Science of the Total Environment 954 (2024) 176587

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Chemical profiling of organic contaminants in rural surface waters combining target and non-target LC-HRMS/MS analysis

Alfonso Fernández-García [a,b], Ana B. Martínez-Piernas [a,b,*], David Moreno-González [a,b], Bienvenida Gilbert-López [a,b], Juan F. García-Reyes [a,b]

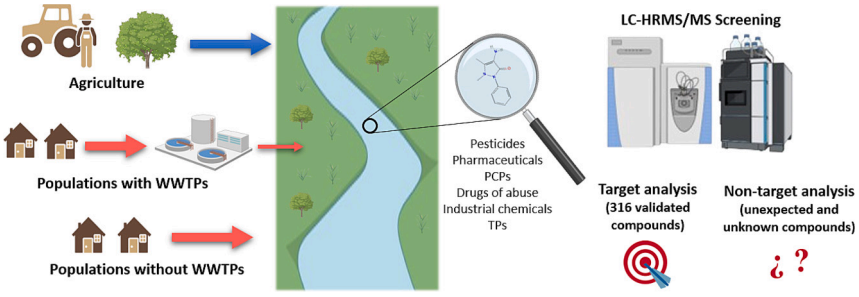
[a] Analytical Chemistry Research Group (FQM 323), Department of Physical and Analytical Chemistry, University of Jaén, Campus Las Lagunillas edif. B3, 23071 Jaén, Spain

[b] University Research Institute for Olives Grove and Olive Oil (INUO), University of Jaén, Jaén, Spain

HIGHLIGHTS

- Comprehensive monitoring of CECs by target and non-target analysis (LC-HRMS)
- Pharmaceuticals were the most detected compounds in surface waters.
- NTA enabled tentative identification of 79 CECs, 12 of them confirmed.
- Even in rural areas most of the pollutants detected are from urban origin.
- The lack of WWTPs in small rural villages showed higher impact on water pollution.

GRAPHICAL ABSTRACT



ARTICLE INFO

Editor: Damià Barceló

Keywords:
Surface water
Emerging contaminants
Target analysis
Non-target screening
High-resolution mass spectrometry

ABSTRACT

The pollution of natural waters by contaminants of emerging concern (CECs) is one of the pressing problems due to their global distribution and potential negative effects on the environment and human health. In rural areas with lower population density and limited industrial development, less contamination is expected. However, the lack of wastewater treatment plants (WWTPs) or their poor removal efficiency can lead to significant input of pollutants. In this context, 11 streams of rural areas in the Guadalquivir River basin, southeast of Spain, were studied over two years to obtain an overview of the origin and distribution of contaminants. A target method using solid-phase extraction and liquid chromatography coupled to high resolution mass spectrometry (LC-HRMS) was developed for the analysis of 316 compounds in surface waters. A total of 78 target analytes were detected, comprising pesticides, pharmaceuticals, personal care products (PCPs), transformation products (TPs), and industrial chemicals. The flame retardant tributyl phosphate (16–3572 ng L^{-1}) was detected in all samples, followed by caffeine (30–8090 ng L^{-1}) and the analgesic tramadol (3–1493 ng L^{-1}). The target approach was combined with a non-target analysis (NTA) strategy to obtain an overall perspective of the chemical profile of unexpected or unknown compounds in the samples. Up to 79 contaminants were tentatively identified, and 12 of them were finally confirmed with standards. Most of the contaminants determined by NTA were pharmaceuticals and their TPs. The results indicated that most of CECs have an urban origin despite traditional agriculture is the main economic activity in this region. Moreover, the absence of WWTPs in small towns is significant, as

* Corresponding author at: Analytical Chemistry Research Group (FQM 323), Department of Physical and Analytical Chemistry, University of Jaén, Campus Las Lagunillas edif. B3, 23071 Jaén, Spain.
E-mail address: apiernas@ujaen.es (A.B. Martínez-Piernas).

https://doi.org/10.1016/j.scitotenv.2024.176587
Received 16 July 2024; Received in revised form 7 September 2024; Accepted 26 September 2024
Available online 27 September 2024

contamination levels at these sites were comparable to or higher than those in larger populations with sewage treatments.

1. Introduction

Nowadays, thousands of organic chemicals are produced and used worldwide. Interest in monitoring many of these compounds, also called contaminants of emerging concern (CECs) or emerging contaminants (ECs), has considerably increased due to their potential risks to aquatic ecosystems and human health (Bletsou et al., 2015; Hollender et al., 2019). CECs, which include pesticides, pharmaceuticals, personal care products (PCPs), and industrial chemicals, among others, are released into surface waters through agricultural runoff and municipal or industrial wastewater discharges (Mofijur et al., 2023). Once in the environment, CECs can be degraded through varied processes into transformation products (TPs), which may occasionally be more mobile, toxic, and persistent than parent compounds (Escher and Fenner, 2011; Zahn et al., 2024). In addition, there is still a significant knowledge gap regarding TPs compared to their parent compounds in terms of structure, physicochemical properties, mobility, and behavior in the environment. Current legislation primarily focuses on parent chemicals, while TPs are rarely addressed. Also, it should be considered that aquatic ecosystems are not isolated compartments and they have an impact on the global environment, including biota and the food chain to humans (Escher et al., 2020). It is therefore necessary to acquire knowledge about the origin, fate, and environmental consequences of the CECs to establish more effective management policies. So far, the deterioration of natural resources caused by the continuous discharge of CECs into water bodies and the changing conditions of aquatic ecosystems due to climate change has prompted the European Union (EU) to establish regulations devoted specifically to the monitoring of CECs in surface waters (EC, 2013, 2022). However, these regulations constrain the number of CECs officially monitored to a few substances included in the denominated Watch List; thus, scientific research in this field is required to provide additional knowledge.

Liquid chromatography coupled to high-resolution mass spectrometry (LC-HRMS) is one of the most efficient tools for monitoring organic micropollutants, making it one of the crucial instrumental techniques for assessing surface water quality (Menger et al., 2020). Nevertheless, conventional and routine analyses are often focused on target strategies, whose scope is restricted to preselected analytes, causing potential pollutants to go unnoticed. The combination of LC-HRMS and non-target analysis (NTA) provides a comprehensive overview of the number of organic micropollutants present in a sample through the identification of unexpected or unknown compounds in the environment, even retrospectively (Tian et al., 2019; Kiefer et al., 2021; Bonnefille et al., 2023). Notwithstanding the evident advantages of NTA, processing the huge amount of HRMS data acquired, the need for feature prioritization steps to enable efficient data reduction of relevant chemical information, and the often-time-consuming manual structural elucidation and compound identification are substantial challenges to the application of NTA strategies. To overcome this, and in addition to vendor software, different open-source data processing algorithms, such as MS-DIAL (Tsugawa et al., 2015), MZmine (Katajamaa et al., 2006), and XCMS (Smith et al., 2006) have been developed in recent decades. Similarly, prioritization strategies, and tentative identification using open-source spectral libraries (MoNA, 2024) have been improved to identify unknown contaminants in environmental samples (Vosough et al., 2024). Besides, recent scientific advancements have proposed semi-quantification approaches for NTA results based on prediction models. These analytical strategies aim to obtain enhanced chemical information, improve the transition from qualitative to quantitative approaches, and overcome the limitations caused by the lack of analytical standards (Vosough et al., 2024).

