**Fluorescence spectroscopy and parallel factor analysis as a dissolved organic monitoring tool to assess treatment performance in drinking water trains**

**M.Vera1, S.Cruz1, M.R. Boleda1, J.Mesa1, J. Martín-Alonso1, S.Casas3, O. Gibert2, 3, J.L. Cortina2,3**

1Aigües de Barcelona, Empresa Metropolitana de Gestió del Cicle Integral de l’Aigua, Gral. Batet 1-7, 08028, Barcelona, Spain

2Polytechnic University of Catalonia, Department of Chemical Engineering, Av. Diagonal 647, 08028, Barcelona, Spain

3CETAQUA, Water Technology Center, Ctra. d’Esplugues 75, 08940, Cornellà de Llobregat, Spain

\*mverac@aiguesdebarcelona.cat

**Abstract**

Fluorescence excitation emission matrix (FEEM) spectroscopy was used to evaluate its applicability as a tool to track dissolved organic matter (DOM) in a drinking water treatment plant (DWTP) that incorporates a conventional line (consisting in ozonation and GAC filtration) and a membrane-based line (consisting in ultrafiltration, reverse osmosis and mineralization) working in parallel. Seven sampling points within the different process stages were characterized monthly during 2014. A global Parallel Factor Analysis (PARAFAC) was used to pull out underlying organic fractions from the fluorescence spectra. Accordingly a five components model was selected to describe the system and the pros and cons of the model were discussed by analysis of the residuals. Among the five fluorescent components, those associated to humic-like matter (C1, C3 and C4) showed a similar season variability in the river water feeding the DWTP (which resembled that of UV254 and TOC), whereas the two components associated to protein-like matter (C2 and C5) exhibited a different behavior. The maximum fluorescence intensity values (Fmax) were used to quantify DOM removals across the plant. Compared to the conventional line, water from the UF/RO membrane-based line showed between 6 and 14 times lower fluorescence intensity signal for the humic-like components and between 1 and 3 for the protein-like components as compared to the conventional line. The differences in DOM composition due to seasonal variations and along the treatment trains point out the suitability of using fluorescence measurements over other parameters such as UV254 as a monitoring tool to help optimize operation conditions of each treatment stage and improve produced water quality in a DWTP.

**Keywords**

Water treatment; fluorescence spectroscopy; Parallel factor analysis (PARAFAC); reverse osmosis (RO); ultrafiltration (UF); dissolved organic matter (DOM)

1. **Introduction**

The Barcelona Metropolitan Area (BMA) drinking water supply has been primarily based on surface water resources. These water resources (e.g. Llobregat River Basin) have historically suffered the effects of salt mining activities and urban and industrial discharges, which has resulted in a deterioration of the quality of the surface water bodies and associated ecosystems. To still aggravate the situation, and as it is typical in Mediterranean areas, periods of drought happen as a cyclic and recurrent scenario, exacerbating even more the scarcity of fresh water in the Llobregat River Basin [1].

In general, the surface water quality in semiarid regions (e.g. the Llobregat River Basin) is poor and advanced treatments (e.g. membrane technologies) are necessary in drinking water treatment plants (DWTPs) to meet the established drinking water quality standards. Reverse Osmosis (RO), with ultrafiltration (UF) as a pre-treatment step, is identified as a robust and flexible technology to improve water quality and taste by removing undesirable compounds such as ions and organic micro pollutants [2]. RO also removes disinfection by-products such as trihalomethanes (THMs) formed in the chlorination stage by reacting with dissolved organic matter (DOM) and bromide ions as it is the case of the Llobregat River. RO produced water sent to the distribution network contains, thus, a total THM concentrations below the legislation threshold of 100 μg/L.

Characterizing DOM along the treatment process can provide a valuable insight for the selection of the optimum operation conditions. Nowadays, total organic carbon (TOC) and UV absorbance at 254 nm are parameters commonly used in industry to monitor the presence of DOM. The analytical techniques are easy to use, fast and economically affordable compared to other techniques. However, they provide limited information about DOM composition. In order to get further insight, more complex and time consuming procedures are necessary such as chromatographic techniques. One of these techniques is the High Performance Size Exclusion Chromatography (HPSEC), which fractionates DOM by the size and molecular weight distribution of the organic compounds [3]. Other techniques such as liquid chromatography coupled to oxygen/carbon detection and nitrogen/carbon detection (LC-OCD-NCD) provide further information about DOM composition. Huber et al. discretized DOM accordingly in biopolymers, humic substances, building blocks, low molecular weight (LMW) neutrals and LMW acids [4].

An alternative technique increasingly used is fluorescence spectroscopy [5]. By measuring the excitation and emission wavelengths at which a family of components fluoresce, it is possible to represent 3D fluorescence excitation and emission matrices (FEEMs) in the form of level curves or contour maps. In this way, complex DOM can be characterized by its fluorescent components (“fluorophores”). Fluorescence spectroscopy has the potential to be implemented easily in both on-line and in-line configurations, it is less expensive than other techniques such as liquid chromatography, it is a noninvasive technique and it clusters DOM fractions based on their chemical properties [6]. However, the quantification of the total amount of organic matter for each fraction is still not solved as it requires previous calibration and it may be complex in systems with a large number of unknown compounds as it is the case of surface waters. Previous studies have correlated the fluorescence spectra with the chemistry behind by using simultaneous techniques such as HPSEC coupled to UV absorbance and fluorescence (HPSEC-UVA-fluorescence-DOC) [7]. However, Peleato et al. found only a limited correlation between the component fluorescence intensity and the LC-OCD fractions analyzed for different water samples [8].

Advanced data analysis techniques for interpretation of the fluorescence data include from simple methods such as peak-picking to more complex methods such as Principal Component Analysis (PCA), Parallel Factor Analysis (PARAFAC) and Self-Organizing Maps (SOM). Peak picking consists in identifying the maximum intensity of the peaks within the spectra. It is seen as a suitable tool to continuously track the dynamics of DOM within a water treatment process but it can be inaccurate as the fluorescence peaks may shift and overlap with each other [9]. PCA makes use of a statistical procedure to reduce the number of variables but preserving most of the relevant information [10]. PARAFAC is also used in advanced data treatment because it decomposes the fluorescence spectra in a set of unique components [11]. Finally, SOM has gained interest recently as an exploratory analysis based on artificial neural networks (ANN) that identify groups of fluorophores, and it has been applied to monitor changes of water quality in real-time [12,13].

FEEM in increasingly being used to gain insight into the nature of DOM in different aspects of water treatment. For instance, it has been applied as a surrogate to BOD5 in monitoring biodegradable organic matter in surface waters [14]; in monitoring DOM removal in a variety of water treatment units (e.g. coagulation, GAC filtration…) [15] and in predicting the formation of disinfection by-products [16,17]. Also, FEEM spectroscopy coupled to PCA has showed effectiveness in anticipating fouling events in stages incorporating membrane technologies [18,19]. Up to date, the majority of the studies that combine FEEM with either PARAFAC or PCA are restricted to a specific water treatment stage and mostly in wastewater systems. On the other hand, studies conducted to monitor DOM changes in full-scale DWTPs are scarce [20]. In these systems, pre-processing and modelling FEEM data with PARAFAC present higher complexity because FEEM spectra are subject to changes in shape due to seasonal variations and along treatment stages in plant. For instance, Shutova et al. needed seven different PARAFAC models to properly capture the information contained in FEEM data from four different DWTPs [21].

