Biodiversity restoration and conservation of inland water ecosystems for environmental and human well-being

BioReset

BiodivRestore-406

2020 - 2021 Joint Call
Joint COFUND Call on "Conservation and restoration of degraded ecosystems and their biodiversity, including a focus on aquatic systems"

Deliverable 3.2

Assessing ecosystem recovery with diatom Raman spectroscopy

Lead Beneficiary	Work package	Delivery month
CIIMAR	**3**	**44**

1. Executive Summary

Diatoms are unicellular microalgae with a porous siliceous cell wall. They can be found in practically all aquatic environments and are well-recognised as bioindicators of ecological water quality. They respond to abiotic stress in short time scales, showing differentiated responses to a variety of aquatic contaminants. Diatoms are the main components of biofilms. These are aggregates of various microorganisms (diatoms and other microalgae, protozoa, fungi and bacteria) embedded in a polymeric matrix. The biofilms are visible macroscopically as greenish to brownish layers colonising different substrates (rocks, pebbles, blocks, shells, clay...).

Diatom sampling is the basis for the evaluation of the ecological water quality with taxonomic quality indices as recommended in the Water Framework Directive. This is a suitable and widely used method, but it is cumbersome and expensive, requiring sample preparation and identification knowledge and experience, among others (Pinto et al., 2021). Various methods have thus been considered for its suitable replacement, though the proposals available are also expensive and require sample preparation significant technical resources and expertise (Pinto et al., 2021). **BioReset** proposes the development of an innovative approach based on Raman Spectroscopy, which does not require sample preparation, can be employed with a minimum number of measurements done by non-specialised personnel, with results analysed using pre-established chemometrics routines. For this, a most relevant step was to develop a tentative index based on the application of Raman spectroscopy to field samples and infer about its potential to discriminate sites under differential pollution pressure. This deliverable reports on the methods and procedures for the development of this tentative index or tool to assess water ecological quality. The samples were collected following the previously established protocol (Deliverable 3.X) in freshwater ecosystems associated to wastewater treatment plants or localised sewage systems. The samples were provided by **BioReset** partners, including CIIMAR, and analysed by the CIIMAR team. The following of the deliverable reports on the samples' diatom composition, the method developed and its first application to better understand ecosystem recovery. Future work will be focused on expanding the results obtained to a larger dataset and relate the results to other community indicators and mechanistic effects at lower biological organisational levels (e.g. molecular). This will allow for testing and improving sensibility and robustness of the draft tool proposed. It will also bring insight on the processes by which diatoms adapt or can develop resilience to environmental stress.

2. Task description

The work was developed within the scope of WP3 (Evaluation of ecosystem conservation and restoration: diatom biofilms), coordinated by CIIMAR. The samplings were carried out by CIIMAR in Portugal (Rio Lis) and Brazil (local lake in Curitiba), UVigo in Spain (local river) and SLU in Sweden (local river). The samples from Brazil were collected by CIIMAR at a local lake within the campus of Universidade Positivo, with the help of local university colleagues. This was done during a visit to the campus within the scope of an ongoing project from CIIMAR. The IFE partner also collected samples in Norway. Unfortunately, they were lost in transit due to unexpected Customs retention in Portugal. The company DHL decided to retain the samples based on insufficient information, though the information provided was all the required by the service, as in the case of Swedish samples. This led to sample loss by deterioration due to the lack of ice and consequent increase in temperature over several days. In all samplings, physico-chemical parameters (pH, conductivity, dissolved oxygen) were measured locally at the time of collection. This was accompanied by the sampling of diatoms biofilm, and sample fixation when necessary, as described previously (Deliverable 3.X). All samples were sent to CIIMAR. At CIIMAR, complementary physico-chemical parameters (nutrient levels, such as phosphate, nitrogen and silica) were determined. In addition, sampling campaigns were extended to eight Portuguese estuarine systems, in collaboration with the monitoring project Ocean3R of CIIMAR (Minho, Lima and Douro estuaries; Ria de Aveiro; Mondego, Tejo and Guadiana estuaries; Ria Formosa). Determination of physico-chemical parameters was also done in the latter. All biofilm samples were characterised for the taxonomic composition in diatoms and taken to Raman spectroscopy. The Raman spectra from Sweden, Spain, Portugal (Lis River) and Brazil were the basis for the development of the proposed draft tool. For this all data was analysed with chemometric methods: Principal Component Analysis, Partial Least Squares analysis and saturated orthogonal multiple linear regression to generate classification models discriminating among the pollution/ecological quality levels. This proposed index will be further tested and refined with the samples from the estuarine systems and other countries with whom

CIIMAR has ongoing collaborations. All methods employed were previously described in Oliva-Teles et al. (2022), Pinto et al. (2022) and Rodrigues et al. (2023). The measurements made in each foreign ecosystem were also organised on specific country Reports with water quality information and provided to the respective Partners or the Brazilian colleagues (included in the Annexes).

3. WP3 – Team members involved in the work development

The following researchers contributed to the samplings and development of the work presented herein:

Name	Organization	Role
Laura Guimarães	CIIMAR	Institutional leader and WP coordinator
António Paulo Carvalho	CIIMAR	Task coordinator
Luís Oliva-Teles	CIIMAR	Task coordinator
Raquel Pinto	CIIMAR	PhD student allocated to the project (FCT grant)
Marco Veludo	CIIMAR	Technician hired through the project
Cristina Matos	REQUIMTE	Institutional leader and WP coordinator
Malin Hultberg	SLU	Institutional leader and WP coordinator
Laura Ferrando-Climent	IFE	Institutional leader and WP coordinator
Angeles Sanroman	UVIGO	Institutional leader and WP coordinator
Cintia Ribas de Oliveira	Universidade Positivo	Collaborator in other projects

4. Developed activities

Concerning this deliverable, the rational for the selection of the sampling sites was as follows. The sampling took place in ecosystems associated with the wastewater treatment plants (WWTP) involved in the project. The specific sampling sites within the ecosystem were chosen to reflect potential contamination gradients, allowing to characterise the input of the WWTP into the system contamination and a possible recovery along the downstream river or system flow, reflecting better conditions. An example was the selection of a site before the WWTP effluent release into the ecosystem, a site immediately after the point of effluent release (a meter at most) and two to three more sampling sites after this one up to 5 to 10 km in distance from the contamination source. Samples from the freshwater lake in Brazil were included to address a very relevant aspect, which is to incorporate natural diatom diversity influenced by different types of contamination and climate conditions. They were collected by CIIMAR team members with the support of colleagues from Universidade Positivo, within the scope of collaboration projects running in parallel to **BioReset**. The lake is located within the university campus and the approach to establish the sampling sites within the lake was also based on collecting nearby a sewage outlet and forward, across a potential dilution gradient. Incorporation of samples from Curitiba (Brazil) allowed to have different climatic regions represented in the data, from subarctic (Sweden system) to the unique Oceanic-Mediterranean transition climate in Spain, the temperate Mediterranean climate in Portugal and the humid subtropical climate (Curitiba, Brazil). The data was used to investigate chemical and ecological water quality in the sampling sites, through conventional water and diatom quality indices. Reports were elaborated and provided to each team leader can be found in the Annexes.

5. Development of a draft index based on diatom Raman spectroscopy

5.1. Sampling sites

The sampling sites in Sweden were located in the Höje River, in the southern part of the country. The Höje River extends over ~35 km, flowing into the Öresund Strait south of Malmö. The catchment is dominated by intensive agriculture, and the river has been recognised as highly impacted by nutrient loading, particularly phosphorus and nitrogen from diffuse agricultural runoff and urban wastewater. The Källby wastewater treatment plant (WWTP), located in the southwestern part of the city of Lund, discharges threated effluents to this system. The WWTP treats the residual water from Lund and its nearby villages serving a population of

about 86,000 people (Larsson et al., 2013). The plant comprises primary (bar screen, grit removal, primary settling), secondary (conventional active sludge with anoxic pre-denitrification) and tertiary (phosphate precipitation with ferric chloride) treatments. Three sampling sites were established in this system, identified as SW1 to SW3. One site was located upstream to the WWTP (SW1), one was within a constructed wetland adjacent to the WWTP (SW2) and one site was located downstream to the WWTP (SW3). The samples were collected and described in Oliva-Teles et al. (2022) and Pinto et al. (2022). Subsequent processing, taxonomic identification and analysis through Raman spectroscopy followed the methods described therein.

The sampling sites in Spain were located in the Muíños River. This is a small, urban-influenced stream that discharges into Playa América. It has local significant recreational relevance. Though, limited information is available about local pollution levels. Two sampling sites were established in this system, identified as SP1 and SP2.

The Lis River is located in central Portugal. The city of Leiria is located on the riverside and the river flows into the Atlantic Ocean (Paíga et al., 2016). This watercourse suffers many anthropogenic impacts related with agriculture, tannery and mineral mining industries, livestock production and urban pollution. Throughout the river there are two WWTPs (Olhalvas and Coimbrão), discharging their effluents into the water (Paíga et al., 2016). Four sampling sites (August 2023) were established in this system, identified as PT1 to PT4. The sites were selected according to a contamination gradient associated to the Olhalvas WWTP: one site upstream (PT1) and 3 sites downstream (PT2, PT3, PT4). The samples were collected and described in Oliva-Teles et al. (2022) and Pinto et al. (2022). Subsequent processing, taxonomic identification and analysis through Raman spectroscopy followed the methods described therein.

The lake system studied in Brazil is located within the campus of Universidade Positivo (Curitiba). It is an artificial water body embedded in an extensive green-space environment. It functions as a recreational area and plays key roles in storm water retention, hydric management of the campus (irrigation, sanitary use). It has a heat-exchange water source, which supports the university's sports facilities. Ecologically, the lake sustains a considerable biodiverse community: several phytoplankton families (Chlorophyta, Cyanobacteria, Euglenophyta e Ochrophyta), fishes, birds, reptiles and mammals including capybaras. Additionally, the surrounding campus vegetation comprises a diversity of arboreal and shrubs species (including fruit-bearing trees). Four sampling sites were established in this system, identified as BR1 to BR4. Site BR1 is in a little impacted area. The samples were collected and described in Oliva-Teles et al. (2022) and Pinto et al. (2022). Subsequent processing, taxonomic identification and analysis through Raman spectroscopy followed the methods described therein.

5.2. Water physico-chemical parameters

Table 1 shows the physico-chemical measurements done in situ, at the time of water sampling, and the nutrient concentrations determined later in the lab. In Sweden, salinity and conductivity were much higher in SW1 and SW2 than in SW3 (6 µS cm -1; 6 ‰). Concentration of oxygen dissolved in the water was much lower in SW2 (1.69 mg L-1) than in SW1 (5.66 mg L-1) and SW3 (4.83 mg L-1) . The pH ranged from 7.7 (SW3) to 8.3 (SW1). Nutrient concentrations varied greatly between sampling sites. Higher nitrite and silicate concentrations were found upstream the WWTP (SW1) and higher nitrate, ammonia and phosphate concentrations were found near the WWTP (SW2).

Table 1. Physico-chemical parameters and nutrient concentrations determined in the samples collected.

Parameter	SW1	SW2	SW3	SP1	SP2	PT1	PT2	PT3	PT4	BR1	BR2	BR3	BR4
Conductivity (µS/cm)	667	787	6	153	265	3150	558	570	1223	585	105	70	58
Salinity (‰)	1.83	2.13	0.02	0.08	0.15	1.72	0.30	0.29	0.56	0.32	0.06	0.04	0.03
Oxygen (mg/L)	5.66	1.69	4.83	8.86	6.97	10.42	8.16	8.46	9.20	4.36	8.33	7.48	8.26
pH	8.30	7.80	7.70	6.66	6.87	7.57	7.62	7.72	7.75	7.27	7.33	7.05	7.26
Nitrite (mg/L)	0.73	0.44	0.06	0.27	0.05	0.21	0.02	<DL	0.22	0.09	0.02	0.20	0.03
Nitrate (mg/L)	0.43	2.30	1.70	<DL	<DL	<DL	2.70	3.00	<DL	<DL	<DL	1.13	<DL
Ammonia (mg/L)	1.69	2.60	0.24	0.96	0.52	0.32	0.48	0.20	0.56	0.76	0.94	1.25	2.26
Phosphate (mg/L)	1.01	3.27	0.15	0.27	0.48	0.68	0.85	0.46	2.50	1.60	2.23	3.97	7.40
Silicate (mg/L)	6.31	6.28	6.16	1.51	1.51	0.85	0.76	0.84	2.00	0.18	0.32	0.42	0.83

In Spain, conductivity and salinity were about the double at SP2 (265 µS cm^{-1} and 0.156‰, respectively) compared to SP1 (153 µS cm^{-1} and 0.084‰). Water dissolved oxygen was higher at SP1 (8.86 mg L^{-1}) than at SP2 (6.97 mg L^{-1}), while pH showed an inversed trend. Nitrite (NO_2^-) and ammonium (NH_4^+) concentrations were higher at SP1 (0.27 mg L^{-1} and 0.96 mg L^{-1}, respectively) than at SP2 (0.05 mg L^{-1} and 0.52 mg L^{-1}). Phosphate (PO_4^{3-}) was higher at SP2 (0.58 mg L^{-1}) compared to SP1 (0.27 mg L^{-1}). Silica (SiO_2) concentrations were similar in SP1 and SP2 (1.51 mg L^{-1}). In Portugal, Temperature was higher in PT4 (28.70 ºC) and lower in PT2 (20.40 ºC). Conductivity and salinity were higher in PT3 (3150 µS/cm and 1.72 ‰, respectively) and lower in PT1 (570 µS/cm and 0.29 ‰, respectively). Water dissolved oxygen varied among sites (from about 10 mg/L in PT3 to about 8 mg/L in PT2), though it was always above the healthy levels. pH values had slight variations across sampling sites. Nitrite, ammonia and phosphate concentrations were higher in samples from PT4 and lower in samples from PT1. Nitrate concentrations were only detected in sites PT2 and PT1, ranging from 2.7 to 3.0 mg/L, respectively. Silicate values were higher in PT4 and lower in PT2.In samples from the Curitiba lake, conductivity values ranged from 58.0 µS cm^{-1} at BR4 to 585 µS cm^{-1} at BR1. Salinity was highest at BR1 (0.320 ‰) and decreased progressively across the sites, reaching 0.030 ‰ at BR4. Water dissolved oxygen (O_2) levels varied between 4.36 mg L^{-1} at BR1 and 8.33 mg L^{-1} at BR2. The pH values were relatively stable, ranging from 7.05 at BR3 to 7.33 at BR2. Regarding nutrients, nitrite (NO_2^-) concentrations were lowest at BR2 (0.02 ± 0.0040 mg L^{-1}) and highest at BR3 (0.20 ± 0.0040 mg L^{-1}). Nitrate (NO_3^-) was only detected at BR3 (1.13 ± 0.0600 mg L^{-1}). Ammonium (NH_4^+) concentrations increased from BR1 (0.76 ± 0.0000 mg L^{-1}) to BR4 (2.26 ± 0.0100 mg L^{-1}). Phosphate (PO_4^{3-}) levels were highest at BR4 (7.40 ± 0.2000 mg L^{-1}) and lowest at BR1 (1.60 ± 0.0000 mg L^{-1}). Finally, silica (SiO_2) concentrations ranged from 0.18 ± 0.0000 mg L^{-1} at BR1 to 0.83 ± 0.0000 mg L^{-1} at BR4.

5.3. Diatom communities

Analysis of the samples collected in Sweden, led to the taxonomic identification of 72 species of diatoms presented in Table 2 and Figure 1. The most abundant species were *P. sulcata* in SW2, *C. placentula* in SW 1 and SW3, and *M. permitis* in SW2. *Cocconeis placentula var. lineata* is a cosmopolitan freshwater species frequently found in neutral to alkaline waters. *Aulacoseira granulata* (SW1) is a widespread taxa tolerant to all trophic conditions but more abundant in eutrophic waters. *Navicula lanceolata* (SW1) is a very common diatom indicative of eutrophic conditions and oligo to mesosaprobic conditions. This diatom is also commonly found in low temperature sites. *Paralia sulcata* is a species typically found in low temperature conditions and high nutrients levels. *Mayamaea permitis* is cosmopolitan and very abundant in Europe. It is usually found in masses in wastewater, sewage plants and other polysaprobic habitats. *Lemnicola hungarica* (SW2) is a widespread diatom frequently found attached to duckweed (*Lemna minor*). *Achnanthidium minutissimum* (SW1 and SW3) is a ubiquitous freshwater diatom with a wide ecological tolerance.

Table 2. Diatom species found in the sampling sites in the Höje River.

Species	SW1	SW2	SW3
Achnanthes adnata Bory 1822	0	0	10
Achnanthidium minutissimum (Kützing) Czarnecki 1994	28	16	30
Amphora liriope Nagumo 2003	0	1	0
Amphora ovalis (Kützing) Kützing 1844	12	0	0
Amphora pediculus (Kützing) Grunow 1875	12	0	0
Aulacoseira granulata (Ehrenberg) Simonsen 1979	40	0	10
Cocconeis placentula var. lineata (Ehrenberg) Van Heurck 1885	93	30	73
Diadesmis confervacea Kützing 1844	0	0	2
Diadesmis contenta (Grunow) D.G.Mann 1990	4	2	0
Diploneis voigtiana Lange-Bertalot & Fuhrmann 2020	8	0	0
Encyonema ventricosum (C.Agardh) Grunow 1875	2	0	4
Eolimna minima (Grunow) Lange-Bertalot & W.Schiller 1997	2	10	0
Fragilaria tenera (W.Smith) Lange-Bertalot 1980	0	0	2
Geissleria kriegeri (Krasske) Lange-Bertalot 1996	10	0	0
Gogorevia exilis (Kütz.) Kulikovskiy and Kociolek 2020	0	16	0
Gomphonema acuminatum Ehrenberg 1832	0	0	4
Gomphonema olivaceum (Hornemann) Brébisson 1838	0	0	6
Gomphonema parvulum (Kützing) Kützing 1849	21	12	4
Gomphonema saprophilum (Lange-Bertalot & E.Reichardt) Abraca, R.Jahn, J.Zimmermann & Enke 2014	0	0	20
Halamphora veneta (Kützing) Levkov 2009	6	0	0
Haslea crucigera (W.Smith) Simonsen 1974	0	0	2
Hippodonta costulata (Grunow) Lange-Bertalot, Metzeltin & Witkowski 1996	2	0	0

Table 2. Diatom species found in the sampling sites in the Höje River (*cont.*).