In the present study, a comprehensive strategy combining cutting-edge analytical tools has been applied to investigate the occurrence of CECs in a set of 11 small streams in rural areas of the Guadalquivir River basin. Although small streams often constitute most of the river network and basin surface, they are commonly underrepresented in surveillance campaigns. These small rivers could be exposed to greater contamination due to a direct connection with the agricultural environment and their lower potential for dilution (Casado et al., 2019; Toth et al., 2024). In addition, some small populations lack wastewater treatment plants (WWTPs) or they exhibit poor removal efficiency. The high expenses linked to construction and maintenance hinder small towns from adopting treatment technologies, and contamination levels of small streams could be comparable to those of large rivers (Fernández-García et al., 2024). Guadalquivir River headwaters are located in the province of Jaén. Therefore, most of its tributaries in this region are small streams subjected to diffuse contamination from agriculture, mainly olive groves, and from urban discharges where small rural populations predominate. Even though this is one of the most important basins in Spain, no monitoring studies specifically focused on CECs have been reported in the last ten years (Masiá et al., 2013; López-Serna et al., 2013; Robles-Molina et al., 2014). In this context, two combined methodologies based on HPLC-HRMS/MS analysis, a wide-scope target method with >300 compounds, and an NTA strategy, were employed for the determination of the contaminants. The integration of both approaches facilitated the identification of potential contamination sources from urban or agricultural origin, offering a comprehensive perspective on the chemical status of the surface waters of rural areas.

2. Material and methods

2.1. Chemicals and reagents

Reference standards of 316 target compounds (Table S1) including pesticides, pharmaceuticals, PCPs, drugs of abuse and some of their more relevant TPs, and 12 isotopically labeled analytical standards (acetamiprid-d_3, atenolol-d_7, bentazone-d_6, caffeine-$^{13}C_3$, chlorfenvinphos-d_{10}, citalopram-d_6, diclofenac-d_4, difenoconazole-d_6, diuron-d_6, imazalil-d_5, isoproturon-d_6, and myclobutanil-d_4) were acquired from Merck Group (Darmstadt, Germany) and Dr. Ehrenstorfer (Augsburg, Germany). Methanol (MeOH), water, and formic acid (FA) (LC-MS grade) were supplied by Merck Group (Darmstadt, Germany). Solid phase extraction (SPE) Oasis HLB cartridges (200 mg, 6 mL) were from Waters (Milford, MA, USA).

2.2. Sampling and pretreatment

The Guadalquivir River is the primary river in southern Spain. It covers an area around 57,000 km^2 with a total length of 657 km. The basin crosses over 12 provinces with a population of 4.4 million inhabitants. Approximately 49 % of the river basin is used for agriculture, while about 1.5 % is for urban use (Confederación Hidrográfica del Guadalquivir, 2024). A total of 36 surface water samples were collected in the upper Guadalquivir River basin located in the province of Jaén (south of Spain) over two years (Fernández-García et al., 2024). Eleven sampling points (P) were monitored in four sampling campaigns in October 2021/2022 (autumn) and April 2022/2023 (spring). Sample collection coincided with the application of phytosanitary products used in olive groves, the primary cultivar of the province. Composite samples were collected in amber glass bottles (1 L) and they were stored at −20 °C until analysis. Most of the points are small streams and tributaries (P1 - P10) of the main Guadalquivir River (P11), as described in

our previous work (Fernández-García et al., 2024). Sampling sites were characterized by high agricultural pressure (mainly olive groves) and low urban influence (headwaters and downstream small villages with <6000 inhabitants, with and without WWTPs). Additional information on sampling points is available in Table S2.

To isolate the CECs from the matrix, the protocol employed was based on our previous studies (Robles-Molina et al., 2014; Fernández-García et al., 2024). Briefly, 200 mL of water sample, previously filtered using cellulose filters (0.45 μm pore size), were loaded in a SPE cartridge (Oasis HLB, 200 mg, 6 mL) preconditioned with 4 mL of MeOH and 8 mL of H_2O (LC-MS grade). After sample loading, the cartridge was dried under vacuum (30 min approx.). Retained analytes were eluted with 2 aliquots of 4 mL of MeOH. Then, the extract was evaporated in a water bath at 37 °C, under nitrogen stream 15 psi until dryness (30 min approx.). The final residue was re-dissolved with 2 mL of MeOH. Finally, an aliquot of 200 μL was diluted with 800 μL of H_2O and spiked with a mix of 12 isotopically labeled internal standards to reach a final concentration of 25 ng L^{-1}. The final sample preconcentration factor was 20.

2.3. LC-HRMS/MS analysis

Instrumental analysis was carried out using the same conditions as previously reported by our group (Fernández-García et al., 2024). Chromatographic separation was conducted in a Thermo Vanquish Flex LC system (Thermo Fisher Scientific, Waltham, MA, USA) using a Zorbax Rapid Resolution High Definition (RRHD) Eclipse Plus C_{18} column (2.1 mm × 50 mm, 1.8 μm) from Agilent (Agilent Technologies Santa Clara, MA, USA). The mobile phase consisted of H_2O (A) and MeOH (B) both with 0.1 % FA. The elution gradient started at 10 % of B; linear ramping to 50 % B from 0.5 to 4 min; linear increased to 95 % B until 17 min; this condition was kept until 25 min; lowered again to 10 % B at 25.1 min, and kept constant until 30 min. The flow rate was 0.3 mL min^{-1}, the injection volume was set at 10 μL, and the column oven temperature at 30 °C. Ionization was carried out with a heated electrospray ionization source (HESI) operating in both positive and negative modes (individual analysis) with a spray voltage of 3.5/−2.5 kV and a capillary temperature of 300 °C. Detection was performed in a Q-Exactive Orbitrap (Thermo Fisher Scientific). Spectral acquisition was performed by MS1 full-scan (FS) (*m/z* 65 to 950) at a resolving power of 70,000 FWHM followed by five data-dependent acquisition (DDA) MS2 spectra with a resolution of 17,500 FWHM for each scan. Details of HESI conditions and measuring parameters are given in Tables S3 and S4.

2.4. Target method validation

The target method validation was carried out using surface water samples. Dynamic linear range, method quantification limits (MQLs), trueness (recovery studies), matrix effect, and precision (expressed in terms of intra-day and inter-day precision as relative standard deviation, RSD) were evaluated. The full information about method validation is provided in the Supplementary material (Text S1 and Table S5). Procedural blanks (reagent H_2O extracted following the sample SPE protocol), and instrumental blanks (H_2O:MeOH (80:20, *v/v*)) were analyzed to check for contamination or carry-over potential effects. Matrix-matched calibration curves prepared in sample extracts were used for the quantitative analysis.

