https://doi.org/10.5194/essd-2025-375
Preprint. Discussion started: 4 March 2026

Open Access Earth System Science Data Discussions

Characterization of dissolved organic matter in surface water and groundwater: a dataset for the Seine River basin (France)

Fulvia Baratelli[1], Edith Parlanti[2], Josette Garnier[1], Gilles Varrault[3], Angélique Goffin[3], Nadège Musabimana[3], Sabrina Guérin Rechdaoui[4], Vincent Rocher[4], Nicolas Flipo[5]

[1]Sorbonne Université, CNRS, EPHE, UMR Metis 7619, F-75005 Paris, France
[2]Univ. Bordeaux, CNRS, Bordeaux INP, EPOC, UMR 5805, F-33600 Pessac, France
[3]LEESU, Univ. Paris Est Creteil, ENPC, Institut Polytechnique de Paris, Creteil, France
[4]SIAAP, Direction Innovation, Colombes, France
[5]Mines ParisTech, PSL University, Centre for Geosciences and Geoengineering, 77300 Fontainebleau, France

Correspondence to: Fulvia Baratelli (fulvia.baratelli@minesparis.psl.eu)

Abstract. Carbon fluxes in river networks represent a major component of the carbon cycle, but they are difficult to estimate at large scale. In particular, the physicochemical properties of organic matter (OM) and its contribution to river carbon fluxes remain poorly understood. In this context, this paper presents a dataset for quantifying and characterizing dissolved organic carbon (DOC) in a regional river basin, the Seine River basin in France (76 000 km^2), which is subject to numerous human pressures. The dataset is the result of several sampling campaigns conducted over a 14-year period (2011–2024). A total of 1047 samples were collected from various water types (surface water, groundwater and treated effluents from wastewater treatment plants, WWTPs) at sites across the basin encompassing diverse land uses. Dissolved organic matter (DOM) was characterized both quantitatively, by measuring the concentration of dissolved organic carbon (DOC), and qualitatively, through analysis of its optical properties using UV-Visible absorbance and excitation–emission matrix (EEM) fluorescence spectroscopy. Additionally, using the 45-day incubation method, the biodegradable fraction of DOC, which plays a particularly significant role in river water quality, was estimated for 27 % of the samples. The content and properties of OM in the Seine River basin vary significantly depending on the site, measurement period, and type of water sampled. DOC concentrations are generally higher in WWTP discharges and gravel pits. The biodegradable fraction of DOC is higher in samples from WWTP discharges and from groundwater in a forested alluvial plain. OM in the basin generally displays a low level of aromaticity but is hydrophilic and characterized by strong biological activity. This biological activity is particularly pronounced in the treated effluents

https://doi.org/10.5194/essd-2025-375
Preprint. Discussion started: 4 March 2026

from small-capacity WWTPs, where OM is mainly of microbial origin, and in gravel pits, where OM consists primarily of protein-type compounds. Groundwater typically contains a mixture of OM of both terrestrial and microbial/biological origins. These data may be useful for future studies of (i) the organic carbon cycle at the regional basin scale, (ii) the characteristics of OM across different compartments of a hydrosystem, and (iii) river metabolism. The dataset is available at https://data.indores.fr/privateurl.xhtml?token=a6b58980-3280-4a1b-9311-a0b401955e75.

1 Introduction

Land-to-ocean transport of carbon through the land-to-ocean aquatic continuum (LOAC; Regnier et al., 2022) is increasingly recognized as a major component of the carbon cycle at both regional and global scales (Battin et al., 2023, 2009). These transfers of carbon occur by means of physical, chemical, and biological processes both within the LOAC, which includes inland waters, estuaries, tidal wetlands, and continental shelf waters, and between the LOAC and the other Earth compartments (atmosphere, biosphere, lithosphere).

These aquatic carbon fluxes are difficult to quantify at large scale and they are often overlooked or only partially considered in global carbon budget assessments and land-surface models. The global carbon budget (Friedlingstein et al., 2023) does not include the metabolism of surface waters (rivers, lakes, and reservoirs) and estuarine waters, which contribute an additional 0.75 GtC yr^{-1} to the atmosphere (Cole et al., 2007). Regarding terrestrial carbon, land-surface models that omit land-to-ocean lateral carbon transfer may lead to an overestimation of 4.5 % in the simulated annual net terrestrial carbon uptake across Europe (Zhang et al., 2022). Further, according to Regnier et al. (2022), ignoring the LOAC carbon fluxes results in an overestimation of carbon storage in terrestrial ecosystems by 0.6 ± 0.4 Pg of carbon per year and an underestimation of sedimentary and oceanic carbon storage. More data are needed to reduce uncertainties in estimating lateral inorganic and organic carbon fluxes and their spatiotemporal variations (Yan et al., 2023; Zhang et al., 2022).

https://doi.org/10.5194/essd-2025-375
Preprint. Discussion started: 4 March 2026

It is thus crucial to improve our understanding of the mechanisms associated with the carbon fluxes in the LOAC and to take them into account when evaluating the global carbon balance and its response to climatic and anthropogenic stresses.

Within the LOAC, river networks in particular can be considered biogeochemical reactors (Battin et al., 2023), where several processes and reactions control carbon fluxes. Carbon in riverine environments is present in different forms: greenhouse gases (CO_2 and CH_4); inorganic carbon from soils and from rock alterations; and organic carbon from organic matter (OM), mainly from terrestrial, riparian, or aquatic vegetation. According to estimates that take into account only permanent waters and in some cases also peripheral flood zones, between 2 and 5 Pg of carbon transits annually through the world's rivers (Battin et al., 2023; Cole et al., 2007; Drake et al., 2018; Raymond et al., 2013). The “active pipe concept” (Cole et al., 2007) shows that only 0.9 Pg of this carbon reaches the ocean, because the continental hydrosphere has three key roles of (i) emitting CO_2 and CH_4 to the atmosphere, (ii) storing organic carbon in river sediments and floodplains, and (iii) exporting to the ocean both bicarbonates produced by rock alterations and slowly biodegradable organic carbon.

Despite its key role in environmental processes, the contribution of OM to river carbon fluxes remains poorly understood. OM is a highly complex and dynamic mixture of organic compounds from both natural and anthropogenic sources (Artifon et al., 2019; Cawley et al., 2012; Bauer et al., 2011). The allochthonous (detritus from terrestrial plants and soils) or autochthonous (*in situ* aquatic production) origin of natural OM in rivers and aquatic ecosystems determines its composition and physicochemical properties, which influence its (photo-)chemical and biological transformation, its reactivity, and its role in the carbon cycle (Besemer et al., 2009; Lambert et al., 2017; Carlson and Hansell, 2015). The concentration and composition of OM also depend on the properties of the surrounding watershed and on the hydrological connectivity between rivers, oxbows, and groundwater, which control the transfer of terrestrial and aquatic OM in riverine ecosystems. Organic carbon is mostly described as dissolved organic carbon (DOC), which is more abundant than particulate organic carbon (Alvarez-Cobelas et al., 2012). Terrestrial organic carbon has long been considered recalcitrant and conservatively transported along river networks, with minimal contribution to aquatic metabolism. Contrary to previous assumptions, several studies carried out over the past two decades have shown that it is much less

https://doi.org/10.5194/essd-2025-375
Preprint. Discussion started: 4 March 2026

refractory and hence more bioavailable than previously thought (Besemer et al., 2009; Battin et al., 2008; Wiegner et al., 2009). Human activities also modify the properties of fluvial OM, with inputs from agricultural and forestry practices as well as from urban, domestic, and industrial discharges (Artifon et al., 2019; Cawley et al., 2012). Biodegradable OM plays a particularly significant role in river water quality. Large inputs of biodegradable OM can disrupt the dissolved oxygen cycle, leading to transient or permanent water anoxia (Tusseau-Vuillemin and Le Reveillé, 2001).

Not only the quantity but also the quality of organic carbon reflects the dynamic interactions between OM sources and biogeochemical processes (Jaffé et al., 2008), and both are likely to have ecological repercussions (Cory et al., 2011; Jaffé et al., 2014). The chemical composition of organic carbon is modified by biogeochemical processes, including bacterial activity. These changes in the composition of organic carbon, in turn, influence heterotrophic activity, CO_2 release, the biogeochemistry of nutrients and pollutants, and the formation of disinfection by-products in drinking water treatment plants (Zhuang and Yang, 2018).

It is thus essential to improve our knowledge of the physicochemical properties of OM and our ability to characterize its composition and reactivity.

In this context, we present a comprehensive dataset to quantify and characterize dissolved organic matter (DOM) in a regional river basin such as the Seine River basin in France, which is subject to numerous human pressures that influence the carbon structure in the various compartments of the hydrosystem. This dataset is the result of several sampling campaigns carried out over a 14-year period (2011–2024) in the framework of the PIREN-Seine research program (https://www.piren-seine.fr/en). The water samples were collected at sites representative of different land uses in the Seine River basin (agricultural and urban) and in different compartments of the hydrosystem: surface water, groundwater, and treated effluents from wastewater treatment plants (WWTPs). The dataset includes variables that characterize DOM both quantitatively and qualitatively using a combination of approaches and methods.

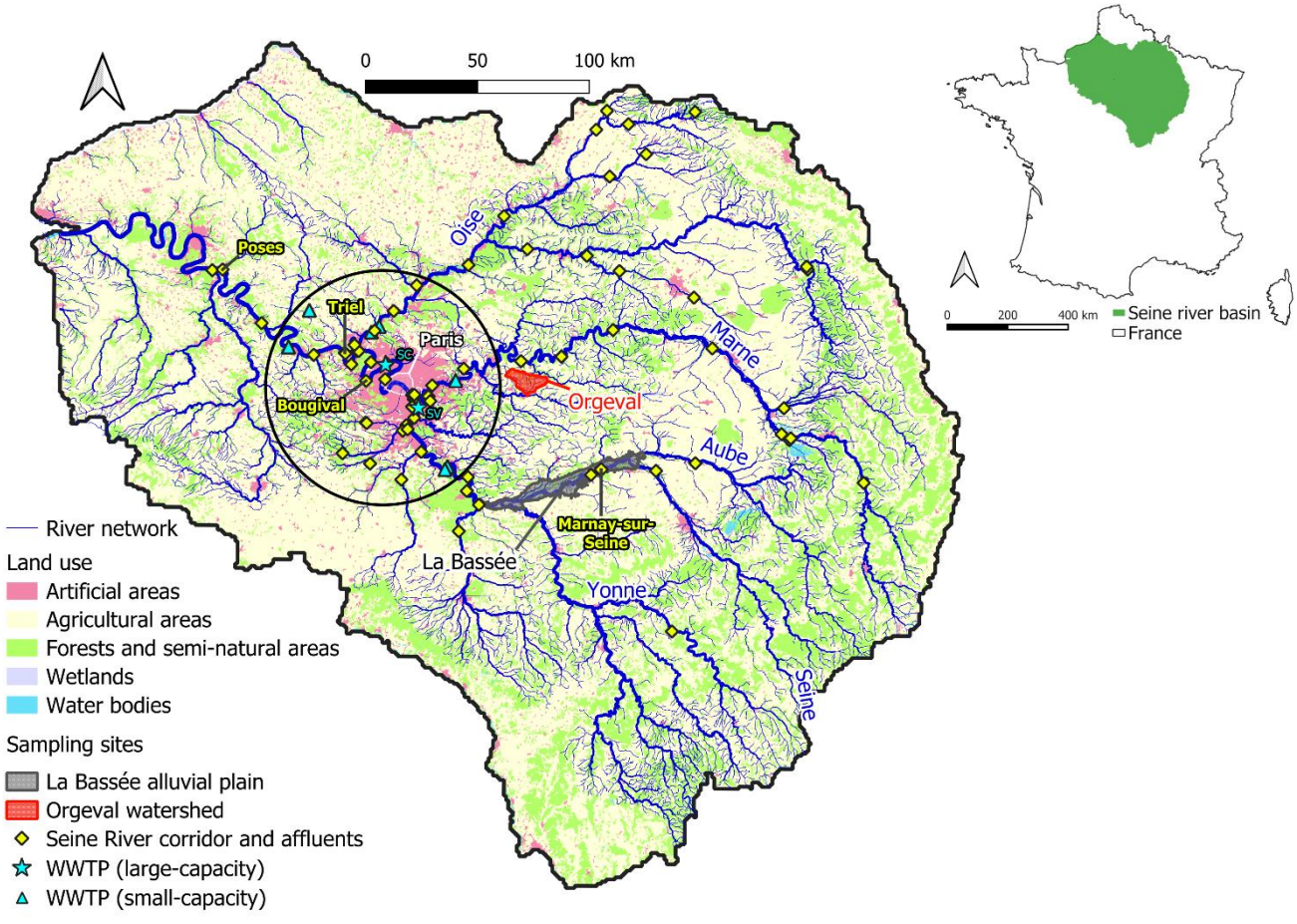
The dataset will remain accessible to the scientific community in order to pose new research questions related to: river metabolism; the sources, dynamics, and transformations of OM in surface water, in groundwater, and at their interface; and the typology of organic carbon in a basin at the regional scale.

https://doi.org/10.5194/essd-2025-375
Preprint. Discussion started: 4 March 2026

These observational data could also be used to parameterize and evaluate land-surface models or models assessing the quality of river networks.

2 Study area

The Seine River basin (76 000 km^2), located in northern France (Fig. 1), is the most urbanized and industrialized basin in France (Billen et al., 2007; Flipo et al., 2021). Given its large population and intensive cereal production, the water resources of the Seine River basin are of high strategic importance. As a major urban center, Paris relies heavily on the Seine River to meet its water needs. However, the city's dense population and industrial activities contribute significantly to the pollution of the river, posing challenges for water quality and ecosystem health.



https://doi.org/10.5194/essd-2025-375
Preprint. Discussion started: 4 March 2026

Figure 1: Seine River basin and sampling sites considered in this study: La Bassée alluvial plain (gray, see Fig. 2); the Orgeval watershed (red, see Fig. 3); the Seine River corridor and its tributaries (yellow diamonds); the small-capacity WWTPs (blue triangles); the two large-capacity WWTPs (blue stars: SA is Seine Aval and SV is Seine Valenton). The area in the black circle is zoomed in Fig. 4. The hydrographic network is represented with a line of increasing thickness according to Strahler's order. Source of land use map: Corine Land Cover 2018, European Union's Copernicus Land Monitoring Service information.