Karayevia clevei (Grunow) Round & Bukhtiyarova 1999	0	21	0
Platessa oblongella (Østrup) C.E.Wetzel, Lange-Bertalot & Ector 2017	2	2	0
Lemnicola hungarica (Grunow) Round & Basson 1997	0	38	0
Luticola goeppertiana (Bleisch) D.G.Mann ex Rarick S.Wu S.S.Lee & Edlund 2017	0	0	4
Mayamaea permitis (Hustedt) K.Bruder & Medlin 2008	0	57	0
Melosira varians C.Agardh 1827	21	0	7
Nanofrustulum cataractarum (Hustedt) C.E.Wetzel, E.Morales & Ector 2019	0	0	2
Navicula capitatoradiata H.Germain ex Gasse 1986	2	0	0
Navicula cryptocephala Kützing 1844	2	16	2
Navicula cryptotenella Lange-Bertalot 1985	2	0	0
Navicula cryptotenelloides Lange-Bertalot 1993	0	0	2
Navicula digitoradiata (Gregory) Ralfs 1861	0	0	8
Navicula exilis Kützing 1844	0	0	6
Navicula expecta VanLandingham 1975	2	0	0
Navicula gregaria Donkin 1861	8	0	2
Navicula lanceolata (C.Agardh) Kützing, 1844	30	0	12
Navicula tenelloides Hustedt 1937	2	0	0
Navicula trivialis Lange-Bertalot 1980	2	0	0
Navicula tripunctata (O.F.Müller) Bory de Saint-Vincent 1822	0	0	16
Nitzschia amphibia Grunow 1862	0	0	4
Nitzschia capitellata Hustedt 1930	4	2	0
Nitzschia dissipata (Kützing) Rabenhorst 1860	0	0	2
Nitzschia filiformis (W.Smith) Van Heurck 1896	12	0	0
Nitzschia fonticola (Grunow) Grunow 1881	0	0	2
Nitzschia inconspicua Grunow 1862	4	8	4
Nitzschia palea var. palea (Kützing) W.Smith 1856	10	2	4
Nitzschia recta Hantzsch 1862	4	0	8
Nitzschia sigma (Kützing) W.Smith 1853	4	0	0
Nitzschia sigmoidea (Nitzsch) W.Smith 1853	0	0	12
Opephora pacifica (Grunow) P.Petit 1889	0	0	2
Paralia sulcata (Ehrenberg) Cleve 1873	20	137	99
Parlibellus E.J.Cox, 1988	2	0	0
Pinnularia halophila Krammer 1992	2	0	0
Pinnularia subcapitata W.Gregory 1856	6	0	0
Planothidium delicatulum (Kützing) Round & Bukhtiyarova 1996	0	0	6
Planothidium frequentissimum (Lange-Bertalot) Lange-Bertalot 1999	4	26	8
Planothidium lanceolatum (Brébisson ex Kützing) Lange-Bertalot 1999	0	0	2
Psamothidium subatomoides (Hustedt) Bukhtiyarova & Round 1996	0	0	3
Rhoicosphenia abbreviata (C.Agardh) Lange-Bertalot 1980	8	4	14
Sellaphora pupula (Kützing) Mereschkovsky 1902	2	0	2
Stauroneis smithii Grunow 1860	0	0	6
Stephanocyclus meneghiniana (Kütz.) Kulikovskiy Genkal and Kociolek 2022	0	0	3
Surirella angusta Kützing 1844	2	0	0
Surirella brebissonii Krammer & Lange-Bertalot 1987	12	0	0
Tabularia tabulata (C.Agardh) Snoeijs 1992	4	4	4
Thalassiosira pseudonana Hasle & Heimdal 1970	2	0	0
Tryblionella acuminata W.Smith 1853	2	0	0
Tryblionella calida (Grunow) D.G.Mann 1990	0	0	2
Tryblionella hungarica (Grunow) Frenguelli 1942	0	0	2
Ulnaria ulna (Nitzsch) Compère 2001	0	0	8

Figure 1 shows the functional groups of the diatoms identified in the sampling sites in Sweden. The diatoms found were predominantly motile and low profile. Sites SW1 and SW3 were dominated by **β**mesosaprobic species. The site SW2 showed a prodominance of α-mesosaprobe and α-polymesosaprobe species. In terms of trophic state, all sampling sites had high numbers of eutrophic species. Moreover, a significantly high number of species were classified as unknown in all functional groups studied.

Analysis of the samples collected in Spain led to the taxonomic identification of 39 species of diatoms (Table 3, Figure 2). The most abundant species were *H. contenta* (SP1), *P. oblongella* (SP1 and SP2), *L. mutica* (SP1 and SP2), *P. subatomoides* (SP1), *N. brevissima* (SP2) and *K. clevei* (SP2). *Humidophila contenta* is a cosmopolitan diatom species. *Platessa oblongella* is also a widespread diatom with a great variety of ecological tolerances. *Luticola mutica* is very frequent in freshwater and brackish waters in Europe. *Nitzschia brevissima* usually lives in freshwater with high levels of dissolved oxygen and low organic pollution. *Karayevia clevei* is usually characteristic of alkaline waters and meso-eutrophic environments.

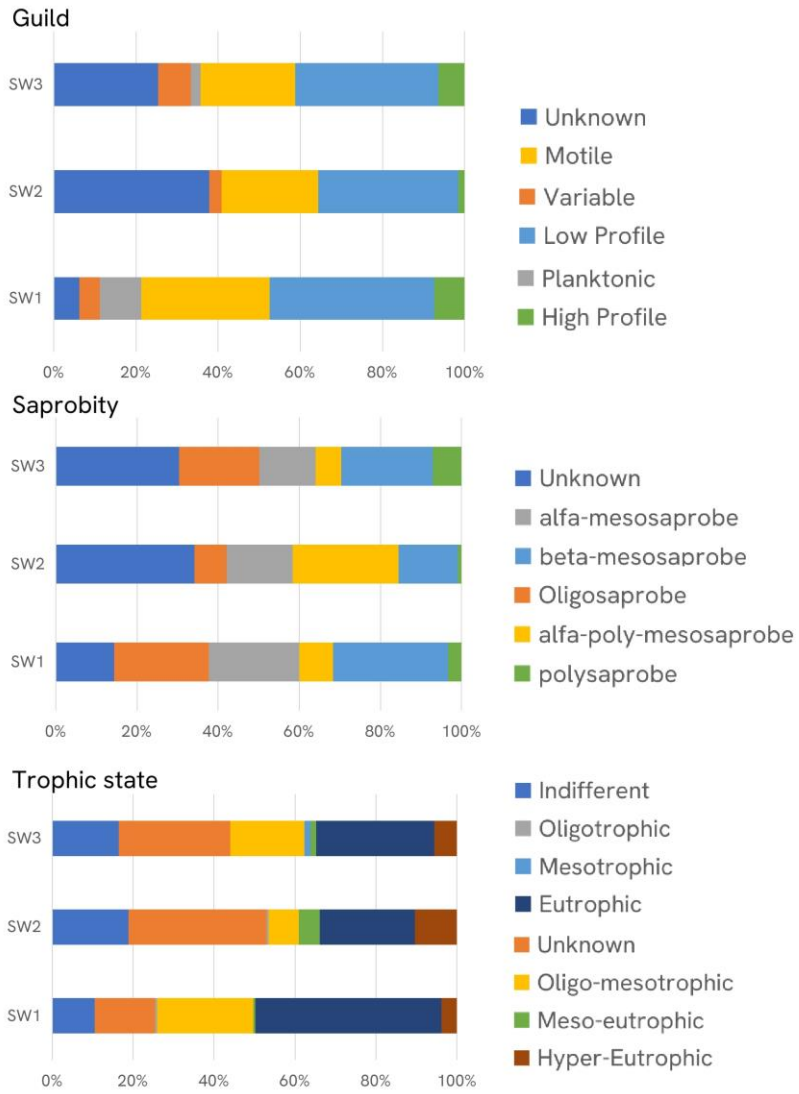

Figure 1. Guild, saprobicity and trophic state, as indicated by the species identified in the sampling sites in the Höje River.

Table 3. Diatom species found in the sampling sites located in River Muiños.

Species	SP1	SP2
Achnanthidium minutissimum (Kützing) Czarnecki 1994	4	0
Aulacoseira granulata (Ehrenberg) Simonsen 1979	2	2
Cosmioneis citriformis R.L.Lowe & A.R.Sherwood 2010	0	2
Denticula creticola (Østrup) Lange-Bertalot & Krammer 1993	2	0
Diploneis chersonensis (Grunow) Cleve 1892	0	2
Diploneis didyma (Ehrenberg) Ehrenberg 1845	0	2
Discostella stelligera (Cleve & Grunow) Houk & Klee 2004	0	6
Eolimna minima (Grunow) Lange-Bertalot & W.Schiller 1997	2	6
Eunotia minor (Kützing) Grunow 1881	4	2
Frustulia vulgaris (Thwaites) De Toni 1891	0	2
Gomphonema parvulum (Kützing) Kützing 1849	8	16
Halamphora coffeiformis (C.Agardh) Mereschkowsky 1903	0	2
Hantzschia amphioxys (Ehrenberg) Grunow 1880	0	2
Humidophila contenta (Grunow) R.L.Lowe, Kociolek, Johansen, Van de Vjiver; Lange-Bertalot & Kopalová 2014	166	14
Karayevia clevei (Grunow) Round & Bukhtiyarova 1999	4	30
Luticola goeppertiana (Bleisch) D.G.Mann ex Rarick S.Wu S.S.Lee & Edlund 2017	6	8
Luticola mutica (Kützing) D.G.Mann 1990	38	22
Melosira nummuloides C.Agardh 1824	2	2
Melosira varians C.Agardh 1827	0	2
Navicula cari Ehrenberg 1836	0	2
Navicula cincta (Ehrenberg) Ralfs 1861	0	2
Navicula cryptocephala Kützing 1844	0	4
Navicula gregaria Donkin 1861	0	28
Navicula rostellata Kützing 1844	0	4

Table 3. Diatom species found in the sampling sites located in River Muiños (*cont.*).

Navicula tripunctata (O.F.Müller) Bory de Saint-Vincent 1822	0	6
Nitzschia brevissima Grunow 1880	12	40
Nitzschia clausii Hantzsch 1860	2	0
Nitzschia inconspicua Grunow 1862	2	6
Nitzschia palea var. palea (Kützing) W.Smith 1856	6	0
Nitzschia sigmoidea (Nitzsch) W.Smith 1853	0	2
Planothidium delicatulum (Kützing) Round & Bukhtiyarova 1996	2	0
Planothidium frequentissimum (Lange-Bertalot) Lange-Bertalot 1999	8	28
Platessa oblongella (Østrup) C.E.Wetzel, Lange-Bertalot & Ector 2017	92	122
Psamothidium subatomoides (Hustedt) Bukhtiyarova & Round 1996	34	12
Rhoicosphenia abbreviata (C.Agardh) Lange-Bertalot 1980	0	2
Sellaphora meridionalis Potapova & Ponader 2008	0	10
Sellaphora subfasciata M.Potapova 2013	2	0
Tabularia tabulata (C.Agardh) Snoeijs 1992	0	6
Thalassiosira pseudonana Hasle & Heimdal 1970	4	4
Ulnaria ulna (Nitzsch) Compère 2001	0	4

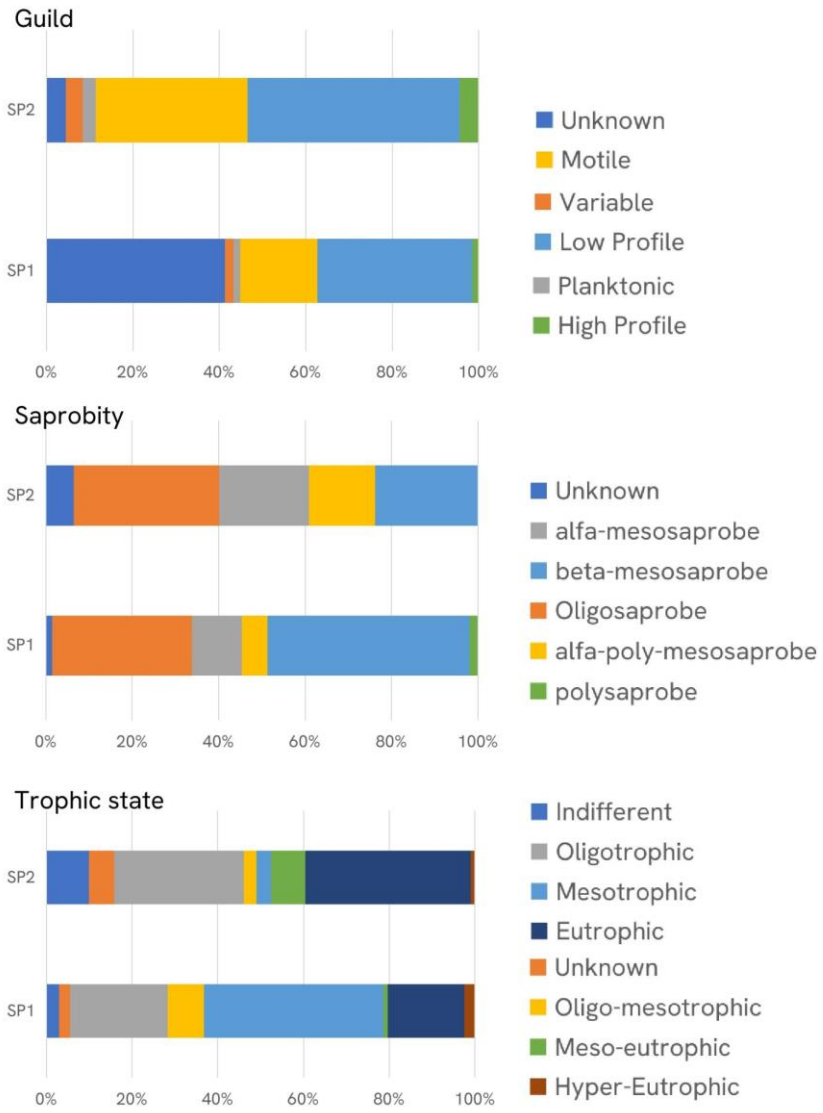

Figure 2. Guild, saprobicity and trophic state, as indicated by the species identified in the sampling sites (SP1 and SP2) in River Muiños.

Figure 2 shows the functional groups of the diatoms identified in the sampling sites in Spain. Motile and low profile diatoms were very abundant in both sampling sites. The guild classification was, however, unknown for a large amount of diatom specimens observed in SP1. Diatoms classified as ß**β**-mesosaprobic were very dominant in SP1 and in general mesosaprobic species of all types were very dominant in SP2. Regarding the trophic state, mesotrophic species dominated SP1 and eutrophic species dominated SP2.

Analysis of the samples collected in the Lis River led to the taxonomic identification of 67 species of diatoms (Table 4, Figure 3). The most abundant species were *A. pediculus* (PT1). *Achnanthidium rivulare* and *N. linearis* were also abundant in PT1. The three most abundant species in PT2 were *A. rivulare*, *C. placentula var. lineata* and *E. minima*. The sampling site PT3 was dominated by *S. pinnata*, *A. pediculus* and *M. varians*. In PT4, the most abundant species were *N. inconspicua*, *N. cryptocephala* and *K. clevei*. *Amphora pediculus* usually associated with low levels of organic pollution. *Achnanthidium rivulare* is adapted to low concentrations of phosphorus and *N. linearis* is characteristic of environments with high luminosity. *Cocconeis placentula var. lineata* is a widespread diatom adapted to alkaline to neutral pH. *Eolimna minima* is a cosmopolitan and mesosaprobic species with high tolerance to metal contamination. *Staurosirella pinnata* and *N. inconspicua* are cosmopolitan species common in environments with moderate to high organic pollution. *Navicula cryptocephala* is a widespread diatom tolerant to a wide range of environmental parameters. *Karayevia clevei* is characteristic of meso to eutrophic water conditions.

Table 4. Diatom species found in the sampling sites located in the Lis River.

Species	PT1	PT2	PT3	PT4
Achnanthidium eutrophilum (Lange-Bert.) Lange-Bert. 1999	0	0	0	24
Achnanthidium minutissimum (Kützing) Czarnecki 1994	6	4	4	0
Achnanthidium rivulare Potapova and Ponader 2004	26	232	18	0
Actinocyclus octonarius Ehrenberg 1837	0	0	1	0
Amphora meridionalis Levkov 2009	23	0	0	0
Amphora ovalis (Kützing) Kützing 1844	10	0	0	0
Amphora pediculus (Kützing) Grunow 1875	110	0	48	2
Anorthoneis excentrica (Donkin) Grunow 1870	0	0	0	1
Aulacoseira ambigua (Grunow) Simonsen 1979	0	5	0	0
Cocconeis pediculus Ehrenberg 1838	2	0	7	1
Cocconeis placentula var. euglypta (Ehrenberg) Cleve 1895	22	0	13	0
Cocconeis placentula var. lineata (Ehrenberg) Van Heurck 1885	8	45	38	4
Cymatopleura elliptica (Brébisson) W.Smith 1851	2	0	0	0
Cymbella subturgidula Krammer 2002	5	0	0	0
Diadesmis confervacea Kützing 1844	2	2	3	0
Diploneis krammeri Lange-Bertalot & E.Reichardt 2000	23	0	0	0
Eolimna minima (Grunow) Lange-Bertalot & W.Schiller 1997	0	47	0	4
Fallacia forcipata (Greville) Stickle & Mann 1990	0	1	0	0
Fallacia subhamulata (Grunow in Van Heurck) D.G.Mann in Round et al. 1990	10	0	0	0
Fistulifera saprophila (Lange-Bertalot & Bonik) Lange-Bertalot 1997	0	1	0	0
Frustulia vulgaris (Thwaites) De Toni 1891	6	0	0	0
Gogorevia exilis (Kütz.) Kulikovskiy and Kociolek 2020	0	0	0	10
Gomphonema exilissimum (Grunow) Lange-Bertalot & E.Reichardt 1996	0	2	0	0
Gomphonema parvulum (Kützing) Kützing 1849	0	6	0	0
Gomphonema truncatum Ehrenberg 1832	4	0	0	0
Gyrosigma acuminatum (Kützing) Rabenhorst 1853	7	2	0	0
Gyrosigma attenuatum (Kützing) Rabenhorst 1853	4	0	0	0
Hippodonta capitata (Ehrenberg) Lange-Bertalot Metzeltin & Witkowski 1996	13	2	0	0
Karayevia amoena (Hustedt) Bukhtiyarova 1999	0	2	9	0
Karayevia clevei (Grunow) Round & Bukhtiyarova 1999	0	0	2	39
Karayevia ploenensis var. gessneri (Hustedt) Bukhtiyarova 1999	0	3	0	0
Luticola celebesica Levkov et al.; Levkov et al. 2013	0	0	0	2
Luticola goeppertiana (Bleisch) D.G.Mann ex Rarick S.Wu S.S.Lee & Edlund 2017	0	8	0	0
Melosira varians C.Agardh 1827	8	0	44	0
Navicula capitatoradiata H.Germain ex Gasse 1986	16	2	0	0
Navicula cryptocephala Kützing 1844	0	4	2	57
Navicula cryptotenella Lange-Bertalot 1985	19	25	4	2
Navicula digitoradiata (Gregory) Ralfs 1861	8	0	0	0
Navicula gregaria Donkin 1861	8	0	4	4
Navicula phyllepta Kützing 1844	0	0	0	2
Navicula tripunctata (O.F.Müller) Bory de Saint-Vincent 1822	0	4	2	0
Navicula veneta Kützing 1844	0	0	0	4
Nitzschia amphibia Grunow 1862	0	6	13	0
Nitzschia dissipata (Kützing) Rabenhorst 1860	0	8	0	0
Nitzschia fonticola (Grunow) Grunow 1881	0	0	6	0
Nitzschia inconspicua Grunow 1862	0	7	35	287
Nitzschia intermedia Hantzsch ex Cleve & Grunow 1880	0	4	0	0
Nitzschia linearis W.Smith 1853	28	0	0	0
Nitzschia microcephala Grunow 1880	0	2	8	0
Nitzschia palea var. palea (Kützing) W.Smith 1856	6	0	6	2
Nitzschia perminuta (Grunow) M.Peragallo 1903	0	0	0	8
Nitzschia sigmoidea (Nitzsch) W.Smith 1853	0	2	0	0
Paralia sulcata (Ehrenberg) Cleve 1873	0	0	3	0

Table 4. Diatom species found in the sampling sites located in the Lis River (*cont.*).

Placoneis clementis (Grunow) E.J.Cox 1987	2	0	0	0
Planothidium engelbrechtii (Cholnoky) Round & L.Bukhtiyarova 1996	0	0	0	1
Planothidium frequentissimum (Lange-Bertalot) Lange-Bertalot 1999	0	9	14	0
Platessa oblongella (Østrup) C.E.Wetzel, Lange-Bertalot & Ector 2017	0	0	1	0
Pseudostaurosira brevistriata (Grunow) D.M.Williams & Round 1988	0	0	2	0
Rhoicosphenia abbreviata (C.Agardh) Lange-Bertalot 1980	21	12	0	0
Sellaphora pupula (Kützing) Mereschkovsky 1902	2	0	0	0
Staurosirella pinnata (Ehrenberg) D.M.Williams & Round 1987	0	13	94	0
Stephanocyclus meneghiniana (Kütz.) Kulikovskiy Genkal and Kociolek 2022	0	0	7	7
Tabularia tabulata (C.Agardh) Snoeijs 1992	0	0	4	0
Thalassiosira pseudonana Hasle & Heimdal 1970	0	0	0	5
Tryblionella acuminata W.Smith 1853	0	0	2	2
Tryblionella coarctata (Grunow) D.G.Mann 1990	4	0	0	0
Ulnaria ulna (Nitzsch) Compère 2001	0	0	10	0

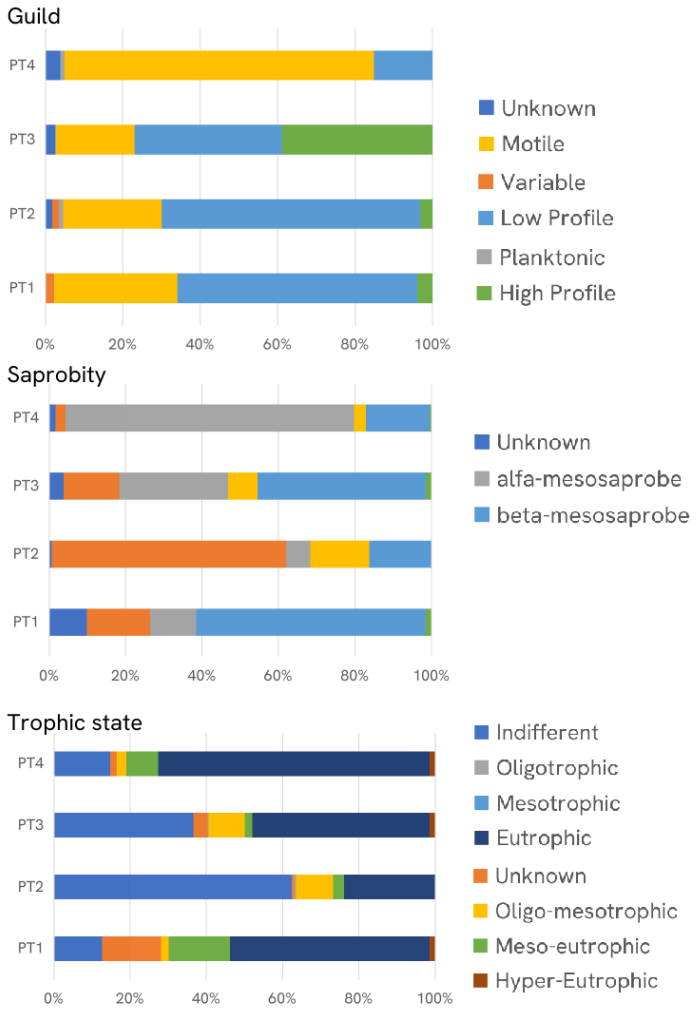

Figure 3. Guild, saprobicity and trophic state, as indicated by the species identified in the sampling sites in the Lis River.