2.5. Processing of target compounds

Xcalibur 3.0 and Trace Finder 3.3 (Thermo Fisher Scientific) software were used for qualitative and quantitative data analysis of the 316 target analytes. The compounds were quantified using matrix-matched calibration curves. The confirmation was based on the following criteria: Rt (±0.1 min), a mass accuracy error of the protonated $[M + H]^+$ or deprotonated molecule $[M\text{-}H]^-$ (< 5 ppm), an isotopic pattern fit of M + 1 and M + 2 (< 5 ppm), and the presence of one or more product ions (< 5 ppm error). The analytes that fulfilled all these criteria were classified at level 1 of confidence (Schymanski et al., 2014). For target compounds without MS2 spectra due to low peak intensity, but that met the requirements for precursor ion and isotopic pattern mass errors, Rt match, and presence of an MS2 spectrum in another sample analyzed in the same batch, they were also considered at level 1.

2.6. Non-target analysis

2.6.1. Preprocessing parameters

HRMS raw data files were firstly converted to .abf format using Reifycs Analysis Base File Converter Reifycs Abf Converter (n.d.). MS-DIAL software (v. 5.1.23) (Tsugawa et al., 2015) was used for data preprocessing with an automated workflow. The optimization of each preprocessing parameter was done individually based on the recall of the target analytes and internal standards. The settings applied for peak picking, spectrum deconvolution, peak alignment, and identification were as follows: mass tolerance for MS1 was set at 0.002 Da and for MS2 at 0.05 Da; minimum peak height 50,000; mass slide width 0.05 Da; minimum peak width 8 scans; smoothing level 4, and sigma window value of 0.5. A retention time (Rt) tolerance of 0.2 min and an MS1 tolerance of 0.002 Da were selected in the alignment. Procedural blanks were used for blank subtraction. The settings for compound annotation using the MassBank of North America (MoNA, 2024) spectral library database included MS1 and MS2 tolerances of 0.002 and 0.05 Da, respectively, with an 80 % of cut off. All detailed parameters applied in MS-DIAL are presented in Table S6.

2.6.2. Compound identification and semi-quantification

For compound annotation, different approaches were implemented. The first strategy involved the use of the open-source MoNA spectral library. Hits with poor peak shape, peak width, and/or low-quality MS2 spectra were discarded after manual revision. Then, compounds with an unequivocal MS2 spectrum match with the library (> 80 % spectral similarity) were assigned to level 2a of confidence, and those with insufficient information for an exact structure proposal were classified at level 3 of confidence, according to Schymanski et al. (2014).

For the remaining unannotated features, a prioritization strategy was implemented to reduce the complexity of the dataset. For this purpose, features with a peak intensity above 1×10^6, present in several samples (at least in the 30 % of the samples), and/or exhibit a recognizable time-trend (a specific feature present at the same sampling point across all sampling campaigns) were considered. For formula assignment, the elements C, H, O, N, P, S, Cl, Br, and F, mass defect, the presence of indicative Br/Cl/S isotopologues, the nitrogen rule, and rings and double bond equivalents (DBE) rule were considered using Xcalibur 3.0 software. Tentative structures were searched in PubChem online database (PubChem, 2024), and a manual MS2 elucidation was performed with the aid of the in-silico webtool Metfrag (Ruttkies et al., 2016; Bolton et al., 2024) to compare the predicted fragments with the acquired ones.

In order to confirm and semi-quantify some of the identified compounds, reference standards were purchased and a semi-quantitative approach using matrix-matched calibration curves was employed. The concentrations of analytes included in identification levels 2 and 3 were estimated using the calibration curve of the closest eluting compound of the target method assuming similar ionization efficiencies (Malm et al., 2021).

3. Results and discussion

3.1. Target analysis

A total of 78 compounds (25 % of the target scope) were detected at least once in the river samples. As shown in Fig. 1, pharmaceuticals were

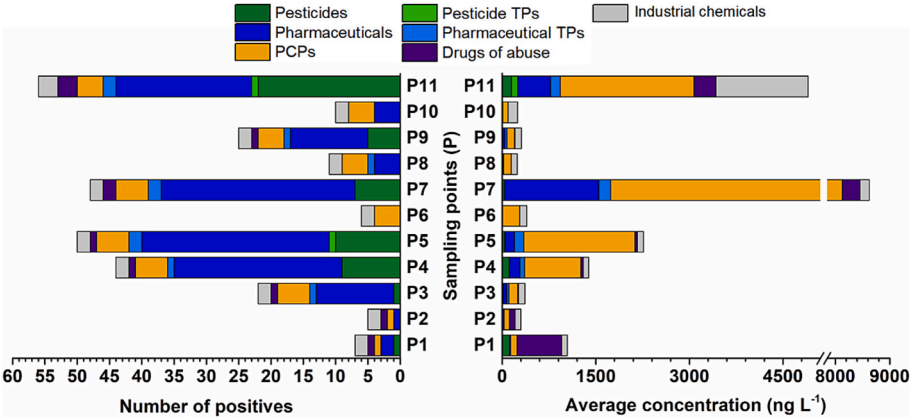


Fig. 1. Number of analytes determined through target analysis (left) and average concentrations in each sampling point during 4 sampling events per compound class (right).

the predominant group with 40 compounds from different therapeutic classes (including 4 TPs) across most of the sampling points. This group was followed by pesticides (25 and 1 TP), drugs of abuse (5), PCPs (5), and industrial chemicals (2, flame retardants (FR) and plasticizers) (Table S7). The profile of CECs detected in this work is consistent with data reported by López-Serna et al. (2013) and Robles-Molina et al. (2014) in the same river basin. Although collected from different sampling points, pharmaceuticals such as antipyrine, atenolol, codeine, metoprolol, naproxen, paracetamol, and sotalol were detected at similar concentration levels. In the same line, pesticides atrazine, DEET, diuron, simazine, and terbuthylazine, and the PCPs caffeine and nicotine were also detected by Robles-Molina et al. (2014) in rivers of the province of Jaén. It should be highlighted that the number of CECs has increased in comparison to the study by Robles-Molina et al. (2014). In the latter, compounds such as carbamazepine, carbendazim, chlorotoluron, diclofenac, flufenamic acid, gemfibrozil, isoproturon, ketoprofen, metformin, or triethyl phosphate were not detected, although they were included in the scope of analysis. Furthermore, the limits of quantification of this previous work were similar or lower than in the present study due to a greater preconcentration step during SPE protocol. In contrast, all of these analytes were detected in the current study, some with frequencies of detection above 30 %. This fact suggests that the occurrence of CECs could have diversified in the last ten years.

Tributyl phosphate was the analyte most frequently detected (100 % of samples) with a maximum concentration of 3572 ng L^{-1}. This FR is widely used in many applications, such as anti-foaming agents, electronic equipment, furniture, and plasticizers, among others. It is commonly detected in aquatic environments (Zhu et al., 2022). Exposure to this compound could cause negative effects on aquatic organisms due to its bioaccumulation in tissues, affecting photosynthesis mechanisms (Song et al., 2016) and posing a potential risk for respiratory diseases like asthma (Coelho et al., 2024). Other compounds with a high frequency of detection were caffeine (97 % of samples, maximum concentration of 8090 ng L^{-1}), and its metabolite theophylline (66 % of samples, maximum concentration of 2316 ng L^{-1}), which agrees with other studies (Robles-Molina et al., 2014).

