



**HAL**  
open science

## Questioning calculation and interpretation of fluorescence indices in natural waters studies

Leïla Serène, Naomi Mazzilli, Christelle Batiot-Guilhe, Christophe Emblanch, Marina Gillon, Milanka Babic, Julien Dupont, Roland Simler, Matthieu Blanc

### ► To cite this version:

Leïla Serène, Naomi Mazzilli, Christelle Batiot-Guilhe, Christophe Emblanch, Marina Gillon, et al.. Questioning calculation and interpretation of fluorescence indices in natural waters studies. *Journal of Hydrology*, 2025, 650, pp.132524. 10.1016/j.jhydrol.2024.132524 . hal-04844802

**HAL Id: hal-04844802**

**<https://hal.science/hal-04844802v1>**

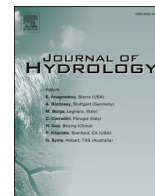
Submitted on 18 Dec 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



## Research papers

## Questioning calculation and interpretation of fluorescence indices in natural waters studies

Leïla Serène<sup>a,b,\*</sup> , Naomi Mazzilli<sup>a</sup> , Christelle Batiot-Guilhe<sup>b</sup> , Christophe Emblanch<sup>a</sup>, Marina Gillon<sup>a</sup>, Milanka Babic<sup>a</sup>, Julien Dupont<sup>a</sup> , Roland Simler<sup>a</sup>, Matthieu Blanc<sup>c</sup>

<sup>a</sup> UMR 1114 EMMAH (AU-INRAE), Université d'Avignon, 84000 Avignon, France

<sup>b</sup> HSM, Univ Montpellier, CNRS, IRD, Montpellier, France

<sup>c</sup> Independent Researcher, Montpellier, France

## A B S T R A C T

Fluorescence indices are particularly well adapted to characterize dissolved organic matter. In environmental studies, the three most often used fluorescence indices are the Fluorescence index (FI), the Humification index (HIX) and the Biological index (BIX). The present paper aims to discuss 4 limitations in the actual use of fluorescence indices: (i) difference between indices values calculated from interpolated and non-interpolated excitation wavelengths; (ii) misfit between the areas and wavelengths from the index definition (HIX, FI, BIX) and the samples (peaks location and curve shape) related to environmental context (e.g. pH, living conditions of organic matter decomposers); (iii) the potential to rethink the 2D-based indices thanks to 3D matrices; (iv) the incidence of recent improvement in organic matter degradation knowledge on the fluorescence indices interpretation. These limitations are illustrated with different types of water samples coming from three different open access datasets: karst groundwaters from Serene et al. (2024), marine waters from Drozdova et al. (2019) and surface water from Sgroi et al. (2019). These results draw the attention of fluorescence indices users to unsupervised calculation and interpretation of 2D-based fluorescence indices. It also greatly encourages the rethinking of fluorescence indices using EEM (Excitation Emission Matrices) in order to adapt the calculation to each purpose and study context and to limit the generalisation of indices interpretation to organic matter structure.

## 1. Introduction

## 1.1. Fluorescence of natural organic matter in environmental studies

Organic molecules that are able to emit a fluorescent signal after being excited are composed of several conjugated bonds or aromatic nuclei (Coble et al., 2014). The required excitation wavelength for molecule to emit a fluorescence signal is longer for complex molecules than simple ones, allowing the distinction of organic matter compounds (Ewald et al., 1988; Zsolnay et al., 1999). Fluorescent compounds are defined thanks to the excitation wavelength ( $\lambda_{ex}$ ) required to excite the molecule and the emission wavelength ( $\lambda_{em}$ ) of the signal emitted by the molecule.

In aquatic environments 4 types of fluorescent compounds are observed, listed from higher to lower emission wavelength and therefore to most complex to most simpler compounds (Coble et al., 2014):

- Humic-like C compounds have the greater emission wavelength and are the most complex because they are composed of lignin degradation residues. Because of their precursors deriving from lignin,

humic-like C presence in water is related to soil influence. This family is composed of 2 peaks named A<sub>C</sub> ( $\lambda_{ex}$  = 260 nm and  $\lambda_{em}$  = [400 to 460] nm) and C ( $\lambda_{ex}$  = [320 to 365] nm and  $\lambda_{em}$  = [420 to 470] nm);

- Humic-like M are complex compounds that result from the aggregation of microbial products. This compounds family is also composed of 2 peaks: A<sub>M</sub> ( $\lambda_{ex}$  = 240 nm and  $\lambda_{em}$  = [350 to 400] nm) and M ( $\lambda_{ex}$  = [290 to 310] nm and  $\lambda_{em}$  = [370 to 420] nm);
- Tryptophan correspond to small and simple organic compounds that directly derive from microbial activity. This family presence is related to the observation of peak T ( $\lambda_{ex}$  = 275 nm and  $\lambda_{em}$  = 340 nm) or A<sub>T</sub> ( $\lambda_{ex}$  = 230 nm and  $\lambda_{em}$  = 340 nm);
- As Tryptophan, Tyrosine is a small and simple molecule observable at two locations: peaks B ( $\lambda_{ex}$  = 275 nm and  $\lambda_{em}$  = 305 nm) and A<sub>B</sub> ( $\lambda_{ex}$  = 230 nm and  $\lambda_{em}$  = 305 nm).

These peaks can be observed in 2D spectra performed at a fixed excitation wavelength, plotting the fluorescence intensity (y-axis) against the emission wavelength (x-axis). 2D spectra can be combined into 3D matrices called Excitation Emission Matrices (EEM), where the x

\* Corresponding author at: UMR 1114 EMMAH (AU-INRAE), Université d'Avignon, 84000 Avignon, France.

E-mail address: [leila.serene@univ-avignon.fr](mailto:leila.serene@univ-avignon.fr) (L. Serène).

and y axes represent the excitation and emission wavelengths and a colour scale represents the fluorescence intensity. Finally, PARAFAC modelling can be performed thanks to an important number of EEM to extract the different compounds present in the set of EEM (Murphy et al., 2013).

In order to facilitate the fluorescence of natural organic matter interpretation in both soil and water, several fluorescence indices were developed (Gabor et al., 2014). The most widely used are the Biological index (BIX, Huguet et al., 2009), the Fluorescence index (FI, McKnight et al., 2001) and the Humification index (HIX, Zsolnay et al., 1999). These three fluorescence indices are most frequently interpreted in combination, with the aim to characterize dissolved organic matter without analytical separation of specific compounds. Indeed, they allow the qualitative assessment of fluorescent compounds sources (BIX: proportion of recently produced organic matter; FI: humic-like C having soil or microbial precursors) and of the aromaticity that has been usually interpreted as the decomposition/humification state (HIX). These parameters vary from one water type to the other according to biological activity, organic matter sources (anthropogenic / natural, soil organic matter derived/ microbiological derived) and residence time (degradation and mineralisation of organic matter). Applications of these indices cover many fields of research. For example, in the field of marine biology, Nelson et al. (2015) show that the signature of submarine springs of Hawaii and other Pacific high islands is depleted in dissolved organic matter and enriched in nutrients, which indicates that groundwater is a key parameter to algal and benthic coral development of nearshore reefs. In hydrogeology, fluorescence indices are often used to assess the contamination of water by organic pollutants. For example, they are used to study the migration and evolution of contamination from livestock farming (Zhang et al., 2021) and from landfill leachates (Jiang et al., 2019) as this contamination enhances microbially derived organic matter and degree of aromaticity. Organic matter is also a key parameter to study non-organic anthropogenic contaminants as it can easily adsorb and desorb contaminants, and therefore facilitate their transport (Shand et al., 2007). Mobilization of organic colloids in a cave was thereby studied by (Liao et al., 2021) with the view to preserve karst groundwater from anthropogenic contamination. Trace metals such as Arsenic are very likely to be sorbed on organic matter, and as organic matter they are particularly sensitive to biological activity which plays an important role in trace metals mobilization (Guo et al., 2019; Mladenov et al., 2010). Therefore, numerous studies also use fluorescence indices in order to assess and understand groundwater contamination by trace metals (e.g. Guo et al., 2019; Kulkarni et al., 2017; Schittich et al., 2018; Yang et al., 2020).