Within this framework, the objective of this study was to monitor the dynamics of DOM by means of FEEM-PARAFAC and assess the treatment performance in a DWTP that incorporates a conventional and a membrane-based treatment trains in parallel. The study presented here will serve as a benchmark to help understand how the organic fractions vary throughout the year and, also, understand the advantages of using a conventional line over a membrane-based line and vice versa. Overall, this study is intended to help implement strategies to improve process efficiencies and water quality by aptly selecting the most appropriate operation conditions.

1. **Materials and Methods**
   1. **Drinking water plant description**

The DWTP located in Sant Joan Despí (Barcelona, Spain) is subjected to high oscillations in water quality in part due to river flow rate fluctuations (3 to 20 m3/s), irregular rainfall events and variable treated wastewater discharges from treatment plants upstream the Llobregat River Basin [22]. The DWTP has a nominal capacity to produce 5.5 m3/s and uptakes water from the lowest part of the Llobregat River basin. Also, it draws water from wells when it is necessary (e.g. when Llobregat water quality is too poor to be treated or when there is overflow by rainfalls). Figure 1 describes the overall process consisting in an initial pre-treatment with di-oxychlorination, coagulation, sedimentation and finally a sand filtration stage. Coming up next, the water is split in two parallel treatment lines: (i) a conventional treatment with an ozonation process and granular activated carbon (GAC) filtration, and (ii) a membrane-based treatment with an ultrafiltration (UF), a reverse osmosis (RO) and, finally, a mineralization step using limestone. Water from both treatment lines is then blended and chlorinated prior to distribution. Since producing water from the membrane-based line is more expensive than through the conventional line (0.2kWh/m3 versus 0.02kWh/m3), the current configuration allows to adjust at any time the flow toward each line so that quality standards in produced water are permanently ensured at an acceptable cost. The different sampling points in which water samples were collected and analyzed through fluorescence spectroscopy are also indicated. There is also additional information about the average coagulant doses in the pre-treatment stage as well as previous to UF. Finally, average pH values across the plant are also indicated.

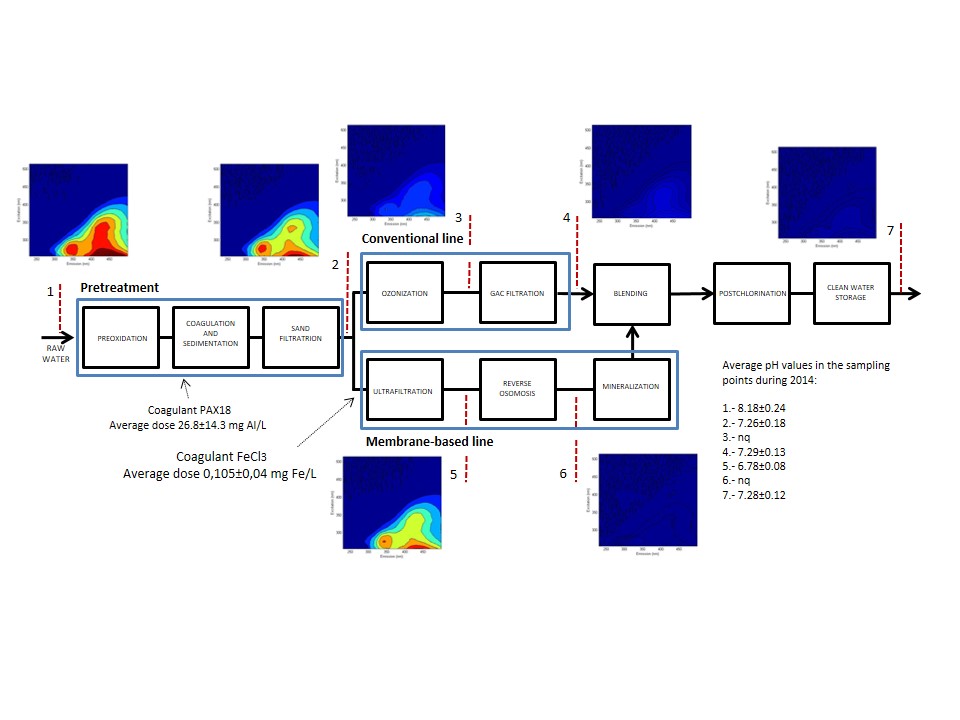


Figure 1.- Schematic description of a DWTP located in Sant Joan Despí. The different sampling points in which water was analyzed through fluorescence spectroscopy are also indicated. The average pH values and the average coagulant doses measured daily during 2014 are also indicated.

**2.2 Sampling procedure and data acquisition**

Water was collected monthly in 100 mL glass bottles from seven different sampling points during 2014 (see Figure 1). Also, 9 additional samples were collected (3 from raw water and an additional from each sampling point) during this period and they were included in the model. For evaluating the dynamics of water quality, samples were collected according to the retention times of each treatment step. They were stored in the fridge at 4ºC and characterized within a maximum of two days after sampling. Within this time frame the stability of DOM was preserved and variations on the measured fluorescence values were below 1%.

All measurements were conducted in the DWTP laboratories with certificates of ISO22000:2005, ISO14001:2004 and OSHAS18001:2007. Further analytical data were acquired using the LIMS database: TOC measurements (Sievers S310), absorbance (Hach DR5000) and turbidity (Hach SS7).

**2.3 Fluorescence measurements**

Fluorescence analyses were done on a Perkin Elmer LS55 System and a 1 cm path length quartz cuvette. The cuvette was rinsed with nitric acid and methanol prior to use and left in the oven for 2 hours. A thermal bath was used to set the temperature at 25ᴼC while analyzing the samples. The equipment was validated with deionized water in which Raman and Rayleigh peaks were measured. Prior to analysis, samples were filtered through 0.45 µm PVDF filters. Spectra were measured at excitation wavelengths of 225-515 nm in 10 nm increments and emission wavelengths of 230-645.5 nm in 0.5 nm increments, using a scan speed of 600 nm min-1, a voltage of 750 V and emission slit widths of 5 nm. An in house made program with MATLAB 2013b was programmed to manage and treat preprocess data. A blank spectra was subtracted from the FEEMs in order to remove most of the Raman scattering peaks. Finally, the FEEM data were normalized by dividing them by the Raman-scatter peaks from the blank to enable comparison within other spectra.