In addition, 6 Priority Substances namely atrazine, diuron, isoproturon, simazine, and terbutryn with median concentrations ranging from 7.3 to 52.8 ng L^{-1} were detected (Table S7). All of them had concentration levels below their respective Environmental Quality Standards (EQS) established in the European regulation (EC, 2013). Furthermore, 9 compounds included in the Watch List (EC, 2022) were determined: azoxystrobin, metformin, *o*-desmethylvenlafaxine, penconazole, sulfamethoxazole, tebuconazole, tetraconazole, trimethoprim, and venlafaxine with median concentrations between 8.3 and 310 ng L^{-1} (Table S7). The presence of these compounds included in European legislation is broadly reported in other Spanish river basins (Fonseca et al., 2019; Čelić et al., 2021; Royano et al., 2023).

In general, pharmaceuticals were the predominant group of compounds found at most of the sampling points (Fig. 1), likely due to their widespread human consumption. Several of the pharmaceuticals detected, including carbamazepine, citalopram, telmisartan, tramadol, or venlafaxine, can bioaccumulate in aquatic invertebrates even at relatively low concentrations (Grabicová et al., 2024). In the present study, 20 % of the pharmaceuticals determined were antibiotics. The main ecotoxicological effects associated with antibiotics include growth inhibition of primary producers and the development of antibiotic resistance (Khan et al., 2020). It is remarkable that although the pharmaceutical metformin was poorly recovered (12.7 %) in SPE due to its strong influence by pH conditions in HLB cartridges (Oertel et al., 2018), it was detected in 61 % of the samples with estimated concentrations ranging between 10.3 and 2031 ng L^{-1}. Gabapentin showed similar results in terms of low recovery rates (14.3 %) and estimated concentrations in the 2.1–95.2 ng L^{-1} range. Recovery values have not been used in the calculation of the concentrations in real samples; consequently, the real concentrations of these polar analytes might be higher. These findings emphasize the need to develop specific methods for the enrichment of very polar compounds to obtain accurate concentrations in aquatic environments.

As expected, and in line with our previous study (Fernández-García et al., 2024), pesticides were one of the main classes of compounds observed. Agriculture is the principal economic activity in the region, primarily focused on olive oil production. In general, a diverse range of pesticides was determined, although they were not found at high concentrations. Only acetamiprid, azoxystrobin, carbendazim, fenuron, fluometuron, tetraconazole, and thiabendazole showed median concentrations above 100 ng L^{-1} (Table S7). However, despite their low concentration levels, mixtures of contaminants could pose a potential ecotoxicological risk for different aquatic ecosystems, as observed with the calculation of the Risk Quotient index (RQ) for pesticides in our previous work conducted in the same river basin (Fernández-García et al., 2024). Regarding PCPs and stimulants, a small number of positive detections were associated with the highest concentration levels (Fig. 1), primarily influenced by the presence of caffeine and theophylline at high concentrations.

The overall CECs detected are consistent with findings from other studies conducted both within and outside of Europe in recent years. The human markers caffeine and nicotine, and some pharmaceuticals, such as atenolol, carbamazepine, diclofenac, or ibuprofen were among the most abundant compounds detected in the north of Italy by Riva et al. (2019) at comparable concentration values. Very similar results were reported by Malnes et al. (2022) in surface water of Sweden, where pharmaceuticals were the predominant group followed by industrial chemicals, PCPs, and pesticides. Numerous pharmaceuticals and PCPs as

well as various TPs were detected in surface and groundwaters of the Danube River Basin by Ng et al. (2023). Pharmaceuticals were also the most detected compounds in Canadian lakes, closely followed by the pesticide atrazine and the metabolite cotinine (Lahens et al., 2024). Likewise, industrial chemicals, among them, plasticizers, flame retardants, and vehicle-related compounds were reported in the Yellow River of China, although the presence of these compounds reached 43 % of the total compounds probably due to a more urbanized area (Cao et al., 2023).

3.2. Non-target analysis

An NTA strategy based on automatic data preprocessing using MS-DIAL open-source software combined with a manual revision of the results was developed to identify compounds out of the target method. Following automatic preprocessing steps, a total of 11,452 and 8905 ion features with assigned MS2 spectra were obtained for positive and negative ionization modes, respectively. The number of features was further reduced to 5338 and 6460 after blank subtraction in each case. The automatic search using the public spectral library (MoNA, 2024) yielded the annotation of 296 preliminary structures in positive ionization mode and 26 structures in negative mode. This list was subsequently refined through manual revision to remove hits with poor peak shape, MS2 spectra with low quality, and compounds detected in the target screening. The final list of candidates using MoNA library included 66 tentative identifications.

For the tentative identification of the remaining unannotated features, a prioritization step based on intensity threshold and occurrence trend across sampling points/events was carried out to reduce data complexity. This feature list consisted of a total of 114 features, where entries with poor chromatographic shape and noisy chromatograms were ruled out. By the examination of their MS1 information and the application of different chemical rules (see Section 2.6.2), up to 16 features could be semi-automatically annotated with a molecular formula. After that, molecular formulas were searched in PubChem website and a manual MS2 spectral elucidation with the aid of Metfrag software was carried out. Up to 13 analytes were tentatively identified using this approach. The tentative identification of the surfactant *N*,*N*-bis(2-hydroxyethyl)octanamide exemplifies the overall process. As depicted in Fig. S1, a feature with *m/z* 232.1908 and Rt 6.35 min was detected in Cambil River (P4), Jandulilla River (P5), and Campillo River (P7) in three samplings. The molecular formula $C_{12}H_{25}NO_3$ was assigned to this ion considering the mass error of $[M + H]^+$ and $[M + Na]^+$ adduct ions, M + 1 and M + 2 isotopologues (< 5 ppm), and DBE rule = 1. The experimental MS2 spectral fragmentation provided by PubChem database using APCI ion source and the in-silico fragmentation obtained using PubChemLite in Metfrag led to its tentative identification in confidence level 3 due to the presence of several fragments with mass errors below 5 ppm.

As shown in Table 1, the final list of candidates included 79 tentative identifications employing both annotation approaches. Among these, 12 compounds were finally confirmed with the reference standards (level 1), 47 were assigned at level 2a of confidence, 11 were classified at level 3, and 9 were included in level 4. Only 3 compounds, diethanolamine, phenacetin, and propham, resulted in false positive identifications due to mismatches with the retention time of the corresponding standard. Notably, these compounds were tentatively identified using the spectral database, meeting all the criteria for their classification at identification level 2a. Furthermore, their MS2 spectra revealed several product ions with mass errors below 5 ppm and relative abundances matching the reference spectra in the library in each case. However, they resulted in false positive identifications despite manual data revision, a fact that underscores the significant complexity involved in the identification process.

Overall, the compounds were classified into: natural products (15), pharmaceuticals (13), TPs (14), industrial chemicals (11), PCPs (14), other organic contaminants (3), and unknown (9). Figs. 2 and S2 illustrate the potential of combining both, targeted and non-targeted approaches, demonstrating how much information can remain hidden if only target methods are employed. Specifically, industrial chemicals, PCPs, and TPs were mainly detected by NTA (Fig. 2), and in the cases of P7 and P11, around 40 % of compounds were discerned using NTA (Fig. S2).