Figure 3 shows the functional groups of the diatoms identified in the sampling sites. Low profile diatoms dominated sampling sites PT1 and PT2. In PT3 the dominant diatoms were classified as low profile and high profile and in PT4 motile diatoms were the most dominant. Polysaprobes dominated PT1 and PT3 sampling sites. The sampling site PT2 had more diatoms classified as oligosaprobes and PT4 had more diatoms classified as α-mesosaprobes. Eutrophic diatoms dominated PT1, PT3 and PT4 sampling sites. In sampling site PT2 most of the diatoms had a unknown classification in terms of trophy.

Analysis of the samples collected in Curitiba led to the taxonomic identification of 26 species of diatoms (Table 5). The most abundant species were *A. eutrophilum* in BR2, *A. catenatum* in BR4, BR3 and BR2, and *A. granulata* in BR1 and BR4. Several species found in BR4 (*D. stelligera*, *A. catenatum* and *A. eutrophilum*) are typical of eutrophic environments; their presence probably reflecting the higher nutrient levels found in this site. Furthermore, *D. stelligera* is a very common species in oligotrophic to mesotrophic lakes all over the world, which is largely influenced by nutrient and light availability. In contrast, *A. granulata*, also abundant in BR4, is a widespread taxa, tolerant to all trophic conditions, though with a tendency for higher abundance in eutrophic waters. *Planothidium frequentissimum*, detected mostly in BR3 and BR4, is considered an indicator of high organic content in the water.

Table 5. Diatom species found in the sampling sites of the Lake system at Universidade Positivo.

Species	BR1	BR2	BR3	BR4
Achnanthidium catenatum (Bilý & Marvan) Lange-Bertalot 1999	78	118	116	166
Achnanthidium eutrophilum (Lange-Bert.) Lange-Bert. 1999	89	168	44	28
Amphora ovalis (Kützing) Kützing 1844	0	0	0	1
Aulacoseira granulata (Ehrenberg) Simonsen 1979	117	68	65	100
Aulacoseira granulata var. angustissima (O.Müller) Simonsen 1979	24	14	24	0
Cocconeis placentula var. lineata (Ehrenberg) Van Heurck 1885	2	0	0	2
Diadesmis contenta (Grunow) D.G.Mann 1990	0	2	2	0
Discostella stelligera (Cleve & Grunow) Houk & Klee 2004	43	37	45	52
Eunotia incisa W.Smith ex W.Gregory 1854	0	0	0	2
Fragilaria crotonensis Kitton 1869	2	4	4	0
Gomphonema parvulum (Kützing) Kützing 1849	0	2	0	0
Gomphonema pumilum (Grunow) E.Reichardt & Lange-Bertalot 1991	0	0	0	6
Mayamaea permitis (Hustedt) K.Bruder & Medlin 2008	0	0	4	0
Navicula cryptocephala Kützing 1844	4	2	26	2
Navicula cryptotenella Lange-Bertalot 1985	2	2	2	4
Navicula oligotraphenta Lange-Bertalot & G.Hofmann 1993	0	0	2	8
Navicula rostellata Kützing 1844	0	0	4	0
Navicula viridula (Kützing) Ehrenberg 1838	0	2	0	0
Nitzschia acicularis (Kützing) W.Smith 1853	4	0	10	0
Nitzschia acidoclinata Lange-Bertalot 1976	0	4	6	2
Nitzschia capitellata Hustedt 1930	2	2	4	2
Nitzschia palea var. palea (Kützing) W.Smith 1856	2	0	2	2
Nitzschia sublinearis Hustedt 1930	14	6	0	6
Planothidium frequentissimum (Lange-Bertalot) Lange-Bertalot 1999	26	0	55	30
Thalassiosira pseudonana Hasle & Heimdal 1970	0	0	4	0
Ulnaria ulna (Nitzsch) Compère 2001	0	0	0	2

Figure 4 shows the functional groups of the diatoms identified in the sampling sites. The diatoms found were predominantly motile and low profile. In some sites, the low profile diatoms outnumbered motile diatoms, such as in BR2, BR3 and BR4. In terms of saprobicity, β-mesosaprobes were dominant in all sampling sites, despite the high number of diatoms with no known classification in BR3 and BR4. Eutrophic diatoms were very abundant in all sampling sites. Diatoms with unknown trophic classification were more abundant in BR3 and BR4.

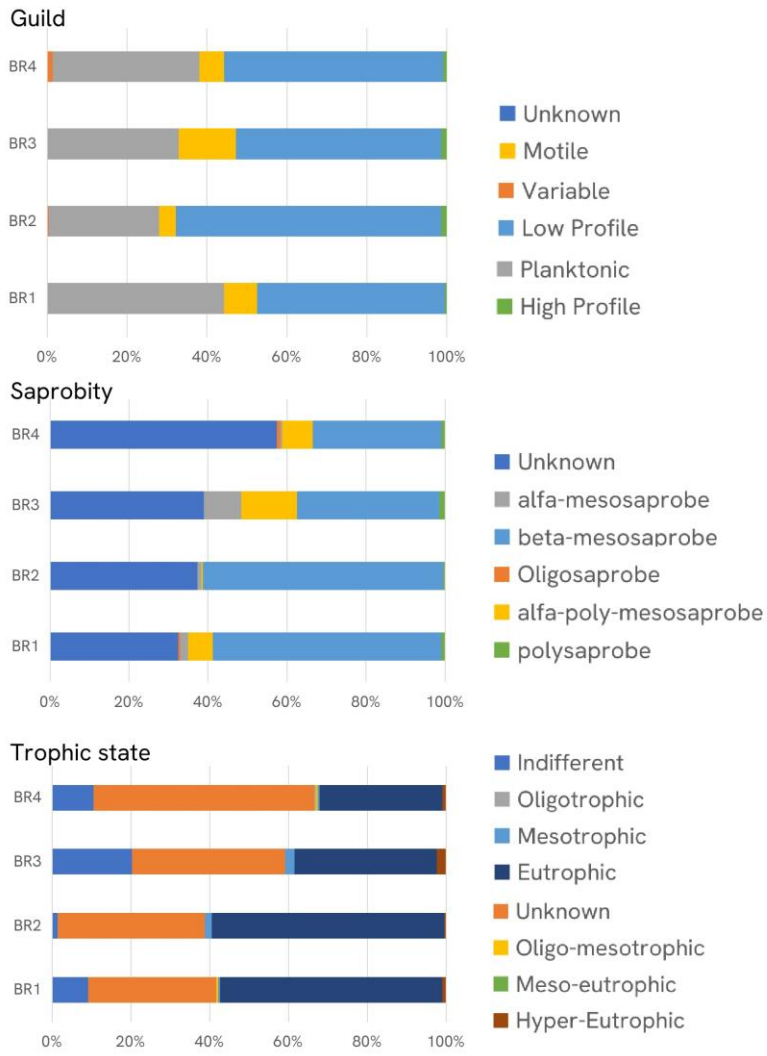

Figure 4. Guild, saprobicity and trophic state, as indicated by the species identified in the sampling sites (BR1 to BR4) in the Lake system at Universidade Positivo.

5.4. Ecological water quality

The Shannon diversity index (H') and the Evenness (J') index were also calculated. The conventional diatom indices provided by Omnidia have a significant drawback, which is that they are region specific. Most of them were developed by integrating functional traits of diatoms from specific regions, though they are used worldwide. Nevertheless, this prevents accurate, meaningful comparisons across sampling sites. The Shannon and Evenness indices are free of such limitations, because their calculation is based on diversity and abundance or richness of species (calculation details provided in the Annexe). A higher H' value indicates higher diversity, while a lower value indicates lower diversity. An H' index of 0 means only one species is present in the community (i.e. no diversity). To better interpret the results, you one can calculate the Shannon Equitability index (or Evenness) to see how close the abundance of species is to a perfectly even distribution.

Equitability is the general ecological principle of how evenly individuals are distributed among different species in a community. The evenness index is the specific, calculated measurement of that concept or, in other words, its metric. The terms are often used interchangeably. A community with high equitability has species that are all represented by roughly equal numbers of individuals. A community with low equitability has one or a few dominant species and many species with very few individuals. This index, thus, improves the interpretation of the Shannon Index. For example, two communities could have the same Shannon index, but one might have many species with low abundance, while the other has fewer species with very high abundance. The Equitability index

helps distinguish between these scenarios. It is a normalised value between 0 and 1 that measures how close the species' abundances are to being perfectly even; a J value closer to 1 indicates that all species are found in roughly equal numbers, while a value closer to 0 means that a few species dominate the community. In general terms, a J' above 0.7 indicates a high to very high evenness or a well-balanced community with nearly equal abundances of all species. A J' value below 0.5 indicates a low to very low evenness, i.e. an unbalanced community with a few species which are noticeably more common than others.

Table 6 shows the results obtained for H' and J' calculated. In Sweden, the sites all showed J' values above 0.7, indicating high to very high evenness. The lowest J' values were found in SP1, PT2, PT4, BR2 and BR4. These results indicate that in these sites the communities were somewhat unbalanced, with some species appearing more common than others. This probably occurs because these sites were more closely associated to WWTP outlets. In Portugal a gradient was found, with PT1 (upstream the WWTP) showing globally higher chemical and ecological quality, which lowers significantly in PT2 (site near the WWTP outlet). The quality levels are subsequently recovered in PT3 as expected, considering the sampling gradient. Finally, in PT4 the ecosystem quality lowers again, reaching the worst condition among all sampling sites investigated. This is apparently related to the presence of other nearby sources discharging into the same river system, including a second WWTP. The site was selected for further investigation to detect and understand the environmental pressures and consider potential mitigation measures.

Table 6. Shannon Diversity (H') and Evenness (J') indices obtained for each sampling site.

Indices	SW1	SW2	SW3	SP1	SP2	PT1	PT2	PT3	PT4	BR1	BR2	BR3	BR4
H'	4.37	3.26	4.28	2.62	3.89	4.05	2.91	3.87	2.19	2.84	2.35	3.17	2.57
J'	0.81	0.77	0.79	0.62	0.76	0.83	0.61	0.8	0.50	0.73	0.62	0.76	0.63

With the physico-chemical data and the Shannon and Evenness indices an integrated quality index (WSE) was calculated and used as global descriptor of ecosystem quality. The WSE was estimated as a weighted sum of these data, following the approach described by Crouzet et al. (1999). The results are presented in Table 7, along with the description of the quality classes. The WSE was used as a descriptor against which the Raman values were modelled to understand and evaluate how would they predict ecosystem quality. Water quality varied from Poor (SW2) to Good (SW3, SP2, PT1 and PT3). The remaining sites showed Fair water chemical quality.

Table 7. Global water quality index (WSE) estimated from classical methods (*left*) and respective quality classes (*right*).

Site	WSE	Quality class
SW1	69	3
SW2	49	4
SW3	80	2
SP1	68	3
SP2	82	2
PT1	79	2
PT2	67	3
PT3	83	2
PT4	57	3
BR1	63	3
BR2	63	3
BR3	63	3
BR4	54	3

Quality class	WSE	Description
1	90–100	Excellent
2	70–90	Good
3	50–70	Fair
4	25–50	Poor
5	0–25	Very Poor

6. Diatom Raman spectroscopy and draft index

All samples collected were also taken to analysis by Raman spectroscopy. Raman spectra were obtained from 468 diatom specimens identified in the samples sites (12 spectra per species, three most abundant species per site, 13 sites in total). Nineteen Raman bands were identified in the spectra acquired, resulting in a total of 8208 bands that were taken into deconvolution into three variables each (i.e. frequency, width and area). Figure 5 shows examples of Raman spectra of the most abundant species in each sampling site and Table 8 shows literature band assignments for the most common bands. The criterion for selection of these bands was their identification in at least 5 spectra of all 12 spectra collected in the three species analysed per sampling site.

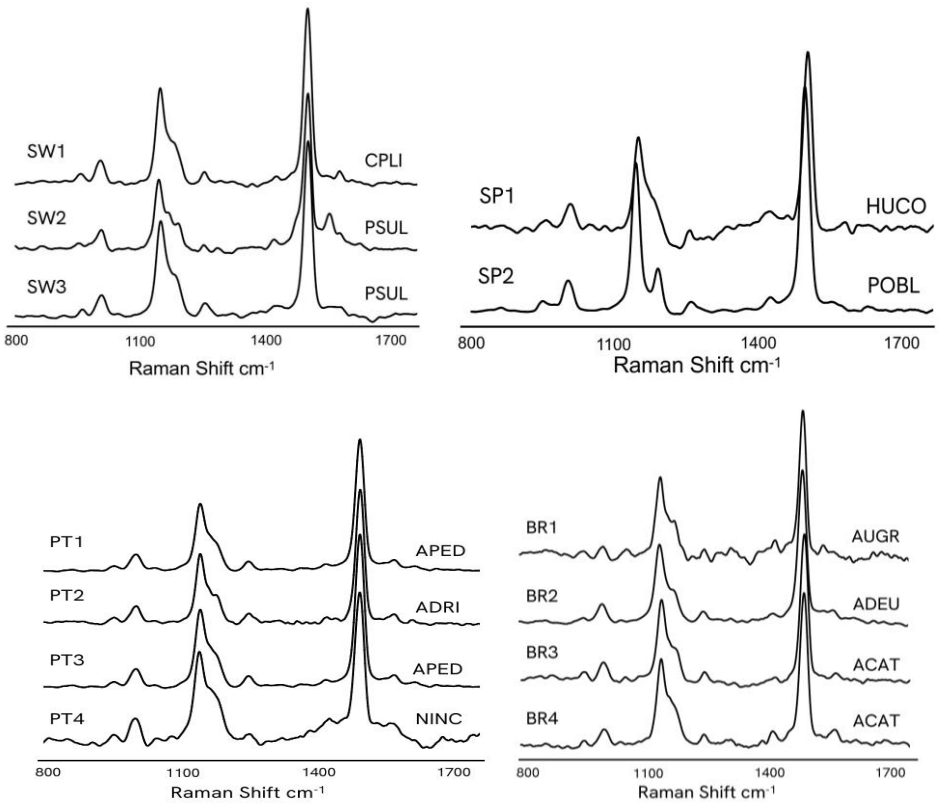

Figure 5. Examples of Raman spectra obtained from the most abundant species in each sampling site; *Cocconeis lineata* (CPLI), *Paralia sulcata* (PSUL), *Humidophila contenta* (HUCO), *Platessa oblongella* (POBL), *Amphora pediculus* (ADED), *Achanthidium rivulare* (ADRI), *Nitzschia inconspicua* (NINC), *Aulacoseira granulata* (AUGR), *Achnanthidium eutrophilum* (ADEU), *Achnanthidium catenatum* (ACAT).

Table 8. Molecular assignments identified for the most frequent bands found in the Raman diatom spectra.

Band (cm^{-1})	Molecular Assignments	Reference
1011	CH3 Stretching modes from carotenoids CC aromatic ring chain from frustule	Alexandre et al. (2014) De Tommasi (2016) De Tommasi et al. (2018) Premvardhan et al. (2009)
1161	C-C stretching modes from carotenoids C=S from frustule	De Tommasi (2016) De Tommasi et al. (2018) Premvardhan et al. (2009) Rüger et al. (2016)
1181	C-H deformational modes from carotenoids C=S from frustule	De Tommasi (2016) De Tommasi et al. (2018) Premvardhan et al. (2009)
1198	N-C stretching modes from Chla C=S frustule	De Tommasi (2016) De Tommasi et al. (2018)
1526	C=C stretching modes from carotenoids	Alexandre et al. (2014) Premvardhan et al. (2009) Rüger et al. (2016)

The first analysis done included all Raman data obtained from the three most abundant species in each sampling site. A relevant question coming out from the previous proof of concept that served as basis for this work (Oliva-Teles et al., 2022), was related to

the best approach to adopt in case no common diatom species are found among sites. In more closed systems, such as the one investigated in Oliva-Teles et al. (2022), variability of diatom communities is lower, such as found in the Brazilian lake, and common species often occur. In open systems, such as rivers and streams, natural diversity of diatom communities is expected to be higher with few or none species in common. The approach adopted was thus to acquired Raman spectra from the three most common species in each site. The line of approach was based on the hypotheses that the most frequent species would reflect local environmental conditions (e.g. biodiversity loss due due to pollution and climate stress) and this would be detected through Raman spectroscopy given that the technique assesses the biochemical profile resulting from these conditions, much as biochemical biomarkers have previously been used as early-warning signals (Rodrigues et al., 2014; Teles et al., 2017; Rodrigues et al., 2019). The results of a PCA (principal components analysis) on all Raman data obtained showed, however, some redundancy or insignificant contribution of some variables. Furthermore, they suggested that 15 variables obtained from five Raman bands (1011, 1161, 1181, 1198 and 1526; Figure 6) of the most common species in each site would be sufficient to distinguish among them. These minimal data were thus employed to develop a predictive model through a saturated orthogonal multiple linear regression (Rodrigues et al., 2023). Here, the WSE was entered in the model as dependent variable and the Raman data (15 variables derived from the five important bands identified) were entered as predictors. The model obtained, proposed as a diagnostic tool, was able to discriminate among less and more impact sites, along the pollution gradients investigated (Figure 7). The highest the values, the more impacted are the sites. The model showed a very high sensitivity with an R2 between the measured and the estimated values of 0.92, which is excellent.

Figure 6. Principal components analysis highlighting the Raman bands showing higher importance to explain the variability in the dataset. The letter associated to the band number indicate the respective frequency (F), width (W) and area (A).

The most interesting result here was the ability to discriminate the sites using a minimum of Raman data derived from five bands obtained from the spectra of the most abundant diatom species in each site. Not only these bands were found to be very sensitive to the environmental conditions, but only one species was enough to distinguish the sites according to their global quality. This is highly relevant and one of the most important aspects to confirm in larger datasets, as it makes the method even more expeditious, easy of application and less time consuming. All relevant aspects when considering routine application. In addition to this, the bands analysed showed an excellent correlation with the integrated global quality descriptor of the ecosystem, but not with each of

its components independently (i.e. common water quality assessment, H' or J'). This suggests that Raman is able to reflect the response to the combined exposure to environmental conditions/stressors. On this regard, it is also of note that the bands able to discriminate among pollution types and levels in the tool developed were previously allocated to carotenoids, chlorophyll a and frustule compounds (Table 8). Moreover, differences in the frequency of a band detected under fixed laser excitation are commonly considered to be related to the presence of different molecular conformations and/or intermolecular interactions. Such changes may thus be indicative of differences in carotenoid molecules and related processes. The band area has been associated to the concentration of the molecules assigned, the smaller the area the lower the concentration. The band width reflects the local environment of the target molecule in the cell, with higher values indicating higher disorder and intermolecular interactions.