**2.4 PARAFAC modelling**

The initial FEEMs containing 832x30 emission intensity readings were analyzed with PARAFAC to decompose them into a set of trilinear terms and a residual array with the aim to estimate the underlying spectra. In order to perform the analysis, the N-way v.3.00 Toolbox for MATLAB was used. Since some of the regions did not provide additional information, 540 emission intensity readings were used instead of 832. Also, 27 excitation readings were used instead of 30 as the lower excitation region provided instability to the PARAFAC models studied. Samples exhibiting abnormal fluorescence patterns (e.g. peaks with intensity values 10 times higher than averaged or with morphologies clearly not describing fluorophores) were identified as outliers and excluded from the model. An exploratory analysis was performed in which preliminary models from 3 to 8 factors were generated. In the modelling stage, the non-negativity constrain was included in all tests. Finally, regarding the pre-treatment steps, the first and second order Rayleigh diagonals were trimmed. The explained variation, the morphology of the contour plots, the contour plots matching from other studies and the residual spectra were the parameters used to select a proper model. Also, half split analysis was used to validate the results, given a set of spectra. The contribution of each component to the sample spectra was measured with the coordinate of maximum intensity value (Fmax) and used to track changes in DOM along the process. The PARAFAC model can be described by Eq. 1 as:

(1)

Where Xijk provides intensity values at specific coordinate points, aif, bjf, ckf. aif contain information about the estimated relative concentrations of each analyte (factor) within the different samples analyzed whereas bjf and ckf are estimated emission and excitation loadings, respectively providing information about the morphology of the analytes. Finally, the F value defines the number of components in the model and the Eijk accounts for the residual variation not explained by the model [23].

1. **Results and discussion**

**3.1 DOM characterization with FEEM-PARAFAC: from raw to treated water**

A total of 93 EEM samples from the different sampling points were used to obtain a global PARAFAC model. The model contained 5 different components whose contour plots are shown in Figure 2. Table 1 shows detailed information about the model. Only, an outlier sample was identified and excluded. Finally, the dataset was split half validated. The results showed similarities above 95% with 99.9% explained variation.

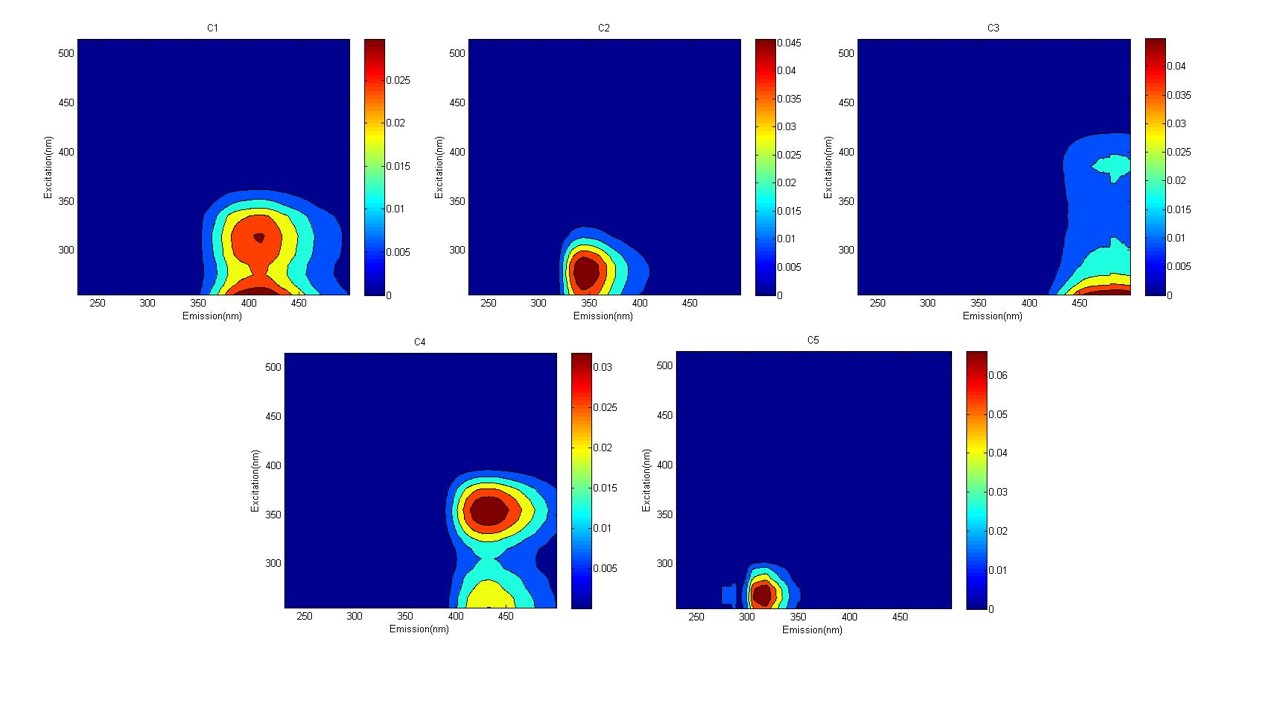


Figure 2.- FEEM contour plots of the five model components (C1, C2, C3, C4 and C5) obtained by using PARAFAC

Table 1.- Detailed parameters from the PARAFAC model

|  |  |
| --- | --- |
| **PARAFAC MODEL** | |
| Nº of components | 5 |
| Nº of samples used | 92 |
| Outlier samples | 1 |
| Explained Variation (%) | >99.9% |
| Split half analysis | 95% similarity |

Component C1 contained two maxima at ex/em 255&315/412 nm. Lu et al demonstrated that both peaks derive from the same emission wavelength derive from the same group of fluorophores [24]. Components C2-C5 exhibited a single maximum at ex/em 275/344 nm, 255/484 nm, 355/432 nm and 265/318 nm, respectively. Chen et al., linked the different regions within the fluorescence spectra with organic fractions of different nature (see Table 2) [25]. Based on this classification, component C1 and C3 fall in the region III associated to fulvic acids, component C2 relates to the presence of microbial by-products (region IV), component C4 falls in the humic acid region and, finally, component C5 is related to aromatic proteins I.

Table 2.- Characterization of different organic fractions in natural waters according to literature [25]

|  |  |  |
| --- | --- | --- |
| **Regions** | **Description** | **Coordinate intervals (ex/em) (nm)** |
| I | Aromatic proteins I | <250/<330 |
| II | Aromatic proteins II | <250/330<em<380 |
| III | Fulvic acids | <250/>380 |
| IV | Microbial by-products | >250/<380 |
| V | Humic acids | >250/>380 |

The Open Fluor database was used to compare the model components found in this study with those found by other studies at the same excitation and emission wavelengths [25]. The database uses the Tucker congruence coefficient to match each component. Table 3 summarizes the results of the comparison. Component C1 and component C2 were best matched with DOM found in the Shark Bay (Australia) [26]. Component C1 was linked to humic-like matter whereas component C2 was related to protein-like matter. Stedmon et al, associated component C2 to almost free tryptophan deriving from terrestrial fluorescent material [27]. Component C3, associated to humic-like, best matched a previous study performed by Shutova et al, who applied a total of seven different PARAFAC models at different treatment stages of a DWTP and identified four components common in all models [21]. Their study support the findings found here as three of their components have been found in the global PARAFAC model obtained. Components C4 and C5 were also identified in a previous work from Yu et al., who related them to humic-like and protein-like (likely tyrosine-like), respectively [29].