To evaluate the impact of the analytes determined by NTA, a semi-quantification strategy was carried out to estimate the concentration of the compounds included in identification levels 1, 2, and 3. Fig. 3 depicts the number of compounds and the estimated concentrations in each sampling point. Although a direct comparison between the concentrations of target analytes and non-target compounds cannot be made because the latter were estimated and potentially subjected to errors, (Malm et al., 2021) the concentration levels at some sampling sites would be equal to or even greater than those calculated for target CECs. This fact was specifically relevant for pharmaceuticals and their TPs, illustrating the analytical potential of combining both strategies.

Some of the pharmaceuticals identified by NTA have previously been reported in other Spanish river basins at similar concentration levels, such as losartan, metoprolol, naproxen, and propyphenazone (Valcárcel et al., 2011; Fonseca et al., 2020). However, the NTA results led to the identification of pharmaceuticals rarely included in monitoring campaigns or mentioned in the literature. This is the case for sitagliptin, an antidiabetic drug with a median estimated concentration 90 ng L^{-1}. It is commonly used when monotherapy with an oral antidiabetic drug, such as metformin, is insufficient for treating type-2 diabetes (Martin et al., 2012). Another example is the immunosuppressant mycophenolic acid (median concentration 436 ng L^{-1}), which is used to prevent organ transplant rejections (Hu et al., 2023). This cytostatic compound has been related to high risks for aquatic biota in surface and wastewater (Gouveia et al., 2019). In addition, most of the TPs identified using NTA are pharmaceutical by-products. Although TPs are often less active and toxic than their parent compounds, in some cases the active substructure can remain intact after transformation, making the TP potentially more toxic in similar concentrations, such as sulfamethoxazole and its TP N4-acetyl-sulfamethoxazole (Lentz et al., 2024). Furthermore, some TPs showed higher estimated median concentrations than their parent compounds: atrazine desisopropyl, 10,11-dihydro-10-hydroxycarbamazepine, carbamazepine-10,11-epoxide, lidocaine N-oxide, *O*-desmethylvenlafaxine, and theobromine. However, precursor compounds were generally detected with greater frequency.

PCPs of different types were detected in river samples. UV filters such as phenylbenzimidazole sulfonic acid or benzophenone-4 are widely used to protect the skin from ultraviolet radiation and are commonly detected in surface, groundwater, and wastewater (Palmiotto et al., 2018). Other compounds used in foaming agents, fragrances, or synthetic musks, such us 4-methyl-5-thiazoleethanol, coumarin, linalool, lauryl diethanolamide, and tonalide were also detected, with median estimated concentrations ranging from 22 to 252 ng L^{-1}. Tonalide, detected in 30 % of samples, is a synthetic musk that can pose a risk to aquatic environments due to its potential toxicity and persistence. It has been detected in other Spanish rivers (Llamas-Dios et al., 2021).

Industrial chemicals such as vehicle-related contaminants 1,3-diphenylguanidine, hexa-methoxymethyl-melamine, and 2-benzothiazolesulfonic acid were detected with median concentrations ranging from 168 to 1170 ng L^{-1}. These compounds used as vulcanization accelerators, rubber components, or crosslinking agents were reported in marine environments (Tian et al., 2019) and surface waters (Finckh et al., 2024). They are associated with toxicological effects in aquatic organisms (Chibwe et al., 2022). Diphenyl isophthalate, N-butylbenzenesulfonamide, and N-ethyl-o-toluenesulfonamide were detected as plasticizers. Despite their frequency of detection was below 14 % in samples, N-ethyl-o-toluenesulfonamide showed one of the highest median concentrations (2872 ng L^{-1}). This industrial chemical is widely used in plastics, dyes, cosmetics, and pesticide formulations (Kandie

Table 1
List of compounds (tentatively) identified by NTA. Retention time (Rt, min), confidence level (CL, according to Schymanski et al., 2014), detection frequency (DF, %), median concentration (ng L^{-1}), concentration range (ng L^{-1}), and standard used for the semi-quantification of analytes included in confidence levels 2 and 3.