The climatic regime of the Seine River basin is pluvial/oceanic, modulated by seasonal variations in evapotranspiration. The average precipitation ranges from 1200 mm yr^{-1} along the coastal shoreline and in the upstream mountainous area to 650 mm yr^{-1} in the central part of the basin, whereas the average precipitation for the entire basin is 800 mm yr^{-1}. The total river network comprises 27 500 km of streams, and the main course of the Seine River is 777 km long (Marescaux et al., 2020). The hydrological regime is pluvial/oceanic, with a high-flow period in winter and a low-flow period in summer. The mean annual discharge at the Poses gauging station, which is located in the most downstream part of the river network at the entrance of the estuary (Fig. 1), is 509 m^3 s^{-1} (average during the period 1990–2006). The hydrological regime is very irregular due to intense floods and droughts. For example, the mean daily discharge at Poses is 179 m^3 s^{-1} during 5-year droughts and 2 590 m^3 s^{-1} during 20-year floods. Land use in the Seine River basin (Fig. 1) is mainly represented by croplands (56.5 %), followed by forested areas (25.6 %), grasslands (9.5 %), and urbanized areas (7.6 %).

The data presented in this paper were obtained from water samples collected at several sites in the Seine River basin, representative of various land uses (Fig. 1): an alluvial plain (La Bassée), a small rural and agricultural watershed (Orgeval), several sites along the Seine River and some of its main tributaries (Marne River, Oise River, Aube River, Yonne River), as well as some minor tributaries. Treated effluents from a number of WWTPs were also sampled (Fig. 1). All these samples are representative of different compartments of the Seine hydrosystem: groundwater from aquifers of various depths (alluvial aquifer, riverbank aquifer, Tertiary Brie aquifer, Upper Cretaceous chalk aquifer), surface water (rivers, gravel pits, springs, oxbow lakes, ponds, agricultural drains), and treated effluents from WWTPs of various capacities.

https://doi.org/10.5194/essd-2025-375
Preprint. Discussion started: 4 March 2026

2.1 La Bassée alluvial plain

La Bassée is a part of the alluvial plain of the Seine River. It is located upstream of Paris, between the confluences of the Seine River with two tributaries: the Aube River and the Yonne River. The alluvial plain is 60 km long and 8 km wide and its surface area is around 320 km^2 (Flipo et al., 2021). The substrate of the alluvial plain is mainly Upper Cretaceous chalk. La Bassée is a territory of major political, economic, and environmental importance. It is a wetland of national importance for its remarkable biodiversity, and it represents an important source of drinking water for the Paris region. Several economic interests coexist here, including a nuclear power plant, a navigation channel, and numerous gravel extraction sites, as well as development projects such as the construction of flood control reservoirs. For these reasons, La Bassée has been studied by the PIREN-Seine research program for some 30 years. Numerous studies have been carried out in a wide range of disciplines, including geology, hydro(geo)logy, geography, geochemistry, and geohistory (Lestel et al., 2018; Flipo et al., 2021).

A network of LOcal MOnitoring Stations (LOMOS, Mouhri et al., 2013) was created in 2015, with the objective of studying the interactions between surface water and groundwater. These stations consist of one or two piezometers and a surface water device, all equipped with sensors to continuously measure pressure and temperature in surface water (river, gravel pit, or oxbow lake), in the associated alluvial aquifer, and, at some sites, also in the underlying chalk aquifer. Surface water and groundwater samples were collected at each LOMOS station. Other surface water bodies were also sampled to enrich the dataset with samples from gravel pits and oxbow lakes. All the sampling points in La Bassée are shown in Fig. 2.

https://doi.org/10.5194/essd-2025-375
Preprint. Discussion started: 4 March 2026

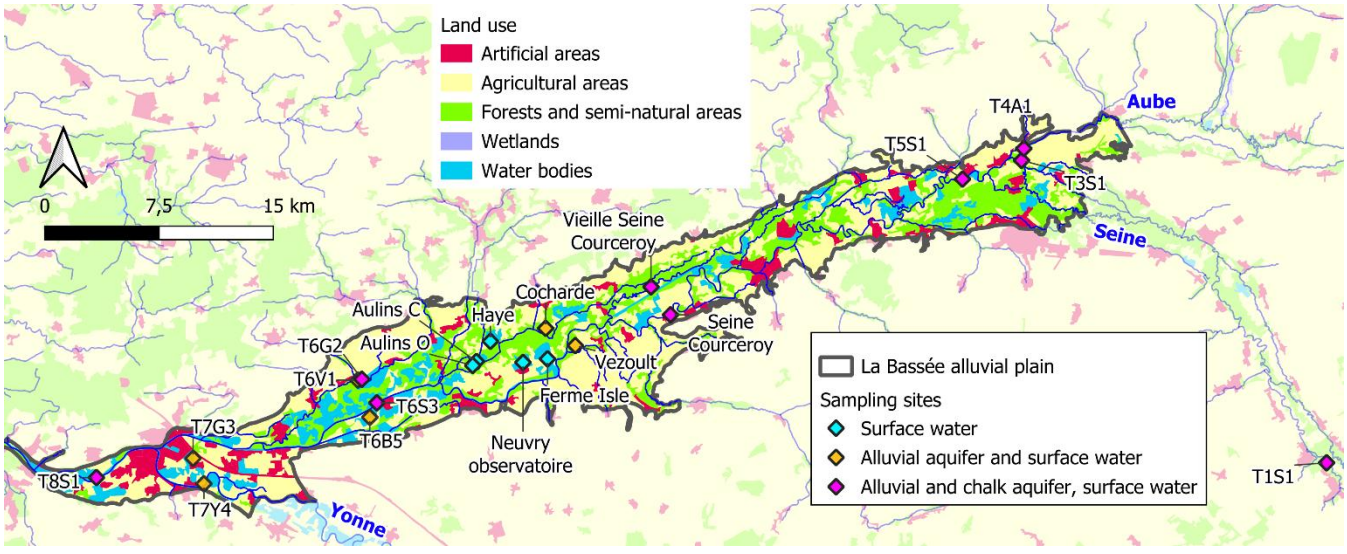


Figure 2: Sites sampled in La Bassée alluvial plain. Source of land use map: Corine Land Cover 2018, European Union's Copernicus Land Monitoring Service information.

2.2 The Orgeval watershed

The 104 km^2 Orgeval watershed is located 70 km east of Paris (Garnier et al., 2014; Mouhri et al., 2013). It is a sub-catchment of the Marne River basin, which is part of the Seine River basin (Fig. 1). The Orgeval watershed aquifer system is composed of two main geological units: the Brie limestone Oligocene aquifer and the deeper Champigny limestone Eocene aquifer. These two aquifer units are separated by a clayey aquitard (green clays and marls) and are covered by a loess layer approx. 2–10 m thick. Land use is mostly agricultural (82 %), dominated by cereal crops (wheat, maize, barley). The remaining surface is covered by woods (17 %) and urban zones or roads (1 % of the surface). The Orgeval watershed is representative of the headwaters of the Seine hydrosystem, characterized by extensive agricultural coverage. It is also representative of a multi-layer aquifer system that exchanges water and nutrients with a river network under temperate climate conditions.

The Orgeval watershed is a critical zone observatory and is part of the French Research Infrastructure OZCAR (Floury et al., 2019; Gaillardet et al., 2018). It has been a long-term experimental observatory and research site for 60 years (https://gisoracle.inrae.fr/). This site has been equipped with monitoring

https://doi.org/10.5194/essd-2025-375
Preprint. Discussion started: 4 March 2026

stations for decades, enabling the measurement of a large number of physicochemical variables. In particular, some LOMOS stations were installed to measure pressure and temperature in the river and in the underlying aquifers, with the aim of understanding the interactions between groundwater and surface water. Surface water and groundwater sampling points (Fig. 3) are representative of different types of water and positions in the watershed. Surface water types include rivers, springs, agricultural drains, and ponds. Groundwater types include colluvium and the Brie limestone Oligocene aquifer.

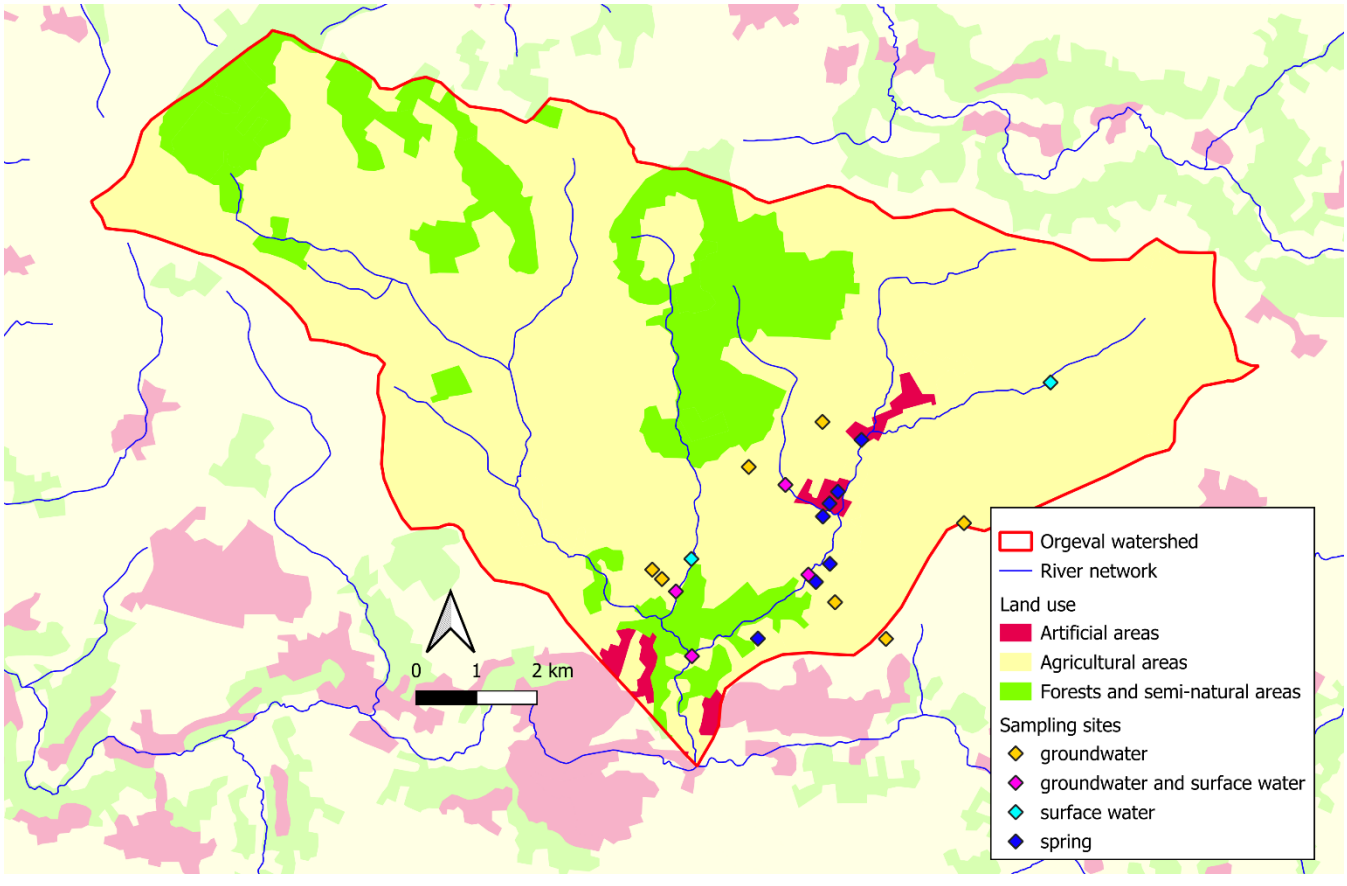


Figure 3: The Orgeval watershed and the sites sampled. Source of land use map: Corine Land Cover 2018, European Union's Copernicus Land Monitoring Service information.

https://doi.org/10.5194/essd-2025-375
Preprint. Discussion started: 4 March 2026


Open Access Earth System Science Data Discussions

2.3 The Seine River corridor and its tributaries

The sampling sites along the Seine River corridor and its tributaries (Figs. 1 and 4) are representative of different levels of urbanization and anthropogenic pressure on the environment. Two types of samples were collected: river water samples and WWTP discharges. The main sampling sites along the Seine River are Marnay-sur-Seine, Bougival, Triel, and Poses. Marnay-sur-Seine is located some 200 km upstream of Paris (Fig. 1), in La Bassée alluvial plain. Bougival and Triel are located in urban areas (Fig. 4): Bougival lies just downstream of Paris and the Seine Centre WWTP, while Triel is located downstream of the Paris metropolitan area, the confluence with the Oise River and the Seine Aval WWTP. Poses is located at the entrance to the Seine estuary (Fig. 1), in a predominantly agricultural area. Other sites were sampled along the Seine River, as well as on some of its major tributaries (Marne and Oise Rivers), and on minor tributaries during specific campaigns.