### 1.2. Definition of the indices

The Humification index HIX is a commonly used parameter in soil science. Humification corresponds to the complexation of organic matter. This process increases the number of conjugated bonds and aromatic nuclei, and therefore lengthens the emission wavelength (Duchaufour et al., 2020). Zsolnay et al. (1999) proposed a way to assess HIX using a 2D spectra at 254 nm excitation wavelength as it was identified to show the higher fluorescence intensities. HIX is thus defined as the ratio of the integral under the intensity curve of emission wavelength corresponding to heavy organic matter (peak C) to the one corresponding to light organic matter (peaks B and T). Even though this method of calculation was calibrated on soil waters, Zsolnay et al. (1999) definition is the most frequently used in water studies (Gabor et al., 2014).

The Fluorescence index FI was first proposed by McKnight et al. (2001) in order to identify the different precursors of humic-like C compounds. It is based on the fact that humic-like substances contain residues of lignin degradation that have a higher carbon to nitrogen ratio and higher aromaticity and therefore longer emission wavelengths than those derived from microbial material (Aiken et al., 1996; McKnight et al., 1994). As HIX, FI is calculated from a 2D spectra

(excitation 370 nm). It corresponds to the ratio of intensity at maximal and minimal emissions wavelengths of peak C (humic-like C compound). It aims to assess whether peak C is shifted towards higher (contain lignin residues) or shorter (contain microbial products) emission wavelengths. This index was calibrated on stream and rivers waters.

The biological index (BIX, initially named freshness index) was developed to identify microbial influence on dissolved organic matter in marine environments (Parlanti et al., 2000). According to these authors, peaks M and B correspond to autochthonous N rich organic compounds produced by biological activity in marine water, while peak C corresponds to allochthonous compounds produced in soil from fresh organic matter such as lignin provided by estuarine waters. The ratio M/C is thus considered as a proxy of the relative quantity of organic matter recently produced by microbial activity in marine and estuarine waters. Low values of M/C ratio are interpreted as resulting from relatively low autochthonous organic matter production, while high values are supposed to stem from relatively high biological activity resulting in greater autochthonous production. Huguet et al. (2009) proposed to calculate BIX thanks to a 2D spectra (excitation wavelength 310 nm) using the ratio of intensities at emission wavelengths corresponding to peaks M and C. Even if peak M is rarely found in natural waters (Coble et al., 2014), BIX is often used in such environments (Cao et al., 2023; Jiang et al., 2019; Li et al., 2019).

### 1.3. Fluorescent indices limits

The present study aims to discuss and illustrate the limitations in the actual use of fluorescence indices. 4 potential limitations are identified.

- (1) Many studies calculate one or three of the fluorescence indices (FI, BIX, HIX) without specifying the metrics (D'Andrilli et al., 2022). However, most of excitation-emission matrices (EEM) have an excitation wavelength step from 3 to 10 nm, that may avoid the required excitation wavelength. To calculate the indices, the required excitation wavelengths may be interpolated. Therefore, the first question investigated in this paper is: does the interpolation affect the calculated indices?
- (2) Serène et al. (2022) show the mismatch of the HIX (Zsolnay et al., 1999) excitation and emission areas with a karst groundwater dataset. This inconsistency was attributed to the application of the method to a different aquatic environment (groundwater) than the one for which it was developed (soil water). The second question investigated in this paper is: Is such mismatch observed for other aquatic environments and for the two other fluorescence indices? If so, where does it come from?
- (3) Thanks to the advent of 3D EEM, recent studies have developed their own fluorescent indices inspired by HIX, BIX and FI (Li et al., 2019; Serène et al., 2022; Yang et al., 2020). Third question: Are the 3D-based fluorescence indices better than the 2D-based indices?
- (4) 16 years after the development of HIX, the Humification model has been reconsidered (Lehmann and Kleber, 2015). The final question investigated in this paper is whether the improved knowledge of organic matter degradation since the first development of fluorescence indices influences their interpretation.

## 2. Datasets and method

### 2.1. Data used

Three open access datasets of several types of water are considered in this paper in order to illustrate the discussion about the limitations of fluorescent indices (Table 1). They were chosen for their accessibility and because they represent various aquatic environments including, surface, marine and underground waters.

The karst groundwaters dataset (Serène et al., 2024) consists in 278

**Table 1**  
Summary of the datasets used in this paper.

Types of water	Samples location	Number of samples used	Reference	Types of data
Karst groundwaters (springs, unsaturated zone flows, loss)	Vaucluse, France	278	Serene et al. (2024)	<ul style="list-style-type: none"> <li>EEM at <math>\lambda_{ex} = [220: 450]</math>, 10 nm interval and <math>\lambda_{em} = [250: 550]</math> with 1 nm interval;</li> <li>2D spectra at <math>\lambda_{ex} = 254</math> nm and <math>\lambda_{em} = [220: 530]</math> with 1 nm interval.</li> </ul>
Surface waters (artificial lake) contaminated by wastewater	Sicily, Italy	16	Sgroi et al. (2019)	EEM at $\lambda_{ex} = [220: 450]$ , 5 nm interval and $\lambda_{em} = [250: 580]$ with 1 nm interval.
Marine waters	Laptev and East Siberian seas, north Russia	2	Drozdova et al. (2019)	2D spectra for: HIX: $\lambda_{ex} = 254$ nm, $\lambda_{em} = [300:480]$ nm; BIX: $\lambda_{ex} = 310$ nm, $\lambda_{em} = [370:550]$ nm; FI: $\lambda_{ex} = 370$ nm, $\lambda_{em} = [430:550]$ nm.

water samples taken from June 2020 to October 2021 corresponding to i) karst groundwaters from the Fontaine de Vaucluse observatory belonging to SNO KARST (St Trinit, Millet, and Nesque springs, Fontaine de Vaucluse main outlet) and unsaturated zone flows accessed thanks to the LSBB galleries (<https://lsbb.cnrs.fr>), and ii) at Nesque losses. These waters represent a panel of karst waters with long and short transit time (long for Fontaine de Vaucluse outlet, short for unsaturated zone flows), draining different carbonated rocks (Cretaceous for all the springs except Nesque spring and losses draining Tertiary carbonates) and with different vulnerability to anthropogenic activities (no influence for unsaturated zone flows, important influence for St Trinit spring).

EEM (Excitation-emission matrices) were performed on unfiltered water samples at the laboratory of HydroSciences Montpellier thanks to a SHIMADZU RF-5301 PC spectrofluorimeter (150 W xenon lamp). Excitation wavelengths vary from 220 to 450 nm with 10 nm interval and emission wavelengths from 250 to 550 nm with 1 nm interval. Slit widths of 15 nm were used for the monochromators with a fast default scan speed. To complete EEM, and with the same unfiltered water, 2D spectra at 254 nm excitation wavelength and emission wavelengths from 220 to 530 nm with 1 nm interval were performed to calculate HIX according to Zsolnay et al. (1999).

The surface waters dataset consists in 16 EEM from Sgroi et al. (2019) dataset. These EEM correspond to dilutions ranging from 37.5 % sample to 100 % of one unfiltered sample of Pozzillo lake which is an artificial lake contaminated by wastewater in Sicily, Italy. A Shimadzu RF-5301PC fluorescence spectrophotometer was used to perform EEM on these samples with excitation wavelength varying from 220 to 450 nm with 5 nm interval and emission wavelength from 250 to 580 nm with 1 nm interval. Slit widths of both excitation and emission wavelength were set at 5 nm.

Finally, two samples from the marine waters dataset Drozdova et al. (2019) were selected to illustrate the discussion about fluorescence indices on marine water samples. Samples 5627\_1 (sampling date: 17 September 2017) and 5612\_47 (sampling date 8 September 2017) were selected as they correspond to two water from different depth (5627\_1 sample was taken at 1 m depth and 5612\_47 at 47 m depth), two different seas from north Russia: Laptev (5627\_1) and East Siberian (5612\_47) seas, and two different distances from the continent (5627\_1 was sampled in the estuary and 5627\_47 several km away from the coast). Filtered samples (0.7  $\mu\text{m}$ ) were analyzed with 2D spectra performed at the required excitation wavelength to calculate HIX, BIX and FI, respectively 254, 310 and 370 nm, thanks to Fluorat-02-Panorama spectrofluorometer (Lumex Instruments, Drozdova et al., 2019).