Table 3.- Matching source of FEEM-PARAFAC components obtained from using the Open Fluor database [28]

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Component** | **Tucker congruence (TC) coefficient** | **Number of times found (>0.95 TC)** | **Nature of component** | **Best matching source** |
| C1 | 0.989 | 33 | Humic-like (terrestrial origin) | [26] |
| C2 | 0.997 | 11 | Protein-like ( Tryptophan-like) | [26] |
| C3 | 0.984 | 26 | Humic-like | [21] |
| C4 | 0.958 | 1 | Humic-like (terrestrial origin) | [29] |
| C5 | 0.988 | 1 | Protein-like (Tyrosine-like) | [29] |

**3.2 Seasonal variation and covariance of FEEM-PARAFAC components and standard water quality parameters (DOC, UV254) of Llobregat river water**

Understanding the dynamics of the river water quality throughout the year in arid and semiarid regions is complex [30]. Figure 3a and 3b show both the average TOC and UV254 measurements (data were taken daily) from the Llobregat River and, also, show their values from the day in which the samples were taken for the FEEM analysis. In both figures, a similar trend was observed in which TOC remained approximately steady from January to July being February and July the months with the lowest TOC concentrations. From August onwards a considerable increase in TOC concentration was observed, with its maximum in September, after which TOC concentration gradually diminished from month to month until February. Figure 3c shows turbidity measurements; the high variability observed is associated mainly to rainfall events and oscillations in river flux during the year [31]. Error bars were calculated by averaging daily data during one month of operation.

Regarding the fluorescence data, the seasonal variation of Fmax values of the five PARAFAC components showed two different trends: one exhibited by humic-like components (C1, C3 and C4) (Figure 3d) and the other one by protein-like components (C2 and C5) (Figure 3e). For humic-like components, the Fmax value for each component remained steady from January to July, when a sharp increase in the Fmax values was observed with their maxima occurring in September. During October and November the Fmax values decreased until they reached approximately constant values again. The observed changes in temperature did not correlate with the variations observed regarding the organic loads (Figure 3f). Instead, the Fmax increments observed for components C1, C3 and C4 may be rather altered by rainfall events (Figure 3f).

On the other hand, the protein-like components (C2 and C5) showed totally different patterns (Figure 3d), with ups and downs during the year with causes not fully understood at present. Also, an analysis of the potential correlation of the Llobregat River flow values with Fmax values of C2 and C5 components, to analyze the potential dilution effect on component concentrations, provided a low correlation coefficient (r2<0.2).

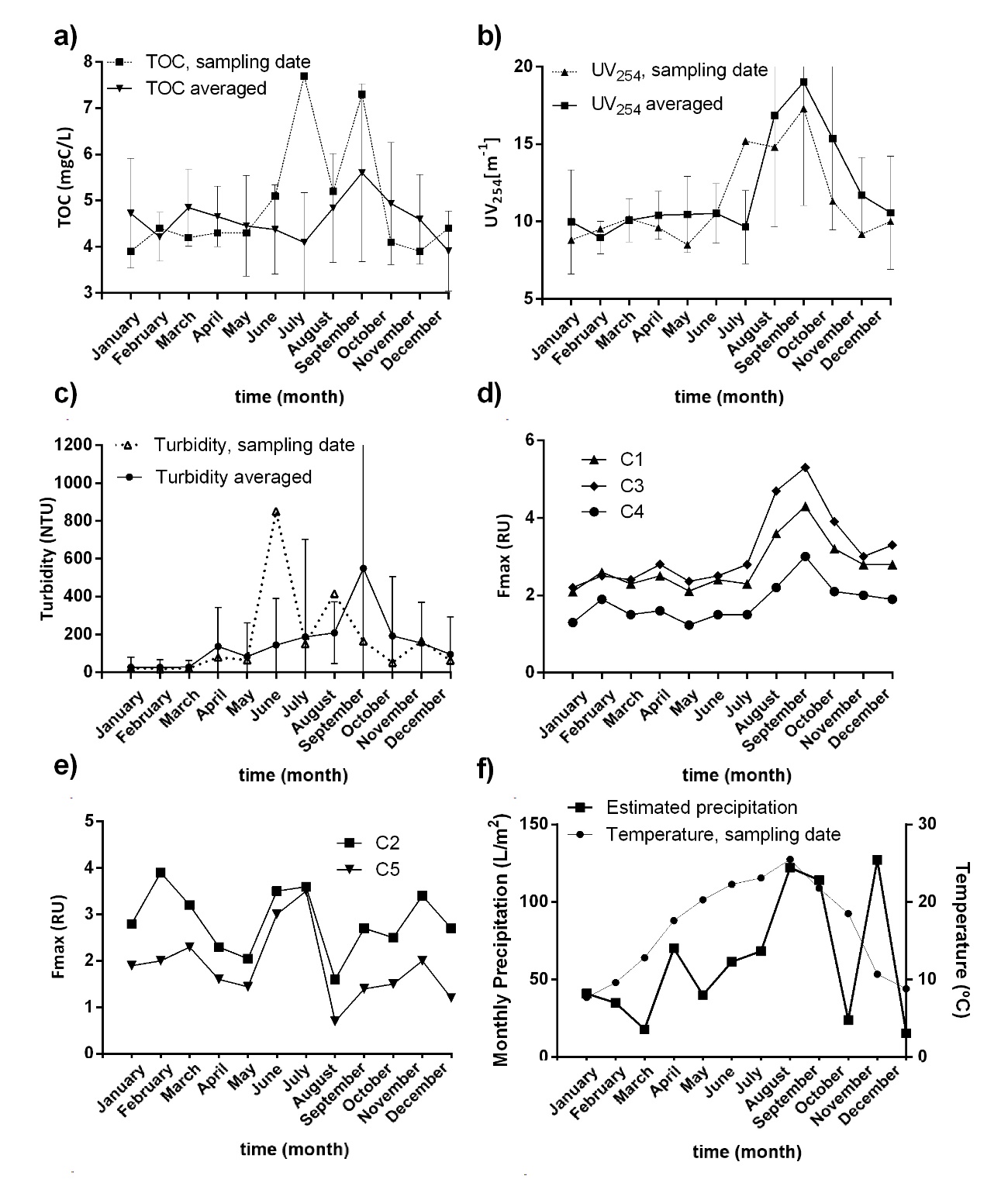


Figure 3.- Llobregat River quality parameters measured during 2014: a) TOC, b) turbidity, c) UV254, d) protein-like component obtained from FEEM-PARAFAC ,e) humic-like components obtained from FEEM-PARAFAC and f) temperature from the river measured the days the sample measurements were conducted (8 am) and precipitation (data acquired from the statistical institute of Catalonia, idescat). The error bars represent the standard deviation from averaged data acquired each day for one month

Table 4 shows the correlations between TOC and UV254 and fluorescence measurements conducted for raw river water. High correlations (r2>0.92) were found within the humic-like components (C1, C3 and C4). Also, the humic-like components showed high correlations with UV254 measurements (r2>0.97) and TOC (r2>0.89). The fluorescence components linked to protein-like (C2 and C5) correlated fairly well with each other (r2>0.77).