Compound	Formula	Rt (min)	*m/z*	Adduct	CL	DF (%) (n = 36)	Median conc. (ng L^{-1})	Range (ng L^{-1})	Class	Quantitative standard (Rt)
1-(3-Carboxypropyl)-3,7-dimethylxanthine	$C_{11}H_{14}N_4O_4$	3.8	267.1088	[M+H]+	2a	6	22	10–33	Pharmaceutical TP	Antipyrine (3.8)
1,3-Diphenylguanidine	$C_{13}H_{13}N_3$	3.5	212.1182	[M+H]+	2a	19	168	18–640	Vulcaniser/catalyst	Atrazine desisopropyl (3.5)
Unknown[a]	$C_{20}H_{26}O_3$	12.3	315.1952	[M+H]+	4	8	–	–	–	–
10,11-Dihydro-10-Hydroxycarbamazepine	$C_{15}H_{14}N_2O_2$	4.9	255.1128	[M+H]+	1	8	298	207–356	Pharmaceutical TP	–
1-Aminoadamantane	$C_{10}H_{17}N$	3.5	152.1434	[M+H]+	2a	8	70	30–109	Antiviral	Fuberidazole (3.6)
2,4-Dihydroxybenzophenone	$C_{13}H_{10}O_3$	7.2	215.0703	[M+H]+	1	3	258	258	UV filter	
Unknown[a]	$C_8H_{10}O_3S$	3.4	185.0278	[M−H]−	4	17	–	–	–	–
2-[4-(Diethylamino)-2-hydroxybenzoyl]benzoic acid	$C_{18}H_{19}NO_4$	8.7	314.1387	[M+H]+	3	3	89	89	UV filter TP	Triadimefon (8.7)
2′-Methylacetanilide	$C_9H_{11}NO$	4.7	150.0913	[M+H]+	3	11	53	35–280	Reagent	Venlafaxine (4.7)
2-Benzothiazolesulfonic acid	$C_7H_5NO_3S_2$	3.2	215.9784	[M+H]+	2a	8	1170	925–1357	Vulcaniser	Thiametoxam (3.2)
Unknown[a]	$C_8H_{18}O_4S$	6.6	209.0852	[M−H]−	4	25	–	–	–	–
2-Methyl-4′-(methylthio)-2-morpholinopropiophenone	$C_{15}H_{21}NO_2S$	4.8	280.1366	[M+H]+	2a	3	35	35	Drug of abuse	Propanolol (4.9)
3-(2-Hydroxyphenyl) propanoic acid	$C_9H_{10}O_3$	4.5	165.0557	[M−H]−	2a	6	100	93–106	Natural product	Atrazine desethyl (4.4)
4-Acetamidoantipyrine	$C_{13}H_{15}N_3O_2$	3.1	246.1237	[M+H]+	2a	64	548	57–6715	Pharmaceutical TP	Trimethoprim (3.1)
4-Aminoantipyrine	$C_{11}H_{13}N_3O$	2.1	204.1131	[M+H]+	1	17	414	136–1179	Pharmaceutical TP	
4-Formylaminoantipyrine	$C_{12}H_{13}N_3O_2$	3.1	232.1081	[M+H]+	2a	56	268	58–2486	Pharmaceutical TP	Trimethoprim (3.1)
Unknown[a]	$C_9H_{12}O_3S$	4.5	199.0436	[M−H]−	4	8	–	–	–	–
4-Methyl-5-thiazoleethanol	C_6H_9NOS	1.1	144.0478	[M+H]+	2a	22	252	51–606	Flavouring agent	Sotalol (1.3)
4-Pyridoxic acid	$C_8H_9NO_4$	0.9	184.0604	[M+H]+	2a	6	254	231–278	Natural product	Thiocyclam (0.9)
6-Phenyl-1,3,5-triazine-2,4-diamine	$C_9H_9N_5$	3.0	188.0931	[M+H]+	2a	8	833	584–892	Cross-linking agent	Lincomycin (2.9)
7-Diethylamino-4-methylcoumarin	$C_{14}H_{17}NO_2$	8.3	232.1332	[M+H]+	2a	3	1819	1819	Brightener/ fluorescent dye	Propyzamid (8.3)
Acetylsalicylic acid[a]	$C_9H_8O_4$	6.9	181.0492	[M+H]+	2a	25	–	–	Pharmaceutical	–
Adrenosterone	$C_{19}H_{24}O_3$	6.0	301.1798	[M+H]+	2a	3	223	223	Hormone	Carbamazepine (6.0)
Androsta-1,4-Dien-3,17-dione	$C_{19}H_{24}O_2$	6.9	285.1838	[M+H]+	2a	6	564	247–880	Steroid	Diuron (6.9)
Atenolol acid	$C_{14}H_{21}NO_4$	3.0	268.1543	[M+H]+	2a	6	95	38–153	Pharmaceutical TP	Trimethoprim (3.1)
Azoxystrobin (free acid)	$C_{21}H_{15}N_3O_5$	6.9	390.1084	[M+H]+	2a	3	92	92	Pesticide TP	Diuron (6.9)
Benzenesulfonamide	$C_6H_7NO_2S$	6.2	158.0270	[M+H]+	2a	3	74	74	Pharmaceutical TP	Fluometuron (6.2)
Benzoic acid	$C_7H_6O_2$	3.3	123.0442	[M+H]+	3	8	41	30–68	Pesticide TP	Thiabendazole (3.3)
Benzophenone-4	$C_{14}H_{12}O_6S$	4.9	307.0282	[M−H]−	1	6	1099	581–1617	UV filter	–
Benzoylecgonine	$C_{16}H_{19}NO_4$	3.7	290.1387	[M+H]+	1	33	419	176–986	Drug of abuse TP	–
Bisoprolol	$C_{18}H_{31}NO_4$	4.6	326.2326	[M+H]+	1	8	114	111–115	Beta-blocker	–
Corticosterone	$C_{21}H_{30}O_4$	7.9	347.2217	[M+H]+	2a	6	37	32–42	Natural product	Terbuthylazine (7.9)
Coumarin	$C_9H_6O_2$	4.5	147.0441	[M+H]+	2a	11	50	28–436	Cleaning product	Atrazine desethyl (4.4)
Cyclo(Leu-Pro)	$C_{11}H_{18}N_2O_2$	4.0	211.1442	[M+H]+	2a	8	207	145–265	Food additive	Lamotrigine (4.0)
Unknown[a]	$C_{12}H_{14}O_4$	6.9	223.0969	[M+H]+	4	17	–	–	–	–
Unknown[a]	$C_8H_{18}O_3$	4.8	163.1330	[M+H]+	4	17	–	–	–	–
Unknown[a]	$C_{16}H_{22}O_4$	11.8	279.1594	[M+H]+	4	8	–	–	–	–
Diphenyl isophthalate	$C_{20}H_{14}O_4$	5.7	319.0965	[M+H]+	3	3	30	30	Plasticizer	Imazalil (5.7)
Enalapril	$C_{20}H_{28}N_2O_5$	5.1	377.2078	[M+H]+	1	11	156	118–302	Antihypertensive	–
Enterolactone	$C_{18}H_{18}O_4$	5.8	299.1281	[M+H]+	2a	6	162	152–171	Natural product	Flumequine (5.8)
Europine	$C_{16}H_{27}NO_6$	2.2	330.1911	[M+H]+	2a	3	1225	1225	Natural product	Theobromine (2.1)
Heliotrine	$C_{16}H_{27}NO_5$	3.1	314.1962	[M+H]+	2a	3	89	89	Natural product	Trimethoprim (3.1)
Heliotrine N-oxide	$C_{16}H_{27}NO_6$	3.4	330.1911	[M+H]+	3	3	240	240	Natural product	Thiabendazole (3.3)
Hexa-methoxymethyl-melamine	$C_{15}H_{30}N_6O_6$	6.1	391.2300	[M+H]+	2a	6	381	222–540	Cross-linking agent/ vulcaniser	Carbamazepine (6.0)
Hordenine	$C_{10}H_{15}NO$	0.7	166.1226	[M+H]+	2a	6	347	211–483	Natural product	Cyromazine (0.8)

(*continued on next page*)

Table 1 (*continued*)