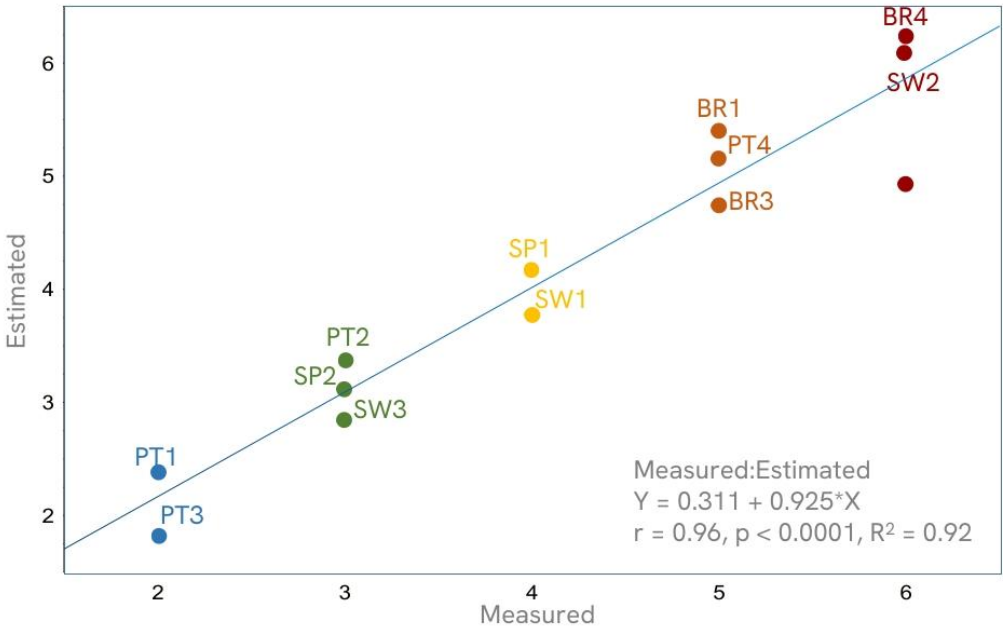

Figure 7. Goodness of fit assessment between the measured values and the values predicted by the model developed (estimated).

Overall, the results obtained thus suggest that survival and/or adaptation of diatoms to challenging environmental stress (pollution and climate conditions) may be mediated by the antioxidant system. In diatoms, carotenoids serve crucial roles such as: i) extending their ability to absorb blue-green light for photosynthesis in light harvesting (e.g. fucoxanthin absorbs light in the blue-green spectrum, transferring energy to chlorophylls), ii) acting in photoprotection (i.e. dissipating excess light energy to prevent damage via the xanthophyll cycle involving diadinoxanthin/diatoxanthin), or iii) bearing antioxidant activity (acting as powerful antioxidants, scavenging harmful free radicals like singlet oxygen and hydroxyl radicals, protecting cells from oxidative stress) (Pinto et al. 2021 and references herein). They also play an important structural role by building and maintaining the stability of the light-harvesting pigment-protein complexes (FCPs) within the cell. Given the well-known role of the antioxidant system on coping with pollutant exposure, the results point out a significant pathway of resilience, also potentially useful as early warning signal on the diagnosis based on Raman spectroscopy. This hypothesis will be investigated in future work set to better understand the molecules driving the adaptation and resilience pathways and serving as diagnostic and follow up biomarkers. Future work will also focuses on testing the tool with the data from the eight other Portuguese systems sampled. Here the data will involve an analytical characterisation of the sites (on course) for relevant emerging contaminants of concern (e.g. pharmaceuticals and their transformation products), in addition to the physico-chemical parameters and the H' and J' indices, as well comparison with other biological effects data (e.g. biomarkers, metabarcoding). This will help refine the tool and improve its diagnostic robustness and

sensitivity for use in other regions. The work presented herein is under preparation for submission to publication in an international indexed journal.

7. Associated indicators

The indicators which led to the development of this work that made or will make use of it are indicated below.

Publications

1. Pinto R, Vilarinho R, Carvalho AP, Moreira JA, Guimarães L, Oliva-Teles L. 2022. Novel Approach to Freshwater Diatom Profiling and Identification Using Raman Spectroscopy and Chemometric Analysis. Water 14:2116. DOI: 10.3390/w14132116
2. Oliva-Teles L, Pinto R, Vilarinho R, Carvalho AP, Moreira JA, Guimarães L. 2022. Environmental diagnosis with Raman Spectroscopy applied to diatoms. Biosensors and Bioelectronics 198:113800. DOI: 10.1016/j.bios.2021.113800

Communications

1. Sousa, M., Pinto, R., Carvalho, A.P., Guimarães, L., Oliva-Teles, L. Diatom taxonomic identification in the Lima River Estuary. Blue Think Conference 2023. Poster presentation | September 13, 2023 | Matosinhos, Portugal
2. Pinto, R., Vilarinho, R., Carvalho, A.P., Agostinho, J., Guimarães, L., Oliva-Teles, L. Raman spectroscopy applied to diatoms as an alternative to conventional water quality assessment. Blue Think Conference 2023, Poster presentation | September 13, 2023 | Matosinhos, Portugal
3. Pinto, R., Vilarinho, R., Carvalho, A.P., Agostinho, J., Guimarães, L., Oliva-Teles, L. Potential of Raman Spectroscopy for Environmental diagnosis using diatoms. Ciência 2023 - Encontro com a Ciência e a Tecnologia - Ciência e Oceano para Além do Horizonte. Poster presentation | July 5-7, 2023 | Aveiro, Portugal
4. Sousa, M., Pinto, R., Carvalho, A.P., Guimarães, L., Oliva-Teles, L. Diversity of Saltmarsh Diatom Communities in the Lima River Estuary. Ciência 2023 - Encontro com a Ciência e a Tecnologia - Ciência e Oceano para Além do Horizonte. Poster presentation | July 5-7, 2023 | Aveiro, Portugal
5. Sousa, M., Pinto, R., Carvalho, A.P., Guimarães, L., Oliva-Teles, L. Community Composition of saltmarsh diatoms in the Lima River Estuary. OCEAN3R Final Workshop and meeting. Poster presentation | June 27, 2023 | Matosinhos, Portugal
6. Pinto, R., Vilarinho, R., Carvalho, A.P., Agostinho, J., Guimarães, L., Oliva-Teles, L. Water Quality assessment using Raman Spectroscopy applied to Diatoms. OCEAN3R Final Workshop and meeting. Poster presentation | June 27, 2023 | Matosinhos, Portugal

PhD thesis

1. Raquel Pinto. Ecosystem assessment and monitoring using diatom Raman spectroscopy. PhD in Biology, University of Porto (Portugal) | 2023 (on course), Fundação para a Ciência e a Tecnologia (FCT) (Portugal) grant 2022.09984.BD

Master dissertation

1. Marco Sousa (2022) Characterisation of saltmarsh diatoms in the Lima River Estuary: Community composition and Raman Spectroscopy applied for environmental diagnosis. Master in Toxicology and Environmental Contamination. Institute of Biomedical Sciences Abel Salazar (ICBAS), University of Porto.

Stakeholder interaction

Based on the work proposed and developed, leader of P1(PT) and P7(PT) were selected to join the BiodivRestore Knowledge Hub (KHub) founded by Biodiversa+ to support nature restoration and the implementation of the Nature Restoration Law. They integrated the Monitoring and Learned Lessons workstream, within the Monitoring & Innovation Task Force. Subsequently (in late 2024), leader of P7(PT) was invited to act as Co-Chair of the Monitoring & Innovation Task Force.

References

Alexandre, M. T., Gundermann, K., Pascal, A. A., van Grondelle, R., Büchel, C., & Robert, B. (2014) Probing the carotenoid content of intact Cyclotella cells by resonance Raman spectroscopy. Photosynthesis Research 119: 273-281.

Crouzet, P., Leonard, J., Nixon, S., Rees, Y., Parr, W., Laffon, L., Bøgestrand, J., Kristensen, P., Lallana, C., Izzo, G., Bokn, T., Bak, J., Lack, T.J. (1999) Nutrients in European ecosystems. In: Thyssen, N. (ed). Environmental Assessment Report No. 4. Copenhagen: European Environmental Agency, 82.

De Tommasi, E. (2016) Light manipulation by single cells: the case of diatoms. Journal of Spectroscopy, 2016: 2490128.

De Tommasi, E., Congestri, R., Dardano, P., De Luca, A. C., Managò, S., Rea, I., & De Stefano, M. (2018) UV-shielding and wavelength conversion by centric diatom nanopatterned frustules. Scientific Reports 8: 16285.

Oliva-Teles, L., Pinto, R., Vilarinho, R., Carvalho, A.P., Moreira, J.A., Guimarães, L. (2022) Environmental diagnosis with Raman Spectroscopy applied to diatoms. Biosensors and Bioelectronics 198: 113800

Paíga, P., Santos, L., Ramos, S., Jorge, S., Silva, J., Delerue-Matos, C. (2016) Presence of pharmaceuticals in the Lis river (Portugal): Sources, fate and seasonal variation. Science of the Total Environment, 573: 164–177

Pielou, E.C. (1966) The Measurement of Diversity in Different Types of Biological Collections. Journal of Theoretical Biology 13: 131-144.

Pinto, R., Vilarinho, R., Carvalho, A.P., Moreira, J.A., Guimarães, L., Oliva-Teles, L. (2021) Raman spectroscopy applied to diatoms (microalgae, Bacillariophyta): Prospective use in the environmental diagnosis of freshwater ecosystems. Water Research 198: 117102

Pinto, R., Vilarinho, R., Carvalho, A.P., Moreira, J.A., Guimarães, L., Oliva-Teles, L. (2022) Novel Approach to Freshwater Diatom Profiling and Identification Using Raman Spectroscopy and Chemometric Analysis. Water, 14 (13): 2116

Premvardhan, L., Bordes, L., Beer, A., Büchel, C., & Robert, B. (2009). Carotenoid structures and environments in trimeric and oligomeric fucoxanthin chlorophyll a/c2 proteins from resonance Raman spectroscopy. The Journal of Physical Chemistry B 113: 12565-12574.

Rodrigues, A.P., Oliva-Teles, T., Mesquita, S.R., Delerue-Matos, C., Guimarães, L. (2014) Integrated biomarker responses of an estuarine invertebrate to high abiotic stress and decreased metal contamination. Marine Environmental Research, 101: 101 - 114

Rodrigues, C., Bio, A., Guimarães, L., Cruz Fernandes, V.C., Delerue Matos, C., Vieira, N. (2019) Assessing the ecological status of fluvial ecosystems employing a macroinvertebrate multi-taxon and multi-biomarker approach. Environmental Monitoring and Assessment, 191: art. no. 503

Rodrigues, P., Guimarães, L., Carvalho, A.P., Oliva-Teles, L. (2023) Carbamazepine, venlafaxine, tramadol, and their main metabolites: Toxicological effects on zebrafish embryos and larvae. Journal of Hazardous Materials, 448: 130909

Rüger, J., Unger, N., Schie, I. W., Brunner, E., Popp, J., & Krafft, C. (2016). Assessment of growth phases of the diatom Ditylum brightwellii by FT-IR and Raman spectroscopy. Algal research 19: 246-252.

Teles, M., Soares, A.M.V.M., Tort, L., Guimarães, L., Oliveira, M. (2017) Linking cortisol response with gene expression in fish exposed to gold nanoparticles. Science of the Total Environment, 584-585: 1004 - 1011

Annexes

Diatoms communities from Höje Rive (Sweden)

Study site

The Höje River located in southern Sweden. It extends over ~35 km, flowing into the Öresund Strait south of Malmö. The catchment is dominated by intensive agriculture, and the river has been recognised as highly impacted by nutrient loading, particularly phosphorus and nitrogen from diffuse agricultural runoff and urban wastewater. The Källby wastewater treatment plant (WWTP), located in the southwestern part of the city of Lund, discharges threated effluents to this system. The WWTP treats the residual water from Lund and its nearby villages serving a population of about 86,000 people (Larsson et al., 2013). The plant comprises primary (bar screen, grit removal, primary settling), secondary (conventional active sludge with anoxic pre-denitrification) and tertiary (phosphate precipitation with ferric chloride) treatments. Three sampling sites were established in this system, identified as SW1 to SW3 (Figure 1). One site was located upstream to the WWTP (SW1), one was within a constructed wetland adjacent to the WWTP (SW2) and one site was located downstream to the WWTP (SW3). The samples were collected and described in Oliva-Teles et al. (2022) and Pinto et al. (2022). Subsequent processing, taxonomic identification and analysis through Raman spectroscopy followed the methods described therein.

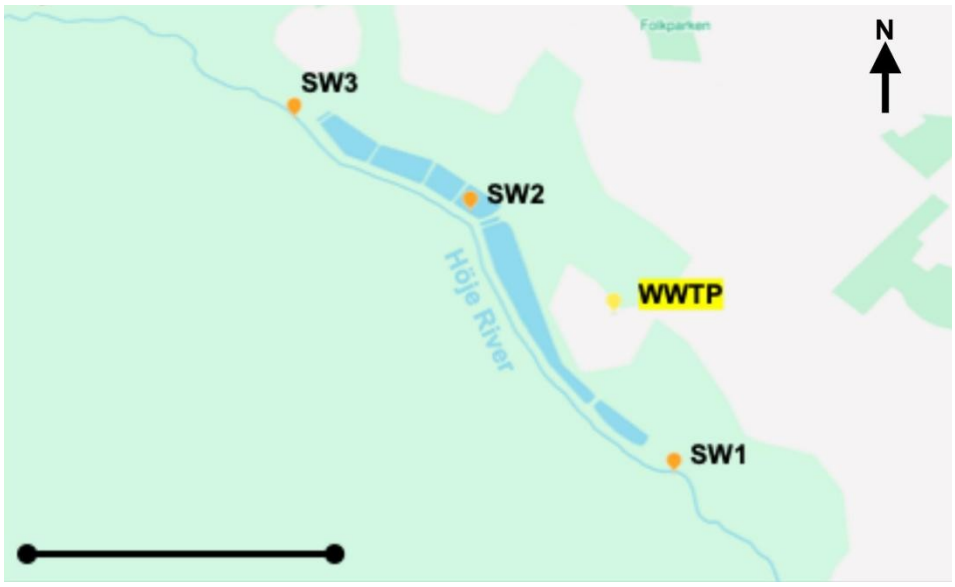

Figure 1. Location of the sampling sites and wastewater treatment plant (WWTP) in Höje River (scale bar 1 Km).

In situ measurements and nutrient concentrations

Table 1 shows the physico-chemical measurements done in situ, at the time of water sampling, and the nutrient concentrations determined later in the lab. Temperature varied between 6.8 °C and 7.3 °C across the sampling sites. Salinity and conductivity were much higher in SW1 and SW2 than in SW3 (6 µS cm -1; 6 ‰). Oxygen concentration was much lower in SW2 (1.69 mg L^{-1}) than in SW1 (5.66 mg L^{-1}) and SW3 (4.83 mg L^{-1}) . The pH ranged from 7.7 (SW3) to 8.3 (SW1). Nutrient concentrations varied greatly between sampling

sites. Higher nitrite and silicate concentrations were found upstream the WWTP (SW1) and higher nitrate, ammonia and phosphate concentrations were found near the WWTP (SW2).

Table 1. Physico-chemical parameters and nutrient concentrations determined in the samples collected.

	SW1	SW2	SW3
Conductivity (µS cm $^{-1}$)	667	787	6
Salinity (‰)	1.830	2.130	0.015
O_2 (mg L^{-1})	5.66	1.69	4.83
pH	8.3	7.8	7.7
NO_2^-(mg L^{-1})	0.73 (±0.0015)	0.44 (± 0.0043)	0.06 (± 0.0096)
NO_3^- (mg L^{-1})	0.43 (±0.0577)	2.30 (±0.0000)	1.70 (±0.1000)
NH_4^+ (mg L^{-1})	1.69 (±0.0058)	2.60 (±0.0000)	0.24 (±0.0058)
PO_4^{3-} (mg L^{-1})	1.01 (±0.0231)	3.27 (±0.0461)	0.15 (±0.0000)
SiO_2 (mg L^{-1})	6.30 (±0.0000)	6.28 (±0.0000)	6.16 (±0.0000)

Diatom communities

Analysis of the samples collected led to the taxonomic identification of 72 species of diatoms (Table 2). The most abundant species were *P. sulcata* in SW2, *C. placentula* in SW 1 and SW3, and *M. permitis* in SW2. *Cocconeis placentula var. lineata* is a cosmopolitan freshwater species frequently found in neutral to alkaline waters. *Aulacoseira granulata* (SW1) is a widespread taxa tolerant to all trophic conditions but more abundant in eutrophic waters. *Navicula lanceolata* (SW1) is a very common diatom indicative of eutrophic conditions and oligo to mesosaprobic conditions. This diatom is also commonly found in low temperature sites. *Paralia sulcata* is a species typically found in low temperature conditions and high nutrients levels. *Mayamaea permitis* is cosmopolitan and very abundant in Europe. It is usually found in masses in wastewater, sewage plants and other polysaprobic habitats. *Lemnicola hungarica* (SW2) is a widespread diatom frequently found attached to duckweed (*Lemna minor*). *Achnanthidium minutissimum* (SW1 and SW3) is a ubiquitous freshwater diatom with a wide ecological tolerance.

Table 2. Diatom species found in the sampling sites in the Höje River.

Species	**SW1**	**SW2**	**SW3**
Achnanthes adnata Bory 1822	0	0	10
Achnanthidium minutissimum (Kützing) Czarnecki 1994	28	16	30
Amphora liriope Nagumo 2003	0	1	0
Amphora ovalis (Kützing) Kützing 1844	12	0	0
Amphora pediculus (Kützing) Grunow 1875	12	0	0
Aulacoseira granulata (Ehrenberg) Simonsen 1979	40	0	10
Cocconeis placentula var. *lineata* (Ehrenberg) Van Heurck 1885	93	30	73

Diadesmis confervacea Kützing 1844	0	0	2
Diadesmis contenta (Grunow) D.G.Mann 1990	4	2	0
Diploneis voigtiana Lange-Bertalot & Fuhrmann 2020	8	0	0
Encyonema ventricosum (C.Agardh) Grunow 1875	2	0	4
Eolimna minima (Grunow) Lange-Bertalot & W.Schiller 1997	2	10	0
Fragilaria tenera (W.Smith) Lange-Bertalot 1980	0	0	2
Geissleria kriegeri (Krasske) Lange-Bertalot 1996	10	0	0
Gogorevia exilis (Kütz.) Kulikovskiy and Kociolek 2020	0	16	0
Gomphonema acuminatum Ehrenberg 1832	0	0	4
Gomphonema olivaceum (Hornemann) Brébisson 1838	0	0	6
Gomphonema parvulum (Kützing) Kützing 1849	21	12	4
Gomphonema saprophilum (Lange-Bertalot & E.Reichardt) Abraca, R.Jahn, J.Zimmermann & Enke 2014	0	0	20
Halamphora veneta (Kützing) Levkov 2009	6	0	0
Haslea crucigera (W.Smith) Simonsen 1974	0	0	2
Hippodonta costulata (Grunow) Lange-Bertalot, Metzeltin & Witkowski 1996	2	0	0
Karayevia clevei (Grunow) Round & Bukhtiyarova 1999	0	21	0
Platessa oblongella (Østrup) C.E.Wetzel, Lange-Bertalot & Ector 2017	2	2	0
Lemnicola hungarica (Grunow) Round & Basson 1997	0	38	0
Luticola goeppertiana (Bleisch) D.G.Mann ex Rarick S.Wu S.S.Lee & Edlund 2017	0	0	4
Mayamaea permitis (Hustedt) K.Bruder & Medlin 2008	0	57	0
Melosira varians C.Agardh 1827	21	0	7
Nanofrustulum cataractarum (Hustedt) C.E.Wetzel, E.Morales & Ector 2019	0	0	2
Navicula capitatoradiata H.Germain ex Gasse 1986	2	0	0
Navicula cryptocephala Kützing 1844	2	16	2
Navicula cryptotenella Lange-Bertalot 1985	2	0	0
Navicula cryptotenelloides Lange-Bertalot 1993	0	0	2
Navicula digitoradiata (Gregory) Ralfs 1861	0	0	8
Navicula exilis Kützing 1844	0	0	6
Navicula expecta VanLandingham 1975	2	0	0
Navicula gregaria Donkin 1861	8	0	2
Navicula lanceolata (C.Agardh) Kützing, 1844	30	0	12
Navicula tenelloides Hustedt 1937	2	0	0
Navicula trivialis Lange-Bertalot 1980	2	0	0
Navicula tripunctata (O.F.Müller) Bory de Saint-Vincent 1822	0	0	16
Nitzschia amphibia Grunow 1862	0	0	4
Nitzschia capitellata Hustedt 1930	4	2	0
Nitzschia dissipata (Kützing) Rabenhorst 1860	0	0	2
Nitzschia filiformis (W.Smith) Van Heurck 1896	12	0	0
Nitzschia fonticola (Grunow) Grunow 1881	0	0	2
Nitzschia inconspicua Grunow 1862	4	8	4
Nitzschia palea var. palea (Kützing) W.Smith 1856	10	2	4
Nitzschia recta Hantzsch 1862	4	0	8
Nitzschia sigma (Kützing) W.Smith 1853	4	0	0
Nitzschia sigmoidea (Nitzsch) W.Smith 1853	0	0	12
Opephora pacifica (Grunow) P.Petit 1889	0	0	2
Paralia sulcata (Ehrenberg) Cleve 1873	20	137	99
Parlibellus E.J.Cox, 1988	2	0	0
Pinnularia halophila Krammer 1992	2	0	0
Pinnularia subcapitata W.Gregory 1856	6	0	0
Planothidium delicatulum (Kützing) Round & Bukhtiyarova 1996	0	0	6
Planothidium frequentissimum (Lange-Bertalot) Lange-Bertalot 1999	4	26	8
Planothidium lanceolatum (Brébisson ex Kützing) Lange-Bertalot 1999	0	0	2

Psamothidium subatomoides (Hustedt) Bukhtiyarova & Round 1996	0	0	3
Rhoicosphenia abbreviata (C.Agardh) Lange-Bertalot 1980	8	4	14
Sellaphora pupula (Kützing) Mereschkovsky 1902	2	0	2
Stauroneis smithii Grunow 1860	0	0	6
Stephanocyclus meneghiniana (Kütz.) Kulikovskiy Genkal and Kociolek 2022	0	0	3
Surirella angusta Kützing 1844	2	0	0
Surirella brebissonii Krammer & Lange-Bertalot 1987	12	0	0
Tabularia tabulata (C.Agardh) Snoeijs 1992	4	4	4
Thalassiosira pseudonana Hasle & Heimdal 1970	2	0	0
Tryblionella acuminata W.Smith 1853	2	0	0
Tryblionella calida (Grunow) D.G.Mann 1990	0	0	2
Tryblionella hungarica (Grunow) Frenguelli 1942	0	0	2
Ulnaria ulna (Nitzsch) Compère 2001	0	0	8

Ecology

Figure 2 shows the functional groups of the diatoms identified in the sampling sites. The diatoms found were predominantly motile and low profile. Sites SW1 and SW3 were dominated by βmesosaprobic species. The site SW2 showed a prodominance of α-mesosaprobe and α-polymesosaprobe species. In terms of trophic state, all sampling sites had high numbers of eutrophic species. Moreover, a significantly high number of species were classified as unknown in all functional groups studied.