WWTPs of various capacities were also sampled in the Paris region (Fig. 4) to characterize the quality of their treated water in terms of OM. Seven relatively low-capacity WWTPs were selected: Boissette, Dammaries, St. Thibault, Butry, Auvers, Magny en Vexin, and Rosny. Two large-capacity WWTPs managed by the SIAAP (Interdepartmental Syndicate for the Sanitation of Greater Paris) were sampled: Seine Centre and Seine Valenton.

https://doi.org/10.5194/essd-2025-375
Preprint. Discussion started: 4 March 2026

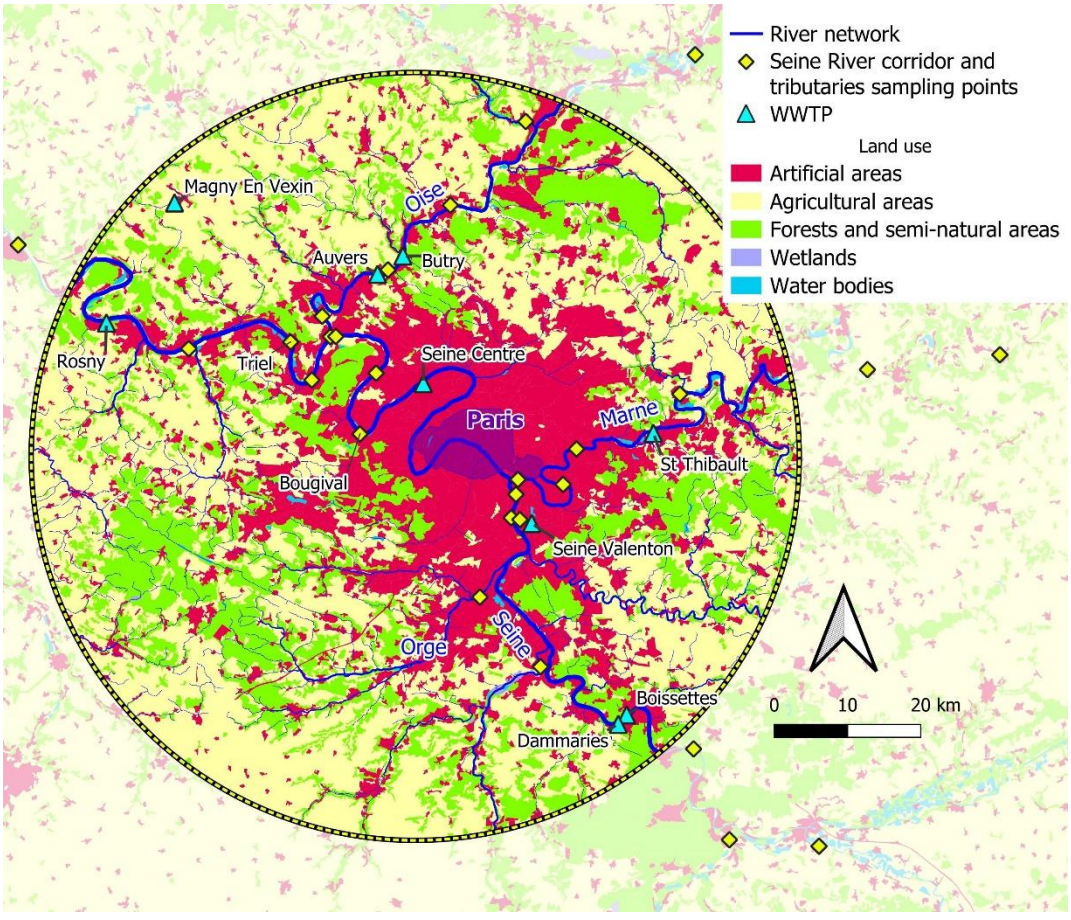


Figure 4: Sites sampled along the Seine River corridor and its tributaries: focus on the area around Paris. Source of land use map: Corine Land Cover 2018, European Union's Copernicus Land Monitoring Service information.

3 Methodology

3.1 The sampling campaigns

Several field campaigns were carried out between 2011 and 2024 in the different sites of the Seine River basin. A total of 1047 samples were collected and analyzed, including 178 groundwater samples, 787 surface water samples, 23 spring samples, and 59 WWTP effluent samples (Table 1). For all these samples, DOM was characterized both quantitatively (DOC concentration measurement) and qualitatively, through analysis of the optical properties of DOM by optical spectroscopy (UV-Visible

https://doi.org/10.5194/essd-2025-375
Preprint. Discussion started: 4 March 2026

absorbance and 3D excitation–emission matrix (EEM) fluorescence spectroscopy). The biodegradable fraction of DOC (BDOC) was also estimated for 27 % of the samples. Standard physicochemical water parameters were measured for 45 % of the samples, when possible, to better understand their relationship with OM properties.

Samples in the database (total: 1047, 27 % with BDOC)								
La Bassée			**Orgeval**			**Seine River corridor and tributaries**		
	Samples analyzed	Samples with BDOC		Samples analyzed	Samples with BDOC		Samples analyzed	Samples with BDOC
Alluvial aquifer	75	27 %	Groundwater	62	84 %	Rivers	649	10 %
Chalk aquifer	41	34 %	Rivers	25	80 %	WWTPs	59	92 %
Rivers	46	20 %	Pond	8	100 %			
Oxbow lakes	17	35 %	Agricultural drain	4	100 %			
Gravel pits	38	61 %	Springs	23	48 %			
Total	**217**	**33 %**	**Total**	**122**	**78 %**	**Total**	**708**	**17 %**

Table 1: Number of samples collected and analyzed for each geographical unit and type of water. The percentage of samples where the biodegradable fraction of dissolved organic carbon was estimated is also indicated.

3.2 Collection and analysis of the water samples

Sampling of surface water (rivers, gravel pits, oxbow lakes, ponds, agricultural drains, and springs) was carried out using a bucket previously rinsed with the respective water to avoid any external pollution. Groundwater was generally sampled with a 1-L bailer; however, at certain sites in the Orgeval watershed a peristaltic pump was used because the piezometer opening was obstructed by probes, preventing the use of a bailer. In all cases, the water level inside the piezometer is first measured, after which the piezometer is purged and allowed to refill with water before sampling. The sampled water (surface water or groundwater) is then poured gently into various types of bottles, which are first rinsed well with the same water, and stored at temperatures between 4°C and 10°C until processing, which occurred on site within a few minutes to 1 h.

https://doi.org/10.5194/essd-2025-375
Preprint. Discussion started: 4 March 2026

Samples from the low-capacity WWTPs were taken by on-site personnel, who collected 2–4 L of treated water from the bottles collected over 24 h using refrigerated samplers. These samples were stored for a few hours in a refrigerator and then transported to the laboratory in a cooler at a temperature below 10°C, where processing occurred within 1–2 h, depending on the distance from the WWTP. In the case of the large-capacity SIAAP WWTPs, samples were obtained by spot sampling to minimize the evolution of OM and its rapidly biodegradable fraction. After sampling, filtration was carried out on site.

3.2.1 Physicochemical water parameters

Measurements of conductivity, pH, dissolved oxygen, and temperature were performed with multiparameter probes directly in the piezometer, after purging, or in the bucket (for surface water and groundwater taken at the pump). Total alkalinity (mmol L^{-1}) was measured on 20 mL of filtered water (GF/F: 0.7 µm) using an automatic titrator (Titrando 905) and HCl (hydrochloric acid, 0.01 M). From several sub-sample analyses (2 to 5) conducted on multiple occasions, we calculated a variation coefficient of 7 %.

3.2.2 Dissolved organic carbon and biodegradability

Water samples were filtered immediately after collection through glass fiber filters (Whatman, GF/F, 0.7 µm, 25 mm diameter) that had been previously calcined at 450°C for 5 h. DOC analysis was carried out by infrared spectroscopy, after oxidation of the samples using either sodium peroxodisulfate or catalytic combustion, depending on the analysis laboratory.

Biodegradability was estimated using the batch incubation method. Filtered samples were re-inoculated with 1 % of non-filtered water and stored in the laboratory at 21°C, in the dark, and with agitation for 45 days, during which time the bacteria consumed the OM aerobically (Servais et al., 1999). The biodegradable fraction was then calculated as the difference between DOC content at the end of the incubation period and that at the beginning.

The uncertainty associated with the DOC and BDOC concentrations was estimated at 5 % and 10 %, respectively. The uncertainty in DOC concentrations was estimated as the coefficient of variation

https://doi.org/10.5194/essd-2025-375
Preprint. Discussion started: 4 March 2026

calculated from quintuplicate analyses of three types of samples: groundwater, river, and pond. Each quintuplicate was incubated for 45 days and sampled three times to estimate BDOC. The uncertainty in BDOC concentrations was estimated as the associated coefficient of variation.

3.2.3 Optical properties of dissolved organic matter

UV-Visible absorbance and EEM fluorescence spectroscopy are rapid, semi-quantitative, and noninvasive techniques that have been widely used for decades (Catalán et al., 2021; Derrien et al., 2019; Huguet et al., 2009; Jaffé et al., 2014; Minor et al., 2014; Parlanti et al., 2000).

Immediately after sampling, the water samples were filtered through pyrolyzed (450°C for 1 h) glass fiber filters (Whatman GF/F, 0.7 µm pore size, 25 or 47 mm diameter). The samples were then stored in glass vials at 4°C in the dark until spectroscopic analysis.

UV–Visible spectroscopic analysis

UV-Visible absorption spectroscopy was used for rapid determination of the general properties of DOM. Absorbance spectra were recorded (between wavelengths of 210 and 700 nm) using Jasco V-560 and V-760, Aqualog (HORIBA Jobin-Yvon), or UviLine 9400 (Secoman) spectrophotometers. Several indices can be used to obtain information on DOM properties from absorbance spectra. Two indices were retained in the database: the spectral slope ratio (SR) and the SUVA (Specific UV-Absorbance). SR is calculated as the ratio of the absorbance spectrum slope for short wavelengths ($S_{275\text{-}295nm}$) to the spectral slope for longer wavelengths ($S_{350\text{-}400nm}$). The SR ratio is negatively correlated with DOM size: as SR increases the molecular weight decreases (Helms et al., 2008). The SUVA index is calculated as the absorbance at 254 nm normalized by the DOC concentration. SUVA is correlated with the percentage of aromaticity of DOM in the aquatic environment. When the SUVA value is high (> 4), DOM is said to be hydrophobic with a strong aromatic character, whereas when it is low (< 3), DOM is said to be hydrophilic (Zsolnay et al., 1999).

EEM fluorescence spectroscopy and PARAFAC decomposition

https://doi.org/10.5194/essd-2025-375
Preprint. Discussion started: 4 March 2026

The fluorescence properties of DOM provide information on its structure and general characteristics. Fluorescence is a highly sensitive technique that enables the characterization of DOM from small-volume aqueous samples without the need for concentration or extraction. EEM fluorescence spectroscopy is generally used to characterize DOM and study its dynamics in aquatic environments (Carstea et al., 2010; Ejarque et al., 2017; Huguet et al., 2009; Jaffé et al., 2014; Tzortziou et al., 2015). EEM spectra highlight the various fluorophores constituting DOM and provide information on its source, chemical composition, state of degradation, and reactivity (Ejarque et al., 2017; Fellman et al., 2010; McKnight et al., 2001; Parlanti et al., 2000).

Fluorescence spectra were recorded using the Jasco FP-8300, Horiba Jobin-Yvon Aqualog, and Fluorolog FL3-22 spectrofluorometers. Spectra were instrumentally corrected for each instrument and fluorescence intensities were normalized to Raman units. However, the acquisition parameters were not identical across instruments, particularly in terms of the wavelength range and the increments used for excitation and emission. The matrices were therefore homogenized to have common wavelength ranges—250–410 nm at excitation with 10-nm increments, and 260–600 nm at emission with 2-nm increments—so that they could be processed as a whole. The semi-quantitative and qualitative data to be taken into account are the intensity (proportional to fluorophore concentration) and position of fluorescence maxima, which vary according to the nature and origin of the samples and depend on the fluorescent molecular species they contain. Fluorescence indices (humification index [HIX], biological activity index [BIX], and fluorescence index [FI]) are determined in order to characterize the sources and degree of maturation of fluorescent DOM.

The HIX (Zsolnay et al., 1999) is used to estimate the degree of aromaticity of DOM. High HIX values (HIX > 12) indicate the presence of highly aromatic OM that has undergone transformation or polymerization/polycondensation processes, while low values (HIX < 4) indicate a more recent and less aromatic DOM (Huguet et al., 2009). The BIX is used to estimate the presence of freshly produced OM in the environment (Huguet et al., 2009): high values (BIX > 0.8) indicate strong biological activity and therefore a recent autochthonous origin (biological or aquatic bacterial) of DOM. The FI (McKnight et al., 2001) identifies the relative contribution of terrestrial (FI < 1.3) or microbial (FI > 1.9) DOM.

https://doi.org/10.5194/essd-2025-375
Preprint. Discussion started: 4 March 2026

All EEM spectra were processed using a multivariate analysis method called "PARAllel FACtor analysis" (PARAFAC), which enables the identification and quantification of fluorophores that vary independently and are referred to as "components." PARAFAC was implemented in Matlab R2021b with the DOMFluor 1.7 toolbox. The number of components explaining the overall fluorescence of the whole dataset was determined using split-half analysis ($S_4C_4T_2$ - Splits: 4, Combinations: 4, Tests: 2) and visual inspection of residuals and component spectra (Murphy et al., 2013).

This PARAFAC statistical modeling approach allowed us to determine and validate a 5-component model (Fig. 5), explaining 99.5 % of the variability in the entire dataset (1047 samples).