## 2.2. Method

In order to explore the first limit exposed in section 1.3, a comparison between HIX, BIX and FI values obtained with and without the interpolation of the required excitation wavelength is done in section 3. For HIX, we used only the Serene et al. (2024) dataset on karst waters as none of the other ones have both the 254 nm excitation wavelength (to

calculate the index according to (Zsolnay et al., 1999 definition) and EEM (to interpolate the required excitation wavelength). For BIX and FI, only the data from Sgroi et al. (2019) dataset can be used as marine water samples from Drozdova et al. (2019) does not allow the interpolation because EEM were not performed; and since the EEM from the Serene et al. (2024) dataset already contains the 310 and 370 nm excitation wavelengths, this must be deleted to obtain interpolated wavelength values, setting the excitation wavelength range from 10 nm to 20 nm, which is too large to obtain comparative data.

Indices obtained with and without the interpolation were calculated in raw data. Non-interpolated HIX was calculated thanks to the area under the intensity curve between 435 and 480 nm emission wavelength (area named H, area close to peak C) and between 300 and 345 nm (area named L, area close to peak A<sub>T</sub>). HIX was then calculated dividing H by the sum of H and L, according to Ohno (2002) definition Zsolnay et al., (1999) definition is H/L and requires a correction that is integrated into Ohno (2002) formula. Interpolated HIX was calculated according to Ohno (2002) definition thanks to R package staRdom, using the eem\_humification\_index function (Pucher et al., 2019).

Same R package was used to calculate non-interpolated BIX and FI using the eem\_biological\_index (for BIX, according to Huguet et al., 2009 definition) and the eem\_fluorescence\_index functions (for FI, according to McKnight et al., 2001 definition). The 310 and 370 nm excitation wavelengths measurements were deleted from the EEM to allow their interpolation for BIX and FI calculation with the same staRdom function of BIX and FI calculated. The removal set the excitation wavelength interval from 5 to 10 nm which is acceptable (similar interval than Serene et al., 2024 dataset).

In order to explore the second limit exposed in section 1.3, the adequacy of fluorescence indices excitation and emission windows is assessed by the comparison of the 2D spectra with theoretical excitation wavelengths from indices definition in section 4. Raw data of the following samples are used: the two marine water samples from Drozdova et al. (2019) dataset, the Pozzillo lake sample non-diluted from Sgroi et al. (2019) dataset and a selection of the most different samples from Serene et al. (2024) karst water dataset: unsaturated zone flow AY (sampling date: 4 October 2021), St Trinit karst spring (16 November 2020), Nesque karst spring (14 April 2021) and Nesque losses (3 May 2021). The sample from surface water was not used for HIX as EEM do not contain the required 254 nm excitation wavelength.

## 3. Incidence of excitation wavelength interpolation on fluorescent indices values

Fluorescence intensity is nowadays often measured thanks to EEM that do not necessarily contain the required excitation wavelengths for HIX, BIX and FI calculation (Pucher et al., 2019). To be able to calculate fluorescence indices from EEM, missing excitation wavelengths can be interpolated using e.g. statistical tools such as staRdom package in R (Pucher et al., 2019). This interpolation may have an impact on the fluorescence indices values. To assess this potential influence, this

section compares HIX, FI and BIX values obtained with and without interpolation.

- **Comparison of interpolated and non-interpolated HIX on karst groundwaters**

Fig. 1 shows that interpolated and non-interpolated HIX are rarely similar as few points are confounded with the 1:1 line. However, most of the samples (67 %) are between the 1:1 line and the  $\pm 10\%$  difference lines, meaning that the interpolation of the wavelength causes less than 10 % difference in the HIX value. For these samples, the influence of the interpolation of the 254 nm excitation wavelength may be neglected. Few samples (12.7 %) are underestimated above  $-10\%$  in comparison with the number of samples overestimated by the interpolation above  $+10\%$  (20.4 %). Moreover, the underestimation does not reach  $-30\%$  while the overestimation exceeds  $+30\%$ . It therefore suggests that the interpolation of the 254 nm excitation wavelength may tend to overestimate HIX. This tendency does not seem to be linked with the fluorescence intensity (higher at Nesque losses than unsaturated zone flows), related to the amount of organic matter, nor the kind of water (surface water at Nesque losses, unsaturated zone flows, springs).

- **Comparison of interpolated and non-interpolated FI and BIX on surface waters**

For FI and BIX, several dilutions of the same sample of the Italian lake contaminated with wastewater is used to compare the influence of interpolation. Dilutions correspond to different amounts of organic matter and thus different fluorescence intensities. Therefore, they can be useful to assess the influence of the amount of organic matter on the shift between the index values obtained with interpolated and non-interpolated excitation wavelengths.

Fig. 2 shows less differences for FI and BIX (respectively  $\pm 4\%$  and  $\pm 1\%$ ) between interpolated and non-interpolated excitation wavelength than Fig. 1 with HIX (reaching  $\pm 30\%$ ). It is of course related

with the number of samples used for the comparison, but the larger range of difference for FI ( $\pm 4\%$ ) than BIX ( $\pm 1\%$ ) suggests that it is indices-dependant, related to the excitation wavelength required for the calculation (310 nm for BIX and 370 nm for FI). As for HIX, no correlation is found with the fluorescence intensity (varying according to dilution) and the difference between interpolated and non-interpolated data.

Therefore, results show that the interpolation of the required excitation wavelength from EEM to calculate indices may provide an overview of indices, but results need to be taken with caution as interpolation may involve a different of indices values reaching 30 % which may provide a misinterpretation when comparing different values. Results also show that the interpolation error is not influenced by the fluorescence intensity and therefore not related to the amount of organic matter. This error is therefore difficult to assess and to correct.

Therefore, it is advised to clearly explain how the indices are calculated, to pay attention to the comparison of indices values obtained with different techniques (in particular with and without wavelength interpolation), and to prioritize the measurement of the required excitation wavelengths according to the index definition instead of interpolating them.

#### 4. Limits of the use of fluorescence indices calculated from pre-established peaks position

Fluorescence indices based on 2D spectra were developed in different aquatic environments, respectively stream/riverine (FI), soil (HIX) and marine (BIX) waters whereas they are frequently used in compartments of the water cycle different from those in which they were developed. This argument was used by Serène et al. (2022) in order to explain the mismatch between observed peak position for a karst groundwater dataset and theoretical excitation and emission wavelengths required to calculate HIX. This section aims to (i) observe whether such mismatch is observed only for different aquatic environments than the one the three indices were developed for; and (ii) discuss about the possible origin of