Ecological quality

A physico-chemical indice (WQI) was also calculated using the measured *in situ* parameters and nutrient concentrations. According to the results the water is generally considered to have a good chemical status (Table 3). Based on the species identified in the sampling sites, the most common diatom indices were also estimated, using the Omnidia software: Diatom Biologic Indice (IBD), Polluosensitivity Indice (IPS) and European Communities Indice (CEE). According to the IPS, SW1 and SW2 were identified as of poor ecological quality (indicated in orange in Table 3).

Table 3. Diatom (IBD, IPS and CEE) and physico-chemical (WQI) water quality indices calculated for the sampling sites.

	SW1	**SW2**	**SW3**
WQI (1-100)	69	49	80
IBD (1-20)	12.1	10.7	11.7
IPS (1-20)	7.3	8.1	9.6
CEE (1-20)	10	10.2	11.9

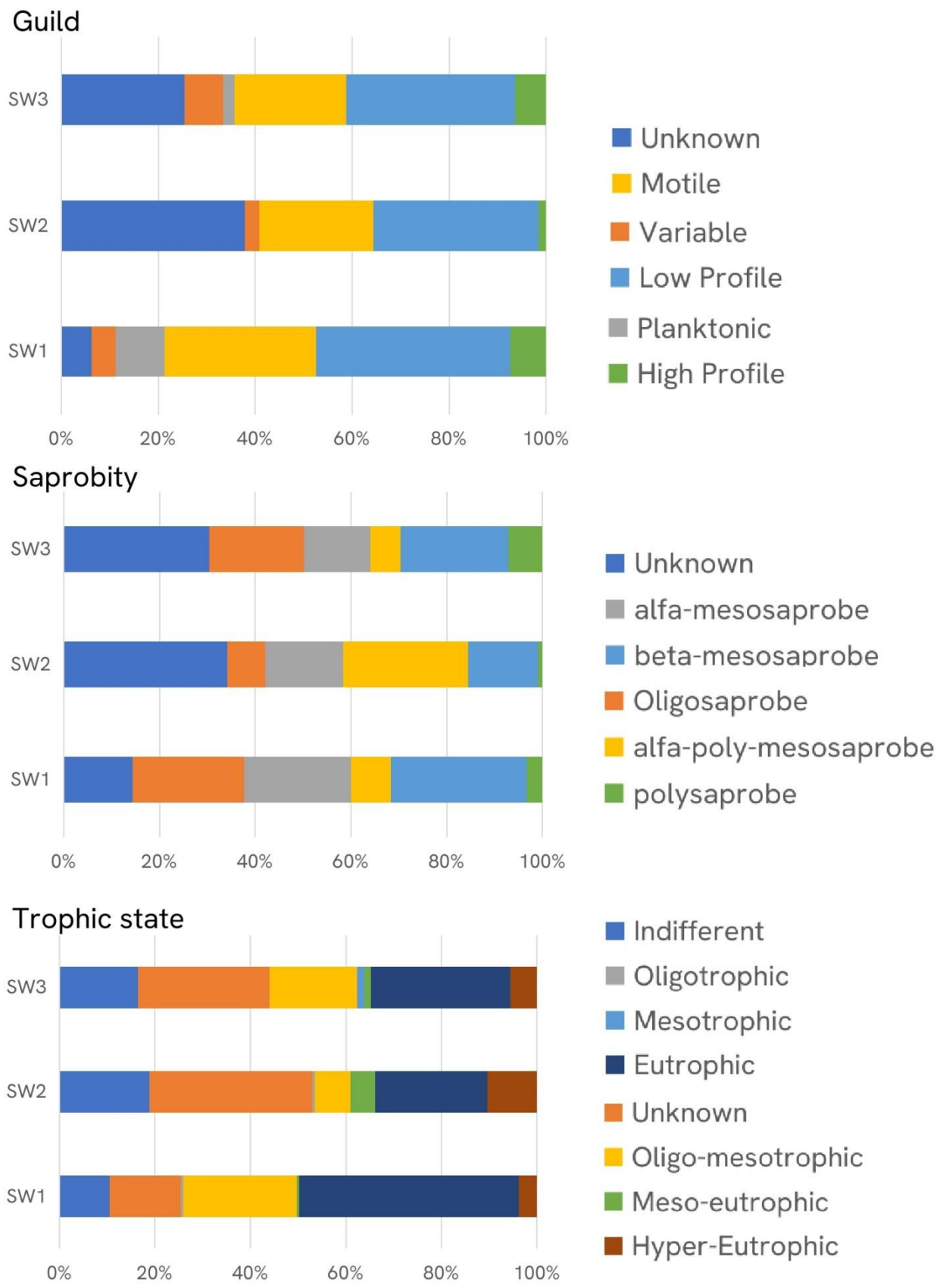

Figure 2. Guild, saprobicity and trophic state, as indicated by the species identified in the sampling sites (SW1 to SW3).

Table 4 shows the results of the Shannon diversity index (H′) and the Evenness index (J) calculated for the sampling sites. The H′ values above 3 indicate high species richness and abundance. Values of the Evenness index (J) are interpreted by how close they are to 1. Values closer to 1 indicating a more balanced community where species are present in nearly equal numbers; values closer to 0 indicating an unbalanced community where one or a few species dominate (further details on the interpretation of these indices are given in the Annexe).

Table 4. Shannon Diversity Index Values (H′) and Evenness (J) of each sampling site.

	SW1	**SW2**	**SW3**
H′	4.37	3.26	4.28
J′	0.81	0.77	0.79

Raman spectroscopy

All samples collected were also taken to analysis by Raman spectroscopy. A total of 108 Raman spectra were collected from the three most abundant species in each sampling site. Each spectra had 19 spectral bands mainly assigned to pigments. Figure 3 shows examples of Raman spectra of the most abundant species in each sampling site and Table 5 shows literature band assignments for the most common bands. The criterion for selection of these bands was their identification in at least 5 spectra of all 12 spectra collected in the each of the three species analysed per sampling site.

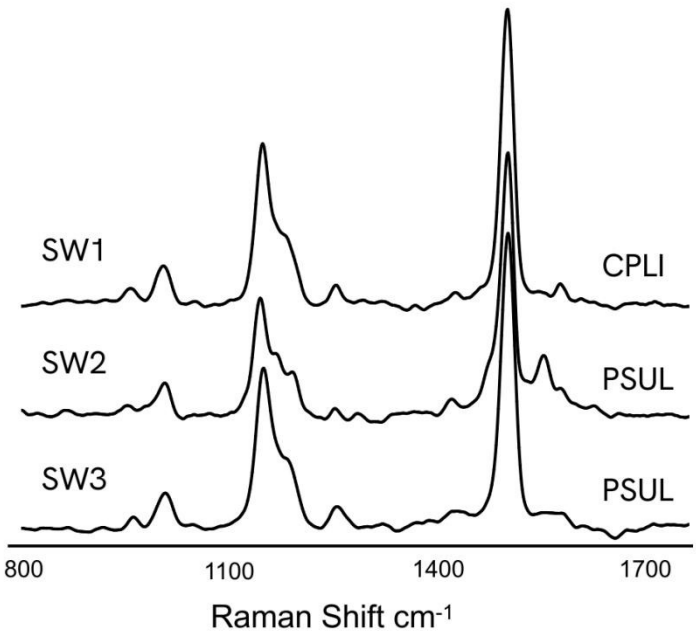

Figure 3. Examples of Raman spectra obtained from the most abundant species in each sampling site; *Cocconeis lineata* (CPLI)*, Paralia sulcata* (PSUL).

Table 5. Molecular assignments identified for the most frequent bands found in the Raman diatom spectra.

Band (cm^{-1})	Molecular Assignments	Reference
1011	CH_3 Stretching modes from carotenoids CC aromatic ring chain from frustule	Alexandre et al. (2014) De Tommasi (2016) De Tommasi et al. (2018) Premvardhan et al. (2009)
1161	C-C stretching modes from carotenoids C=S from frustule	De Tommasi (2016) De Tommasi et al. (2018) Premvardhan et al. (2009) Rüger et al. (2016)
1181	C-H deformational modes from carotenoids C=S from frustule	De Tommasi (2016) De Tommasi et al. (2018) Premvardhan et al. (2009)
1198	N-C stretching modes from Chl_a C=S frustule	De Tommasi (2016) De Tommasi et al. (2018)
1526	C=C stretching modes from carotenoids	Alexandre et al. (2014) Premvardhan et al. (2009) Rüger et al. (2016)

Literature supporting the toxonomic identification and interpretation

Alexandre, M. T., Gundermann, K., Pascal, A. A., van Grondelle, R., Büchel, C., & Robert, B. (2014) Probing the carotenoid content of intact Cyclotella cells by resonance Raman spectroscopy. Photosynthesis Research 119: 273-281.

De Tommasi, E. (2016) Light manipulation by single cells: the case of diatoms. Journal of Spectroscopy, 2016: 2490128.

De Tommasi, E., Congestri, R., Dardano, P., De Luca, A. C., Managò, S., Rea, I., & De Stefano, M. (2018) UV-shielding and wavelength conversion by centric diatom nanopatterned frustules. Scientific Reports 8: 16285.

Larsson, E., Rabayah, A., Jönsson, J.A. (2013) Sludge Removal of Nonsteroidal Anti-Inflammatory Drugs during Wastewater Treatment Studied by Direct Hollow Fiber Liquid Phase Microextraction. Journal of Environmental Protection 2013: 946-955.

Oliva-Teles, L., Pinto, R., Vilarinho, R., Carvalho, A.P., Moreira, J.A., Guimarães, L.* Environmental diagnosis with Raman Spectroscopy applied to diatoms (2022) Biosensors and Bioelectronics 198: 113800

Pielou, E.C. (1966) The Measurement of Diversity in Different Types of Biological Collections. Journal of Theoretical Biology 13: 131-144.

Pinto, R., Vilarinho, R., Carvalho, A.P., Moreira, J.A., Guimarães, L.*, Oliva-Teles, L. Novel Approach to Freshwater Diatom Profiling and Identification Using Raman Spectroscopy and Chemometric Analysis (2022) Water, 14: 2116

Pinto, R., Vilarinho, R., Carvalho, A.P., Moreira, J.A., Guimarães, L.*, Oliva-Teles, L. Raman spectroscopy applied to diatoms (microalgae, Bacillariophyta): Prospective use in the environmental diagnosis of freshwater ecosystems (2021) Water Research 198: 117102

Premvardhan, L., Bordes, L., Beer, A., Büchel, C., & Robert, B. (2009). Carotenoid structures and environments in trimeric and oligomeric fucoxanthin chlorophyll a/c2 proteins from resonance Raman spectroscopy. The Journal of Physical Chemistry B 113: 12565-12574.

Rüger, J., Unger, N., Schie, I. W., Brunner, E., Popp, J., & Krafft, C. (2016). Assessment of growth phases of the diatom Ditylum brightwellii by FT-IR and Raman spectroscopy. Algal research 19: 246-252.

Annex

Interpretation of the Shannon and the Evenness indices

Shannon diversity index

A higher Shannon index (H') value indicates higher diversity, while a lower value indicates lower diversity. An index of 0 means only one species is present; the maximum value of the index depends on the number of species present and their relative abundances. To better interpret the results, you can calculate the Shannon Equitability index to see how close the abundance of species is to a perfectly even distribution.

Interpreting the Shannon index value

Higher index value = Higher diversity: a higher number means a greater number of species and/or a more even distribution of individuals among those species.

Lower index value = Lower diversity: a lower number indicates fewer species or a few species are present that are much more common than others.

Index value of 0: this means there is only one species in the community (no diversity).

Maximum value: the maximum possible value is not a fixed number but depends on the number of species (k) in the community. If all species are equally abundant, the Shannon index will reach its maximum possible value, which is ln (k) (natural logarithm).

Equitability and Evenness index (Pielou 1966)

Equitability is the general ecological principle of how evenly individuals are distributed among different species in a community. The evenness index (as calculated by Pielou, for example) is the specific, calculated measurement of that concept or the metric. The terms are often used interchangeably. The evennesss index (J') is calculated from diversity and richness values.

A community with high equitability has species that are all represented by roughly equal numbers of individuals. A community with low equitability has one or a few dominant species and many species with very few individuals.

This index improves the interpretation of the Shannon Index. It is a normalised value between 0 and 1 that measures how close the species' abundances are to being perfectly even.

Calculation: $J' = H / \ln(k)$

where H is the Shannon index and k is the total number of species.

Interpretation: a J' value closer to 1 indicates that all species are found in roughly equal numbers, while a value closer to 0 means that a few species dominate the community.

Comparing the Shannon index (H) to the Equitability index (E) can provide further context. For example, two communities could have the same Shannon index, but one might have many species with low abundance, while the other has fewer species with very high abundance. The Equitability index helps distinguish between these scenarios.

For interpretation:
Evenness index of 0.90 to 1.00: indicates a very high evenness; the community is extremely balanced with nearly equal abundances of all species.
Evenness index of 0.70 to 0.89: indicates high evenness; the community is well-balanced.
Evenness of 0.50 to 0.69: indicates moderate evenness; the community is somewhat unbalanced, with some species appearing more common than others.
Evenness index of 0.25 to 0.49: indicates low evenness; the community is unbalanced, and a few species are noticeably more common than others.
Evenness index of 0.00 to 0.24: indicates very low evenness; the community is highly dominated by only one or a few species.

Quality classes of the physico-chemical water quality index (WQI) estimated

Quality class	WQI	Description
1	90–100	Excellent
2	70–90	Good
3	50–70	Fair
4	25–50	Poor
5	0–25	Very Poor

Diatoms communities from Muiños River (Vigo, Spain)

Study site

Muíños River is a small, urban-influenced stream that discharges into Playa América. It has local significant recreational relevance. Though, limited information is available about local pollution levels. Two sampling sites were established in this system, identified as SP1 and SP2 (Figure 1). The samples were collected and described in Oliva-Teles et al. (2022) and Pinto et al. (2022). Subsequent processing, taxonomic identification and analysis through Raman spectroscopy followed the methods described therein.

Figure 1. Location of the sampling sites in the Muiños River (scale bar 1 Km).

In situ measurements and nutrient concentrations

Table 1 shows the physico-chemical measurements done in situ, at the time of water sampling, and the nutrient concentrations determined later in the lab. Conductivity and salinity were about the double at SP2 (265 µS cm^{-1} and 0.156‰, respectively) compared to SP1 (153 µS cm^{-1} and 0.084‰). Dissolved oxygen was higher at SP1 (8.86 mg L^{-1}) than at SP2 (6.97 mg L^{-1}), while pH showed an inversed trend. Nitrite (NO_2^-) and ammonium (NH_4^+) concentrations were higher at SP1 (0.27 mg L^{-1} and 0.96 mg L^{-1}, respectively) than at SP2 (0.05 mg L^{-1} and 0.52 mg L^{-1}). Phosphate (PO_4^{3-}) was higher at SP2 (0.58 mg L^{-1}) compared to SP1 (0.27 mg L^{-1}). Silica (SiO_2) concentrations were similar in SP1 and SP2 (1.51 mg L^{-1}).

Table 1. Physico-chemical parameters and nutrient concentrations determined in the samples collected. concentrations of nitrites were below the limit of detection (LD).

	SP1	SP2
Conductivity (µS cm $^{-1}$)	153	265
Salinity (‰)	0.084	0.156
O_2 (mg L^{-1})	8.86	6.97
pH	6.66	6.87
NO_2^-(mg L^{-1})	0.27 (±0.0000)	0.05 (± 0.0020)
NO_3^- (mg L^{-1})	< LD	< LD
NH_4^+ (mg L^{-1})	0.96 (±0.0058)	0.52 (±0.0000)
PO_4^{3-} (mg L^{-1})	0.27 (±0.0010)	0.58 (±0.0010)
SiO_2 (mg L^{-1})	1.51 (±0.0000)	1.51 (±0.0000)

Diatom communities

Analysis of the samples collected led to the taxonomic identification of 39 species of diatoms (Table 2). The most abundant species were *H. contenta* (SP1), *P. oblongella* (SP1 and SP2), *L. mutica* (SP1 and SP2), *P. subatomoides* (SP1), *N. brevissima* (SP2) and *K. clevei* (SP2). *Humidophila contenta* is a cosmopolitan diatom species. *Platessa oblongella* is also a widespread diatom with a great variety of ecological tolerances. *Luticola mutica* is very frequent in freshwater and brackish waters in Europe. *Nitzschia brevissima* usually lives in freshwater with high levels of dissolved oxygen and low organic pollution. *Karayevia clevei* is usually characteristic of alkaline waters and meso-eutrophic environments.

Table 2. Diatom species found in the sampling sites located in River Muiños.