Table 4.- Pearson’s correlation matrix of river water quality parameters

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **C1** | **C2** | **C3** | **C4** | **C5** | **UV254** | **TOC** |  |
| 1 | -0.31 | 0.98 | 0.98 | -0.52 | 0.96 | 0.93 | **C1** |
|  | 1 | -0.44 | -0.15 | 0.77 | -0.48 | -0.42 | **C2** |
|  |  | 1 | 0.92 | -0.54 | 0.97 | 0.92 | **C3** |
|  |  |  | 1 | -0.46 | 0.91 | 0.89 | **C4** |
|  |  |  |  | 1 | -0.63 | -0.64 | **C5** |
|  |  |  |  |  | 1 | 0.97 | **UV254** |
|  |  |  |  |  |  | 1 | **DOC** |

**3.3 Monitoring DOM changes along treatment trains by using FEEM measurements**

Figure 4a shows averaged Fmax values from each fluorescence component along the water treatment process. The Fmax values do not necessarily mean higher concentrations as the fluorescence intensity depends, among other parameters, on the molar absorption coefficient. In the ideal case in which fluorescence conforms to Beer’s Law the variations observed are linear and can be linked to DOM removal [32]. However, the signal could also be partially affected by other factors such as chemical changes, quenching phenomena, complexation of metallic coagulants used (Fe(III) and Al(III) salts) with DOM, and/or changes in temperature and pH [33-35]. Also, Figure 4b shows variations in Fmax values expressed as percentage of fluorescence signal reduction. In the sand filtration, the fluorescence signal intensity was reduced by 21±3%, 37±6% and 22±5% for C1, C3 and C4 (humic-like compounds), respectively. Regarding the protein-like components, little variations were observed for component C2 (4±6%) and the signal reduction for component C5 was 20±11%. These findings are consistent with those reported by Baghoth et al. who stated fluorescence intensity reductions ranging between 5-50% during coagulation and below 10% reductions after sand filtration (being tryptophan the least component removed as well) [22].

In the conventional line, even though ozonation was not expected to result in any DOC decrease, it did result in a considerable decrease of the fluorescence signal for all components. Fmax values for components C1-C5 were reduced by 66±16%, 77±16%, 58±16%, 72±16% and 59±11%, respectively. The reduction was attributed to the oxidation of the different compounds, which oxidize them down into smaller molecular size compounds, affecting the fluorophores. Also, GAC filtration showed effectiveness reducing Fmax for all five components: the signal reduction for components C1-C5 was 53±15%, 70±17%, 64±11%, 65±13% and 66±18%, respectively. Unlike in the pre-treatment process, the results showed higher preferences to reduce component C2. Matilainen et al. showed that GAC filtration provided higher affinities to adsorb compounds with low molecular weight although they found there was an optimal range [36].

Moving to the membrane-based line, the UF stage hardly rejected any DOM as it acts mostly to reduce residual Al(III) by micro-coagulation with Fe(III) salts. Humbert et al, showed that coagulation and UF had little effect on the rejection of medium and low molecular weight compounds [37]. On the contrary, the RO stage was responsible for practically all DOM rejection, showing rejections of 98±1%, 99±1% and 99±1% for components C1, C3 and C4 and rejections of 98±2% and 91±11% for components C2 and C5, respectively. Component C5 had lower rejection rates than the rest of the components and, also, showed higher variability.

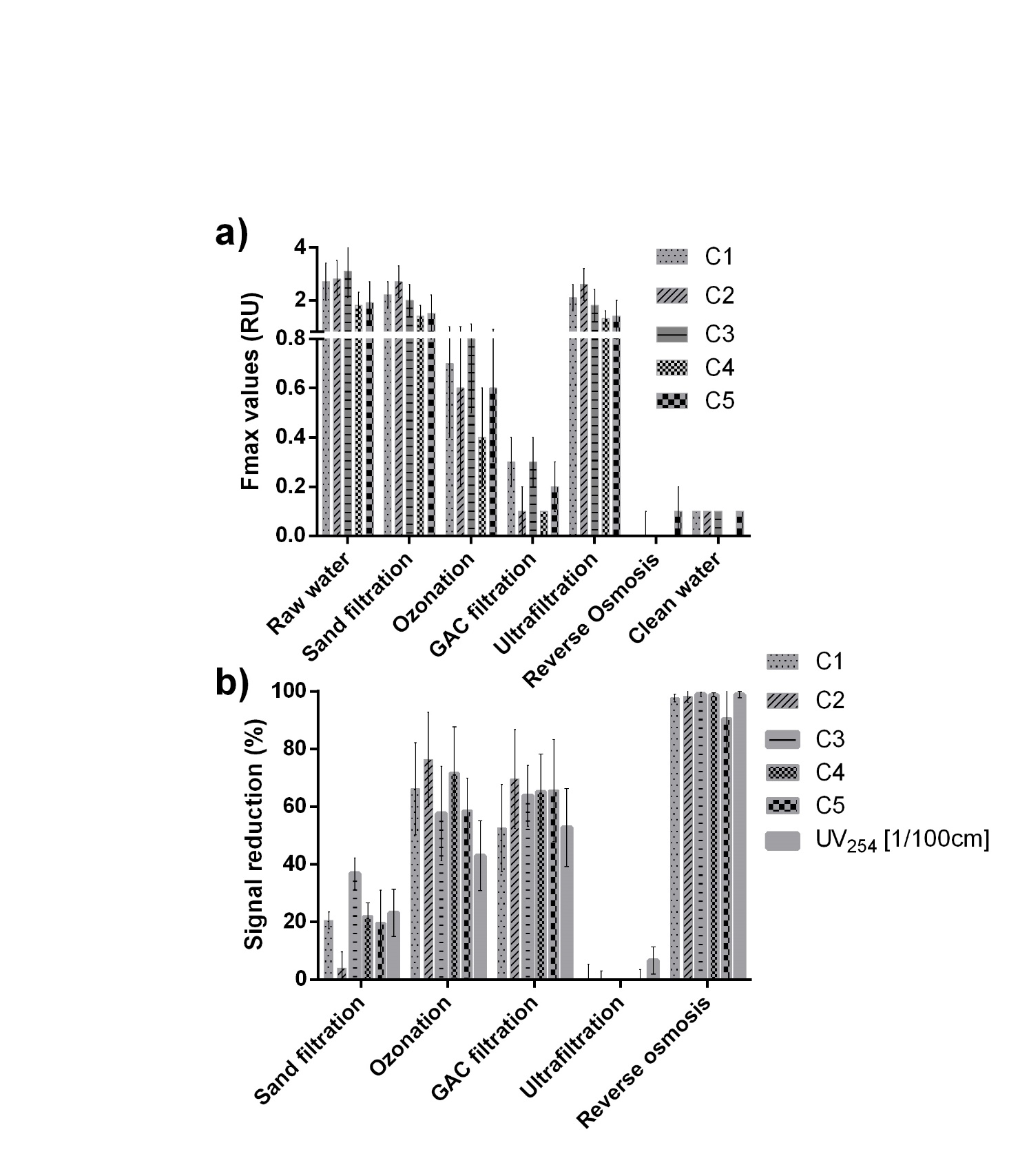


Figure 4.- a) Fmax values and b) Fluorescence signal reduction for all five fluorescence components measured during 2014 along the water treatment processes. The error bars represent the standard deviation from data acquired monthly for one year

Figure 5 exhibits the overall performance of both the conventional and the membrane-based lines. Regarding the humic-like components, component C1 showed the lower preference to be removed among the three (C1, C3 and C4). The differences are less visible in the membrane-based line than in the conventional line. Regarding the protein-like components, component C2 showed the highest rejection ratios among all five components in the conventional line (95±3%) whereas in the membrane-based line its rejection was 98±2%. Furthermore, component C5 showed the lowest rejection ratios in both the conventional and the membrane-based lines, with the singularity of showing over four times higher variability than the other components in the RO: this could indicate that the rejection of component C5 is subjected to the operation conditions of the treatment units. Also, it is possible that component C5 is related to species of lower apparent molecular mass, which may explain its greater capability to permeate through the polyamide RO membrane [38]. Overall, the components treated in the membrane-based line emitted between 6 and 14 times lower fluorescence signal for the humic-like components and between 1 and 3 for the protein-like components than in the conventional line.

