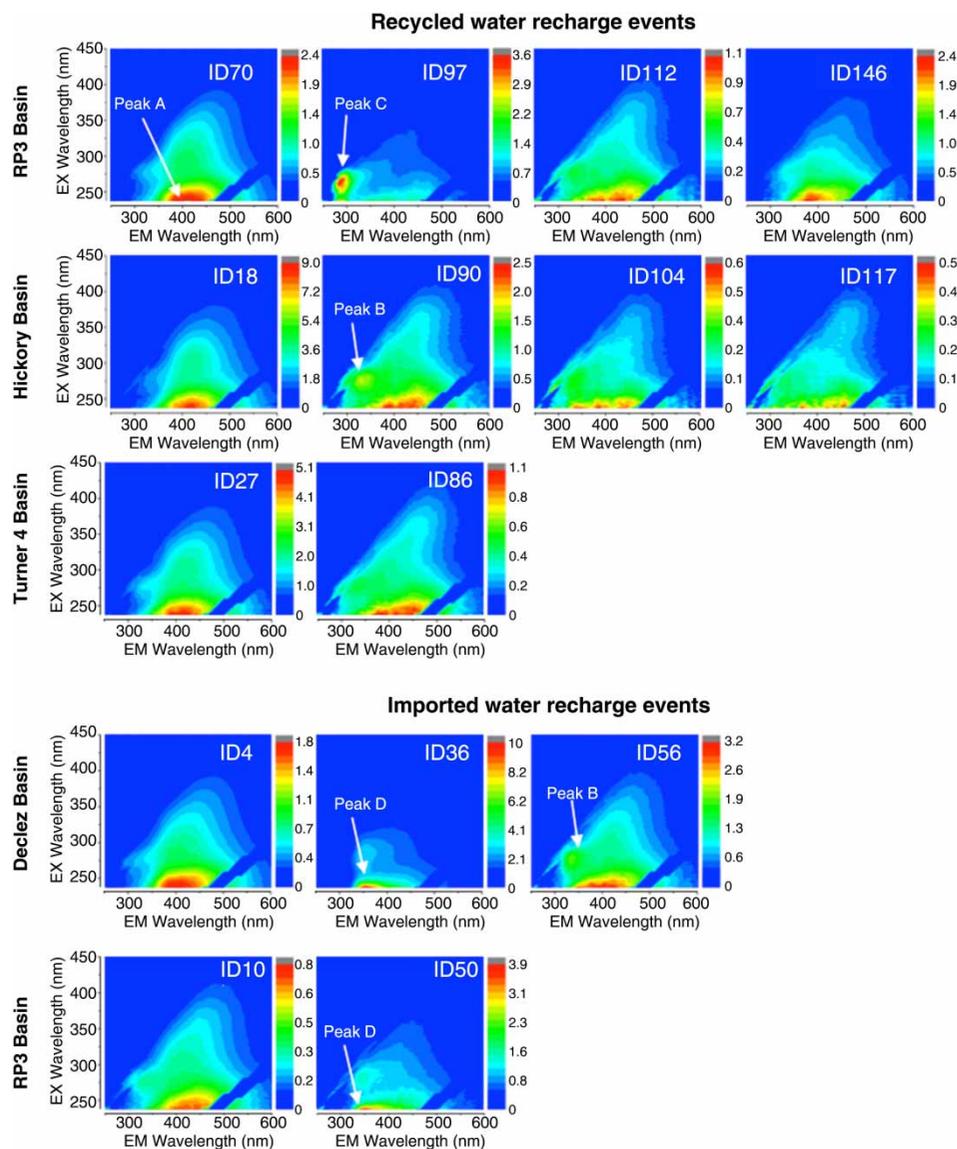


Figure 6 | EEMs of recharge basins RP3, Hickory, Turner 4, and Declez for imported water and recycled water recharge events. Peaks A ($\lambda_{\text{ex/em}}$: 240/400–425 nm), B ($\lambda_{\text{ex/em}}$: 276/335 nm), C ($\lambda_{\text{ex/em}}$: 270/290 nm), and D ($\lambda_{\text{ex/em}}$: <250/335–390 nm) most likely originated from anthropogenic chemicals; peak A was expected to result from the herbicide Habitat Herbicide, peak B from the herbicide DuPont Oust XP, peak C from the pesticide carbofuran, and peak D from the PAH phenanthrene.

Herbicide on 3D-fluorescence spectroscopy is shown for sample ID18 (Hickory basin) and sample ID70 (RP3 basin) during recycled water recharge. Both samples (ID18 and ID70) revealed significantly higher F_{max} values for each component compared to other recharge basins receiving the same type of recycled water (Table S1, Supplementary Information). Furthermore, total fluorescence intensity, DOC concentration, and $UV_{254\text{nm}}$ absorbance were significantly higher. Additionally, sample

ID70 exhibited the same peak (peak A) in the EEM (maxima at $\lambda_{\text{ex/em}}$: 240/400–425 nm) like the Habitat Herbicide. Indeed, the fluorescence spectrum of Habitat Herbicide could overlap with each component identified by PARAFAC analysis.