	SP1	SP2
Achnanthidium minutissimum (Kützing) Czarnecki 1994	4	0
Aulacoseira granulata (Ehrenberg) Simonsen 1979	2	2
Cosmioneis citriformis R.L.Lowe & A.R.Sherwood 2010	0	2
Denticula creticola (Østrup) Lange-Bertalot & Krammer 1993	2	0
Diploneis chersonensis (Grunow) Cleve 1892	0	2
Diploneis didyma (Ehrenberg) Ehrenberg 1845	0	2
Discostella stelligera (Cleve & Grunow) Houk & Klee 2004	0	6
Eolimna minima (Grunow) Lange-Bertalot & W.Schiller 1997	2	6
Eunotia minor (Kützing) Grunow 1881	4	2
Frustulia vulgaris (Thwaites) De Toni 1891	0	2
Gomphonema parvulum (Kützing) Kützing 1849	8	16
Halamphora coffeiformis (C.Agardh) Mereschkowsky 1903	0	2
Hantzschia amphioxys (Ehrenberg) Grunow 1880	0	2
Humidophila contenta (Grunow) R.L.Lowe, Kociolek, Johansen, Van de Vjiver; Lange-Bertalot & Kopalová 2014	166	14
Karayevia clevei (Grunow) Round & Bukhtiyarova 1999	4	30

Luticola goeppertiana (Bleisch) D.G.Mann ex Rarick S.Wu S.S.Lee & Edlund 2017	6	8
Luticola mutica (Kützing) D.G.Mann 1990	38	22
Melosira nummuloides C.Agardh 1824	2	2
Melosira varians C.Agardh 1827	0	2
Navicula cari Ehrenberg 1836	0	2
Navicula cincta (Ehrenberg) Ralfs 1861	0	2
Navicula cryptocephala Kützing 1844	0	4
Navicula gregaria Donkin 1861	0	28
Navicula rostellata Kützing 1844	0	4
Navicula tripunctata (O.F.Müller) Bory de Saint-Vincent 1822	0	6
Nitzschia brevissima Grunow 1880	12	40
Nitzschia clausii Hantzsch 1860	2	0
Nitzschia inconspicua Grunow 1862	2	6
Nitzschia palea var. palea (Kützing) W.Smith 1856	6	0
Nitzschia sigmoidea (Nitzsch) W.Smith 1853	0	2
Planothidium delicatulum (Kützing) Round & Bukhtiyarova 1996	2	0
Planothidium frequentissimum (Lange-Bertalot) Lange-Bertalot 1999	8	28
Platessa oblongella (Østrup) C.E.Wetzel, Lange-Bertalot & Ector 2017	92	122
Psamothidium subatomoides (Hustedt) Bukhtiyarova & Round 1996	34	12
Rhoicosphenia abbreviata (C.Agardh) Lange-Bertalot 1980	0	2
Sellaphora meridionalis Potapova & Ponader 2008	0	10
Sellaphora subfasciata M.Potapova 2013	2	0
Tabularia tabulata (C.Agardh) Snoeijs 1992	0	6
Thalassiosira pseudonana Hasle & Heimdal 1970	4	4
Ulnaria ulna (Nitzsch) Compère 2001	0	4

Ecology

Figure 2 shows the functional groups of the diatoms identified in the sampling sites. Motile and low profile diatoms were very abundant in both sampling sites. The guild classification was, however, unknown for a large amount of diatom specimens observed in SP1. Diatoms classified as ßβ-mesosaprobic were very dominant in SP1 and in general mesosaprobic species of all types were very dominant in SP2. Regarding the trophic state, mesotrophic species dominated SP1 and eutrophic species dominated SP2.

Chemical and ecological quality

A physico-chemical indice (WQI) was also calculated using the measured *in situ* parameters and nutrient concentrations. According to the results the water is considered to have a good chemical status (Table 3). Based on the species identified in the sampling sites, the most common diatom indices were also estimated, using the Omnidia software: Diatom Biologic Indice (IBD), Polluosensitivity Indice (IPS) and European Communities Indice (CEE). According to the indices IBD and IPS the water is in good ecological quality. The CEE index showed the most conservative results, classifying both sampling sites as of poor ecological quality.

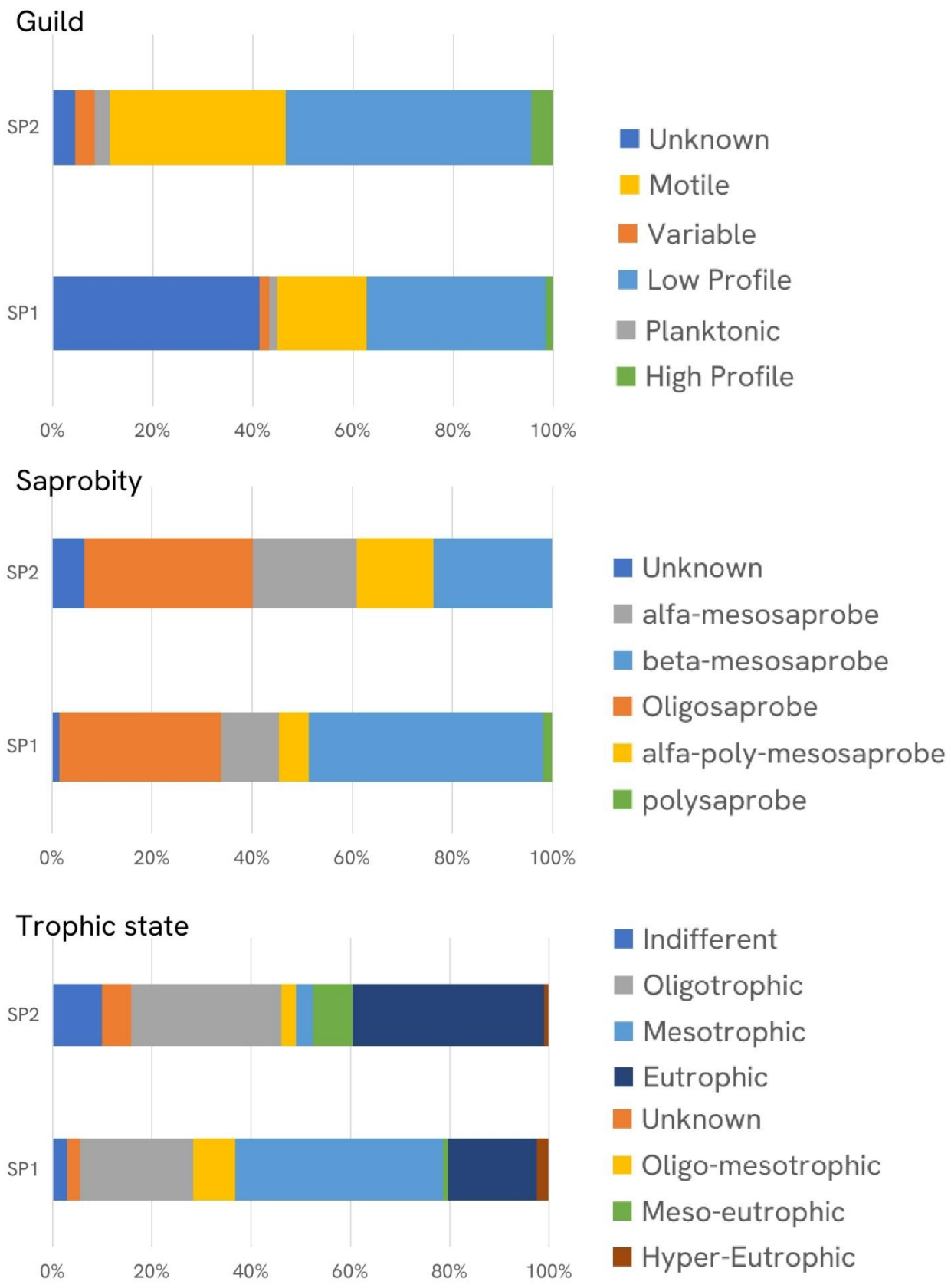

Figure 2. Guild, saprobicity and trophic state, as indicated by the species identified in the sampling sites (SP1 and SP2).

Table 3. Diatom (IBD, IPS and CEE) and physico-chemical (WQI) water quality indices calculated for the sampling sites.

	SP1	SP2
WQI (1-100)	92	88
IBD (1-20)	13.8	12.8
IPS (1-20)	12.6	11.3
CEE (1-20)	8.9	7.6

Table 4 shows the results of the Shannon diversity index (H') and the Evenness index (J) calculated for the sampling sites. The H' index ranges from 0.5 to 5, with values below 2 indicating a low biodiversity ecosystem. Conversely, values above 3 indicate high species richness and abundance. Values of the Evenness index are interpreted by how close they are to 1. Values closer to 1 indicating a more balanced community where species are present in nearly equal numbers; values closer to 0 indicating an unbalanced community where one or a few species dominate (further details on the interpretation of these indices are given in the Annexe).

Table 4. Shannon Diversity Index Values (H') and Evenness (J) of each sampling site.

	SP1	SP2
H'	2.62	3.89
J'	0.62	0.76

Raman spectroscopy

All samples collected were also taken to analysis by Raman spectroscopy. A total of 144 Raman spectra were collected from the three most abundant species in each sampling site. Each spectra had 19 spectral bands mainly assigned to pigments. Figure 3 shows examples of Raman spectra of the most abundant species in each sampling site and Table 5 shows literature band assignments for the most common bands. The criterion for selection of these bands was their identification in at least 5 spectra of all 12 spectra collected in the each of the three species analysed per sampling site.

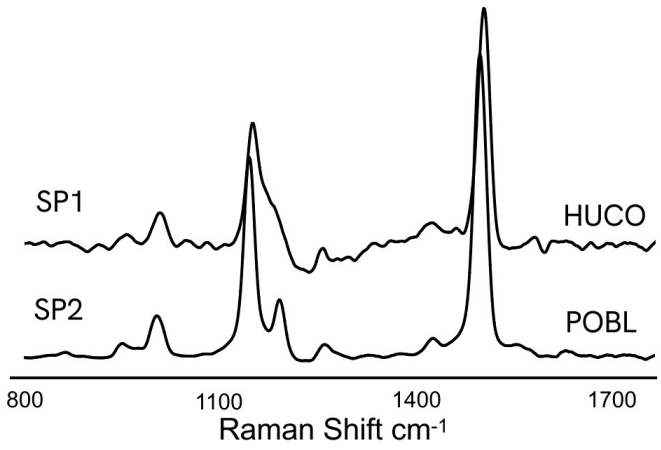

Figure 3. Examples of Raman spectra obtained from the most abundant species in each sampling site; *Humidophila contenta* (HUCO)*, Platessa oblongella* (POBL).

Table 5. Molecular assignments identified for the most frequent bands found in the Raman diatom spectra.

Band (cm^{-1})	**Molecular Assignments**	**Reference**
1011	CH_3 Stretching modes from carotenoids CC aromatic ring chain from frustule	Alexandre et al. (2014) De Tommasi (2016) De Tommasi et al. (2018) Premvardhan et al. (2009)
1161	C-C stretching modes from carotenoids C=S from frustule	De Tommasi (2016) De Tommasi et al. (2018) Premvardhan et al. (2009) Rüger et al. (2016)
1181	C-H deformational modes from carotenoids C=S from frustule	De Tommasi (2016) De Tommasi et al. (2018) Premvardhan et al. (2009)
1198	N-C stretching modes from Chl_a C=S frustule	De Tommasi (2016) De Tommasi et al. (2018)
1526	C=C stretching modes from carotenoids	Alexandre et al. (2014) Premvardhan et al. (2009) Rüger et al. (2016)

Literature supporting the toxonomic identification and interpretation

Alexandre, M. T., Gundermann, K., Pascal, A. A., van Grondelle, R., Büchel, C., & Robert, B. (2014) Probing the carotenoid content of intact Cyclotella cells by resonance Raman spectroscopy. Photosynthesis Research 119: 273-281.

De Tommasi, E. (2016) Light manipulation by single cells: the case of diatoms. Journal of Spectroscopy, 2016: 2490128.

De Tommasi, E., Congestri, R., Dardano, P., De Luca, A. C., Managò, S., Rea, I., & De Stefano, M. (2018) UV-shielding and wavelength conversion by centric diatom nanopatterned frustules. Scientific Reports 8: 16285.

Oliva-Teles, L., Pinto, R., Vilarinho, R., Carvalho, A.P., Moreira, J.A., Guimarães, L.* Environmental diagnosis with Raman Spectroscopy applied to diatoms (2022) Biosensors and Bioelectronics 198: 113800

Pielou, E.C. (1966) The Measurement of Diversity in Different Types of Biological Collections. Journal of Theoretical Biology 13: 131-144.

Pinto, R., Vilarinho, R., Carvalho, A.P., Moreira, J.A., Guimarães, L.*, Oliva-Teles, L. Novel Approach to Freshwater Diatom Profiling and Identification Using Raman Spectroscopy and Chemometric Analysis (2022) Water, 14: 2116

Pinto, R., Vilarinho, R., Carvalho, A.P., Moreira, J.A., Guimarães, L.*, Oliva-Teles, L. Raman spectroscopy applied to diatoms (microalgae, Bacillariophyta): Prospective use in the environmental diagnosis of freshwater ecosystems (2021) Water Research 198: 117102

Premvardhan, L., Bordes, L., Beer, A., Büchel, C., & Robert, B. (2009). Carotenoid structures and environments in trimeric and oligomeric fucoxanthin chlorophyll a/c2 proteins from resonance Raman spectroscopy. The Journal of Physical Chemistry B 113: 12565-12574.

Rüger, J., Unger, N., Schie, I. W., Brunner, E., Popp, J., & Krafft, C. (2016). Assessment of growth phases of the diatom Ditylum brightwellii by FT-IR and Raman spectroscopy. Algal research 19: 246-252.

Annex

Interpretation of the Shannon and the Evenness indices

Shannon diversity index

A higher Shannon index (H') value indicates higher diversity, while a lower value indicates lower diversity. An index of 0 means only one species is present; the maximum value of the index depends on the number of species present and their relative abundances. To better interpret the results, you can calculate the Shannon Equitability index to see how close the abundance of species is to a perfectly even distribution.

Interpreting the Shannon index value

Higher index value = Higher diversity: a higher number means a greater number of species and/or a more even distribution of individuals among those species.

Lower index value = Lower diversity: a lower number indicates fewer species or a few species are present that are much more common than others.

Index value of 0: this means there is only one species in the community (no diversity).

Maximum value: the maximum possible value is not a fixed number but depends on the number of species (k) in the community. If all species are equally abundant, the Shannon index will reach its maximum possible value, which is ln (k) (natural logarithm).

Equitability and Evenness index (Pielou 1966)

Equitability is the general ecological principle of how evenly individuals are distributed among different species in a community. The evenness index (as calculated by Pielou, for example) is the specific, calculated measurement of that concept or the metric. The terms are often used interchangeably. The evennesss index (J') is calculated from diversity and richness values.

A community with high equitability has species that are all represented by roughly equal numbers of individuals. A community with low equitability has one or a few dominant species and many species with very few individuals.

This index improves the interpretation of the Shannon Index. It is a normalised value between 0 and 1 that measures how close the species' abundances are to being perfectly even.

Calculation: $J' = H / \ln(k)$

where H is the Shannon index and k is the total number of species.

Interpretation: a J' value closer to 1 indicates that all species are found in roughly equal numbers, while a value closer to 0 means that a few species dominate the community.

Comparing the Shannon index (H) to the Equitability index (E) can provide further context. For example, two communities could have the same Shannon index, but one might have many species with low abundance, while the other has fewer species with very high abundance. The Equitability index helps distinguish between these scenarios.

For interpretation:
Evenness index of 0.90 to 1.00: indicates a very high evenness; the community is extremely balanced with nearly equal abundances of all species.
Evenness index of 0.70 to 0.89: indicates high evenness; the community is well-balanced.
Evenness of 0.50 to 0.69: indicates moderate evenness; the community is somewhat unbalanced, with some species appearing more common than others.
Evenness index of 0.25 to 0.49: indicates low evenness; the community is unbalanced, and a few species are noticeably more common than others.
Evenness index of 0.00 to 0.24: indicates very low evenness; the community is highly dominated by only one or a few species.

Quality classes of the physico-chemical water quality index (WQI) estimated

Quality class	WQI	Description
1	90–100	Excellent
2	70–90	Good
3	50–70	Fair
4	25–50	Poor
5	0–25	Very Poor

Diatoms communities from Lis River (Leiria, Portugal)

Study site

Lis River is located in central Portugal. The city of Leiria is located in the riverside and the river flows into the Atlantic Ocean (Paíga et al., 2016). This watercourse suffers many anthropogenic impacts related with agriculture, tannery and mineral mining industries, livestock production and urban pollution. Throughout the river there are two Wastewater Treatment Plants (Olhalvas and Coimbrão), discharging their effluents to the water (Paíga et al., 2016). Four sampling sites were established in this system, identified as PT1 to PT4 (Figure 1). One site was located upstream to the WWTP (SW1), one was within a constructed wetland adjacent to the WWTP (SW2) and one site was located downstream to the WWTP (SW3). The samples were collected and described in Oliva-Teles et al. (2022) and Pinto et al. (2022). Subsequent processing, taxonomic identification and analysis through Raman spectroscopy followed the methods described therein. In this context, sampling was carried out in August 2023. The sites were selected according to a contamination gradient associated to Olhalvas WWTP: one site upstream (PT1) and 3 sites downstream (PT2, PT3, PT4) (Figure 1).

Figure 1. Location of the sampling sites in Rio Lis (scale bar 1 Km).

In situ measurements and nutrient concentrations

Table 1 shows the physico-chemical measurements done in situ, at the time of water sampling, and the nutrient concentrations determined later in the lab. Temperature was higher in PT4 (28.70 °C) and lower in PT2 (20.40 °C). Conductivity and salinity were higher in PT3 (3150 µS/cm and 1.72 ‰, respectively) and

lower in PT1 (570 µS/cm and 0.29 ‰, respectively). Water dissolved oxygen varied among sites (from about 10 mg/L in PT3 to about 8 mg/L in PT2), though it was always above the healthy levels. pH values had slight variations across sampling sites. Nitrite, ammonia and phosphate concentrations were higher in samples from PT4 and lower in samples from PT1. Nitrate concentrations were only detected in sites PT2 and PT1, ranging from 2.7 to 3.0 mg/L, respectively. Silicate values were higher in PT4 and lower in PT2.

Table 1. Physico-chemical parameters and nutrient concentrations determined in the samples collected. concentrations of nitrites were below the limit of detection (LD).

	SP1	**SP2**
Conductivity (µS cm $^{-1}$)	153	265
Salinity (‰)	0.084	0.156
O_2 (mg L^{-1})	8.86	6.97
pH	6.66	6.87
NO_2^-(mg L^{-1})	0.27 (±0.0000)	0.05 (± 0.0020)
NO_3^- (mg L^{-1})	< LD	< LD
NH_4^+ (mg L^{-1})	0.96 (±0.0058)	0.52 (±0.0000)
PO_4^{3-} (mg L^{-1})	0.27 (±0.0010)	0.58 (±0.0010)
SiO_2 (mg L^{-1})	1.51 (±0.0000)	1.51 (±0.0000)

Diatom communities

Analysis of the samples collected led to the taxonomic identification of 67 species of diatoms (Table 2). The most abundant species were *A. pediculus* (PT1). *Achnanthidium rivulare* and *N. linearis* were also abundant in PT1. The three most abundant species in PT2 were *A. rivulare*, *C. placentula var. lineata* and *E. minima*. The sampling site PT3 was dominated by *S. pinnata*, *A. pediculus* and *M. varians*. In PT4, the most abundant species were *N. inconspicua*, *N. cryptocephala* and *K. clevei*. *Amphora pediculus* usually associated with low levels of organic pollution. *Achnanthidium rivulare* is adapted to low concentrations of phosphorus and *N. linearis* is characteristic of environments with high luminosity. *Cocconeis placentula var. lineata* is a widespread diatom adapted to alkaline to neutral pH. *Eolimna minima* is a cosmopolitan and mesosaprobic species with high tolerance to metal contamination. *Staurosirella pinnata* and *N. inconspicua* are cosmopolitan species common in environments with moderate to high organic pollution. *Navicula cryptocephala* is a widespread diatom tolerant to a wide range of environmental parameters. *Karayevia clevei* is characteristic of meso to eutrophic water conditions.

Table 2. Diatom species found in the sampling sites located in River Lis.