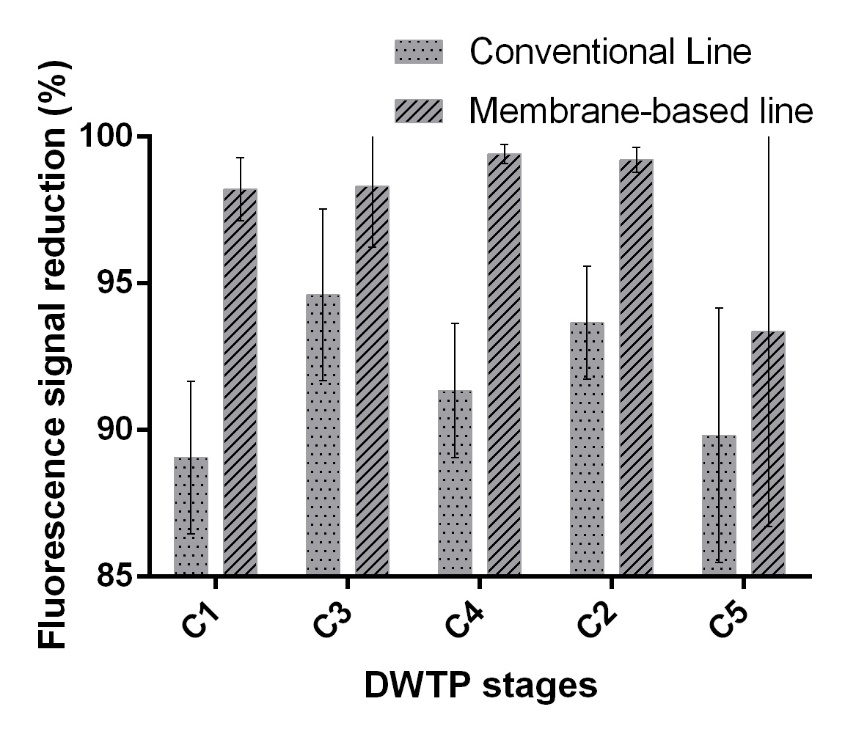


Figure 5.- Comparison of fluorescence signal reduction between the conventional and the membrane-base line for all five fluorescence components measured during 2014

**3.4 Spectral evaluation of FEEM residuals**

The information provided by the spectra not explained by FEEM-PARAFAC was analyzed separately as it can provide further information regarding DOM [39]. For instance, Peleato and Andrews compared PARAFAC, PCA and peak picking and found that PARAFAC skipped to identify a protein-like compound responsible for trihalomethane formation [40]. The residual FEEM data in each treatment stage was averaged to visualize the morphology of the fluorescence contour plots. Figure 6 shows the residual FEEM data for all seven sampling points within the DWTP. There is a region located in the interval of 250-260/400-450 (ex/em) exhibiting a pattern related to the presence of a specific compound. However, the signal intensity only represents 7% of the signal emitted by component C1 (located in the same region). The residual plots indicate that the PARAFAC model was successful in capturing the information provided with the FEEM data.

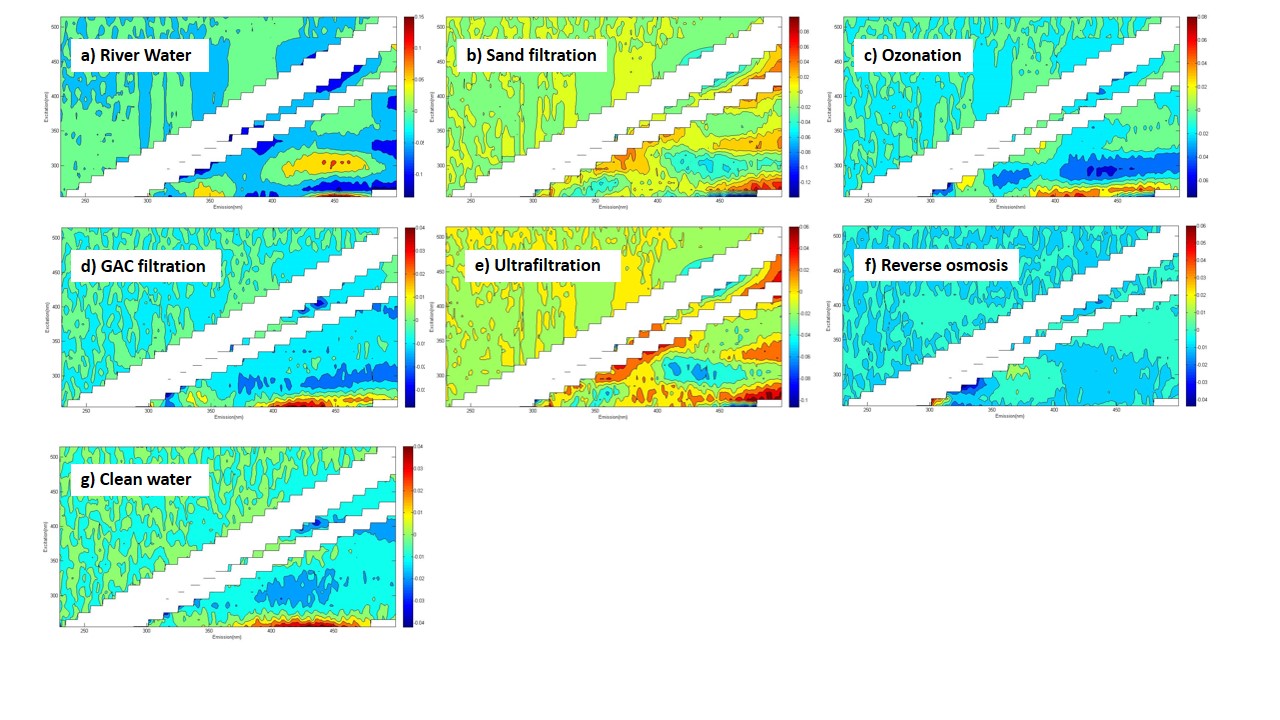
****

Figure 6.- Averaged EEM data from the information not captured with PARAFAC a) River water samples, b) Sand filtered samples, c) samples from ozonation, d) GAC filtration, e) Ultrafiltration, f) Reverse Osmosis and g) clean water

1. **Discussion**

The different observed trends regarding the organic loads in plant during the year showed that not only the amount of DOM (measured with TOC and UV254 analysis) varies but also its composition (measured with FEEM-PARAFAC analysis).

The acquired data can be valuable in managing and improving the plant efficiency in real-time as the organic fractions have a direct impact in process performance (providing thus a benefit from an operation point of view). Besides, the obtained information can also be beneficial in the removal of disinfection by-product precursors (providing thus a benefit from a water quality point of view) by targeting specific fractions as it was stated earlier in the introduction.