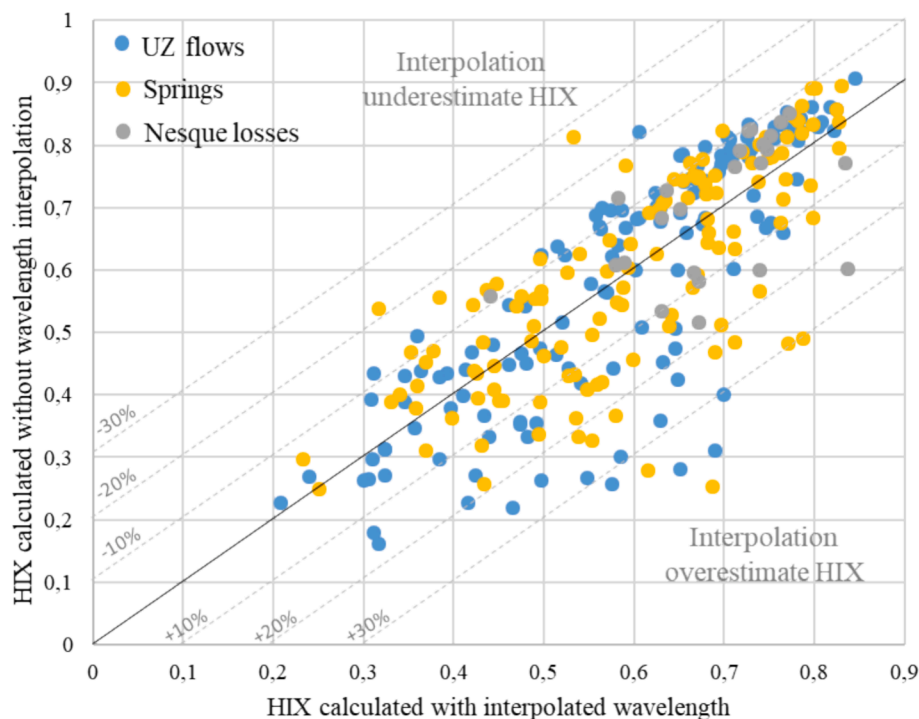
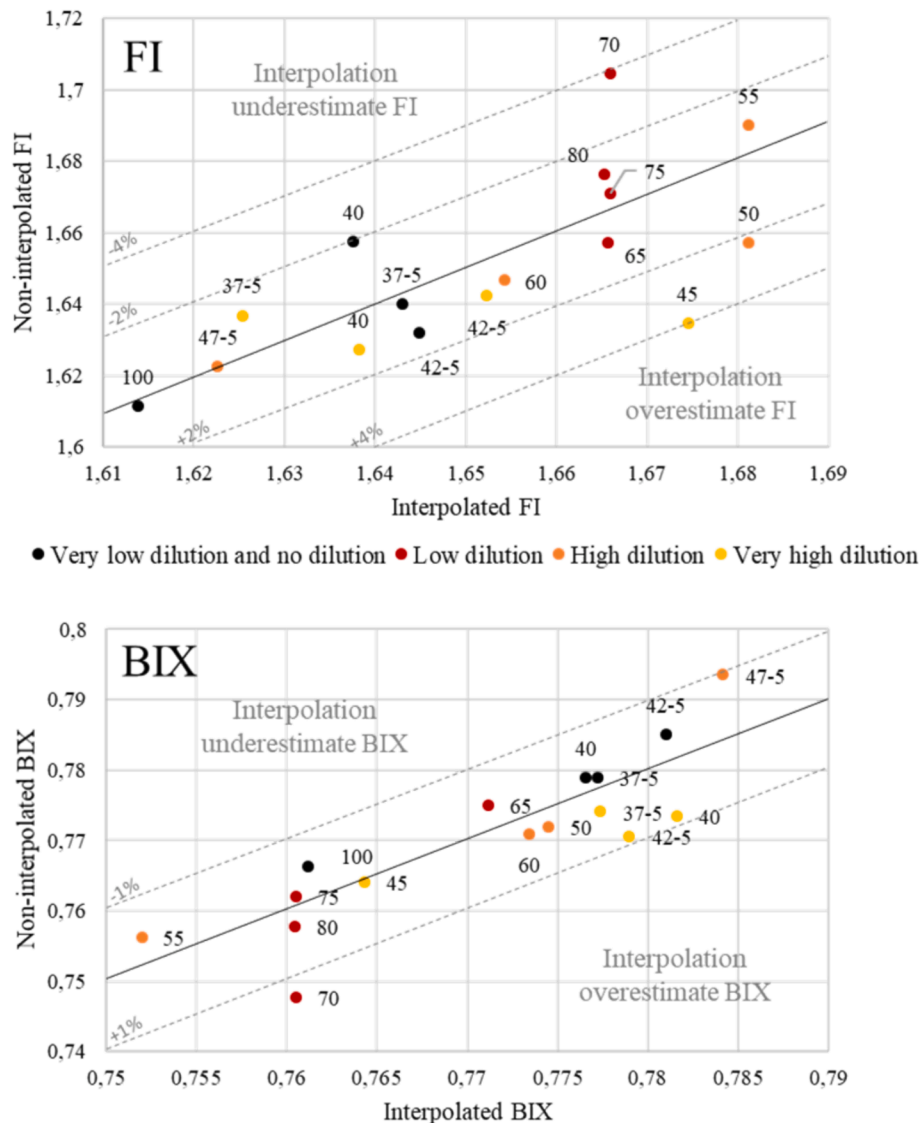


Fig. 1. HIX calculated from a 254 nm excitation wavelength 2D spectra as recommended by the index definition (y-axis), compared with HIX calculated thanks to staRdom that interpolates the 254 nm excitation wavelength from the same excitation wavelengths of EEM (x-axis), for 278 karst groundwater samples: Nesque losses (24 samples), unsaturated zone flows (132 samples) and 4 springs including the main outlet (124 samples).



**Fig. 2.** Comparison of BIX and FI calculated with the required excitation wavelengths (respectively 310 and 370 nm, y-axis, non-interpolated) and using staRdom interpolation of the excitation wavelengths from the same excitation wavelengths EEM (x-axis, interpolated), for 16 samples of an Italian lake contaminated with wastewater. It is the same sample with several dilution mentioned as labels (e.g. 100 is 100 % sample, label 55 is 55 % sample and 45 % of ultrapure water).

this mismatch and its consequences for the use of fluorescence indices.

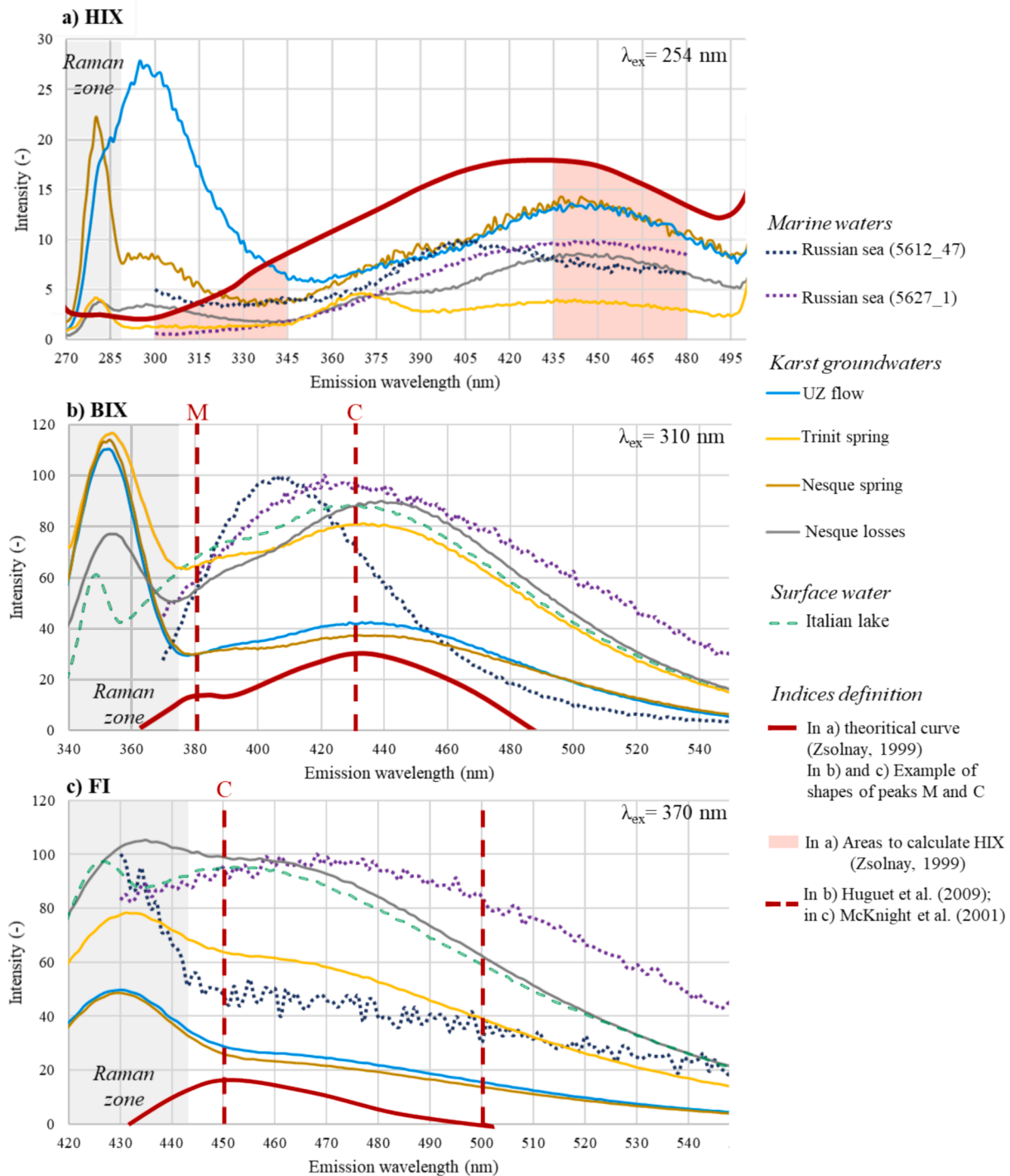
#### • Questioning the transferability of fluorescence indices