Species	PT1	PT2	PT3	PT4
Achnanthidium eutrophilum (Lange-Bert.) Lange-Bert. 1999	0	0	0	24
Achnanthidium minutissimum (Kützing) Czarnecki 1994	6	4	4	0
Achnanthidium rivulare Potapova and Ponader 2004	26	232	18	0
Actinocyclus octonarius Ehrenberg 1837	0	0	1	0
Amphora meridionalis Levkov 2009	23	0	0	0
Amphora ovalis (Kützing) Kützing 1844	10	0	0	0
Amphora pediculus (Kützing) Grunow 1875	110	0	48	2
Anorthoneis excentrica (Donkin) Grunow 1870	0	0	0	1
Aulacoseira ambigua (Grunow) Simonsen 1979	0	5	0	0
Cocconeis pediculus Ehrenberg 1838	2	0	7	1
Cocconeis placentula var. euglypta (Ehrenberg) Cleve 1895	22	0	13	0
Cocconeis placentula var. lineata (Ehrenberg) Van Heurck 1885	8	45	38	4
Cymatopleura elliptica (Brébisson) W.Smith 1851	2	0	0	0
Cymbella subturgidula Krammer 2002	5	0	0	0
Diadesmis confervacea Kützing 1844	2	2	3	0
Diploneis krammeri Lange-Bertalot & E.Reichardt 2000	23	0	0	0
Eolimna minima (Grunow) Lange-Bertalot & W.Schiller 1997	0	47	0	4
Fallacia forcipata (Greville) Stickle & Mann 1990	0	1	0	0
Fallacia subhamulata (Grunow in Van Heurck) D.G.Mann in Round et al. 1990	10	0	0	0
Fistulifera saprophila (Lange-Bertalot & Bonik) Lange-Bertalot 1997	0	1	0	0
Frustulia vulgaris (Thwaites) De Toni 1891	6	0	0	0
Gogorevia exilis (Kütz.) Kulikovskiy and Kociolek 2020	0	0	0	10
Gomphonema exilissimum (Grunow) Lange-Bertalot & E.Reichardt 1996	0	2	0	0
Gomphonema parvulum (Kützing) Kützing 1849	0	6	0	0
Gomphonema truncatum Ehrenberg 1832	4	0	0	0
Gyrosigma acuminatum (Kützing) Rabenhorst 1853	7	2	0	0
Gyrosigma attenuatum (Kützing) Rabenhorst 1853	4	0	0	0
Hippodonta capitata (Ehrenberg) Lange-Bertalot Metzeltin & Witkowski 1996	13	2	0	0
Karayevia amoena (Hustedt) Bukhtiyarova 1999	0	2	9	0
Karayevia clevei (Grunow) Round & Bukhtiyarova 1999	0	0	2	39
Karayevia ploenensis var. gessneri (Hustedt) Bukhtiyarova 1999	0	3	0	0
Luticola celebesica Levkov et al.; Levkov et al. 2013	0	0	0	2
Luticola goeppertiana (Bleisch) D.G.Mann ex Rarick S.Wu S.S.Lee & Edlund 2017	0	8	0	0
Melosira varians C.Agardh 1827	8	0	44	0
Navicula capitatoradiata H.Germain ex Gasse 1986	16	2	0	0
Navicula cryptocephala Kützing 1844	0	4	2	57
Navicula cryptotenella Lange-Bertalot 1985	19	25	4	2
Navicula digitoradiata (Gregory) Ralfs 1861	8	0	0	0

Navicula gregaria Donkin 1861	8	0	4	4
Navicula phyllepta Kützing 1844	0	0	0	2
Navicula tripunctata (O.F.Müller) Bory de Saint-Vincent 1822	0	4	2	0
Navicula veneta Kützing 1844	0	0	0	4
Nitzschia amphibia Grunow 1862	0	6	13	0
Nitzschia dissipata (Kützing) Rabenhorst 1860	0	8	0	0
Nitzschia fonticola (Grunow) Grunow 1881	0	0	6	0
Nitzschia inconspicua Grunow 1862	0	7	35	287
Nitzschia intermedia Hantzsch ex Cleve & Grunow 1880	0	4	0	0
Nitzschia linearis W.Smith 1853	28	0	0	0
Nitzschia microcephala Grunow 1880	0	2	8	0
Nitzschia palea var. palea (Kützing) W.Smith 1856	6	0	6	2
Nitzschia perminuta (Grunow) M.Peragallo 1903	0	0	0	8
Nitzschia sigmoidea (Nitzsch) W.Smith 1853	0	2	0	0
Paralia sulcata (Ehrenberg) Cleve 1873	0	0	3	0
Placoneis clementis (Grunow) E.J.Cox 1987	2	0	0	0
Planothidium engelbrechtii (Cholnoky) Round & L.Bukhtiyarova 1996	0	0	0	1
Planothidium frequentissimum (Lange-Bertalot) Lange-Bertalot 1999	0	9	14	0
Platessa oblongella (Østrup) C.E.Wetzel, Lange-Bertalot & Ector 2017	0	0	1	0
Pseudostaurosira brevistriata (Grunow) D.M.Williams & Round 1988	0	0	2	0
Rhoicosphenia abbreviata (C.Agardh) Lange-Bertalot 1980	21	12	0	0
Sellaphora pupula (Kützing) Mereschkovsky 1902	2	0	0	0
Staurosirella pinnata (Ehrenberg) D.M.Williams & Round 1987	0	13	94	0
Stephanocyclus meneghiniana (Kütz.) Kulikovskiy Genkal and Kociolek 2022	0	0	7	7
Tabularia tabulata (C.Agardh) Snoeijs 1992	0	0	4	0
Thalassiosira pseudonana Hasle & Heimdal 1970	0	0	0	5
Tryblionella acuminata W.Smith 1853	0	0	2	2
Tryblionella coarctata (Grunow) D.G.Mann 1990	4	0	0	0
Ulnaria ulna (Nitzsch) Compère 2001	0	0	10	0

Ecology

Figure 2 shows the functional groups of the diatoms identified in the sampling sites. Low profile diatoms dominated sampling sites PT1 and PT2. In PT3 the dominant diatoms were classified as low profile and high profile and in PT4 motile diatoms were the most dominant. Polysaprobes dominated PT1 and PT3 sampling sites. The sampling site PT2 had more diatoms classified as oligosaprobes and PT4 had more diatoms classified as α-mesosaprobes. Eutrophic diatoms dominated PT1, PT3 and PT4 sampling sites. In sampling site PT2 most of the diatoms had a unknown classification in terms of trophy.

Chemical and ecological quality

A physico-chemical indice (WQI) was also calculated using the measured *in situ* parameters and nutrient concentrations. According to the results the water is considered to have a good chemical status (Table 3). The ecological status of all sampling sites was classified as moderate according to IPS and the physico-chemical status was classified as good according to WQI. According to IBD, PT1 and PT3 were classified as good, PT2 was classified as good and PT4 was classified as moderate. According to CEE, PT1 was classified as good, PT2 and PT3 as moderate and PT4 as poor.

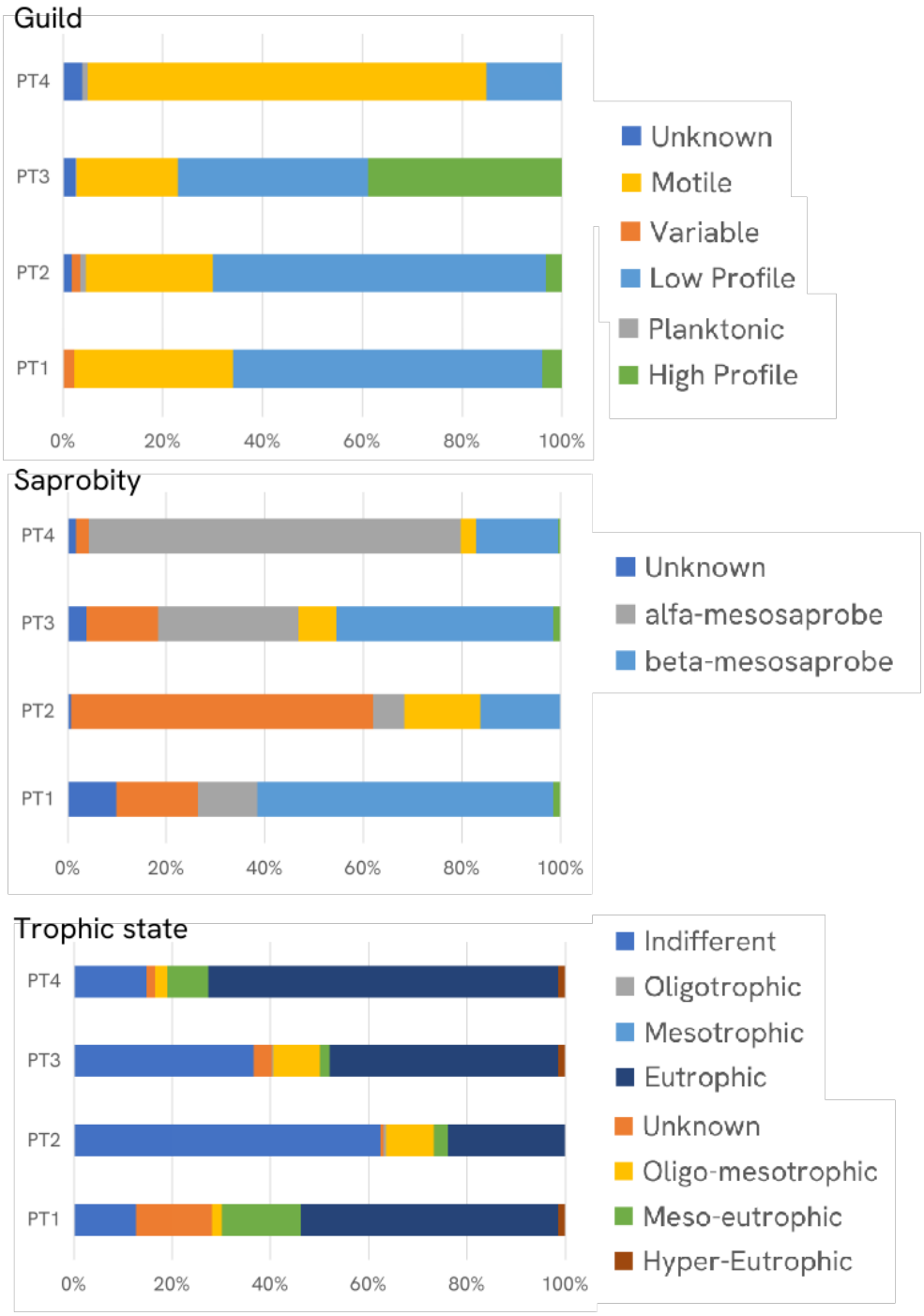

Figure 2. Guild, saprobicity and trophic state, as indicated by the species identified in the sampling sites.

Table 3. Diatom (IBD, IPS and CEE) and physico-chemical (WQI) water quality indices calculated for the sampling sites.

	PT1	PT2	PT3	PT4
WQI (1-100)	77	76	79	71
IBD (1-20)	15.0	17.7	14.2	11.2
IPS (1-20)	12.1	13.1	11.9	10.9
CEE (1-20)	14.5	9.9	12.3	5.6

Table 4 shows the results of the Shannon diversity index (H') and the Evenness index (J) calculated for the sampling sites. The H' index ranges from 0.5 to 5, with values below 2 indicating a low biodiversity ecosystem. Conversely, values above 3 indicate high species richness and abundance. Values of the Evenness index are interpreted by how close they are to 1. Values closer to 1 indicating a more balanced community where species are present in nearly equal numbers; values closer to 0 indicating an unbalanced community where one or a few species dominate (further details on the interpretation of these indices are given in the Annexe).

Table 4. Shannon Diversity Index Values (H') and Evenness (J) of each sampling site.

	PT1	PT2	PT3	PT4
H'	4.05	2.91	3.87	2.19
J'	0.83	0.61	0.80	0.50

Raman spectroscopy

All samples collected were also taken to analysis by Raman spectroscopy. A total of 144 Raman spectra were collected from the three most abundant species in each sampling site. Each spectra had 19 spectral bands mainly assigned to pigments. Figure 3 shows examples of Raman spectra of the most abundant species in each sampling site and Table 5 shows literature band assignments for the most common bands. The criterion for selection of these bands was their identification in at least 5 spectra of all 12 spectra collected in the each of the three species analysed per sampling site.

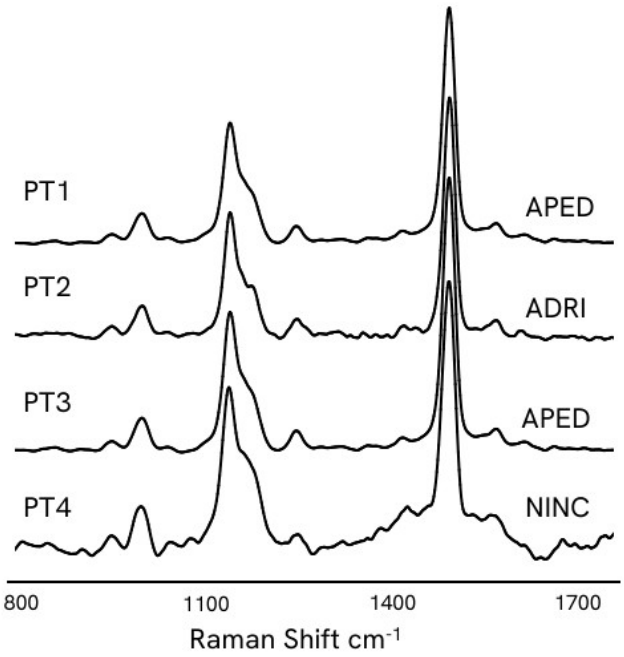

Figure 3. Examples of Raman spectra obtained from the most abundant species in each sampling site; *Amphora pediculus* (ADED), *Achanthidium rivulare* (ADRI), *Nitzschia inconspicua* (NINC).

Table 5. Molecular assignments identified for the most frequent bands found in the Raman diatom spectra.

Band (cm^{-1})	**Molecular Assignments**	**Reference**
1011	CH_3 Stretching modes from carotenoids CC aromatic ring chain from frustule	Alexandre et al. (2014) De Tommasi (2016) De Tommasi et al. (2018) Premvardhan et al. (2009)
1161	C-C stretching modes from carotenoids C=S from frustule	De Tommasi (2016) De Tommasi et al. (2018) Premvardhan et al. (2009) Rüger et al. (2016)
1181	C-H deformational modes from carotenoids C=S from frustule	De Tommasi (2016) De Tommasi et al. (2018) Premvardhan et al. (2009)
1198	N-C stretching modes from Chl_a C=S frustule	De Tommasi (2016) De Tommasi et al. (2018)
1526	C=C stretching modes from carotenoids	Alexandre et al. (2014) Premvardhan et al. (2009) Rüger et al. (2016)

Literature supporting the toxonomic identification and interpretation

Alexandre, M. T., Gundermann, K., Pascal, A. A., van Grondelle, R., Büchel, C., & Robert, B. (2014) Probing the carotenoid content of intact Cyclotella cells by resonance Raman spectroscopy. Photosynthesis Research 119: 273-281.

De Tommasi, E. (2016) Light manipulation by single cells: the case of diatoms. Journal of Spectroscopy, 2016: 2490128.

De Tommasi, E., Congestri, R., Dardano, P., De Luca, A. C., Managò, S., Rea, I., & De Stefano, M. (2018) UV-shielding and wavelength conversion by centric diatom nanopatterned frustules. Scientific Reports 8: 16285.

Oliva-Teles, L., Pinto, R., Vilarinho, R., Carvalho, A.P., Moreira, J.A., Guimarães, L.* Environmental diagnosis with Raman Spectroscopy applied to diatoms (2022) Biosensors and Bioelectronics 198: 113800

Pielou, E.C. (1966) The Measurement of Diversity in Different Types of Biological Collections. Journal of Theoretical Biology 13: 131-144.

Pinto, R., Vilarinho, R., Carvalho, A.P., Moreira, J.A., Guimarães, L.*, Oliva-Teles, L. Novel Approach to Freshwater Diatom Profiling and Identification Using Raman Spectroscopy and Chemometric Analysis (2022) Water, 14: 2116

Pinto, R., Vilarinho, R., Carvalho, A.P., Moreira, J.A., Guimarães, L.*, Oliva-Teles, L. Raman spectroscopy applied to diatoms (microalgae, Bacillariophyta): Prospective use in the environmental diagnosis of freshwater ecosystems (2021) Water Research 198: 117102

Premvardhan, L., Bordes, L., Beer, A., Büchel, C., & Robert, B. (2009). Carotenoid structures and environments in trimeric and oligomeric fucoxanthin chlorophyll a/c2 proteins from resonance Raman spectroscopy. The Journal of Physical Chemistry B 113: 12565-12574.

Rüger, J., Unger, N., Schie, I. W., Brunner, E., Popp, J., & Krafft, C. (2016). Assessment of growth phases of the diatom Ditylum brightwellii by FT-IR and Raman spectroscopy. Algal research 19: 246-252.

Annex

Interpretation of the Shannon and the Evenness indices

Shannon diversity index

A higher Shannon index (H') value indicates higher diversity, while a lower value indicates lower diversity. An index of 0 means only one species is present; the maximum value of the index depends on the number of species present and their relative abundances. To better interpret the results, you can calculate the Shannon Equitability index to see how close the abundance of species is to a perfectly even distribution.

Interpreting the Shannon index value

Higher index value = Higher diversity: a higher number means a greater number of species and/or a more even distribution of individuals among those species.

Lower index value = Lower diversity: a lower number indicates fewer species or a few species are present that are much more common than others.

Index value of 0: this means there is only one species in the community (no diversity).

Maximum value: the maximum possible value is not a fixed number but depends on the number of species (k) in the community. If all species are equally abundant, the Shannon index will reach its maximum possible value, which is ln (k) (natural logarithm).

Equitability and Evenness index (Pielou 1966)

Equitability is the general ecological principle of how evenly individuals are distributed among different species in a community. The evenness index (as calculated by Pielou, for example) is the specific, calculated measurement of that concept or the metric. The terms are often used interchangeably. The evennesss index (J') is calculated from diversity and richness values.

A community with high equitability has species that are all represented by roughly equal numbers of individuals. A community with low equitability has one or a few dominant species and many species with very few individuals.

This index improves the interpretation of the Shannon Index. It is a normalised value between 0 and 1 that measures how close the species' abundances are to being perfectly even.

Calculation: $J' = H / \ln(k)$

where H is the Shannon index and k is the total number of species.

Interpretation: a J' value closer to 1 indicates that all species are found in roughly equal numbers, while a value closer to 0 means that a few species dominate the community.

Comparing the Shannon index (H) to the Equitability index (E) can provide further context. For example, two communities could have the same Shannon index, but one might have many species with low abundance, while the other has fewer species with very high abundance. The Equitability index helps distinguish between these scenarios.

For interpretation:
Evenness index of 0.90 to 1.00: indicates a very high evenness; the community is extremely balanced with nearly equal abundances of all species.
Evenness index of 0.70 to 0.89: indicates high evenness; the community is well-balanced.
Evenness of 0.50 to 0.69: indicates moderate evenness; the community is somewhat unbalanced, with some species appearing more common than others.
Evenness index of 0.25 to 0.49: indicates low evenness; the community is unbalanced, and a few species are noticeably more common than others.
Evenness index of 0.00 to 0.24: indicates very low evenness; the community is highly dominated by only one or a few species.

Quality classes of the physico-chemical water quality index (WQI) estimated

Quality class	WQI	Description
1	90–100	Excellent
2	70–90	Good
3	50–70	Fair
4	25–50	Poor
5	0–25	Very Poor

Diatoms communities from Curitiba (Brazil)

Study site

The lake system studied is located within the campus of Universidade Positivo (Curitiba, Brazil). It is an artificial water body embedded in an extensive green-space environment. It functions as a recreational area and plays key roles in stormwater retention, hydric management of the campus (irrigation, sanitary use). It has a heat-exchange water source, which supports the university's sports facilities. Ecologically, the lake sustains a considerable biodiverse community: several phytoplankton families (Chlorophyta, Cyanobacteria, Euglenophyta e Ochrophyta), fishes, birds, reptiles and mammals including capybaras. Additionally, the surrounding campus vegetation comprises a diversity of arboreal and shrubs species (including fruit-bearing trees). Four sampling sites were established in this system, identified as BR1 to BR4 (Figure 1). Site BR1 is in a little impacted area. The samples were collected and described in Oliva-Teles et al. (2022) and Pinto et al. (2022). Subsequent processing, taxonomic identification and analysis through Raman spectroscopy followed the methods described therein.

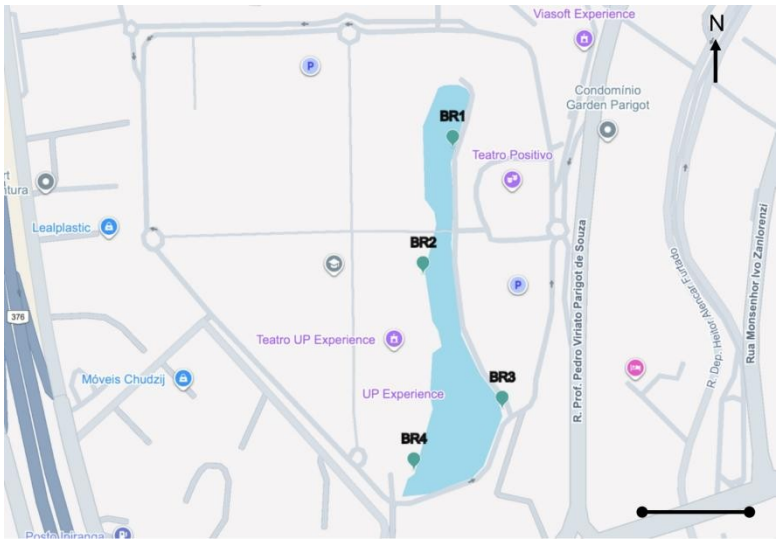

Figure 1. Location of the sampling sites in the lake system at Universidade Positivo (scale bar 200m).