The measurements conducted across the DWTP also showed selectivity of the different treatment stages toward the different organic fractions. According to the results obtained, it would be recommended to prioritize the membrane-based line over the conventional line when removing humic-like matter (i.e. August, September and October). On the other hand, it would be better to avoid using the membrane-based line during periods of high protein-like matter intake as it has been reported to affect membrane performance due to fouling (i.e. February, June and July) [41].

The presented study will serve as a benchmark for future experiments that will further evaluate the appropriateness of using fluorescence spectroscopy in real-time process control.

1. **Conclusions**

In this study, fluorescence spectroscopy was used to track the dynamics of DOM and assess treatment performance in a complex DWTP composed by a conventional and a membrane-based treatment lines working in parallel.

* A global FEEM-PARAFAC model of five components was best fitted to describe the EEM data acquired from seven different strategic sampling points from a DWTP treating blends of surface and groundwater and located in a semi-arid region.
* The fluorescence signal of humic-like components (C1, C3 and C4) showed similar seasonal variations along the year and correlated well with DOC and UV254 measurements. On the other hand, the fluorescence signal of protein like components C2 and C5 showed covariance with each other. However, their variations throughout the year need yet to be deeply analyzed using complementary analytical techniques.
* All components were partially removed during the pre-treatment stage although the percentage of removal for component C2 was below 4%.
* In the conventional treatment line, ozonation reduced the fluorescence intensity values for all five components. After GAC filtration, the removal of all five components ranged between 89-95%.
* Regarding the membrane-based treatment line, the UF stage did not show any rejection within the different fluorescence components. However, the RO showed rejections ranging from 93% to almost 100%.
* Regarding the protein-like fractions, component C5 was not so effectively rejected across the treatment stages as compared to component C2. In fact, component C5 showed lower Fmax values than component C2 in initial stages and greater Fmax values along the subsequent treatment stages. Also, the rejection rates for component C5 in the RO could be subjected to operation because Fmax data from permeate showed high variations as compared to the other components.
* Globally, the Fmax values from the humic-like components were between 6-14 times lower in the membrane-based treatment line than in the conventional treatment line and only between 1-3 times lower regarding the protein-like components.
* The average residual contour plots showed little fluorescence signal left which supported to validate the model employed here.
* The different trends observed within the organic loads during the year (time-lapse) as well as the selectivity observed within the treatment processes (position-lapse) coupled to the ability to alternate different treatments demonstrate the suitability of using fluorescence to manage and control operation and improve the plant efficiency in terms of quality and production.

1. **Acknowledgements**

This work was conducted under the financial support of the Agència de Gestió d’Ajuts Universitaris i de Recerca (AGAUR) and Aigües de Barcelona (AB) through the Industrial PhD Program supporting M. Vera (2013-DI-026). Authors would like to thank our partners from the laboratory for their assistance during the sampling episodes at the Sant Joan Despí DWTP. The authors are also grateful to three anonymous reviewers for their comments that greatly improved the manuscript.

1. **References**

[1] R. F. Bangash, A. Passuello, M. Sanchez-Canales, M. Terrado, A. López, F. J. Elorza, G. Ziv, V. Acuña, and M. Schuhmacher, “Ecosystem services in Mediterranean river basin: Climate change impact on water provisioning and erosion control,” Sci. Total Environ., vol. 458–460, pp. 246–255, 2013.

[2] C. Martinez, V. Gomez, D. Dubert, K. Majamaa, E. Pocurull, and F. Borrull, “Study of reverse osmosis treatment for micropollutants rejection in advanced water reuse applications,” Desalin. Water Treat., vol. 55, no. May 2014, p. Ahead of Print, 2014.

[3] F. C. Wu, R. D. Evans, P. J. Dillon, and Y. R. Cai, “Rapid quantification of humic and fulvic acids by HPLC in natural waters,” Appl. Geochem, vol. 22, no. 8 SPEC. ISS., pp. 1598–1605, 2007.

[4] S. A. Huber, A. Balz, M. Abert, and W. Pronk, “Characterisation of aquatic humic and non-humic matter with size-exclusion chromatography - organic carbon detection - organic nitrogen detection (LC-OCD-OND),” Water Res., vol. 45, no. 2, pp. 879–885, 2011.

[5] E. E. Lavonen, D. N. Kothawala, L. J. Tranvik, M. Gonsior, P. Schmitt-Kopplin, and S. J. Köhler, “Tracking changes in the optical properties and molecular composition of dissolved organic matter during drinking water production,” Water Res., vol. 85, pp. 286–294, 2015.

[6] J. P. Ritson, N. J. D. Graham, M. R. Templeton, J. M. Clark, R. Gough, and C. Freeman, “The impact of climate change on the treatability of dissolved organic matter (DOM) in upland water supplies: A UK perspective,” Sci. Total Environ., vol. 473–474, pp. 714–730, 2014.

[7] N. Her, G. Amy, D. McKnight, J. Sohn, and Y. Yoon, “Characterization of DOM as a function of MW by fluorescence EEM and HPLC-SEC using UVA, DOC, and fluorescence detection.,” Water Res., vol. 37, no. 17, pp. 4295–303, Oct. 2003.

[8] N. M. Peleato, M. McKie, L. Taylor-Edmonds, S. A. Andrews, R. L. Legge, and R. C. Andrews, “Fluorescence spectroscopy for monitoring reduction of natural organic matter and halogenated furanone precursors by biofiltration,” Chemosphere, vol. 153, pp. 155–161, 2016.

[9] J. E. Birdwell and A. S. Engel, “Characterization of dissolved organic matter in cave and spring waters using UV–Vis absorbance and fluorescence spectroscopy,” Org. Geochem., vol. 41, no. 3, pp. 270–280, Mar. 2010.

[10] R. K. Henderson, a Baker, K. R. Murphy, a Hambly, R. M. Stuetz, and S. J. Khan, “Fluorescence as a potential monitoring tool for recycled water systems: a review.,” Water Res., vol. 43, no. 4, pp. 863–81, Mar. 2009.

[11] R. Bro and M. Vidal, “EEMizer: Automated modeling of fluorescence EEM data,” Chemom. Intell. Lab. Syst., vol. 106, no. 1, pp. 86–92, Mar. 2011.

[12] M. Bieroza, A. Baker, and J. Bridgeman, “Exploratory analysis of excitation–emission matrix fluorescence spectra with self-organizing maps—A tutorial,” Educ. Chem. Eng., vol. 7, no. 1, pp. e22–e31, Jan. 2012.

[13] E. M. Carstea, A. Baker, M. Bieroza, and D. Reynolds, “Continuous fluorescence excitation-emission matrix monitoring of river organic matter.,” Water Res., vol. 44, no. 18, pp. 5356–66, Oct. 2010.

[14] N. Hudson, A. Baker, D. Ward, D. M. Reynolds, C. Brunsdon, C. Carliell-Marquet, and S. Browning, “Can fluorescence spectrometry be used as a surrogate for the Biochemical Oxygen Demand (BOD) test in water quality assessment? An example from South West England.,” Sci. Total Environ., vol. 391, no. 1, pp. 149–58, Feb. 2008.