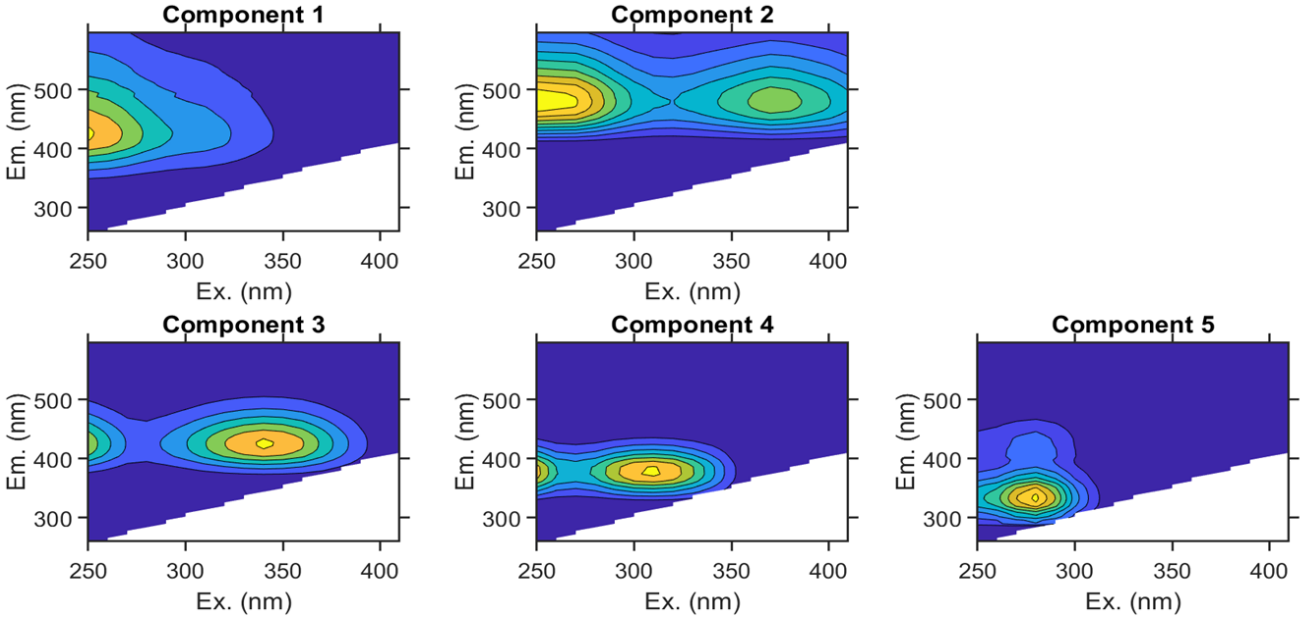


Figure 5. Contour plots of the 5 components determined by PARAFAC.

Components C1, C2, C3, C4, and C5 were then compared with those described in the literature and with those recorded in the OpenFluor database (Murphy et al., 2014b), generating matches (Tucker correlation coefficient > 0.95) with the previously reported spectra of 72, 119, 11, 28, and 98 studies, respectively (as of April 2025).

https://doi.org/10.5194/essd-2025-375
Preprint. Discussion started: 4 March 2026

Component C1 has been frequently observed in freshwater, forest, and agricultural areas (Calderó-Pascual et al., 2022; Marcé et al., 2021; Stedmon and Markager, 2005) and is attributed to DOM of terrestrial, soil, or plant origin, produced by photo-transformation or microbial degradation (DeFrancesco and Guéguen, 2021; Eder et al., 2022; Moona et al., 2021; Murphy et al., 2014b). This component has been described as corresponding to oxidized quinones (Cory and McKnight, 2005; Groeneveld et al., 2020), a higher proportion of which is possibly due in urban areas to the accumulation of microbe-derived quinones induced by anthropogenic activity (Chai et al., 2019).

The literature commonly describes components C2 and C3 as predominantly terrigenous compounds, which are commonly found in freshwater and are generally associated with high-molecular-weight, aromatic, and phenolic materials (DeFrancesco and Guéguen, 2021; Menendez and Tzortziou, 2024; Murphy et al., 2014b; Peleato et al., 2017; Yamashita et al., 2015; Yang et al., 2019).

Component C2 corresponds to compounds of terrigenous origin that are commonly found in freshwater and are generally associated with high molecular weights and aromatic material (Dainard et al., 2015; DeFrancesco and Guéguen, 2021; Du et al., 2021; Menéndez and Tzortziou, 2024; Peleato et al., 2017). C2 has been attributed to highly degraded, plant-derived DOM (Podgorski et al., 2018) and reduced quinones of terrestrial origin in forested areas, agricultural watersheds, and wetlands (Cory and McKnight, 2005; Eder et al., 2022; Graeber et al., 2012; Yamashita et al., 2011). Additionally, C2 has been reported as a product or intermediate of the photochemical degradation of terrestrial OM (Eder et al., 2022; Murphy et al., 2014b).

Component C3 has a similar emission spectrum to that of a lignin degradation product, syringaldehyde, and has been observed in forested areas (Eder et al., 2022; Murphy et al., 2014b; Peleato et al., 2017). Additionally, it has been found in agricultural areas and nutrient-rich environments, as well as in areas exposed to discharges from WWTPs (Batista-Andrade et al., 2023; Graeber et al., 2012; Jutaporn et al., 2020; Murphy et al., 2011; Yu et al., 2015); it has also been associated with a fluorophore of microbial origin (Maurischat et al., 2022).

The fourth component (C4) has been described in the literature as a product of microbial degradation and/or remineralization. It is also associated with biological activity, microbial growth, and algal growth and production (Kurek et al., 2022; Meilleur et al., 2023; Parlanti et al., 2000; Zhuang et al., 2021).

https://doi.org/10.5194/essd-2025-375
Preprint. Discussion started: 4 March 2026

364 However, it has also been linked to anthropogenic sources, such as WWTP effluents (Paradina-Fernández
365 et al., 2023).
366 Component C5 corresponds to protein-like material and is associated with biological activity in the
367 environment (Hambly et al., 2015; Menendez and Tzortziou, 2024; Parlanti et al., 2000; Wünsch and
368 Murphy, 2021). It has been reported as bioavailable DOM in highly productive waters (Gao and Guéguen,
369 2017; Hambly et al., 2015) and in rivers subject to anthropogenic influences (Ouyang et al., 2024).
370
371

372 # 4 Description of the dataset

373 ## 4.1 Structure of the dataset

374 The structure of the dataset is thoroughly described in the file "README.pdf". For each sample, the
375 dataset includes the following variables: temperature, electrical conductivity, pH, dissolved oxygen
376 concentration, alkalinity, concentration of dissolved organic carbon (DOC) and its biodegradable fraction
377 (BDOC), UV-Visible absorbance indices (among them: SUVA (Weishaar et al., 2003) and SR (Helms et
378 al., 2008)), fluorescence indices (HIX (Zsolnay et al., 1999), BIX (Huguet et al., 2009), and FI (McKnight
379 et al., 2001)) and the PARAFAC components (Stedmon and Bro, 2008). For each of the fluorescence
380 spectra of the samples, the excitation-emission matrix is also included.

381 ## 4.2 General characteristics of dissolved organic matter in the Seine River basin

382 DOM can be characterized by several variables, such as DOC concentration and the indices derived from
383 spectroscopic analysis: the HIX humification index linked to aromaticity, the BIX biological activity
384 index, the FI index, which discriminates between terrestrial and microbial sources, the SR index linked
385 to OM size, and the SUVA index, associated with aromaticity percentage. Added to these variables are
386 the components determined by the PARAFAC multiway analysis, representing the various fluorophores
387 contributing to the total fluorescence of the sample and of the whole dataset.

https://doi.org/10.5194/essd-2025-375
Preprint. Discussion started: 4 March 2026

The physicochemical properties of water, such as dissolved oxygen concentration, conductivity, and pH, can also be studied to better understand their relationship with the properties of DOM.

The Seine River basin dataset shows thatsurface water samples are well oxygenated (around 10 mg L^{-1}), while groundwater is characterized by lower oxygen concentrations (1–2.5 mg L^{-1}), particularly in La Bassée alluvial plain, where some samples are almost anoxic (Fig. 6). WWTP discharges have a median conductivity of 1065 µS cm^{-1}, which is significantly higher than that of natural water samples (534 µS cm^{-1}). The conductivity of groundwater is higher than that of surface water, with median values of 609 µS cm^{-1} and 498 µS cm^{-1}. pH is close to neutral in groundwater and spring water and is basic in surface water (median 8.0). Particularly high values (up to 9.6) are measured in the gravel pits of La Bassée, due to photosynthesis activity, which consumes CO_2.

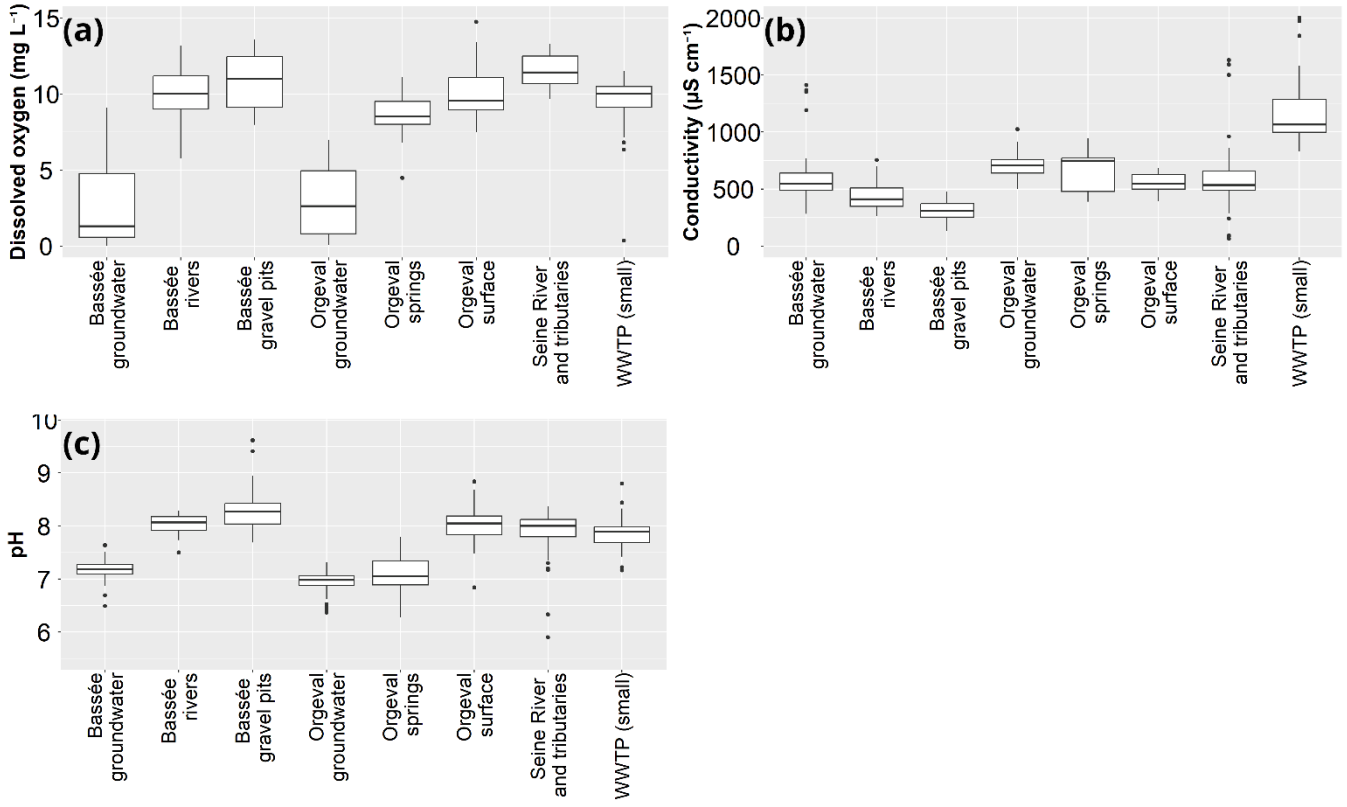


Figure 6: Box plots of the main physicochemical parameters for the different sites and water types: (a) dissolved oxygen; (b) conductivity; (c) pH. The horizontal lines within the boxes indicate the median. The boxes denote the 25th and 75th percentile. The whiskers indicate the minimum and maximum data points, excluding outliers. The points outside the whiskers represent potential outliers.

https://doi.org/10.5194/essd-2025-375
Preprint. Discussion started: 4 March 2026

The DOC concentration in the Seine River basin varies depending on the site, the measurement period, and the type of water sampled. Discharges from WWTPs have a higher DOC content than natural waters (median values of 6.8 mgC L^{-1} and 2.9 mgC L^{-1}, respectively) (Fig. 7, Table 2). Overall, concentrations are higher for small-capacity WWTPs than for large-capacity WWTPs. Among natural water samples, surface waters are generally richer in DOC than groundwater is. This is particularly evident in La Bassée alluvial plain, where the median DOC concentrations are 1.4 mgC L^{-1} in aquifers and 2.6 mgC L^{-1} in samples from rivers and oxbows. The numerous gravel pits in La Bassée are particularly rich in DOC (median 5 mgC L^{-1}), as in these water bodies the intense primary production activity is a source of OM. In the Orgeval watershed, DOC values are more similar between surface water and groundwater (Fig. 7, Table 2). This is probably due to the fact that surface water in the Orgeval watershed is represented by small rivers, which are characterized by low primary production. Moreover, this is an essentially agricultural region with few wetlands and therefore few sources of OM, unlike La Bassée alluvial plain. Samples collected from the Seine River and from its tributaries show a wide variability in DOC concentrations (Fig. 7), due to both the diversity of hydrological sampling conditions and the heterogeneity of these samples, ranging from small streams to large rivers.