A shift of the emission wavelengths of some fluorescent compounds is commonly observed (at least for peaks C and A<sub>C</sub>, Fig. 3) between waters from different aquatic compartments of the environment (soil, surface, ground, marine waters; Baker and Genty, 1999; Conmy et al., 2009). It suggests a possible lack of transferability of indices definition as they are based on specific excitation and emission wavelengths. This shift is partly related to environmental physico-chemical parameters such as pH (Lakowicz, 2006; Osburn et al., 2014) which is known to affect both fluorescence intensity and spectral shapes of fluorescent compounds (Osburn et al., 2014). Indeed, pH controls the ionization potential in phenolic and carboxylic groups that are present in most organic fluorescent compounds, thereby influencing the metal-ligand complexation that causes fluorescence quenching. In addition, pH variation can cause intramolecular rearrangements that affect both fluorescence intensity and peak shape and position (Osburn et al., 2014). However, pH range vary from one aquatic compartment to the other (inland water: 5.5 to 7.5 Groeneveld et al., 2022; marine waters: 8 to 8.5

Garcia-Soto et al., 2021; soil: 3 to 10 Penn and Camberato, 2019) but they also overlap. It therefore suggests that fluorescence indices calculation based on peaks position cannot be generalized for the entire environment and also cannot be specified to each compartment of the environment (e.g. soil, surface, ground, marine waters) because of pH variation at these scales; and especially since climate change leads to local soils and surface water acidification (Osburn et al., 2014). In addition, some fluorescence indices have been developed with samples at pH different from most natural waters, such as HIX developed at a pH of 2 (Zsolnay et al., 1999).

As fluorescence peaks position is related to organic matter structure, itself mainly affected by organic matter degradation by microorganisms, degradation of organic matter is an important parameter that impacts peaks position. The soil continuum model of organic matter proposed by Lehmann and Kleber (2015) suggests that the degradation of organic matter in soil depends a lot on decomposer organisms and their living conditions (e.g. oxygenation, availability of nutrients and organic matter), involving important differences from one soil to the other, and therefore from one water to the other that may even come from the same water body (e.g. Mediterranean Sea, aquifer, spring).

Finally, it appears that some environmental parameters (e.g. pH,



**Fig. 3.** 2D emission spectra at different fixed  $\lambda_{Ex}$  from raw EEM in order to calculate a) HIX, b) BIX and c) FI. In the figures, samples considered are: the karst unsaturated zone flow AY (4 october 2021), St Trinit karst spring (16 November 2020), Nesque karst spring (14 april 2021), Nesque losses (3 may 2021), Italian lake contaminated by waste water (Pozzillo 100 UF, Sgroi et al., 2019); Marine water (Russian sea 5612\_47 and 5627\_1, Drozdova et al., 2019).

type of microorganisms, living conditions of microorganisms) can affect peaks position, involving that picking the intensity at precise location of a 2D spectra to calculate indices as are the historical definition of fluorescence indices should be questioned.

• **Assessing the transferability of fluorescence indices**

The adequacy of excitation and emission wavelengths required to calculate fluorescence indices according to definitions is assessed thanks

to 2D spectra from the karst groundwater and the marine waters datasets for HIX, FI and BIX, and with the surface waters dataset for BIX and FI. Only representative spectra of these 3 datasets are illustrated in Fig. 3 in order to keep readability.

According to Zsolnay et al. (1999) definition, the lower emission wavelength area used to calculate HIX (from 300 to 345 nm) correspond to the beginning of the peak and the higher emission wavelength area begins at the maximal intensity of the same peak (see red curve and areas, Fig. 3a). Only the marine water sample 5627\_1 spectrum shows

consistent shape with the theoretical one from Zsolnay et al. (1999). The other marine water sample (5612\_47) spectrum is completely offset from the theoretical curve, showing a maximal intensity between the two areas, around 400 nm. It suggests that the selected areas to calculate HIX are not adapted to this sample whereas it is for the other marine water sample. The consistency of the areas therefore does not appear to be only related to the type of water (e.g. marine water, groundwater, surface water) as the two marine water samples matches differently with the theoretical curve. For the 4 other samples from karst groundwaters (UZ flow, Nesque losses, Nesque and Trinit spring), the second area appears to be quite consistent, beginning close to the maximal intensity of the peak. However, the first area does not correspond at all to the beginning of the peak whose maximum appears in the second area.

HIX emission wavelengths areas are therefore poorly adapted to 5 to 6 of the samples shown in Fig. 3a. These areas probably match better aqueous soil extracts or samples at low pH (calibrated for pH = 2, Zsolnay et al., 1999) but they may give inconsistent HIX values for most marine and groundwater samples.

BIX is calculated as the ratio of the intensity at 380 nm (peak M) to the intensity at 430 nm (peak C; Huguet et al., 2009; Fig. 3b). For spectra allowing the observation of peak M (all except Russian marine samples), the maximal intensity appears to be offset, around 390 nm. The 430 nm intensity corresponds well to the peak C maximal intensity of the Italian lake and Trinit spring. But it does not match with the 5 other samples. The maximal peak C intensity wavelength is slightly longer for Nesque spring and losses, and UZ flow (around 440 nm) and shorter for the two marine water samples.

BIX was developed for marine water but the theoretical emission wavelengths do not correspond to the maximal intensity of peaks M and C. Moreover, peak M is not systematically present as it cannot be observed on these two samples which may compromise the accuracy of the final index value.

FI is calculated as the ratio of peak C maximal intensity (450 nm) to the peak C minimum intensity (500 nm, McKnight et al., 2001). The 450 nm emission wavelength does not match peak C maximal intensity of the 7 selected samples. Italian lake and Nesque losses are the closer, around 455 nm, the Russian sea (5627\_1) peak C is more offset around 470 nm. For the 4 other samples, peak C is hardly visible at this excitation wavelength whereas its presence is clear at shorter excitation wavelength (Fig. 3).

Fig. 3 therefore show a misfit for most of the samples between the areas and emission wavelength from the index definition (HIX, FI, BIX) and areas and emission wavelength observed for the samples (peaks location and curve shape). This misfit is not related to the use of the index on a different water type than the one it was developed for, as it was observed on the two marine samples for BIX developed on marine waters; and on the springs for FI developed on stream water. The transferability of the method from one aquatic environment to another is therefore less questioned than from some environmental conditions to others.

#### • Suggestions for improving the use of fluorescence indices

The observation of this misfit strongly encourages the check of the accuracy of the wavelength with the samples as done in Fig. 3 before the calculation and the interpretation of fluorescence indices. For example, BIX may not be calculated in the absence of peak M which is rarely observed in surface and groundwaters (Coble et al., 2014). Moreover, for FI, the emission wavelengths selected (450 and 500 nm, McKnight et al., 2001) do not match the peak C emissions wavelengths commonly referenced in aquatic environments (420 to 460 nm on the Fig. 3c).

A good alternative to pre-established excitation and emission wavelengths is to identify a range for emission wavelength as proposed by Wilson and Xenopoulos (2009) for the BIX from Huguet et al. (2009), in order to adapt the calculation of the indices to each environmental context. Wilson and Xenopoulos (2009) proposed to select the maximal

intensity between 420 and 435 nm instead of the emission at 430 nm advised by Huguet et al. (2009). The range of emission wavelength needs to be defined according to degradation parameters (e.g. number and types of decomposers organisms, availability of nutrients, oxygen and organic matter; Lehmann and Kleber, 2015) and other environmental conditions (e.g. pH) allowing its adaptation to each study site or water sample; or peaks position should be directly identified by the fluorescence indices user.

It would be also interesting to define excitation wavelength ranges as well, as the one required to calculate FI does not allow peak C observation for most of the tested samples whereas it is clearly observed at shorter excitation wavelength.