The contribution of the herbicide DuPont Oust XP to the fluorescence signal (PARAFAC component 5) became obvious in samples ID90 (Hickory basin, recycled water recharge) and ID56 (Declez basin, imported water

recharge), where a clear peak (peak B) was detected in the EEM that was very similar to the peak of DuPont Oust XP (Figure 2).

In sample ID97 (RP3 basin, recycled water recharge) F_{\max} of component 6 was significantly higher (3.38 R.U.). This was consistent with visual investigation of the EEM (Figure 6, peak C), suggesting that the peak in the EEM of sample ID97 was very similar to PARAFAC component 6 and possibly the pesticide carbofuran.

High F_{\max} standard deviations were also identified for component 3 in basin samples during imported water recharge (Declez and RP3). The following imported water recharge samples revealed increased F_{\max} values of component 3 and total fluorescence intensity: samples ID36, ID56 (both Declez basin), and ID50 (RP3 basin). F_{\max} of component 3 in these samples were 9.78, 1.47, and 3.30 R.U., respectively. This was consistent with visual investigation of the EEMs (peak D in Figure 6). However, F_{\max} of component 3 was almost zero in the source water (imported water, Figure 4) but significantly higher in the recharge basin samples (Figure 5), which suggests that component 3 was either already present in the basins or was discharged in some other way into the basin (e.g. surface runoff, dry deposition).

In summary, samples ID18, ID36, ID50, ID56, ID70, and ID97 revealed not only significantly higher F_{\max} values compared to the other samples, but also different fluorescence shapes in the EEMs and/or increased fluorescence intensity, DOC concentrations and UV_{254nm} absorbance.

In addition, some components exhibited a higher abundance at specific lysimeters in the subsurface. For example, all samples of Declez lysimeter 10.7 m (ID9, ID61, and ID69) during imported water recharge revealed a fluorescence shape in the EEM that was very unique and looked very similar to component 4 (Figure 7). The fact that sample ID170 (Declez lysimeter 10.7 m) during stormwater recharge revealed the same unique shape supports the conclusion that this shape in the EEM was not originating from a specific source water type but was rather characteristic for this specific lysimeter. Besides, the F_{\max} rate percent difference between those samples was below 5% for components 1, 2, and 6 and below 15% for components 3 and 4. However, component 5 exhibited a rate percent difference of 33%, as it was not present in sample ID170.

CONCLUSIONS

Monitoring the fate of chromophoric DOM during water recharge with 3D-fluorescence spectroscopy proved to be a sensitive and powerful tool. PARAFAC analysis was applied not only to identify and isolate different chromophoric DOM components, but also to derive semi-quantitative assessments and illustrate the decrease of chromophoric DOM in the subsurface over time and travel distance during groundwater recharge. In establishing a representative source water specific fingerprint as a baseline for MAR applications, some limitations were encountered, for

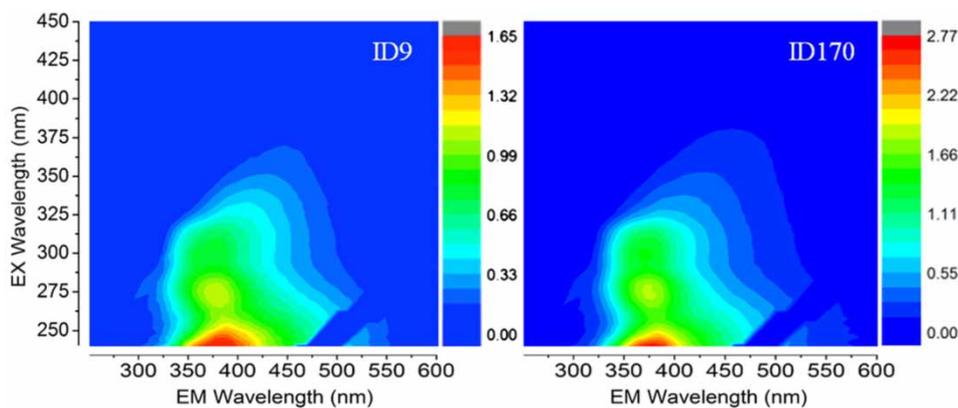


Figure 7 | 3D-EEMs of samples that were collected from lysimeter 10.7 m depth at the Declez basin for imported water (sample ID9) and stormwater (sample ID170) recharge event.

example significant impact on UV and fluorescence signals of source water unspecific anthropogenic chemicals from non-point sources that are introduced to the recharge facilities (i.e. application of weed and vector control in the vicinity, wet and dry deposition at the basins). However, PARAFAC could identify, extract, and factor out at least one herbicide with chromophoric features from surface and groundwater EEMs.