These data shows that the highest DOC concentrations logically occur in WWTP followed by concentrations in surface water, with a high range of values, in the large rivers, possibly receptacles of treated effluents from WWWT. DOC is also high in stagnant system where the organic load is mainly due to algal biomass.

https://doi.org/10.5194/essd-2025-375
Preprint. Discussion started: 4 March 2026

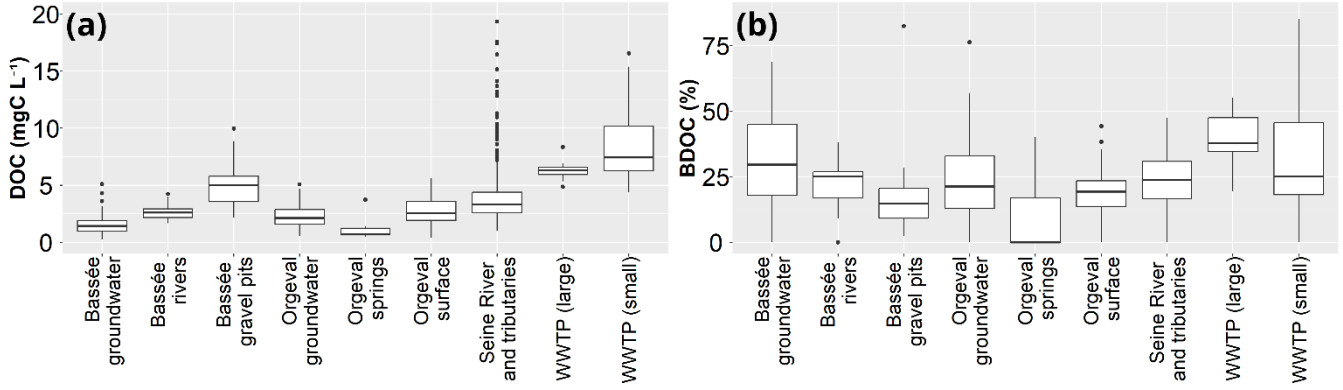


Figure 7: Box plots of (a) the dissolved organic carbon concentration and (b) its biodegradable fraction for the different sites and water types. The horizontal lines within the boxes indicate the median. The boxes denote the 25th and 75th percentile. The whiskers indicate the minimum and maximum data points, excluding outliers. The points outside the whiskers represent potential outliers.

	DOC (mgC L^{-1})	BDOC (mgC L^{-1})	BDOC/DOC (%)
La Bassée: groundwater	1.4	0.4	30
La Bassée: rivers	2.6	0.5	24
La Bassée: gravel pits	5.0	0.9	15
Orgeval: groundwater	2.1	0.5	21
Orgeval: springs	0.7	undetactable	-
Orgeval: surface water	2.6	0.5	19
Seine River and tributaries	3.4	0.6	25
WWTP (large)	6.3	2.4	38
WWTP (small)	7.6	1.6	25

Table 2: Median values for DOC, BDOC, and their ratio.

The biodegradable fraction of DOC in the Seine River basin is highly variable, with values ranging from below the limit of quantification to 85 % (Fig. 7, Table 2). The variability in BDOC among the different samples is high for each type of water, in particular for WWTPs, for groundwater in La Bassée alluvial plain and, to a lesser extent, for groundwater of the Orgeval (Fig. 7). This variability is due to the characteristics of the measurement sites as well as to the hydrological and meteorological conditions of

https://doi.org/10.5194/essd-2025-375
Preprint. Discussion started: 4 March 2026

437 the sampling campaigns. DOC inputs from riparian vegetation, or from tree leaf, could be rather
438 biodegradable. It would be interesting to better investigate this variability setting up in situ and/or
439 laboratory experiments. BDOC fractions are generally higher in large-capacity WWTPs (median 38 %)
440 than in small-capacity WWTPs (median 25 %). Natural waters in the basin generally have lower BDOC
441 values (median 21 %), with the exception of groundwater in La Bassée.
442 Specifically, in La Bassée, BDOC fractions are higher in groundwater than in rivers (30 % vs 24 %
443 respectively), and the lowest values are measured in gravel pits (15 %). This configuration could indicate
444 significant bacterial activity in the surface waters of La Bassée, particularly in gravel pits, where much of
445 the produced OM is degraded. Biodegradable fractions in the Orgeval watershed are slightly higher in
446 groundwater than in surface water, which also supports significant bacterial activity (Fig. 7, Table 2). OM
447 in spring water generally contains a lower biodegradable fraction than in other compartments of the Seine
448 hydrosystem. In the Seine River and its tributaries, BDOC fractions range from 0 % to 47 %, with a
449 median of 25 %.
450 DOM in the Seine River basin is relatively low in aromaticity and rather more hydrophilic. Indeed, HIX
451 values are mostly between 3 and 10 for all compartments of the hydrosystem, and the median value in the
452 basin is 5.7 (Fig. 8). This range of HIX values indicates the presence of a significant recent autochthonous
453 contribution of OM. The aromatic character is slightly higher in groundwater (median HIX = 6.9) than in
454 surface water (median HIX = 5.7). La Bassée gravel pits and WWTP discharges are characterized by very
455 low levels of aromatic DOM (median HIX = 2.7 and 3.1, respectively), indicating that the recent
456 autochthonous contribution is particularly significant in these two compartments. Samples collected in
457 the Seine River and in its tributaries show considerable variability in HIX values, as already observed for
458 DOC. BIX values are high for all sites and all water types (Fig.8), revealing high to very high biological
459 activity in all compartments. Biological activity is particularly high in WWTP effluents, especially small-
460 capacity WWTPs (median BIX = 1.1), in the Orgeval springs (BIX = 1), and in groundwater and gravel
461 pits in La Bassée (BIX = 0.9). With the exception of gravel pits, surface waters are characterized by
462 slightly lower BIX than groundwater.
463

https://doi.org/10.5194/essd-2025-375
Preprint. Discussion started: 4 March 2026

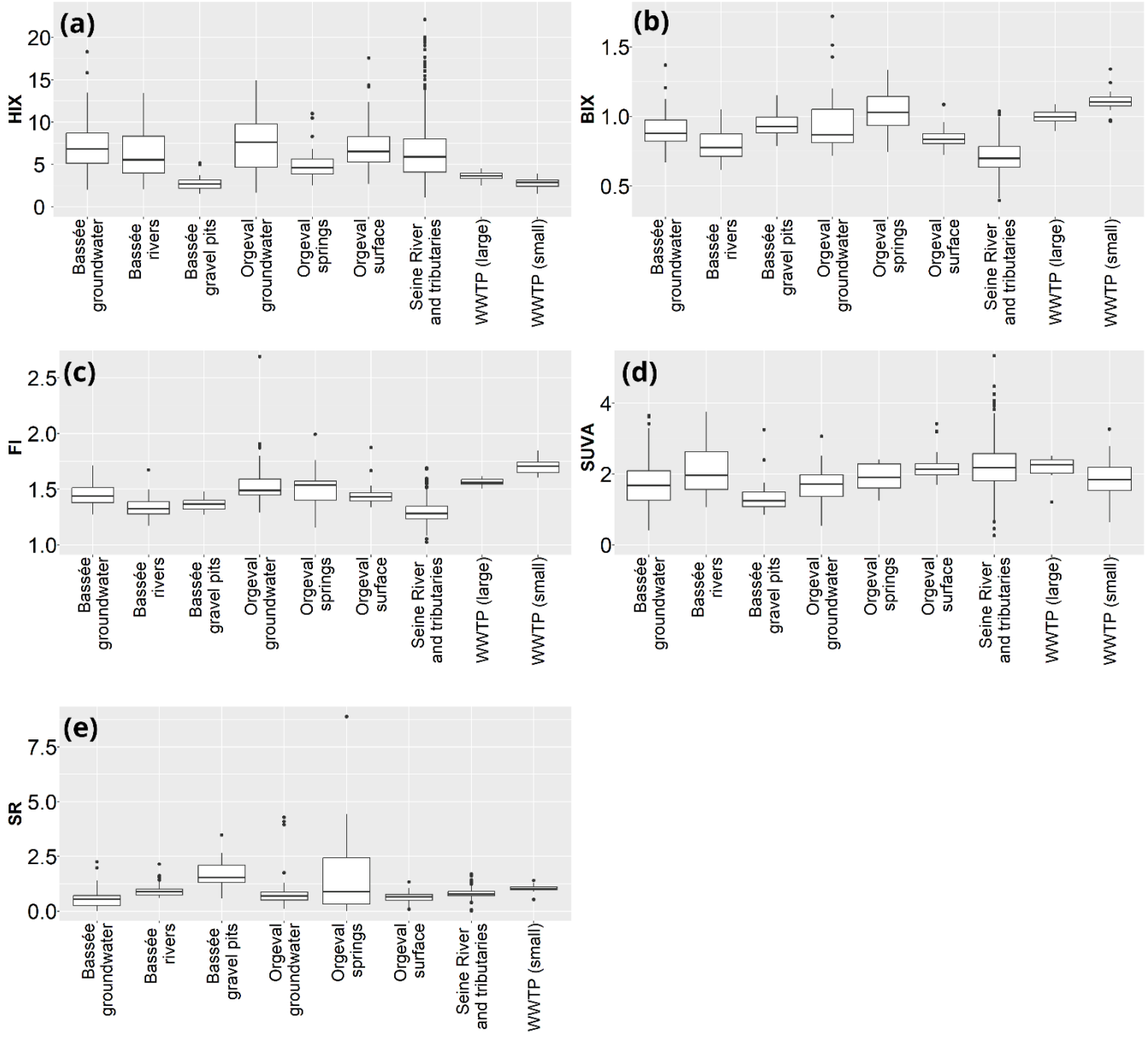


Figure 8: Box plots of the main indices characterizing DOM by means of optical spectroscopy (UV-Visible absorbance and EEM fluorescence). (a) HIX, humification index; (b) BIX, biological activity index; (c) FI, fluorescence index; (d) SUVA, index, associated with aromaticity percentage; (e) SR, index linked to organic matter size. The horizontal lines within the boxes indicate the median. The boxes denote the 25th and 75th percentile. The whiskers indicate the minimum and maximum data points, excluding outliers. The points outside the whiskers represent potential outliers.

https://doi.org/10.5194/essd-2025-375
Preprint. Discussion started: 4 March 2026

The median FI index (Fig. 8) is close to 1.5 for La Bassée and Orgeval groundwater and surface water samples, indicating a mixture of higher plants and microbial sources. Effluents from WWTP are characterized by a more important microbial origin, as indicated by the higher FI values (1.7 median), particularly for small-capacity WWTPs. On the other hand, in the Seine River and its tributaries, the FI indicates a predominantly terrestrial origin (median FI = 1.3). SUVA values are relatively low for groundwater and gravel pits, and slightly higher for river water, spring water, and WWTP effluents. However, most of the values are low (below 3), indicating that DOM is mostly hydrophilic, with a weak aromatic character and a low molecular weight. The smallest molecule sizes (highest SR values) are found in La Bassée gravel pits and in WWTP effluents.

These box plots, constructed using simple descriptive statistics, show a typical general trend for DOC and BDOC contents, and highlight the spatio-temporal variability of DOM optical properties in different water bodies in the Seine River watershed. They call for a more in-depth analysis of carbon behavior at the surface water and groundwater interfaces.

A principal component analysis was carried out (Fig. 9) with the following variables: conductivity, pH, dissolved oxygen concentration, temperature, DOC, BDOC, HIX, BIX, FI, SUVA, SR, and the relative fluorescence of the five PARAFAC components (expressed as the percentage of fluorescence of each component relative to the sum of the fluorescence intensities of the five components for each sample).

The number of individuals was 1047. PCA was performed in R using the FactoMineR and FactoExtra packages. The first plane explains 50 % of the total variability of the dataset. Considering the third axis, the explained variability increases to 60 %.

The principal component analysis shows that axis 1 (35 %) is mainly defined in the positive direction by BIX, FI, and the relative fluorescence of component C4, which is of microbial/bacterial type. In the negative direction, axis 1 is defined by the relative fluorescence of components C1 and C2, which are of terrestrial origin, and by the indices HIX and SUVA. This axis therefore seems to be associated with DOM quality: in the positive direction, DOM is autochthonous, freshly produced, degradable, with a strong biological and microbial signature. In the negative direction, DOM is more aromatic, mature, hydrophobic, and of terrestrial origin. Axis 2 (15 %) is mainly represented, in the positive direction, by pH and dissolved oxygen concentration. Data variability seems then to be mainly related to DOM type

https://doi.org/10.5194/essd-2025-375
Preprint. Discussion started: 4 March 2026

and to environmental variables (pH and dissolved oxygen), rather than to DOC content and river temperature. These results emphasize the value of this dataset, which encompasses both the quantitative and qualitative characterization of DOM alongside environmental data.

WWTP samples are located on the right side in the first plane of the graph of individuals (Fig. 9). They are grouped into two distinct clusters, depending on whether they originate from the effluents of small (Fig. 9b, green asterisks) or large-capacity (Fig. 9b, pink squares) WWTPs. Samples from small-capacity WWTPs are characterized by high biological activity and degradable OM of microbial origin, represented by microbial/bacterial component C4. Samples from large-capacity WWTPs have properties similar to those of certain samples from the Seine and its tributaries: the microbial origin of the OM and the biological activity are lower than for small WWTPs. Gravel pits are well oxygenated, have relatively basic pH values, and contain DOM represented mainly by protein-like compounds characteristic of biological activity (C5). Groundwater is characterized by high conductivity, fairly high biological activity (BIX, FI), and DOM represented mainly by the bacterially/microbially degraded C3 terrestrial component. Samples from the Seine River and its tributaries show OM that is rather aromatic and of terrestrial origin. However, the cloud of observations is quite scattered, indicating that the properties of these samples are highly variable depending on the site and measurement period.