## 5. Towards fluorescence indices update thanks to 3D excitation-emission matrices (EEM)

EEM correspond to 3D data recording fluorescence intensity according to emission and excitation wavelengths whereas 2D spectra only records intensity according to emission wavelength (excitation wavelength is fixed). Since EEM are a combination of 2D spectra, they are more accurate in identifying the location of the fluorescent compound peaks (excitation and emission wavelengths). Peak location can be identified by hand to manually adjust the fluorescence indices calculation, observing which excitation wavelength is more representative, and the same for the emission wavelength, thus overcoming the problems identified in section 4 by allowing the adaptation of the calculation method to any environmental context. PARAFAC models can also be calculated with a great number of EEMs in order to automatize the identification of the peak's location for an entire dataset. PARAFAC components can be extracted from these models and used to calculate fluorescence indices. Recent studies calculated fluorescence indices using EEM, and more particularly PARAFAC components, inspired from HIX (e.g. the protein to humic index (PHI) from Li et al. (2019) and Yang et al. (2020); TTI from Serène et al., 2022), and BIX (microbial to terrestrial index (MTI) from Li et al. (2019) and Yang et al. (2020)), putting into relief the potential of EEM and of PARAFAC models to improve fluorescent indices calculation.

Therefore, the rise of EEM comes with a rethinking of the 2D-based fluorescent indices as EEM appears to have the potential to rethink the way fluorescence indices are calculated by providing an accurate way to identify the fluorescent compounds location (excitation and emission wavelengths), overcoming the problems of 2D-based indices discussed in section 4. Moreover, the comparability of these indices is not called into question by the use of EEM rather than 2D emission spectra, since PARAFAC and EEM are only a means of identifying the families of compounds involved in calculating the indices, and since indices are defined thanks to families of compounds (Humic C, Tryptophan and Tyrosine for HIX, Humic C and M for BIX and Humic C for FI).

## 6. Questioning the interpretation of fluorescence indices

Knowledge about organic matter has improved a lot since the development of the first fluorescent indices. Main difference is that the emergent view supposes the potential reuse of simpler compounds to produce complex ones; whereas traditional view supposes the degradation of organic matter to smaller compounds until recalcitrant or inorganic compounds (Lehmann and Kleber, 2015). This change of view may influence the way HIX, BIX, FI can be interpreted (summarized in Fig. 4).

For HIX, this change is particularly important as one of the interpretations of this index is that it is an assessment of the humification process (Ohno, 2002; Zsolnay et al., 1999), that leads to the production of recalcitrant organic matter (mature organic matter), which existence is reconsidered nowadays (Lehmann and Kleber, 2015). However, HIX first interpretation by Kalbitz et al. (2000) is still accurate as it is related to the structure of organic matter, it is the ratio of complex molecules



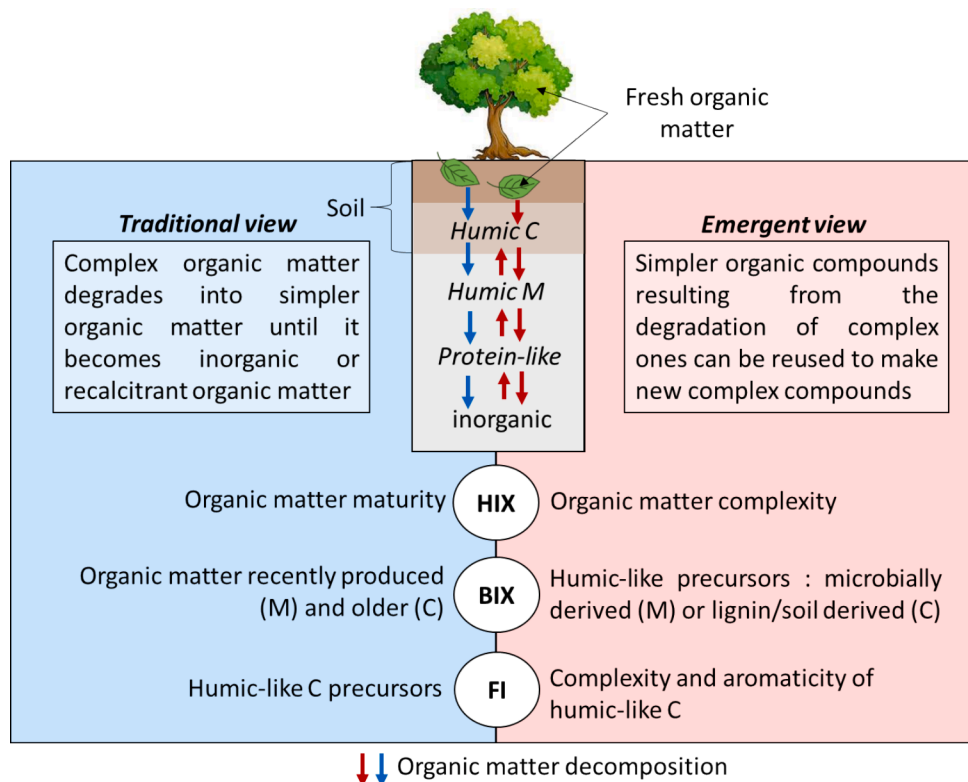


Fig. 4. Incidence of the model of evolution of organic matter decomposition on the interpretation of fluorescence indices. Traditional and Emergent view are summarized from Lehmann and Kleber (2015).

(numerous aromatic nuclei, heavy organic matter, humic-like) to simple molecules (few aromatic nuclei, light organic matter, protein-like). Therefore, HIX can still be used to provide information about the complexity of organic matter as it was found to be correlated with the carbon to nitrogen ratio, the aromaticity and polycondensation of organic matter molecules; but HIX cannot be interpreted in terms of organic maturity or humification degree (Fig. 4).

BIX is commonly interpreted as the proportion of recently produced dissolved organic matter, based on the hypothesis that humic-like M compounds are recently created in comparison with humic-like C that are more mature organic matter (Gabor et al., 2014). This interpretation of BIX needs to be reconsidered as “recently created” and “mature” organic matter is not consistent with the emergent view of organic matter decomposition. BIX interpretation proposed by Parlanti et al. (2000) assumes autochthonous production of peak M and allochthonous production of peak C in estuarine waters. However, humic M may be produced in contaminated surface water (Coble et al., 2014), and then be considered as allochthonous in a groundwater body with recharge from these contaminated surface water. Therefore, this interpretation cannot be generalized as it depends a lot on organic matter sources, fates, and to water mixes.

Finally, BIX can only be interpreted as an indicator of humic-like origin: microbially (humic M) or lignin/ soil derived (humic-like C) in accordance with the molecular structure of these fluorescent compounds. Moreover, BIX cannot be calculated in the absence of humic-like M, its presence needs to be checked before the calculation particularly if unsupervised method such as automatic peak picking is used, to guarantee the accuracy of indices value.

FI usual interpretation is not reconsidered by the improvement of organic matter degradation knowledge as it is based on the structure of humic-like C molecules, identifying if its precursors are more microbially or terrestrially derived. However, this interpretation may lead to a confusion with BIX that use the two extremes of humic-like compounds family: terrestrially (humic-like C) and microbially derived (humic-like

M). FI objective is to describe the variability of humic-like C structure. Some of humic-like C molecules have higher emission wavelength, and are therefore more complex because they have more terrestrially derived precursors (lignin), in comparison with others humic-like C molecules that have lower emission wavelength and are simpler because they have more microbially derived precursors (Coble et al., 2014). Therefore, FI can still be interpreted as an indicator of humic-like C precursors, but it could also be used to assess the complexity and aromaticity of humic-like C compounds.

Transferable interpretation of FI, BIX, HIX and other fluorescent indices from one environmental context to another is therefore limited to organic matter structure. To go further into the interpretation, it requires more information about organic matter sources and fates at the study site scale or in the specific environmental context as organic matter decomposition depend a lot on the accessibility of organic matter to the decomposer’s organisms (Lehmann and Kleber, 2015).

## 7. Conclusion

Fluorescence indices are frequently used and are particularly adapted to characterize the composition and sources of dissolved organic matter (Gabor et al., 2014). The present study aims to discuss several limitations in the present use of fluorescence indices and to illustrate them with different types of water samples coming from three different open access datasets: karst waters from Serène et al. (2024), marine waters from Drozdova et al. (2019) and surface water from Sgroi et al. (2019).

First, results show a difference of indices values calculated from interpolated and non-interpolated excitation wavelengths that may reach more than 30 % (for HIX). This difference is not influenced by the fluorescence intensity and therefore not related to the amount of organic matter, involving that it cannot be predicted. As a consequence, it is recommended to clearly mention how indices are calculated and to promote the measurement of the required excitation wavelength to not

have to interpolate it.