ACKNOWLEDGEMENTS

This study was partially supported by the Metropolitan Water District of Southern California. We also acknowledge co-funding through the National Science Foundation (NSF) Engineering Research Center for Reinventing the Nation's Water Infrastructure (ReNUWI) under cooperate agreement EEC-1028968.

REFERENCES

- Alves, S. R. C., Severino, P. C., Ibbotson, D. P., da Silva, A. Z., Lopes, F. R. A. S., Sáenz, L. A. & Bainy, A. C. D. 2002 Effects of furadan in the brown mussel *Perna perna* and in the mangrove oyster *Crassostrea rhizophorae*. *Mar. Environ. Res.* **54**, 241–245.
- Andersen, C. M. & Bro, R. 2003 Practical aspects of PARAFAC modeling of fluorescence excitation-emission data. *J. Chemom.* **17**, 200–215.
- APHA 2012 *Standard Methods for the Examination of Water and Wastewater*. American Public Health Association, Washington, DC.
- Baghoth, S. A., Sharma, S. K. & Amy, G. L. 2011 Tracking natural organic matter (NOM) in a drinking water treatment plant using fluorescence excitation–emission matrices and PARAFAC. *Water Res.* **45**, 797–809.
- Bro, R. 1997 PARAFAC. Tutorial and applications. *Chemom. Intell. Lab. Syst.* **38**, 149–171.
- Bro, R. 1998 *Multi-way analysis in the food industry: models, algorithms, and applications*. Doctoral Thesis, University of Copenhagen, Denmark.
- Chen, J., LeBoeuf, E. J., Dai, S. & Gu, B. 2003 Fluorescence spectroscopic studies of natural organic matter fractions. *Chemosphere* **50**, 639–647.
- Dahm, K. G., Van Straaten, C. M., Munakata-Marr, J. & Drewes, J. E. 2013 Identifying well contamination through the use of 3-D fluorescence spectroscopy to classify coalbed methane produced water. *Environ. Sci. Technol.* **47**, 649–656.
- Drewes, J. E. 2009 Groundwater replenishment with recycled water–water quality improvements during managed aquifer recharge. *Groundwater* **47**, 502–505.
- Drewes, J. E. & Khan, S. J. 2011 Water reuse for drinking water augmentation. In: *Water Quality & Treatment: A Handbook on Drinking Water* (Edzwald, J.K., ed.). McGraw-Hill Professional, New York, pp. 16.1–16.48.
- Fellman, J. B., Miller, M. P., Cory, R. M., D'Amore, D. V. & White, D. 2009 Characterizing dissolved organic matter using PARAFAC modeling of fluorescence spectroscopy: a comparison of two models. *Environ. Sci. Technol.* **43**, 6228–6234.
- Gilmore, A. M. 2011 *Water quality measurements with HORIBA Jobin Yvon fluorescence instrumentation*. Technical Report No. 38, Horiba Yobin Ivon, Edison, NJ.
- Guo, W., Xu, J., Wang, J., Wen, Y., Zhuo, J. & Yan, Y. 2010 Characterization of dissolved organic matter in urban sewage using excitation emission matrix fluorescence spectroscopy and parallel factor analysis. *J. Environ. Sci.* **22**, 1728–1734.
- Hambly, A. C., Henderson, R. K., Storey, M. V., Baker, A., Stuetz, R. M. & Khan, S. J. 2010 Fluorescence monitoring at a recycled water treatment plant and associated dual distribution system–Implications for cross-connection detection. *Water Res.* **44**, 5323–5333.
- Hao, R., Ren, H., Li, J., Ma, Z., Wan, H., Zheng, X. & Cheng, S. 2012 Use of three-dimensional excitation and emission matrix fluorescence spectroscopy for predicting the disinfection by-product formation potential of reclaimed water. *Water Res.* **46**, 5765–5776.
- Henderson, R. K., Baker, A., Murphy, K. R., Hambly, A., Stuetz, R. M. & Khan, S. J. 2009 Fluorescence as a potential monitoring tool for recycled water systems: a review. *Water Res.* **43**, 863–881.
- Hoppe-Jones, C., Oldham, G. & Drewes, J. E. 2010 Attenuation of total organic carbon and unregulated trace organic chemicals in U.S. riverbank filtration systems. *Water Res.* **44**, 4643–4659.
- Hua, B., Veum, K., Koirala, A., Jones, J., Clevenger, T. & Deng, B. 2007 Fluorescence fingerprints to monitor total trihalomethanes and N-nitrosodimethylamine formation potentials in water. *Environ. Chem. Lett.* **5**, 73–77.
- Hudson, N., Baker, A. & Reynolds, D. 2007 Fluorescence analysis of dissolved organic matter in natural, waste and polluted waters—a review. *River Res. Appl.* **23**, 631–649.
- Jiji, R. D., Cooper, G. A. & Booksh, K. S. 1999 Excitation-emission matrix fluorescence based determination of carbamate pesticides and polycyclic aromatic hydrocarbons. *Anal. Chim. Acta* **397**, 61–72.
- Kompany-Zareh, Y. A. M. 2012 Tucker core consistency for validation of restricted Tucker3 models. *Anal. Chim. Acta* **723**, 18–26.
- Kowalczyk, P., Durako, M. J., Young, H., Kahn, A. E., Cooper, W. J. & Gonsior, M. 2009 Characterization of dissolved organic matter fluorescence in the South Atlantic Bight with use of PARAFAC model: interannual variability. *Mar. Chem.* **113**, 182–196.
- Lawaetz, A. J. & Stedmon, C. A. 2009 Fluorescence intensity calibration using the Raman scatter peak of water. *Appl. Spectrosc.* **63**, 936–940.