Compound	Formula	Rt (min)	*m/z*	Adduct	CL	DF (%) (n = 36)	Median conc. (ng L^{-1})	Range (ng L^{-1})	Class	Quantitative standard (Rt)
Indoleacetic acid	$C_{10}H_9NO_2$	4.6	176.0706	[M+H]+	2a	8	369	228–559	Plant growth regulator	Codeine (2.2)
Kynurenic acid	$C_{10}H_7NO_3$	3.1	190.0499	[M+H]+	2a	17	156	30–1062	Natural product	Trimethoprim (3.1)
Lasiocarpine N-oxide	$C_{21}H_{33}NO_8$	4.8	428.2279	[M+H]+	2a	3	39	39	Natural product	Venlafaxine (4.7)
Lauryl diethanolamide	$C_{16}H_{33}NO_3$	11.7	288.2533	[M+H]+	2a	14	4925	427–35,835	Foaming agent	Difenoconazole (11.6)
Lidocaine N-oxide	$C_{14}H_{22}N_2O_2$	3.7	251.1754	[M+H]+	2a	6	71	53–89	Pharmaceutical TP	Antipyrine (3.8)
Linalool	$C_{10}H_{18}O$	5.6	155.1430	[M+H]+	2a	6	81	60–102	Fragrance	Hexazinone (5.5)
Loliolide	$C_{11}H_{16}O_3$	4.4	197.1170	[M+H]+	2a	33	168	12–579	Natural product	Fluconazole (4.3)
Losartan	$C_{22}H_{23}ClN_6O$	7.2	423.1700	[M+H]+	2a	33	56	15–221	Antihypertensive	Spiroxamine (7.1)
Lumichrome	$C_{12}H_{10}N_4O_2$	5.2	243.0877	[M+H]+	2a	28	1645	262–5639	Natural product	Carbadox (5.2)
Methyl trans-styryl ketone	$C_{10}H_{10}O$	6.4	147.0804	[M+H]+	2a	8	2661	1599–3646	Flavouring agent	metobromuron (6.3)
Metoprolol	$C_{15}H_{25}NO_3$	3.9	268.1907	[M+H]+	1	3	202	202	Beta-blocker	–
Mycophenolic acid	$C_{17}H_{20}O_6$	7.5	321.1333	[M+H]+	2a	11	436	129–1045	Immunosuppresant	Irbesartan (7.5)
N-(2,4-Dimethylphenyl) formamide	$C_9H_{11}NO$	5.0	150.0913	[M+H]+	3	3	898	898	Pesticide TP	citalopram (5.0)
N4-Acetylsulfamethoxazole	$C_{12}H_{13}N_3O_4S$	4.4	296.0700	[M+H]+	2a	3	72.9	73	Pharmaceutical TP	Fluconazole (4.3)
N6-(delta2-Isopentyl)-adenine	$C_{10}H_{13}N_5$	4.1	204.1244	[M+H]+	2a	6	74	69–80	Natural product	Lamotrigine (4.0)
N,*N*-Bis(2-hydroxyethyl) decanamide[a]	$C_{14}H_{29}NO_3$	9.2	260.2226	[M+H]+	3	8	–	–	PCP/surfactant	–
N,*N*-Bis(2-hydroxyethyl) octanamide[a]	$C_{12}H_{25}NO_3$	6.4	232.1905	[M+H]+	3	8	–	–	Surfactant	–
Unknown[a]	$C_{13}H_{21}N$	2.2	192.1744	[M+H]+	4	8	–	–	–	–
N,*N*-Dimethyldodecylamine N-oxide[a]	$C_{14}H_{31}NO$	9.4	230.2479	[M+H]+	3	8	–	–	PCP/surfactant	–
Naproxen	$C_{14}H_{14}O_3$	7.8	231.1016	[M+H]+	1	8	1043	866–2276	Anti-inflammatory	
N-Butylbenzenesulfonamide	$C_{10}H_{15}NO_2S$	6.2	214.0896	[M+H]+	2a	14	53	37–343	Plasticizer	Fluometuron (6.2)
N-Ethyl-4-menthane-3-carboxamide	$C_{13}H_{25}NO$	8.9	212.2009	[M+H]+	2a	14	756	71–891	Flavouring agent	Chloroxuron (8.9)
N-Ethyl-o-toluenesulfonamide	$C_9H_{13}NO_2S$	5.2	200.0740	[M+H]+	2a	6	2872	1876–3867	Plasticizer	Carbadox (5.2)
Unknown[a]	$C_{14}H_{29}NO_2$	4.4	244.2266	[M+H]+	4	8	–	–	–	–
N-Nitrosomorpholine	$C_4H_8N_2O_2$	1.0	117.0659	[M+H]+	1	6	3348	381–6316	Nitrosamine	–
Pentoxifylline	$C_{13}H_{18}N_4O_3$	4.5	279.1446	[M+H]+	1	11	252	134–359	Hemorrhologic agent	–
Phenylbenzimidazole sulfonic acid	$C_{13}H_{10}N_2O_3S$	3.0	275.0485	[M+H]+	2a	22	90	44–339	UV filter	Trimethoprim (3.1)
Piperine	$C_{17}H_{19}NO_3$	9.0	286.1438	[M+H]+	2a	3	994	994	Natural product	Chloroxuron (8.9)
Propyphenazone	$C_{14}H_{18}N_2O$	5.9	231.1492	[M+H]+	1	8	175	173–236	Analgesic	–
Sitagliptin	$C_{16}H_{15}F_6N_5O$	4.0	408.1254	[M+H]+	2a	11	90	59–98.8	Antidiabetic	Lamotrigine (4.0)
Tiapride	$C_{15}H_{24}N_2O_4S$	2.3	329.1530	[M+H]+	2a	6	143	113–173	Antidepressant	Codeine (2.2)
Tonalide	$C_{18}H_{26}O$	14.5	259.2056	[M+H]+	3	31	22	18–71	Synthetic musk	Fenazaquin (14.5)
Triphenylphosphine oxide	$C_{18}H_{15}OP$	7.6	279.0933	[M+H]+	2a	8	2134	2009–2660	Catalyst	Irbesartan (7.5)
Ursocholic acid	$C_{24}H_{40}O_5$	12.0	426.3214	[M+NH4]+	3	3	136	136	Natural product	Prosulfocarb (12.0)

[a] Compounds tentatively identified by manual annotation.

et al., 2020), and it is considered a neurotoxic compound that can affect the kidneys and liver (Marrocco et al., 2015).

3.3. Spatial distribution of CECs

Surface water is a highly dynamic environment, meaning each sampling event provides only a snapshot of the pollution of the sampled area. A total of 157 analytes were determined in samples, a fact that reflects the high structural diversity and usage of the contaminants found. Fig. 4 depicts the distribution of different classes of compounds identified by both target and NTA, as well as the total concentration considering the median values of the four sampling campaigns.

Pharmaceuticals were the predominant group determined at almost all catchment points, representing around 50 % of the contamination in P3, P4, P5, P7, P9, and P11, mainly affected by urban discharges. PCPs were detected in all sampling points representing approximately between 15 and 30 % of total concentration. Caffeine, nicotine, and their by-products were among the most ubiquitous compounds. Despite the predominant agricultural practice of olive groves in the basin, pesticides represented an average of 20 % of the total concentrations. This indicates that although agriculture is the main economic activity of the region, it is not the primary source of most of the determined pollution. TP had an important role in samples, particularly in those catchments close to urban areas, where they represented between 8 and 23 % of compounds. Industrial chemicals were detected in all sampling points, mainly flame retardants and plasticizers globally used.

As expected, headwaters and sampling points close to natural areas (P1, P2, P6, P10) exhibited a lower number of detected compounds. The absence of agricultural activity and human disturbance is reflected in low concentration levels and a less varied range of compounds. A few kilometers downstream from headwaters; in the Arbuniel River (P3, downstream P1) and the Cambil River (P4, downstream P2), the number of compounds determined was significantly higher than upstream in the case of P4 and more diverse at both sampling points (Fig. 4). While the

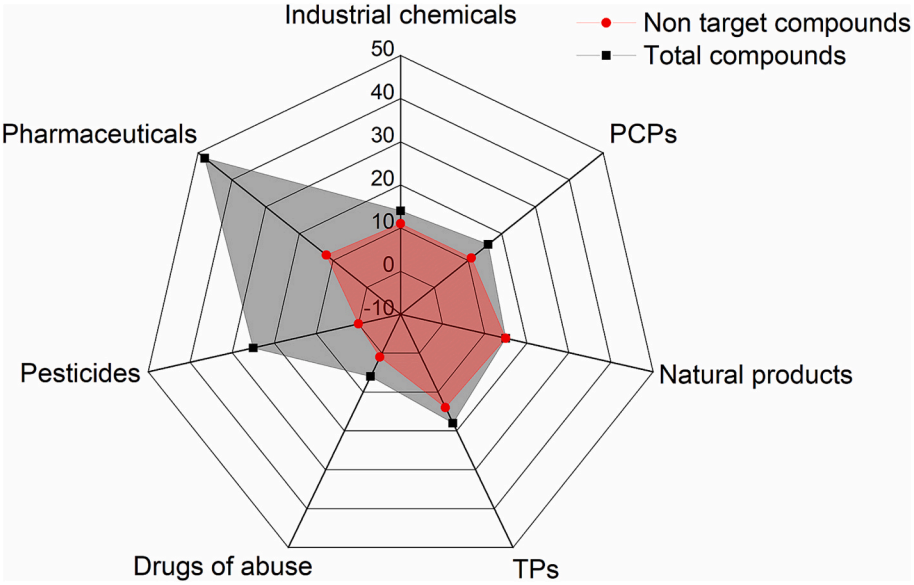


Fig. 2. Number of compounds identified by non-target analysis and total compounds detected.