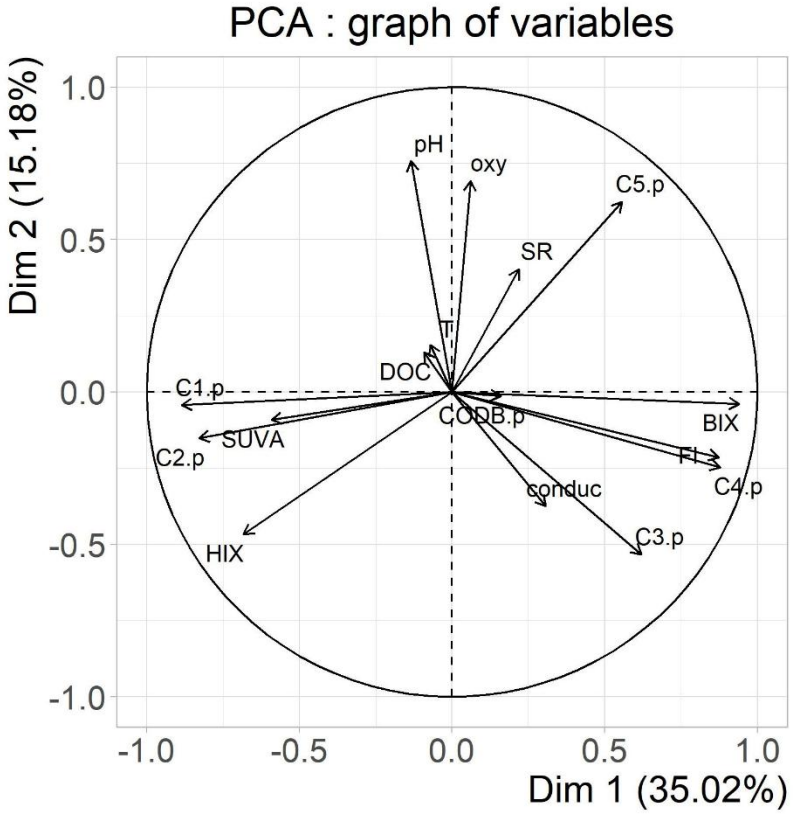
To our knowledge, such a dataset, which combines DOC and BDOC concentrations with UV-Visible absorption and EEM/PARAFAC fluorescence properties of DOM, covering a regional watershed over a 14-year period, is being proposed to the scientific community for the first time. There is no doubt that due to these widely accessible approaches today, OM characteristics will be made available in other sites than temperate one, giving insights on the microbial processes at work.

Overall, the PCA allows us to clearly distinguish between the various types of water in the Seine hydrosystem (surface water, groundwater, WWTP effluent) and geographical units (e.g., La Bassee, Orgeval) and to state that the type of DOM is the main factor affecting the variability of the dataset. PCA analysis is typically recommended for a general overview of data. These data could be further analysed in more detail, perhaps by splitting the dataset into hydrological compartments or seasons, or by adding similar data from other watersheds.

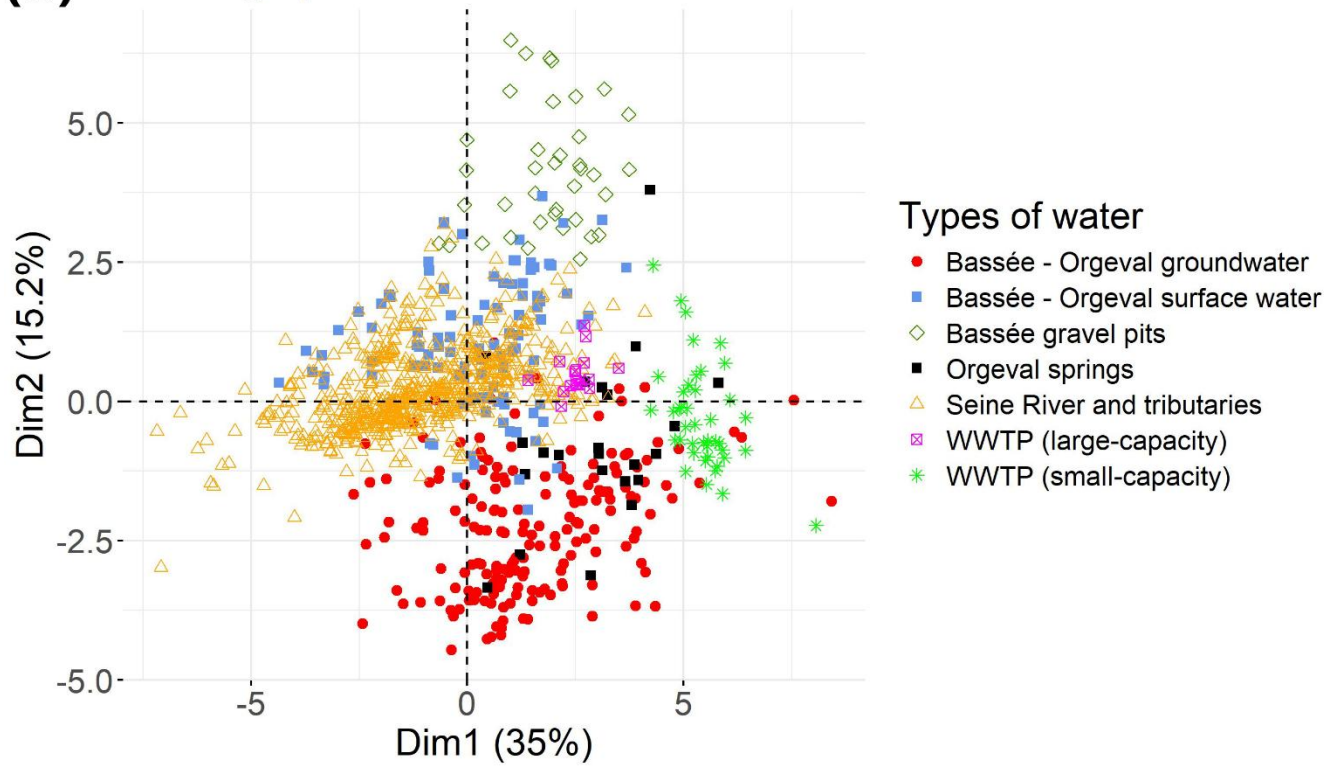
https://doi.org/10.5194/essd-2025-375
Preprint. Discussion started: 4 March 2026

Open Access Earth System Science Data Discussions

(a)



(b)



527

https://doi.org/10.5194/essd-2025-375
Preprint. Discussion started: 4 March 2026

Figure 9: Principal component analysis: a) graphs of variables (environmental data (pH, conductivity, dissolved oxygen, temperature); DOC content and percentage of biodegradable fraction (CODB.p); HIX, BIX, and FI fluorescence indices; SUVA and SR absorbance indices; relative fluorescence (%) of the five components determined by PARAFAC (C1.p to C5.p)) and b) graphs of individuals for the first two dimensions.

Data availability

All data presented in this paper are freely available under the agreement CC BY 4.0 at data.Indores repository: https://data.indores.fr/privateurl.xhtml?token=a6b58980-3280-4a1b-9311-a0b401955e75 (Baratelli et al., 2025).

The dataset consists of the geographic file "Sampling_sites.txt" containing the coordinates and the main characteristics of all the sampling sites; the data file "Data_carbon_Seine_basin.xlsx" containing the variables measured on each sample; a folder with 1047 files with the fluorescence spectra of all the samples; and the metadata file "Metadata_carbon_Seine_basin.xslx", which contains the description of all the columns in the data file and in the geographic file.

Conclusions

A dataset is presented to characterize DOM in the Seine River basin both qualitatively and quantitatively. The dataset includes dissolved organic carbon concentrations as well as the results of the analysis of the optical properties of DOM by UV-Visible absorbance and EEM fluorescence spectroscopy. The biodegradable fraction of DOC was also estimated for 27 % of the samples.

This dataset is remarkable because water samples were collected from multiple sites across a basin at regional scale, i.e., the Seine River basin in France, which is subject to significant anthropogenic forcing. The sampling sites are representative of different land uses in the Seine River basin and of different compartments of the hydrosystem: surface water, groundwater, and WWTP discharges. Moreover, the dataset covers a 14-year period as well as various hydrological and meteorological conditions.

This dataset is available to the scientific community. It can be used to help understand the sources, dynamics, and transformations of OM in the different compartments of a hydrosystem, and particularly at the interface between surface water and groundwater. The influence of various hydrological/meteorological conditions and of the type of site (land use, geology, aquifer residence times)

https://doi.org/10.5194/essd-2025-375
Preprint. Discussion started: 4 March 2026

on the DOM characterization can be investigated. These data, measured in the Seine River basin, could serve as a comparison for studies in other regions and to validate new indices characterizing the sources of OM. The dataset can also be used to deepen our understanding of the organic carbon cycle at the regional basin scale, even though studies have already shown that DOC concentrations are about 10 times lower than those of dissolved inorganic carbon (Marescaux et al., 2020; Garnier et al., 2025). Nevertheless, DOM is central for studying river metabolism and additional combined O_2, CO_2, and DOM high-frequency measurements should also offer more in-depth knowledge. In addition to the suggestions provided in the text to encourage the use of this dataset, these observational data could be also used to parameterize and evaluate land-surface models or models assessing the quality of river networks. The model RIVE developed in the framework of the PIREN-Seine program (Billen et al., 1994; Garnier and Billen, 1994; Wang et al., 2024) and describing the biogeochemical processes in aquatic systems is a good candidate for revisiting the DOC-BDOC formalization, using a detailed DOM characterization. This modelling tool pyRive is now deposited as a v3.2 version (https://gitlab.in2p3.fr/rive/pyrive). The model ProSe-PA (Wang et al., 2022) allowed to improve water quality simulations using a data assimilation technique. Its application at the scale of the Seine River basin pointed out the need of a finer description of the biodegradable component of the organic matter flux at the system's boundary. Our dataset could then contribute to further improve river metabolism simulation with this approach, especially during low flow without algae bloom.

The findings from this dataset therefore provide essential information for regional ecosystem modelling, the management of aquatic environments, and restoration strategies that target the impact of human activity on these environments.

Author contribution

Measurements of dissolved organic carbon and its biodegradable fraction were carried out at the METIS laboratory (Sorbonne University) and at LEESU (Paris-Est Créteil University). The investigation of the DOM sources and stages of evolution/degradation was carried out by studying its optical properties at the EPOC laboratory (University of Bordeaux), at LEESU, and at SIAAP (Innovation Department).

https://doi.org/10.5194/essd-2025-375
Preprint. Discussion started: 4 March 2026

Open Access Earth System Science Data Discussions

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgments

We would like to thank Gwendoline Bourdon, Yuzhe Guo, Zoé Hayet, Benjamin Mercier, Anun Martinez, Romane Nespoulet, Phuong Thanh Nguyen, and Mahaut Sourzac for collecting and/or analyzing water samples.

We are grateful to SIAAP, Veolia, Suez, and SAUR for organizing the sampling of the WWTPs and to the MeSeine observatory program (SIAAP) for its help in collecting samples in the Seine Basin. We would also like to thank Jean-Marie Mouchel and Sophie Guillon for contributing to the sampling of some springs in the Orgeval watershed and for providing the coordinates of the sampling sites for the 2018 flood campaign. We are also grateful to the farmers in the Orgeval watershed for giving us access to the sampling stations, and to the colleagues at INRAE for giving access to the Boissy experimental station.

The PIREN-Seine program (6^{th}, 7^{th}, 8^{th} Phase: https://piren-seine.fr/) is greatly acknowledged for its interdisciplinary framework and financial support.

References

Alvarez-Cobelas, M., Angeler, D. G., Sánchez-Carrillo, S., Almendros, G., 2012. A worldwide view of organic carbon export from catchments. Biogeochemistry 107(1), 275–293.

Artifon, V., Zanardi-Lamardo, E., Fillmann, G., 2019. Aquatic organic matter: Classification and interaction with organic microcontaminants. Science of the Total Environment 649, 1620–1635.

Baratelli, F., Parlanti, E., Garnier, J., Varrault, G., Goffin, A., Musabimana, N., Guérin Rechdaoui, S., Rocher, V., Flipo, N., 2025. Dissolved organic carbon concentrations and optical properties of dissolved organic matter in the Seine River basin. Data.InDoRES, https://data.indores.fr/privateurl.xhtml?token=a6b58980-3280-4a1b-9311-a0b401955e75.

Batista-Andrade, J.A., Diaz, E., Iglesias Vega, D., Hain, E., Rose, M.R., Blaney, L., 2023. Spatiotemporal analysis of fluorescent dissolved organic matter to identify the impacts of failing sewer infrastructure in urban streams. Water Research 229, 119521. https://doi.org/10.1016/j.watres.2022.119521

Battin, T. J., Kaplan, L. A., Findlay, S., Hopkinson, C. S., Marti, E., Packman, A. I., Newbold, J. D., Sabater, F., 2008. Biophysical controls on organic carbon fluxes in fluvial networks. Nature Geoscience 1(2), 95–100.

Battin, T.J., Luyssaert, S., Kaplan, L.A., Aufdenkampe, A.K., Richter, A., Tranvik, L.J., 2009. The boundless carbon cycle. Nature Geosciences 2, 598–600. https://doi.org/10.1038/ngeo618

https://doi.org/10.5194/essd-2025-375
Preprint. Discussion started: 4 March 2026

Battin, T.J., Lauerwald, R., Bernhardt, E.S., Bertuzzo, E., Gener, L.G., Hall, R.O., Hotchkiss, E.R., Maavara, T., Pavelsky, T.M., Ran, L., Raymond, P., Rosentreter, J.A., Regnier, P., 2023. River ecosystem metabolism and carbon biogeochemistry in a changing world. Nature 613, 449–459. https://doi.org/10.1038/s41586-022-05500-8

Bauer, J. E., Bianchi, T. S., 2011. Dissolved Organic Carbon Cycling and Transformation. In Treatise on Estuarine and Coastal Science, Volume 5, pp. 7–67. Elsevier.

Besemer, K., Luef, B., Preiner, S., Eichberger, B., Agis, M., Peduzzi, P., 2009. Sources and composition of organic matter for bacterial growth in a large European river floodplain system (Danube, Austria). Organic Geochemistry 40(3), 321–331.

Billen, G., Garnier, J., Hanset, P., 1994. Modelling phytoplankton development in whole drainage networks: the RIVERSTRAHLER model applied to the Seine River system. Hydrobiologia (289), 119–137. https://doi.org/10.1007/978-94-017-2670-2_11.