- Li, W.-T., Chen, S.-Y., Xu, Z.-X., Li, Y., Shuang, C.-D. & Li, A.-M. 2014 Characterization of dissolved organic matter in municipal wastewater using fluorescence PARAFAC analysis and chromatography multi-excitation/emission scan: a comparative study. *Environ. Sci. Technol.* **48**, 2603–2609.
- Murphy, K. R., Stedmon, C. A., Waite, T. D. & Ruiz, G. M. 2008 Distinguishing between terrestrial and autochthonous organic matter sources in marine environments using fluorescence spectroscopy. *Mar. Chem.* **108**, 40–58.
- Murphy, K. R., Stedmon, C. A., Graeber, D. & Bro, R. 2013 Fluorescence spectroscopy and multi-way techniques. *PARAFAC. Anal. Methods* **5**, 6557–6566.
- Nir, I. 2013 *HORIBA Aqualog User Manual*. Horiba Yobin Ivon, Edison, NJ.
- Parsekian, A. D., Regnery, J., Wing, A. D., Knight, R. & Drewes, J. E. 2014 Geophysical and hydrochemical identification of flow paths with implications for water quality at an ARR site. *Groundwater Monit. Remediat.* **34**, 105–116.
- Prasad, M. B. K., Kumar, A., Datta, D. K. & Ramanathan, L. 2014 Spectrofluorometric analysis of organic matter in the Sundarban mangrove, Bangladesh. *Indian J. Geo-Mar. Sci.* **43**, 999–1006.
- Rauch-Williams, T., Hoppe-Jones, C. & Drewes, J. E. 2010 The role of organic matter in the removal of emerging trace organic chemicals during managed aquifer recharge. *Water Res.* **44**, 449–460.
- Ray, C., Grischek, T., Hubbs, S., Drewes, J., Haas, D. & Darnault, C. 2008 Riverbank filtration for drinking water supply. In: *Encyclopedia of Hydrological Sciences* (Anderson, M. G., ed.). John Wiley & Sons, Ltd., London.
- Singh, S., Inamdar, S. & Scott, D. 2013 Comparison of two PARAFAC models of dissolved organic matter fluorescence for a Mid-Atlantic forested watershed in the USA. *J. Ecosyst.* **2013**, 532–424.
- Stedmon, C. A., Markager, S. & Bro, R. 2003 Tracing dissolved organic matter in aquatic environments using a new approach to fluorescence spectroscopy. *Mar. Chem.* **82**, 239–254.
- Stedmon, C. A., Seredyńska-Sobecka, B., Boe-Hansen, R., Le Tallec, N., Waul, C. K. & Arvin, E. 2011 A potential approach for monitoring drinking water quality from groundwater systems using organic matter fluorescence as an early warning for contamination events. *Water Res.* **45**, 6030–6038.
- Wang, H., Zhang, Y. & Xiao, X. 2010 Quantification of polycyclic aromatic hydrocarbons in water: a comparative study based on three-dimensional excitation-emission matrix fluorescence. *Anal. Sci. Int. J. Jpn. Soc. Anal. Chem.* **26**, 1271–1276.
- Weishaar, J. L., Aiken, G. R., Bergamaschi, B. A., Fram, M. S., Fujii, R. & Mopper, K. 2003 Evaluation of specific ultraviolet absorbance as an indicator of the chemical composition and reactivity of dissolved organic carbon. *Environ. Sci. Technol.* **37**, 4702–4708.

First received 15 August 2015; accepted in revised form 3 October 2015. Available online 20 November 2015