[15] J.-M. Siéliéchi, B. S. Lartiges, G. J. Kayem, S. Hupont, C. Frochot, J. Thieme, J. Ghanbaja, J. B. d’Espinose de la Caillerie, O. Barrès, R. Kamga, P. Levitz, and L. J. Michot, “Changes in humic acid conformation during coagulation with ferric chloride: implications for drinking water treatment.,” Water Res., vol. 42, no. 8–9, pp. 2111–23, Apr. 2008.

[16] A. D. Pifer and J. L. Fairey, “Improving on SUVA 254 using fluorescence-PARAFAC analysis and asymmetric flow-field flow fractionation for assessing disinfection byproduct formation and control,” Water Res., vol. 46, no. 9, pp. 2927–2936, 2012.

[17] J. Hur, B. M. Lee, S. Lee, and J. K. Shin, “Characterization of chromophoric dissolved organic matter and trihalomethane formation potential in a recently constructed reservoir and the surrounding areas - Impoundment effects,” J. Hydrol., vol. 515, pp. 71–80, 2014.

[18] R. H. Peiris, M. Jaklewicz, H. Budman, R. L. Legge, and C. Moresoli, “Assessing the role of feed water constituents in irreversible membrane fouling of pilot-scale ultrafiltration drinking water treatment systems.,” Water Res., vol. 47, no. 10, pp. 3364–74, Jun. 2013.

[19] R. H. Peiris, C. Hallé, H. Budman, C. Moresoli, S. Peldszus, P. M. Huck, and R. L. Legge, “Identifying fouling events in a membrane-based drinking water treatment process using principal component analysis of fluorescence excitation-emission matrices.,” Water Res., vol. 44, no. 1, pp. 185–94, Jan. 2010.

[20] S. A. Baghoth, S. K. Sharma, and G. L. Amy, “Tracking natural organic matter (NOM) in a drinking water treatment plant using fluorescence excitation-emission matrices and PARAFAC.,” Water Res., vol. 45, no. 2, pp. 797–809, Jan. 2011.

[21] Y. Shutova, A. Baker, J. Bridgeman, and R. K. Henderson, “Spectroscopic characterization of dissolved organic matter changes in drinking water treatment: From PARAFAC analysis to online monitoring wavelengths,” Water Res., vol. 54, pp. 159–169, 2014.

[22] A. Momblanch, J. Paredes-Arquiola, A. Munné, A. Manzano, J. Arnau, and J. Andreu, “Managing water quality under drought conditions in the Llobregat River Basin,” Sci. Total Environ., vol. 503–504, pp. 300–318, 2015.

[23] C. M. Andersen and R. Bro, “Review Practical aspects of PARAFAC modeling of Fluorescence excitation-emission data,” pp. 200–215, 2003.

[24] F. Lu, C. H. Chang, D. J. Lee, P. J. He, L. M. Shao, and A. Su, “Dissolved organic matter with multi-peak fluorophores in landfill leachate,” Chemosphere, vol. 74, no. 4, pp. 575–582, 2009.

[25] W. Chen, P. Westerhoff, J. a Leenheer, and K. Booksh, “Fluorescence excitation - Emission matrix regional integration to quantify spectra for dissolved organic matter,” Environ. Sci. Technol., vol. 37, no. 24, pp. 5701–5710, 2003.

[26] K. M. Cawley, Y. Ding, J. Fourqurean, and R. Jaff́, “Characterizing the sources and fate of dissolved organic matter in Shark Bay, Australia: A preliminary study using optical properties and stable carbon isotopes,” Mar. Freshw. Res., vol. 63, no. 11, pp. 1098–1107, 2012.

[27] C. A. Stedmon and S. Markager, “Resolving the variability of dissolved organic matter fluorescence in a temperate estuary and its catchment using PARAFAC analysis,” Limnol. Oceanogr., vol. 50, no. 2, pp. 686–697, 2005.

[28] K. R. Murphy, C. A. Stedmon, P. Wenig, and R. Bro, “OpenFluor- an online spectral library of auto-fluorescence by organic compounds in the environment,” Anal. Methods, vol. 6, no. 3, pp. 658–661, 2014.

[29] H. Yu, H. Liang, F. Qu, Z. Han, S. Shao, H. Chang, and G. Li, “Impact of dataset diversity on accuracy and sensitivity of parallel factor analysis model of dissolved organic matter fluorescence excitation-emission matrix.,” Sci. Rep., vol. 5, no. October 2014, p. 10207, 2015.

[30] S. Huang, N. Voutchkov, and S. C. Jiang, “Investigation of environmental influences on membrane biofouling in a Southern California desalination pilot plant,” Desalination, vol. 319, pp. 1–9, 2013.

[31] A. M. Hurst, M. J. Edwards, M. Chipps, B. Jefferson, and S. A. Parsons, “The impact of rainstorm events on coagulation and clarifier performance in potable water treatment,” Sci. Total Environ., vol. 321, no. 1–3, pp. 219–230, 2004.

[32] R. Kathleen, A. Methods, and D. Version, “Analytical Methods,” 2013.

[33] E. M. Carstea, A. Baker, M. Bieroza, D. M. Reynolds, and J. Bridgeman, “Characterization of dissolved organic matter fluorescence properties by PARAFAC analysis and thermal quenching,” Water Res., vol. 61, no. 0, pp. 152–161, 2014.

[34] N. Patel-sorrentino, S. Mounier, and J. Y. Benaim, “Excitation-emission fluorescence matrix to study pH influence on organic matter fluorescence in the Amazon basin rivers,” Water Res., vol. 36, pp. 2371–2581, 2002.

[35] D. M. Reynolds and S. R. Ahmad, “The effect of metal ions on the fluorescence of sewage wastewater,” Water Res., vol. 29, no. 9, pp. 2214–2216, 1995.

[36] A. Matilainen, N. Vieno, and T. Tuhkanen, “Efficiency of the activated carbon filtration in the natural organic matter removal,” Environ. Int., vol. 32, no. 3, pp. 324–331, 2006.

[37] H. Humbert, H. Gallard, V. Jacquemet, and J. P. Croué, “Combination of coagulation and ion exchange for the reduction of UF fouling properties of a high DOC content surface water,” Water Res., vol. 41, no. 17, pp. 3803–3811, 2007.

[38] W. T. Li, Z. X. Xu, A. M. Li, W. Wu, Q. Zhou, and J. N. Wang, “HPLC/HPSEC-FLD with multi-excitation/emission scan for EEM interpretation and dissolved organic matter analysis,” Water Res., vol. 47, no. 3, pp. 1246–1256, 2013.

[39] K. Fish, A. M. Osborn, and J. Boxall, “Characterizing and Understanding the Impact of Microbial Biofilms and the Extracellular Polymeric Substance (EPS) Matrix in Drinking Water Distribution Systems,” Environ. Sci. Water Res. Technol., vol. 2, pp. 614–630, 2016.

[40] N. M. Peleato and R. C. Andrews, “Comparison of three-dimensional fluorescence analysis methods for predicting formation of trihalomethanes and haloacetic acids,” J. Environ. Sci. (China), vol. 27, no. C, pp. 159–167, 2015.

[41] S. Shao, H. Liang, F. Qu, H. Yu, K. Li, and G. Li, “Fluorescent natural organic matter fractions responsible for ultrafiltration membrane fouling: Identification by adsorption pretreatment coupled with parallel factor analysis of excitation-emission matrices,” J. Memb. Sci., vol. 464, pp. 33–42, 2014.