Second, an important misfit for almost all samples between the areas and wavelengths from the index definition (HIX, FI, BIX) and the areas and wavelengths from the samples (peaks location and curve shape) was observed. These results show that the excitation and emission wavelengths used in these indices are poorly transferable to other study site than the one it was developed for. This misfit is mainly related to the different environmental conditions such as pH and the living conditions of organic matter decomposers that affect both the fluorescence intensity and peaks position. To overcome this issue, it is proposed to define ranges of excitation and emission wavelengths associated with the fluorescence indices in order to adapt the calculation method to each purpose and study.

Third, discussion about the rise of EEM analyses puts into relief the potential to rethink the 2D-based indices thanks to 3D matrices. Direct extraction of maximal intensity of fluorescent components on EEM or peaks extraction obtained thanks to PARAFAC models as proposed by Li et al. (2019), Serène et al. (2022), and Yang et al. (2020) appear to be a nice and accurate way to overcome the previously cited limits of 2D-based indices.

Fourth, the incidence of recent improvement in organic matter knowledge on the fluorescence indices interpretation is highlighted, putting into relief that the interpretation of fluorescence indices should be limited to molecular structure. More specific interpretation should be thought sample by sample, study by study because the source and fate of compounds can hardly be generalized as they depend a lot on the accessibility of organic matter to the decomposer's organisms (Lehmann and Kleber, 2015) and other environmental conditions.

To conclude, the present study draws the attention of fluorescence indices users to unsupervised calculation and interpretation of 2D-based fluorescence indices. It also greatly encourages the rethinking of fluorescence indices using EEM, to be able to adapt the calculation to each purpose and study. Moreover, it puts into relief the need of a conceptual model of the source and fate of fluorescent compounds to go deeper into fluorescence indices interpretation.

## 8. Code availability

The code is not publicly accessible as it was built following the steps of the package developer available here: [https://cran.r-project.org/web/packages/staRdom/vignettes/PARAFAC\\_analysis\\_of\\_EEM.html](https://cran.r-project.org/web/packages/staRdom/vignettes/PARAFAC_analysis_of_EEM.html) (last access: 07 December 2023).

## 9. Financial support

This research has been supported by the Total SA project "ALBION2" (grant no. 187737) and a doctoral fellowship from the GAIA doctoral school of the University of Montpellier (GAIA competition 2019).

## CRediT authorship contribution statement

**Leïla Serène:** Writing – original draft, Software, Methodology, Formal analysis, Data curation, Conceptualization. **Naomi Mazzilli:** Writing – review & editing, Supervision, Conceptualization. **Christelle Batiot-Guilhe:** Writing – review & editing, Supervision, Conceptualization. **Christophe Emblanch:** Writing – review & editing, Supervision. **Marina Gillon:** Writing – review & editing. **Milanka Babic:** Methodology. **Julien Dupont:** Methodology. **Roland Simler:** Methodology. **Matthieu Blanc:** Methodology.

## Author contribution

Matthieu Blanc and Leïla Serène took water samples from the karst groundwater dataset that were analyzed for major elements, TOC, and water stable isotopes by Milanka Babic, Julien Dupont, and Roland Simler and for fluorescence of organic matter by Leïla Serène. Christelle

Batiot-Guilhe, Christophe Emblanch, Naomi Mazzilli, and Leïla Serène provided critical feedback and helped to shape the research and the analysis. Marina Gillon reviewed the manuscript. Leïla Serène prepared the manuscript with contributions from all co-authors.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgement

Starting from the karst groundwater dataset, this work was performed within the framework of the FDV/LSBB observation site, which is part of OZCAR (French network of Critical Zone observatories), SNO KARST (French observatory network, <https://sokarst.org/>) initiative of INSU/CNRS, which seeks to support knowledge sharing and promote cross-disciplinary research on karst systems, and of the H + observatory network (French observatory network, <https://hplus.ore.fr/>). Thanks to these different structures and to the AQUA department of INRAE for their support of the observatory. The authors would like to express their gratitude to the LSBB team for their technical and logistic help, especially to Michel Auguste who helped us a lot by sampling LSBB flows for us. The authors are also grateful to the SMBS (Syndicat Mixte du Bassin des Sorgues) for their help in sampling the Fontaine de Vaucluse water and to the Town Hall of Fontaine de Vaucluse for their collaboration in instrumenting the site. A special acknowledgement is given to Yves Perrette for his conceptual contributions to this paper.

## Data availability

The raw excitation-emission matrices of organic matter fluorescence are available at <https://doi.org/10.15148/8d6104e1-ae78-4b4e-8e50-198ccc5b19c9#2024> (Serene et al., 2024) for the karst water dataset, Sgroi et al., 2019 <https://doi.org/10.17632/4zss49jycj.1> for the Italian lake data, and Drozdova et al., 2019. <https://doi.org/10.1016/j.dib.2019.104562> for marine water data.

## References

- Aiken, G., McKnight, D., Harnish, R., Wershaw, R., 1996. Geochemistry of aquatic humic substances in the Lake Fryxell Basin, Antarctica. *Biogeochemistry* 34. <https://doi.org/10.1007/BF00000900>.
- Baker, A., Genty, D., 1999. Fluorescence wavelength and intensity variations of cave waters. *J. Hydrol.* 217, 19–34. [https://doi.org/10.1016/S0022-1694\(99\)00010-4](https://doi.org/10.1016/S0022-1694(99)00010-4).
- Cao, X., He, W., Fan, M., He, W., Shi, Y., An, T., Chen, X., Zhang, Z., Liu, F., Zhao, Y., Zhou, P., Chen, C., He, J., 2023. Novel insights into source apportionment of dissolved organic matter in aquifer affected by anthropogenic groundwater recharge: Applicability of end-member mixing analysis based optical indices. *Sci. Total Environ.* 863, 160885. <https://doi.org/10.1016/j.scitotenv.2022.160885>.
- Coble, P.G., Spencer, R.G.M., Baker, A., Reynolds, D.M., 2014. Aquatic 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. 75–122. <https://doi.org/10.1017/CBO9781139045452.006>.
- Conmy, R.N., Coble, P.G., Cannizzaro, J.P., Heil, C.A., 2009. Influence of extreme storm events on West Florida Shelf CDOM distributions. *J. Geophys. Res.* 114, G00F04. <https://doi.org/10.1029/2009JG000981>.
- D'Andrilli, J., Silverman, V., Buckley, S., Rosario-Ortiz, F.L., 2022. Inferring Ecosystem Function from Dissolved Organic Matter Optical Properties: A Critical Review. *Environ. Sci. Technol.* 56, 11146–11161. <https://doi.org/10.1021/acs.est.2c04240>.
- Drozdova, A.N., Puiman, M.S., Krylov, I.N., Patsaeva, S.V., Shatravin, A.V., 2019. Dataset on optical characteristics and spectroscopic indices of dissolved organic matter of the Kara, Laptev, and East Siberian seas in August–September 2017. *Data Brief* 26, 104562. <https://doi.org/10.1016/j.dib.2019.104562>.
- Duchaufour, P., Faivre, P., Poulencard, J., Gury, M.: Introduction à la science du sol - Sol, végétation, environnement, 7e édition. ed, Sciences Sup. Dunod, ISBN : 978-2-10-081992-8, 2020.
- Ewald, M., Berger, P., Visser, S.A., 1988. UV-visible absorption and fluorescence properties of fulvic acids of microbial origin as functions of their molecular weights. *Geoderma* 43, 11–20. [https://doi.org/10.1016/0016-7061\(88\)90051-1](https://doi.org/10.1016/0016-7061(88)90051-1).