Billen, G., Garnier, J., Mouchel, J.-M., Silvestre, M., 2007. The Seine system: Introduction to a multidisciplinary approach of the functioning of a regional river system. Science of The Total Environment, Human activity and material fluxes in a regional river basin: the Seine River watershed 375, 1–12. https://doi.org/10.1016/j.scitotenv.2006.12.001

Calderó-Pascual, M., Yıldız, D., Yalçın, G., Metin, M., Yetim, S., Fiorentin, C., Andersen, M.R., Jennings, E., Jeppesen, E., Ger, K.A., Beklioğlu, M., McCarthy, V., 2022. The importance of allochthonous organic matter quality when investigating pulse disturbance events in freshwater lakes: a mesocosm experiment. Hydrobiologia 849, 3905–3929.

Carlson, C.A., Hansell, D.A., 2015. DOM Sources, Sinks, Reactivity, and Budgets, in: Biogeochemistry of Marine Dissolved Organic Matter. Elsevier, pp. 65–126. https://doi.org/10.1016/B978-0-12-405940-5.00003-0

Carstea, E.M., Baker, A., Bieroza, M., Reynolds, D., 2010. Continuous fluorescence excitation–emission matrix monitoring of river organic matter. Water Res 44, 5356–5366.

Catalán, N., Pastor, A., Borrego, C.M., Casas-Ruiz, J.P., Hawkes, J.A., Gutiérrez, C., Schiller, D., Marcé, R., 2021. The relevance of environment vs. composition on dissolved organic matter degradation in freshwaters. Limnol Oceanogr 66, 306–320.

Cawley, K. M., Butler, K. D., Aiken, G. R., Larsen, L. G., Huntington, T. G., McKnight, D. M., 2012. Identifying fluorescent pulp mill effluent in the Gulf of Maine and its watershed. Marine Pollution Bulletin 64(8), 1678–1687.

Chai, L., Huang, M., Fan, H., Wang, J., Jiang, D., Zhang, M., Huang, Y., 2019. Urbanization altered regional soil organic matter quantity and quality: Insight from excitation emission matrix (EEM) and parallel factor analysis (PARAFAC). Chemosphere 220, 249–258.

Cole, J.J., Prairie, Y.T., Caraco, N.F., McDowell, W.H., Tranvik, L.J., Striegl, R.G., Duarte, C.M., Kortelainen, P., Downing, J.A., Middelburg, J.J., Melack, J., 2007. Plumbing the Global Carbon Cycle: Integrating Inland Waters into the Terrestrial Carbon Budget. Ecosystems 10, 172–185. https://doi.org/10.1007/s10021-006-9013-8

Cory, R. M., McKnight, D. M., 2005. Fluorescence spectroscopy reveals ubiquitous presence of oxidized and reduced quinones in dissolved organic matter. Environmental Science & Technology 39(21), 8142–8149.

Cory, R. M., Boyer, E. W., McKnight, D. M., 2011. Spectral Methods to Advance Understanding of Dissolved Organic Carbon Dynamics in Forested Catchments, pp. 117–135. Dordrecht: Springer Netherlands.

Dainard, P.G., Guéguen, C., McDonald, N., Williams, W.J., 2015. Photobleaching of fluorescent dissolved organic matter in Beaufort Sea and North Atlantic Subtropical Gyre. Marine Chemistry 177, 630–637.

DeFrancesco, C., Guéguen, C., 2021. Long-term trends in dissolved organic matter composition and its relation to sea ice in the Canada Basin, Arctic Ocean (2007–2017). Journal of Geophysical Research: Oceans 126(2).

Derrien, M., Brogi, S.R., Gonçalves-Araujo, R., 2019. Characterization of aquatic organic matter: Assessment, perspectives and research priorities. Water Research 163, 114908. https://doi.org/10.1016/j.watres.2019.114908

Drake, T.W., Raymond, P.A., Spencer, R.G.M., 2018. Terrestrial carbon inputs to inland waters: A current synthesis of estimates and uncertainty. Limnology and Oceanography Letters 3, 132–142. https://doi.org/10.1002/lol2.10055

Du, Y., Lu, Y., Roebuck, J.A., Liu, D., Chen, F., Zeng, Q., Xiao, K., He, H., Liu, Z., Zhang, Y., Jaffé, R., 2021. Direct versus indirect effects of human activities on dissolved organic matter in highly impacted lakes. Science of The Total Environment 752, 141839.

Eder, A., Weigelhofer, G., Pucher, M., Tiefenbacher, A., Strauss, P., Brandl, M., Blöschl, G., 2022. Pathways and composition of dissolved organic carbon in a small agricultural catchment during base flow conditions. Ecohydrology & Hydrobiology 22, 96–112.

https://doi.org/10.5194/essd-2025-375
Preprint. Discussion started: 4 March 2026

Ejarque, E., Freixa, A., Vazquez, E., Guarch, A., Amalfitano, S., Fazi, S., Romaníe, A.M., Butturini, A., 2017. Quality and reactivity of dissolved organic matter in a Mediterranean river across hydrological and spatial gradients. Science of the Total Environment 599–600, 1802–1812.

Fellman, J.B., Hood, E., Spencer, R.G.M., 2010. Fluorescence spectroscopy opens new windows into dissolved organic matter dynamics in freshwater ecosystems: A review. Limnol. Oceanogr 55, 2452–2462.

Flipo, N., Labadie, P., Lestel, L. (Eds.), 2021. The Seine River Basin, The Handbook of Environmental Chemistry. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-030-54260-3

Floury, P., Gaillardet, J., Tallec, G., Ansart, P., Bouchez, J., Louvat, P., Gorge, C., 2019. Chemical weathering and CO_2 consumption rate in a multilayered-aquifer dominated watershed under intensive farming: The Orgeval Critical Zone Observatory, France. Hydrological Processes 33, 195–213. https://doi.org/10.1002/hyp.13340

Friedlingstein, P., O'Sullivan, M., Jones, M.W., Andrew, R.M., Bakker, D.C.E., Hauck, J., Landschützer, P., Le Quéré, C., Luijkx, I.T., Peters, G.P., Peters, W., Pongratz, J., Schwingshackl, C., Sitch, S., Canadell, J.G., Ciais, P., Jackson, R.B., Alin, S.R., Anthoni, P., Barbero, L., Bates, N.R., Becker, M., Bellouin, N., Decharme, B., Bopp, L., Brasika, I.B.M., Cadule, P., Chamberlain, M.A., Chandra, N., Chau, T.-T.-T., Chevallier, F., Chini, L.P., Cronin, M., Dou, X., Enyo, K., Evans, W., Falk, S., Feely, R.A., Feng, L., Ford, D.J., Gasser, T., Ghattas, J., Gkritzalis, T., Grassi, G., Gregor, L., Gruber, N., Gürses, Ö., Harris, I., Hefner, M., Heinke, J., Houghton, R.A., Hurtt, G.C., Iida, Y., Ilyina, T., Jacobson, A.R., Jain, A., Jarníková, T., Jersild, A., Jiang, F., Jin, Z., Joos, F., Kato, E., Keeling, R.F., Kennedy, D., Klein Goldewijk, K., Knauer, J., Korsbakken, J.I., Körtzinger, A., Lan, X., Lefèvre, N., Li, H., Liu, J., Liu, Z., Ma, L., Marland, G., Mayot, N., McGuire, P.C., McKinley, G.A., Meyer, G., Morgan, E.J., Munro, D.R., Nakaoka, S.-I., Niwa, Y., O'Brien, K.M., Olsen, A., Omar, A.M., Ono, T., Paulsen, M., Pierrot, D., Pocock, K., Poulter, B., Powis, C.M., Rehder, G., Resplandy, L., Robertson, E., Rödenbeck, C., Rosan, T.M., Schwinger, J., Séférian, R., Smallman, T.L., Smith, S.M., Sospedra-Alfonso, R., Sun, Q., Sutton, A.J., Sweeney, C., Takao, S., Tans, P.P., Tian, H., Tilbrook, B., Tsujino, H., Tubiello, F., Van Der Werf, G.R., Van Ooijen, E., Wanninkhof, R., Watanabe, M., Wimart-Rousseau, C., Yang, D., Yang, X., Yuan, W., Yue, X., Zaehle, S., Zeng, J., Zheng, B., 2023. Global Carbon Budget 2023. Earth Syst. Sci. Data 15, 5301–5369. https://doi.org/10.5194/essd-15-5301-2023

Gaillardet, J., Braud, I., Hankard, F., Anquetin, S., Bour, O., Dorfliger, N., de Dreuzy, J. r., Galle, S., Galy, C., Gogo, S., Gourcy, L., Habets, F., Laggoun, F., Longuevergne, L., Le Borgne, T., Naaim-Bouvet, F., Nord, G., Simonneaux, V., Six, D., Tallec, T., Valentin, C., Abril, G., Allemand, P., Arènes, A., Arfib, B., Arnaud, L., Arnaud, N., Arnaud, P., Audry, S., Comte, V.B., Batiot, C., Battais, A., Bellot, H., Bernard, E., Bertrand, C., Bessière, H., Binet, S., Bodin, J., Bodin, X., Boithias, L., Bouchez, J., Boudevillain, B., Moussa, I.B., Branger, F., Braun, J.J., Brunet, P., Caceres, B., Calmels, D., Cappelaere, B., Celle-Jeanton, H., Chabaux, F., Chalikakis, K., Champollion, C., Copard, Y., Cotel, C., Davy, P., Deline, P., Delrieu, G., Demarty, J., Dessert, C., Dumont, M., Emblanch, C., Ezzahar, J., Estèves, M., Favier, V., Faucheux, M., Filizola, N., Flammarion, P., Floury, P., Fovet, O., Fournier, M., Francez, A.J., Gandois, L., Gascuel, C., Gayer, E., Genthon, C., Gérard, M.F., Gilbert, D., Gouttevin, I., Grippa, M., Gruau, G., Jardani, A., Jeanneau, L., Join, J.L., Jourde, H., Karbou, F., Labat, D., Lagadeuc, Y., Lajeunesse, E., Lastennet, R., Lavado, W., Lawin, E., Lebel, T., Le Bouteiller, C., Legout, C., Lejeune, Y., Le Meur, E., Le Moigne, N., Lions, J., Lucas, A., Malet, J.P., Marais-Sicre, C., Maréchal, J.C., Marlin, C., Martin, P., Martins, J., Martinez, J.M., Massei, N., Mauclerc, A., Mazzilli, N., Molénat, J., Moreira-Turcq, P., Mougin, E., Morin, S., Ngoupayou, J.N., Panthou, G., Peugeot, C., Picard, G., Pierret, M.C., Porel, G., Probst, A., Probst, J.L., Rabatel, A., Raclot, D., Ravanel, L., Rejiba, F., René, P., Ribolzi, O., Riotte, J., Rivière, A., Robain, H., Ruiz, L., Sanchez-Perez, J.M., Santini, W., Sauvage, S., Schoeneich, P., Seidel, J.L., Sekhar, M., Sengtaheuanghoung, O., Silvera, N., Steinmann, M., Soruco, A., Tallec, G., Thibert, E., Lao, D.V., Vincent, C., Viville, D., Wagnon, P., Zitouna, R., 2018. OZCAR: The French Network of Critical Zone Observatories. Vadose Zone Journal 17, 180067. https://doi.org/10.2136/vzj2018.04.0067

Gao, Z., Guéguen, C., 2017. Size distribution of absorbing and fluorescing DOM in Beaufort Sea, Canada Basin. Deep Sea Research Part I: Oceanographic Research Papers 121, 30–37. https://doi.org/10.1016/j.dsr.2016.12.014

Garnier, J., Billen, G., 1994. Ecological interactions in a shallow sand-pit lake (Lake Créteil, Parisian Basin, France): a modelling approach. Hydrobiologia 275/276, 97–114. https://doi.org/10.1007/978-94-017-2460-9_9.

Garnier, J., Billen, G., Vilain, G., Benoit, M., Passy, P., Tallec, G., Tournebize, J., Anglade, J., Billy, C., Mercier, B., Ansart, P., Azougui, A., Sebilo, M., Kao, C., 2014. Curative vs. preventive management of nitrogen transfers in rural areas:

https://doi.org/10.5194/essd-2025-375
Preprint. Discussion started: 4 March 2026

Lessons from the case of the Orgeval watershed (Seine River basin, France). Journal of Environmental Management 144, 125–134. https://doi.org/10.1016/j.jenvman.2014.04.030

Garnier, J., Guillon, S., Hénine, H., Billen, G., Escoffier, N., Mercier, B., Martinez A., Mouchel, J. M., 2025. Sources and fate of CO_2 along the soil–aquifer–stream–atmosphere continuum (the Orgeval headwater catchment, France). Geoderma, 458, 117297. https://doi.org/10.1016/j.geoderma.2025.117297

Graeber, D., Gelbrecht, J., Pusch, M.T., Anlanger, C., Von Schiller, D., 2012. Agriculture has changed the amount and composition of dissolved organic matter in Central European headwater streams. Science of The Total Environment 438, 435–446.

Groeneveld, M., Catalán, N., Attermeyer, K., Hawkes, J., Einarsdóttir, K., Kothawala, D., Bergquist, J., Tranvik, L., 2020. Selective Adsorption of Terrestrial Dissolved Organic Matter to Inorganic Surfaces Along a Boreal Inland Water Continuum. JGR Biogeosciences 125, e2019JG005236.

Hambly, A.C., Arvin, E., Pedersen, L.-F., Pedersen, P.B., Seredyńska-Sobecka, B., Stedmon, C.A., 2015. Characterising organic matter in recirculating aquaculture systems with fluorescence EEM spectroscopy. Water Research 83, 112–120. https://doi.org/10.1016/j.watres.2015.06.037

Helms, J.R., Stubbins, A., Ritchie, J.D., Minor, E.C., Kieber, D.J., Mopper, K., 2008. Absorption spectral slopes and slope ratios as indicators of molecular weight, source, and photobleaching of chromophoric dissolved organic matter. Limnol. Oceanogr 53, 955–969.