In situ measurements and nutrient concentrations

Table 1 shows the physico-chemical measurements done in situ, at the time of water sampling, and the nutrient concentrations determined later in the lab. Conductivity values ranged from 58.0 µS cm^{-1} at BR4 to 585 µS cm^{-1} at BR1. Salinity was highest at BR1 (0.320 ‰) and decreased progressively across the sites, reaching 0.030 ‰ at BR4. Dissolved oxygen (O_2) levels varied between 4.36 mg L^{-1} at BR1 and 8.33 mg L^{-1} at BR2. The pH values were relatively stable, ranging from 7.05 at BR3 to 7.33 at BR2.

Regarding nutrients, nitrite (NO_2^-) concentrations were lowest at BR2 (0.02 ± 0.0040 mg L^{-1}) and highest at BR3 (0.20 ± 0.0040 mg L^{-1}). Nitrate (NO_3^-) was only detected at BR3 (1.13 ± 0.0600 mg L^{-1}). Ammonium (NH_4^+) concentrations increased from BR1 (0.76 ± 0.0000 mg L^{-1}) to BR4 (2.26 ± 0.0100 mg L^{-1}). Phosphate

(PO_4^{3-}) levels were highest at BR4 (7.40 ± 0.2000 mg L^{-1}) and lowest at BR1 (1.60 ± 0.0000 mg L^{-1}). Finally, silica (SiO_2) concentrations ranged from 0.18 ± 0.0000 mg L^{-1} at BR1 to 0.83 ± 0.0000 mg L^{-1} at BR4.

Table 1. Physico-chemical parameters and nutrient concentrations determined in the samples collected; concentrations of nitrites were below the limit of detection (LD) in some sampling points.

	BR1	BR2	BR3	BR4
Conductivity (µS cm $^{-1}$)	585	105	70.0	58.0
Salinity (‰)	0.320	0.060	0.040	0.030
O_2 (mg L^{-1})	4.36	8.33	7.48	8.26
pH	7.27	7.33	7.05	7.26
NO_2^-(mg L^{-1})	0.09 (±0.0010)	0.02 (± 0.0040)	0.20 (± 0.0040)	0.03 (± 0.0010)
NO_3^- (mg L^{-1})	< LD	< LD	1.13 (±0.0600)	< LD
NH_4^+ (mg L^{-1})	0.76 (±0.0000)	0.94 (±0.0100)	1.25 (±0.0000)	2.26 (±0.0100)
PO_4^{3-} (mg L^{-1})	1.60 (±0.0000)	2.23 (±0.0600)	3.97 (±0.0600)	7.40 (±0.2000)
SiO_2 (mg L^{-1})	0.18 (±0.0000)	0.32 (±0.0000)	0.42 (±0.0000)	0.83 (±0.0000)

Diatom communities

Analysis of the samples collected led to the taxonomic identification of 26 species of diatoms (Table 2). The most abundant species were *A. eutrophilum* in BR2, *A. catenatum* in BR4, BR3 and BR2, and *A. granulata* in BR1 and BR4. Several species found in BR4 (*D. stelligera*, *A. catenatum* and *A. eutrophilum*) are typical of eutrophic environments; their presence probably reflecting the higher nutrient levels found in this site. Furthermore, *D. stelligera* is a very common species in oligotrophic to mesotrophic lakes all over the world, which is largely influenced by nutrient and light availability. In contrast, *A. granulata*, also abundant in BR4, is a widespread taxa, tolerant to all trophic conditions, though with a tendency for higher abundance in eutrophic waters. *Planothidium frequentissimum*, detected mostly in BR3 and BR4, is considered an indicator of high organic content in the water.

Table 2. Diatom species found in the sampling sites of the Lake system at Universidade Positivo.

Species	BR1	BR2	BR3	BR4
Achnanthidium catenatum (Bílý & Marvan) Lange-Bertalot 1999	78	118	116	166
Achnanthidium eutrophilum (Lange-Bert.) Lange-Bert. 1999	89	168	44	28
Amphora ovalis (Kützing) Kützing 1844	0	0	0	1
Aulacoseira granulata (Ehrenberg) Simonsen 1979	117	68	65	100
Aulacoseira granulata var. angustissima (O.Müller) Simonsen 1979	24	14	24	0
Cocconeis placentula var. lineata (Ehrenberg) Van Heurck 1885	2	0	0	2
Diadesmis contenta (Grunow) D.G.Mann 1990	0	2	2	0
Discostella stelligera (Cleve & Grunow) Houk & Klee 2004	43	37	45	52
Eunotia incisa W.Smith ex W.Gregory 1854	0	0	0	2
Fragilaria crotonensis Kitton 1869	2	4	4	0

Gomphonema parvulum (Kützing) Kützing 1849	0	2	0	0
Gomphonema pumilum (Grunow) E.Reichardt & Lange-Bertalot 1991	0	0	0	6
Mayamaea permitis (Hustedt) K.Bruder & Medlin 2008	0	0	4	0
Navicula cryptocephala Kützing 1844	4	2	26	2
Navicula cryptotenella Lange-Bertalot 1985	2	2	2	4
Navicula oligotraphenta Lange-Bertalot & G.Hofmann 1993	0	0	2	8
Navicula rostellata Kützing 1844	0	0	4	0
Navicula viridula (Kützing) Ehrenberg 1838	0	2	0	0
Nitzschia acicularis (Kützing) W.Smith 1853	4	0	10	0
Nitzschia acidoclinata Lange-Bertalot 1976	0	4	6	2
Nitzschia capitellata Hustedt 1930	2	2	4	2
Nitzschia palea var. palea (Kützing) W.Smith 1856	2	0	2	2
Nitzschia sublinearis Hustedt 1930	14	6	0	6
Planothidium frequentissimum (Lange-Bertalot) Lange-Bertalot 1999	26	0	55	30
Thalassiosira pseudonana Hasle & Heimdal 1970	0	0	4	0
Ulnaria ulna (Nitzsch) Compère 2001	0	0	0	2

Ecology

Figure 2 shows the functional groups of the diatoms identified in the sampling sites. The diatoms found were predominantly motile and low profile. In some sites, the low profile diatoms outnumbered motile diatoms, such as in BR2, BR3 and BR4. In terms of saprobicity, β-mesosaprobes were dominant in all sampling sites, despite the high number of diatoms with no known classification in BR3 and BR4. Eutrophic diatoms were very abundant in all sampling sites. Diatoms with unknown trophic classification were more abundant in BR3 and BR4.

Chemical and ecological quality

A physico-chemical indice (WQI) was also calculated using the measured *in situ* parameters and nutrient concentrations. According to the results the water is considered to have a good chemical status (Table 3). Based on the species identified in the sampling sites, the most common diatom indices were also estimated, using the Omnidia software: Diatom Biologic Indice (IBD), Polluosensitivity Indice (IPS) and European Communities Indice (CEE). According to the IBD and IPS all sampling sites were classified as in good ecological status. The Indice CEE classified BR1, BR2 and BR3 as poor (indicated in yellow in Table 3).

Table 3. Diatom (IBD, IPS and CEE) and physico-chemical (WQI) water quality indices calculated for the sampling sites.

	BR1	**BR2**	**BR3**	**BR4**
WQI (1-100)	73	88	80	79
IBD (1-20)	13.7	14.9	13.8	15.2
IPS (1-20)	13.4	14.0	13.7	15.2
CEE (1-20)	7.9	6.8	8.5	13.1

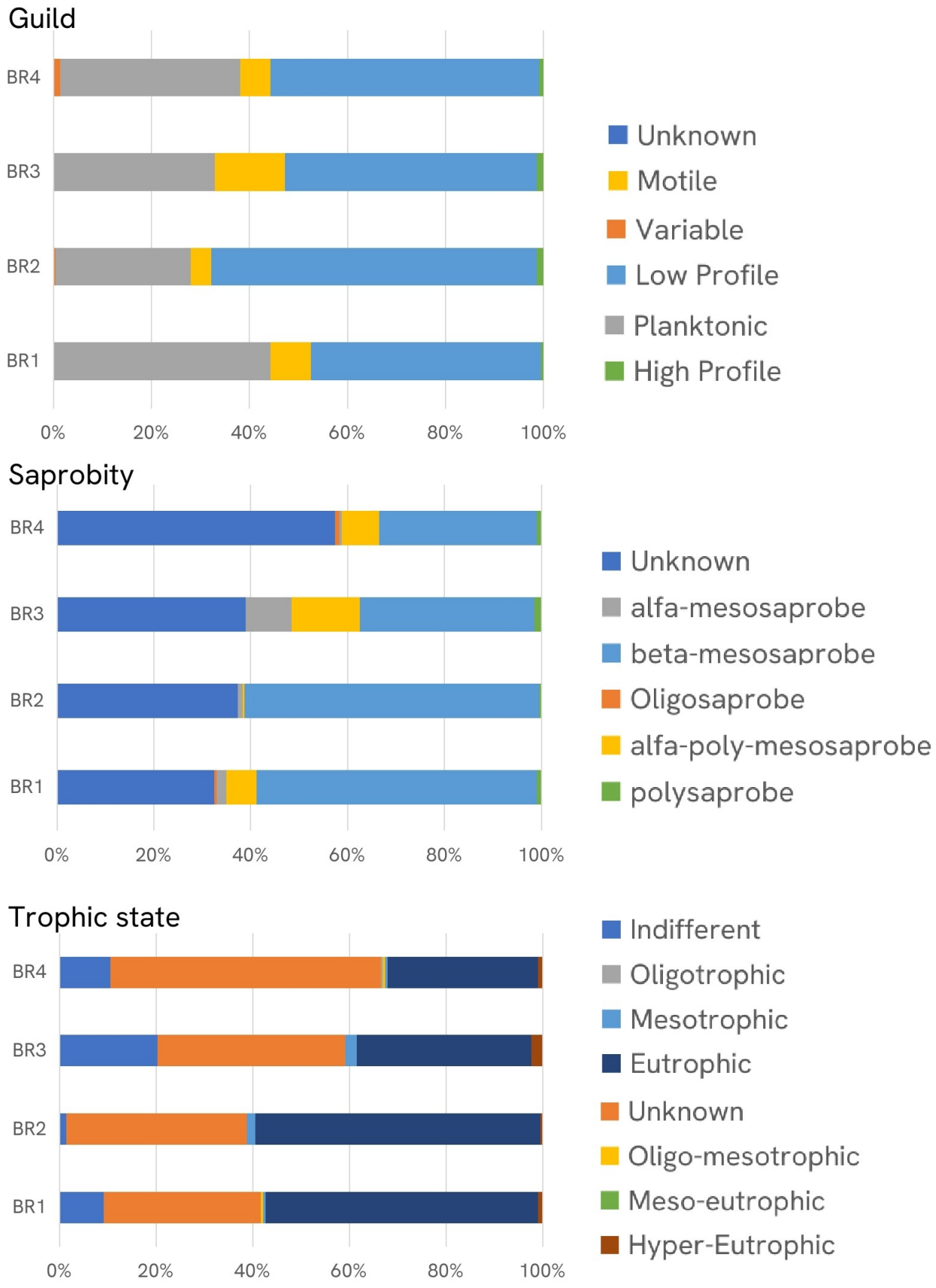

Figure 2. Guild, saprobicity and trophic state, as indicated by the species identified in the sampling sites (BR1 to BR4).

Table 4 shows the results of the Shannon diversity index (H') and the Evenness index (J) calculated for the sampling sites. The H' index ranges from 0.5 to 5, with values below 2 indicating a low biodiversity ecosystem. Conversely, values above 3 indicate high species richness and abundance. Values of the Evenness index are interpreted by how close they are to 1. Values closer to 1 indicating a more balanced community where species are present in nearly equal numbers; values closer to 0 indicating an unbalanced community where one or a few species dominate (further details on the interpretation of these indices are given in the Annexe).

Table 4. Shannon Diversity Index Values (H') and Evenness (J) of each sampling site.

	BR1	**BR2**	**BR3**	**BR4**
H'	2.84	2.35	3.17	2.57
J	0.73	0.62	0.76	0.63

Raman spectroscopy

All samples collected were also taken to analysis by Raman spectroscopy. A total of 144 Raman spectra were collected from the three most abundant species in each sampling site. Each spectra had 19 spectral bands mainly assigned to pigments. Figure 3 shows examples of Raman spectra of the most abundant species in each sampling site and Table 5 shows literature band assignments for the most common bands. The criterion for selection of these bands was their identification in at least 5 spectra of all 12 spectra collected in the each of the three species analysed per sampling site.

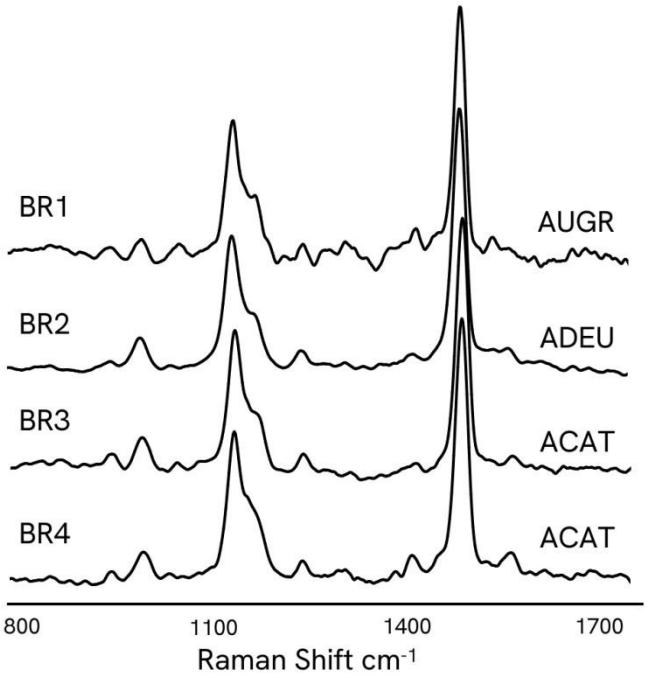

Figure 3. Examples of Raman spectra obtained from the most abundant species in each sampling site; *Aulacoseira granulata* (AUGR), *Achnanthidium eutrophilum* (ADEU), *Achnanthidium catenatum* (ACAT).

Table 5. Molecular assignments identified for the most frequent bands found in the Raman diatom spectra.

Band (cm^{-1})	Molecular Assignments	Reference
1011	CH_3 Stretching modes from carotenoids CC aromatic ring chain from frustule	Alexandre et al. (2014) De Tommasi (2016) De Tommasi et al. (2018) Premvardhan et al. (2009)
1161	C-C stretching modes from carotenoids C=S from frustule	De Tommasi (2016) De Tommasi et al. (2018) Premvardhan et al. (2009) Rüger et al. (2016)
1181	C-H deformational modes from carotenoids C=S from frustule	De Tommasi (2016) De Tommasi et al. (2018) Premvardhan et al. (2009)
1198	N-C stretching modes from Chl_a C=S frustule	De Tommasi (2016) De Tommasi et al. (2018)
1526	C=C stretching modes from carotenoids	Alexandre et al. (2014) Premvardhan et al. (2009) Rüger et al. (2016)

Literature supporting the toxonomic identification and interpretation

Alexandre, M. T., Gundermann, K., Pascal, A. A., van Grondelle, R., Büchel, C., & Robert, B. (2014) Probing the carotenoid content of intact Cyclotella cells by resonance Raman spectroscopy. Photosynthesis Research 119: 273-281.

De Tommasi, E. (2016) Light manipulation by single cells: the case of diatoms. Journal of Spectroscopy, 2016: 2490128.

De Tommasi, E., Congestri, R., Dardano, P., De Luca, A. C., Managò, S., Rea, I., & De Stefano, M. (2018) UV-shielding and wavelength conversion by centric diatom nanopatterned frustules. Scientific Reports 8: 16285.

Oliva-Teles, L., Pinto, R., Vilarinho, R., Carvalho, A.P., Moreira, J.A., Guimarães, L.* Environmental diagnosis with Raman Spectroscopy applied to diatoms (2022) Biosensors and Bioelectronics 198: 113800

Pielou, E.C. (1966) The Measurement of Diversity in Different Types of Biological Collections. Journal of Theoretical Biology 13: 131-144.

Pinto, R., Vilarinho, R., Carvalho, A.P., Moreira, J.A., Guimarães, L.*, Oliva-Teles, L. Novel Approach to Freshwater Diatom Profiling and Identification Using Raman Spectroscopy and Chemometric Analysis (2022) Water, 14: 2116

Pinto, R., Vilarinho, R., Carvalho, A.P., Moreira, J.A., Guimarães, L.*, Oliva-Teles, L. Raman spectroscopy applied to diatoms (microalgae, Bacillariophyta): Prospective use in the environmental diagnosis of freshwater ecosystems (2021) Water Research 198: 117102

Premvardhan, L., Bordes, L., Beer, A., Büchel, C., & Robert, B. (2009). Carotenoid structures and environments in trimeric and oligomeric fucoxanthin chlorophyll a/c2 proteins from resonance Raman spectroscopy. The Journal of Physical Chemistry B 113: 12565-12574.

Rüger, J., Unger, N., Schie, I. W., Brunner, E., Popp, J., & Krafft, C. (2016). Assessment of growth phases of the diatom Ditylum brightwellii by FT-IR and Raman spectroscopy. Algal research 19: 246-252.

Annex

Interpretation of the Shannon and the Evenness indices

Shannon diversity index

A higher Shannon index (H') value indicates higher diversity, while a lower value indicates lower diversity. An index of 0 means only one species is present; the maximum value of the index depends on the number of species present and their relative abundances. To better interpret the results, you can calculate the Shannon Equitability index to see how close the abundance of species is to a perfectly even distribution.

Interpreting the Shannon index value

Higher index value = Higher diversity: a higher number means a greater number of species and/or a more even distribution of individuals among those species.

Lower index value = Lower diversity: a lower number indicates fewer species or a few species are present that are much more common than others.

Index value of 0: this means there is only one species in the community (no diversity).

Maximum value: the maximum possible value is not a fixed number but depends on the number of species (k) in the community. If all species are equally abundant, the Shannon index will reach its maximum possible value, which is ln (k) (natural logarithm).

Equitability and Evenness index (Pielou 1966)

Equitability is the general ecological principle of how evenly individuals are distributed among different species in a community. The evenness index (as calculated by Pielou, for example) is the specific, calculated measurement of that concept or the metric. The terms are often used interchangeably. The evennesss index (J') is calculated from diversity and richness values.

A community with high equitability has species that are all represented by roughly equal numbers of individuals. A community with low equitability has one or a few dominant species and many species with very few individuals.

This index improves the interpretation of the Shannon Index. It is a normalised value between 0 and 1 that measures how close the species' abundances are to being perfectly even.

Calculation: $J' = H / \ln(k)$

where H is the Shannon index and k is the total number of species.

Interpretation: a J' value closer to 1 indicates that all species are found in roughly equal numbers, while a value closer to 0 means that a few species dominate the community.

Comparing the Shannon index (H) to the Equitability index (E) can provide further context. For example, two communities could have the same Shannon index, but one might have many species with low abundance, while the other has fewer species with very high abundance. The Equitability index helps distinguish between these scenarios.

For interpretation:
Evenness index of 0.90 to 1.00: indicates a very high evenness; the community is extremely balanced with nearly equal abundances of all species.
Evenness index of 0.70 to 0.89: indicates high evenness; the community is well-balanced.
Evenness of 0.50 to 0.69: indicates moderate evenness; the community is somewhat unbalanced, with some species appearing more common than others.
Evenness index of 0.25 to 0.49: indicates low evenness; the community is unbalanced, and a few species are noticeably more common than others.
Evenness index of 0.00 to 0.24: indicates very low evenness; the community is highly dominated by only one or a few species.

Quality classes of the physico-chemical water quality index (WQI) estimated

Quality class	WQI	Description
1	90–100	Excellent
2	70–90	Good
3	50–70	Fair
4	25–50	Poor
5	0–25	Very Poor