- Gabor, R.S., Baker, A., McKnight, D.M., Miller, M.P., 2014. Fluorescence indices and their interpretation. *Aquatic Organic Matter Fluorescence*. <https://doi.org/10.1017/CBO9781139045452.015>.
- García-Soto, C., Cheng, L., Caesar, L., Schmidt, S., Jewett, E.B., Cheripka, A., Rigor, I., Caballero, A., Chiba, S., Báez, J.C., Zielinski, T., Abraham, J.P., 2021. An Overview of Ocean Climate Change Indicators: Sea Surface Temperature, Ocean Heat Content, Ocean pH, Dissolved Oxygen Concentration, Arctic Sea Ice Extent, Thickness and Volume, Sea Level and Strength of the AMOC (Atlantic Meridional Overturning Circulation). *Front. Mar. Sci.* 8, 642372. <https://doi.org/10.3389/fmars.2021.642372>.
- Groeneveld, M., Catalán, N., Einarsdóttir, K., Bravo, A.G., Kothawala, D.N., 2022. The influence of pH on dissolved organic matter fluorescence in inland waters. *Anal. Methods* 14, 1351–1360. <https://doi.org/10.1039/D1AY01702K>.
- Guo, H., Li, X., Xiu, W., He, W., Cao, Y., Zhang, D., Wang, A., 2019. Controls of organic matter bioreactivity on arsenic mobility in shallow aquifers of the Hetao Basin, P.R. China. *J. Hydrol.* 571, 448–459. <https://doi.org/10.1016/j.jhydrol.2019.01.076>.
- 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, 706–719. <https://doi.org/10.1016/j.orggeochem.2009.03.002>.
- Jiang, Y., Li, R., Yang, Y., Yu, M., Xi, B., Li, M., Xu, Z., Gao, S., Yang, C., 2019. Migration and evolution of dissolved organic matter in landfill leachate-contaminated groundwater plume. *Resour. Conserv. Recycl.* 151, 104463. <https://doi.org/10.1016/j.resconrec.2019.104463>.
- Kalbitz, K., Solinger, S., Park, J.-H., Michalzik, B., Matzner, E., 2000. CONTROLS ON THE DYNAMICS OF DISSOLVED ORGANIC MATTER IN SOILS: A REVIEW. *Soil Sci.* 165, 277–304. <https://doi.org/10.1097/00010694-200004000-00001>.
- Kulkarni, H.V., Mladenov, N., Johannesson, K.H., Datta, S., 2017. Contrasting dissolved organic matter quality in groundwater in Holocene and Pleistocene aquifers and implications for influencing arsenic mobility. *Appl. Geochem.* 77, 194–205. <https://doi.org/10.1016/j.apgeochem.2016.06.002>.
- Lakowicz, J.R.: *Principles of fluorescence spectroscopy*, 3rd ed. ed. Springer, New York, ISBN : 978-0-387-31278-1, 2006.
- Lehmann, J., Kleber, M., 2015. The contentious nature of soil organic matter. *Nature* 528, 60–68. <https://doi.org/10.1038/nature16069>.
- Li, X., Guo, H., Zheng, H., Xiu, W., He, W., Ding, Q., 2019. Roles of different molecular weights of dissolved organic matter in arsenic enrichment in groundwater: Evidences from ultrafiltration and EEM-PARAFAC. *Appl. Geochem.* 104, 124–134. <https://doi.org/10.1016/j.apgeochem.2019.03.024>.
- Liao, J., Hu, C., Li, X., Ruan, J.: Drying increases organic colloidal mobilization in the karst vadose zone: Evidence from a 15-year cave-monitoring study. *Hydrol. Process.* 35, 10.1002/hyp.14163, 2021.
- McKnight, D.M., Andrews, E.D., Spaulding, S.A., Aiken, G.R., 1994. Aquatic fulvic acids in algal-rich antarctic ponds. *Limnol. Oceanogr.* 39, 1972–1979. <https://doi.org/10.4319/lo.1994.39.8.1972>.
- 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. <https://doi.org/10.4319/lo.2001.46.1.0038>.
- Mladenov, N., Zheng, Y., Miller, M.P., Nemergut, D.R., Legg, T., Simone, B., Hageman, C., Rahman, M.M., Ahmed, K.M., McKnight, D.M., 2010. Dissolved Organic Matter Sources and Consequences for Iron and Arsenic Mobilization in Bangladesh Aquifers. *Environ. Sci. Technol.* 44, 123–128. <https://doi.org/10.1021/es901472g>.
- Murphy, K.R., Stedmon, C.A., Graeber, D., Bro, R., 2013. Fluorescence spectroscopy and multi-way techniques. *PARAFAC. Anal. Methods* 5, 6557. <https://doi.org/10.1039/c3ay41160e>.
- Nelson, C.E., Donahue, M.J., Dulaiova, H., Goldberg, S.J., La Valle, F.F., Lubarsky, K., Miyano, J., Richardson, C., Silbiger, N.J., Thomas, F.I.M., 2015. Fluorescent dissolved organic matter as a multivariate biogeochemical tracer of submarine groundwater discharge in coral reef ecosystems. *Mar. Chem.* 177, 232–243. <https://doi.org/10.1016/j.marchem.2015.06.026>.
- Ohno, T., 2002. Fluorescence Inner-Filtering Correction for Determining the Humification Index of Dissolved Organic Matter. *Environ. Sci. Technol.* 36, 742–746. <https://doi.org/10.1021/es0155276>.
- Osburn, C.L., Del Vecchio, R., Boyd, T.J., 2014. Physicochemical Effects on Dissolved Organic Matter Fluorescence in Natural Waters. In: Coble, P.G., Lead, J., Baker, A., Reynolds, D.M., Spencer, R.G.M. (Eds.), *Aquatic Organic Matter Fluorescence*. Cambridge University Press, pp. 233–277. <https://doi.org/10.1017/CBO9781139045452.012>.
- Parlanti, E., Wörz, 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. [https://doi.org/10.1016/S0146-6380\(00\)00124-8](https://doi.org/10.1016/S0146-6380(00)00124-8).
- Penn, C., Camberato, J., 2019. A Critical Review on Soil Chemical Processes that Control How Soil pH Affects Phosphorus Availability to Plants. *Agriculture* 9, 120. <https://doi.org/10.3390/agriculture9060120>.
- Pucher, M., Wunsch, U., Weigelhofer, G., Murphy, K., Hein, T., Graeber, D., 2019. starDom: Versatile Software for Analyzing Spectroscopic Data of Dissolved Organic Matter in R. *Water* 11, 2366. <https://doi.org/10.3390/w11112366>.
- Schittich, A.-R., Wunsch, U.J., Kulkarni, H.V., Battistel, M., Bregnhøj, H., Stedmon, C.A., McKnight, U.S., 2018. Investigating Fluorescent Organic-Matter Composition as a Key Predictor for Arsenic Mobility in Groundwater Aquifers. *Environ. Sci. Technol.* 52, 13027–13036. <https://doi.org/10.1021/acs.est.8b04070>.
- Serène, L., Batiot-Guilhe, C., Mazzilli, N., Emblanch, C., Babic, M., Dupont, J., Simler, R., Blanc, M., Massonnat, G., 2022. Transit Time index (TTI) as an adaptation of the humification index to illustrate transit time differences in karst hydrosystems: application to the karst springs of the Fontaine de Vaucluse system (southeastern France). *Hydrol. Earth Syst. Sci.* 26, 5035–5049. <https://doi.org/10.5194/hess-26-5035-2022>.
- Sgroi, M., Gagliano, E., Vagliasindi, F., Roccaro, P.: Data on the inner filter effect, suspended solids and nitrate interferences in fluorescence measurements of wastewater organic matter. DOI: 10.17632/4zss49jycj.1, 2019.
- Shand, P., Edmunds, W.M., Lawrence, A.R., Smedley, P., Burke, S.: The natural (baseline) quality of groundwater in England and Wales. British Geological Survey, Keyworth, Nottingham, ISBN 978 085272595 5, 2007.
- Wilson, H.F., Xenopoulos, M.A., 2009. Effects of agricultural land use on the composition of fluvial dissolved organic matter. *Nat. Geosci.* 2, 37–41. <https://doi.org/10.1038/ngeo391>.
- Yang, Y., Yuan, X., Deng, Y., Xie, X., Gan, Y., Wang, Y., 2020. Seasonal dynamics of dissolved organic matter in high arsenic shallow groundwater systems. *J. Hydrol.* 589, 125120. <https://doi.org/10.1016/j.jhydrol.2020.125120>.
- Zhang, Y., Liu, Y., Zhou, A., Zhang, L., 2021. Identification of groundwater pollution from livestock farming using fluorescence spectroscopy coupled with multivariate statistical methods. *Water Res.* 206, 117754. <https://doi.org/10.1016/j.watres.2021.117754>.
- Zsolnay, A., Baigar, E., Jimenez, M., Steinweg, B., Saccomandi, F., 1999. Differentiating soil fluorescence spectroscopy sources of dissolved organic matter in soils subjected to drying. *Chemosphere* 38, 45–50. [https://doi.org/10.1016/S0045-6535\(98\)00166-0](https://doi.org/10.1016/S0045-6535(98)00166-0).