Huguet, A., Vacher, L., Relexans, S., Saubusse, S., Froidefond, J.-M., Parlanti, E., 2009. Properties of fluorescent dissolved organic matter in the Gironde Estuary. Org. Geochem 40, 706.

Jaffé, R., McKnight, D., Maie, N., Cory, R., McDowell, W.H., Campbell, J.L., 2008. Spatial and temporal variations in DOM composition in ecosystems: The importance of long-term monitoring of optical properties. J. Geophys. Res. 113, 2008JG000683.

Jaffé, R., Cawley, K.M., Yamashita, Y., 2014. Applications of Excitation Emission Matrix Fluorescence with Parallel Factor Analysis (EEM-PARAFAC) in Assessing Environmental Dynamics of Natural Dissolved Organic Matter (DOM) in Aquatic Environments: A Review. Advances in the Physicochemical Characterization of Dissolved Organic Matter: Impact on Natural and Engineered Systems. Chapter 3, 27–73.

Jutaporn, P., Armstrong, M.D., Coronell, O., 2020. Assessment of C-DBP and N-DBP formation potential and its reduction by MIEX® DOC and MIEX® GOLD resins using fluorescence spectroscopy and parallel factor analysis. Water Research 172, 115460. https://doi.org/10.1016/j.watres.2019.115460

Kurek, M.R., Frey, K.E., Guillemette, F., Podgorski, D.C., Townsend-Small, A., Arp, C.D., Kellerman, A.M., Spencer, R.G.M., 2022. Trapped Under Ice: Spatial and Seasonal Dynamics of Dissolved Organic Matter Composition in Tundra Lakes. JGR Biogeosciences 127, e2021JG006578. https://doi.org/10.1029/2021JG006578

Lambert, T., Bouillon, S., Darchambeau, F., Morana, C., Roland, F. A. E., Descy, J. P., Borges, A. V., 2017. Effects of human land use on the terrestrial and aquatic sources of fluvial organic matter in a temperate river basin (The Meuse River, Belgium). Biogeochemistry 136.

Lestel, L., Eschbach, D., Steinmann, R., Gastaldi, N., 2018. ArchiSEINE : une approche géohistorique du bassin de la Seine, Fascicule #18 du PIREN-Seine, ISBN : 978-2-490463-07-7, ARCEAU-IdF, 64 p.

McKnight, D.M., Boyer, E.W., Westerhoff, P.K., Doran, P.T., Kulbe, T., Andersen, D.T., 2001. Spectrofluorometric characterization of dissolved organic matter for indication of precursor organic material and aromaticity. Limnol. Oceanogr 46, 38–48.

Marcé, R., Verdura, L., Leung, N., 2021. Dissolved organic matter spectroscopy reveals a hot spot of organic matter changes at the river–reservoir boundary. Aquat Sci 83, 67.

Marescaux, A., Thieu, V., Gypens, N., Silvestre, M., Garnier, J., 2020. Modeling inorganic carbon dynamics in the Seine River continuum in France. Hydrology and Earth System Sciences, 24(5), 2379-2398. https://doi.org/10.5194/hess-24-2379-2020

Maurischat, P., Lehnert, L., Zerres, V.H.D., Tran, T.V., Kalbitz, K., Rinnan, Å., Li, X.G., Dorji, T., Guggenberger, G., 2022. The glacial–terrestrial–fluvial pathway: A multiparametrical analysis of spatiotemporal dissolved organic matter variation in three catchments of Lake Nam Co, Tibetan Plateau. Science of The Total Environment 838, 156542. https://doi.org/10.1016/j.scitotenv.2022.156542

https://doi.org/10.5194/essd-2025-375
Preprint. Discussion started: 4 March 2026

Meilleur, C., Kamula, M., Kuzyk, Z.A., Guéguen, C., 2023. Insights into surface circulation and mixing in James Bay and Hudson Bay from dissolved organic matter optical properties. Journal of Marine Systems 238, 103841.

Menendez, A., Tzortziou, M., 2024. Driving factors of colored dissolved organic matter dynamics across a complex urbanized estuary. Science of The Total Environment 921, 171083. https://doi.org/10.1016/j.scitotenv.2024.171083

Minor, E.C., Swenson, M.M., Mattson, B.M., Oyler, A.R., 2014. Structural characterization of dissolved organic matter: a review of current techniques for isolation and analysis. Environ. Sci.: Processes Impacts 16, 2064–2079. https://doi.org/10.1039/C4EM00062E

Moona, N., Holmes, A., Wünsch, U. J., Pettersson, T. J. R., Murphy, K. R., 2021. Full-scale manipulation of the empty bed contact time to optimize dissolved organic matter removal by drinking water biofilters. ACS ES&T Water 1(5), 1117–1126.

Mouhri, A., Flipo, N., Rejiba, F., de Fouquet, C., Bodet, L., Kurtulus, B., Tallec, G., Durand, V., Jost, A., Ansart, P., Goblet, P., 2013. Designing a multi-scale sampling system of stream–aquifer interfaces in a sedimentary basin. Journal of Hydrology 504, 194–206. https://doi.org/10.1016/j.jhydrol.2013.09.036

Murphy, K.R., Hambly, A., Singh, S., Henderson, R.K., Baker, A., Stuetz, R., Khan, S.J., 2011. Organic Matter Fluorescence in Municipal Water Recycling Schemes: Toward a Unified PARAFAC Model. Environ. Sci. Technol. 45, 2909–2916. https://doi.org/10.1021/es103015e

Murphy, K. R., Stedmon, C. A., Graeber, D., Bro, R., 2013. Fluorescence spectroscopy and multi-way techniques. PARAFAC. Analytical Methods 5(23), 6557.

Murphy, K.R., Bro, R., Stedmon, C.A., 2014a. Chemometric Analysis of Organic Matter Fluorescence, in: Coble, P.G., Lead, J., Baker, A., Reynolds, D.M., Spencer, R.G.M. (Eds.), Aquatic Organic Matter Fluorescence. Cambridge University Press, pp. 339–375.

Murphy, K.R., Stedmon, C.A., Wenig, P., Bro, R., 2014b. OpenFluor– an online spectral library of auto-fluorescence by organic compounds in the environment. Anal. Methods 6, 658.

Ouyang, T., McKenna, A.M., Wozniak, A.S., 2024. Storm-driven hydrological, seasonal, and land use/land cover impact on dissolved organic matter dynamics in a mid-Atlantic, USA coastal plain river system characterized by 21 T FT-ICR mass spectrometry. Front. Environ. Sci. 12, 1379238. https://doi.org/10.3389/fenvs.2024.1379238

Paradina-Fernández, L., Wünsch, U., Bro, R., Murphy, K., 2023. Direct Measurement of Organic Micropollutants in Water and Wastewater Using Fluorescence Spectroscopy. ACS EST Water 3, 3905–3915. https://doi.org/10.1021/acsestwater.3c00323

Parlanti, E., Worz, K., Geoffroy, L., Lamotte, M., 2000. Dissolved organic matter fluorescence spectroscopy as a tool to estimate biological activity in a coastal zone submitted to anthropogenic inputs. Org. Geochem 31, 1765–1781.

Peleato, N. M., Sidhu, B. S., Legge, R. L., Andrews, R. C., 2017. Investigation of ozone and peroxone impacts on natural organic matter character and biofiltration performance using fluorescence spectroscopy. Chemosphere 172, 225–233.

Podgorski, D.C., Zito, P., McGuire, J.T., Martinovic-Weigelt, D., Cozzarelli, I.M., Bekins, B.A., Spencer, R.G.M., 2018. Examining Natural Attenuation and Acute Toxicity of Petroleum-Derived Dissolved Organic Matter with Optical Spectroscopy. Environ. Sci. Technol. 52, 6157–6166. https://doi.org/10.1021/acs.est.8b00016

Raymond, P.A., Hartmann, J., Lauerwald, R., Sobek, S., McDonald, C., Hoover, M., Butman, D., Striegl, R., Mayorga, E., Humborg, C., Kortelainen, P., Dürr, H., Meybeck, M., Ciais, P., Guth, P., 2013. Global carbon dioxide emissions from inland waters. Nature 503, 355–359. https://doi.org/10.1038/nature12760

Regnier, P., Resplandy, L., Najjar, R.G., Ciais, P., 2022. The land-to-ocean loops of the global carbon cycle. Nature 603, 401–410. https://doi.org/10.1038/s41586-021-04339-9

Servais, P., Garnier, J., Demarteau, N., Brion, N., Billen, G., 1999. Supply of organic matter and bacteria to aquatic ecosystems through waste water effluents. Water Research 33, 3521–3531. https://doi.org/10.1016/S0043-1354(99)00056-1

Stedmon, C. A., Markager, S., 2005. Resolving the variability in dissolved organic matter fluorescence in a temperate estuary and its catchment using PARAFAC analysis. Limnology and Oceanography 50(2), 686–697.

Stedmon, C., Bro, R., 2008. Characterizing dissolved organic matter fluorescence with parallel factor analysis: a tutorial. Limnol. Ocean. Methods 6, 572–579.

Tusseau-Vuillemin, M. H., Le Reveillé, G., 2001. Le carbone organique biodégradable dans les eaux traitées des stations d'épuration du bassin de la seine. Ingénieries eau-agriculture-territoires 25, 3–12.

https://doi.org/10.5194/essd-2025-375
Preprint. Discussion started: 4 March 2026

808 Tzortziou, M., Zeri, C., Dimitriou, E., Ding, Y., Jaffé, R., Anagnostou, E., Pitta, E., Mentzafou, A., 2015. Colored dissolved
809 organic matter dynamics and anthropogenic influences in a major transboundary river and its coastal wetland. Limnol.
810 Oceanogr 60, 1222–1240.
811 Wang, S., Flipo, N., Romary, T., Hasanyar, M., 2022. Particle filter for high frequency oxygen data assimilation in river
812 systems. Environmental Modelling & Software, 151:105382. ISSN 1364-8152.
813 https://doi.org/10.1016/j.envsoft.2022.105382
814 Wang, S., Thieu, V., Billen, G., Garnier, J., Silvestre, M., Marescaux, A., Yan, X., Flipo, N., 2024. The community-centered
815 freshwater biogeochemistry model unified RIVE v1. 0: a unified version for water column. Geoscientific Model
816 Development, 17(1), 449-476. https://doi.org/10.5194/gmd-17-449-2024
817 Weishaar, G., Aiken, R., Bergamschi, B.A., Fram, M.S., Fugii, R., Mopper, K. 2003. Evaluation of Specific Ultraviolet
818 Absorbance as an Indicator of the Chemical Composition and Reactivity of Dissolved Organic Carbon.
819 Environmental Science and Technology, 37(20), 4702-4708, doi:10.1021/es030360x.
820 Wiegner, T. N., Tubal, R. L., MacKenzie, R.A., 2009. Bioavailability and export of dissolved organic matter from a tropical
821 river during base- and stormflow conditions. Limnology and Oceanography, 54(4):1233–1242. ISSN 00243590. doi:
822 10.4319/lo.2009.54.4.1233.
823 Wünsch, U.J., Murphy, K., 2021. A simple method to isolate fluorescence spectra from small dissolved organic matter datasets.
824 Water Research 190, 116730. https://doi.org/10.1016/j.watres.2020.116730
825 Yamashita, Y., Kloeppel, B.D., Knoepp, J., Zausen, G.L., Jaffé, R., 2011. Effects of Watershed History on Dissolved Organic
826 Matter Characteristics in Headwater Streams. Ecosystems 14, 1110–1122.
827 Yan, Y., Lauerwald, R., Wang, X., Regnier, P., Ciais, P., Ran, L., Gao, Y., Huang, L., Zhang, Y., Duan, Z., Papa, F., Yu, B.,
828 Piao, S., 2023. Increasing riverine export of dissolved organic carbon from China. Global Change Biology 29, 5014–
829 5032. https://doi.org/10.1111/gcb.16819
830 Yang, L., Chen, W., Zhuang, W.-E., Cheng, Q., Li, W., Wang, H., Guo, W., Chen, C.-T.A., Liu, M., 2019. Characterization
831 and bioavailability of rainwater dissolved organic matter at the southeast coast of China using absorption spectroscopy
832 and fluorescence EEM-PARAFAC. Estuarine, Coastal and Shelf Science 217, 45–55.
833 https://doi.org/10.1016/j.ecss.2018.11.002
834 Yu, H., Liang, H., Qu, F., Han, Z., Shao, S., Chang, H., Li, G., 2015. Impact of dataset diversity on accuracy and sensitivity
835 of parallel factor analysis model of dissolved organic matter fluorescence excitation-emission matrix. Sci Rep 5,
836 10207. https://doi.org/10.1038/srep10207
837 Zhang, H., Lauerwald, R., Regnier, P., Ciais, P., Van Oost, K., Naipal, V., Guenet, B., Yuan, W., 2022. Estimating the lateral
838 transfer of organic carbon through the European river network using a land surface model. Earth Syst. Dynam. 13,
839 1119–1144. https://doi.org/10.5194/esd-13-1119-2022
840 Zhuang, W.-E., Yang, L., 2018. Impacts of global changes on the biogeochemistry and environmental effects of dissolved
841 organic matter at the land-ocean interface: a review. Environ Sci Pollut Res 25, 4165–4173.
842 https://doi.org/10.1007/s11356-017-1027-6
843 Zhuang, W.-E., Chen, W., Cheng, Q., Yang, L., 2021. Assessing the priming effect of dissolved organic matter from typical
844 sources using fluorescence EEMs-PARAFAC. Chemosphere 264, 128600.
845 https://doi.org/10.1016/j.chemosphere.2020.128600
846 Zsolnay, A., Baigar, E., Jimenez, M., Steinweg, B., Saccomandi, F., 1999. Differentiating with fluorescence spectroscopy the
847 sources of dissolved organic matter in soils subjected to drying. Chemosphere 38, 45–50.