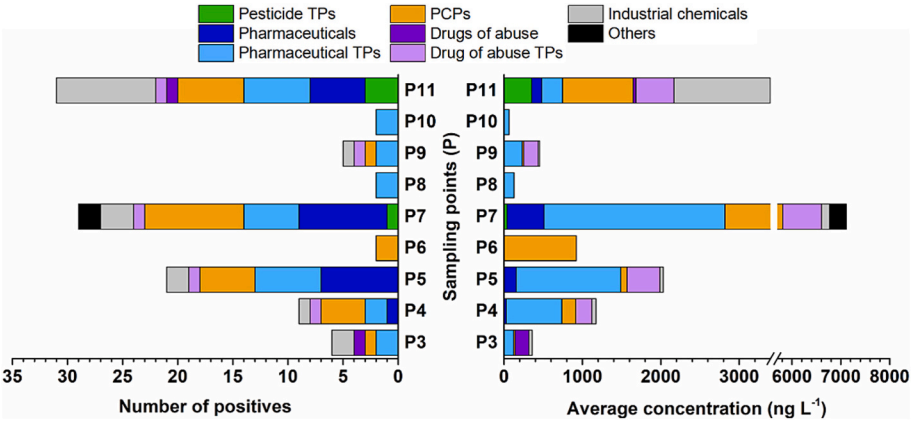


Fig. 3. Number of positive determinations through NTA (left) and estimated average concentrations of compounds in each sampling point (right).

total concentrations found in P1 and P3 were quite similar, an accumulation effect was observed when comparing P2 and P4. The increased concentration in P4 could be primarily attributed to discharges from nearby small villages, as indicated by the concentration levels found for pharmaceuticals and PCPs (Table S2) (Fig. 4), as well as higher agricultural pressure evidenced by the important contribution of pesticides.

Sampling points close to small villages lacking WWTPs or having inadequately upgraded treatment facilities showed higher contaminant emissions. This fact was observed when comparing the Jandulilla River (P5, downstream from a population of 6000 inhabitants with a WWTP) to Campillo River (P7, downstream from a village of 1700 inhabitants without a WWTP) (Fig. 4). Even though the contaminant emission profiles in both areas were equivalent according to the distribution and categories of positive detections, significant differences in total CEC concentration loads were observed (Fig. S3). It is important to mention the case of acetaminophen (paracetamol), with an average concentration of 280 ng L^{-1} in P5 and 29,505 ng L^{-1} in P7 (Fig. S3). It has been proven that the removal efficiency of this compound in municipal WWTPs is around 95 % (Wu et al., 2023). P7 exhibited the highest median concentration level (1635 ng L^{-1}) followed by the largest river (Guadalquivir River, P11) and the second maximum number of CECs (88) (Fig. S2). This fact underscores the importance of the implementation of sewage treatment systems, even in small populations, and the substantial influence that urban discharges have on water courses (Mandaric et al., 2018).

Downstream P7, a different contamination pattern was observed. In Guadalbullón River (P8, downstream P7), the number of positive detections and median concentrations decreased considerably from 1635 (P7) to 172 ng L^{-1} (P8), likely due to dilution effects. Nonetheless, pollution levels slightly increased again downstream in P9 (Jaén River, 320 ng L^{-1}), possibly as a result of a higher agricultural pressure and a larger urban population nearby.

The most polluted point in terms of the highest number of positives (92) was P11. This fact agrees with our previous findings focusing only on pesticides and TPs (Fernández-García et al., 2024). Although this is the largest river sampled and a dilution effect could be expected, the catchment site was located at the confluence of three tributaries, which means it collects the main part of the runoff of the province. The presence of pesticides used in agricultural practices for crops different than olive groves and the influence of industrial chemicals due to the discharges of an industrial area upstream may explain the substantial contribution to the total concentration observed at this point.

4. Conclusions

The chemical status of rural streams with low industrial development

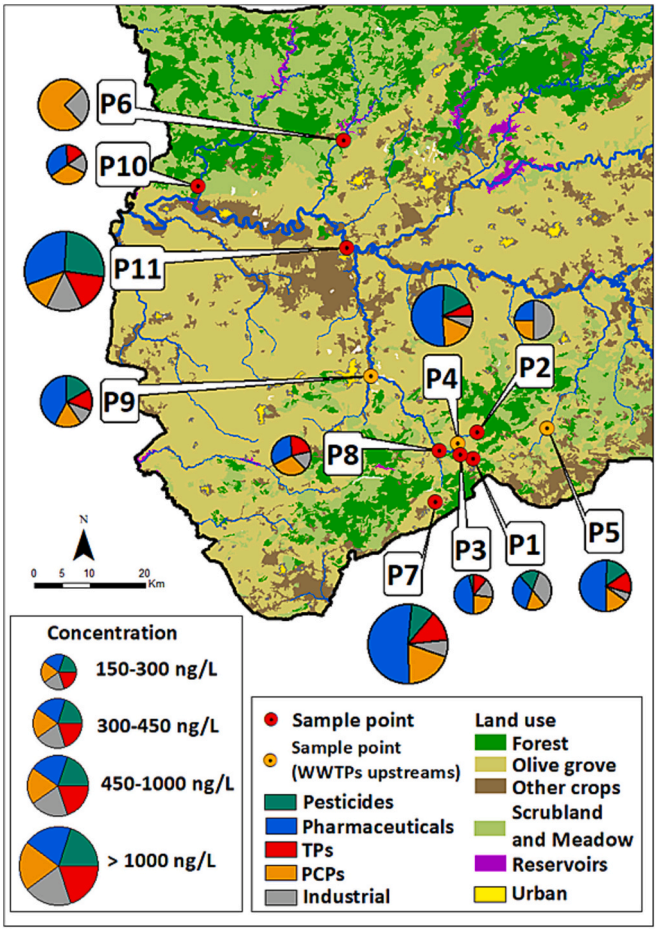


Fig. 4. Map of compounds detected by target and non-target screening and total concentrations expressed as the sum of the medians during the four sampling campaigns. Concentrations of natural products were not included.

in the Guadalquivir River basin was successfully characterized using a combination of target and NTA through LC-HRMS/MS analysis. A total of 157 compounds were identified comprising both strategies. At least five different compounds were detected in all samples studied, which shows the ubiquity of CECs, even in headwaters. The target method revealed that pharmaceuticals (40) were the main class of compounds detected, followed by pesticides (25), while PCPs contributed with the highest concentrations. This demonstrates that contamination has mostly an urban origin, despite agriculture is the main sector in this area. Up to 6 Priority Substances and 9 compounds included in the Watch List were detected, all of them with concentrations below the maximum allowable levels. The application of the NTA approach revealed the presence of 79 additional analytes of different classes. Up to 12 of these compounds were confirmed with their analytical standards, and 47 were tentatively identified at a high confidence identification level (2a). At some sampling sites, the concentration levels of the CECs determined through NTA were similar to those found for target compounds. These results highlight the importance of including non-target strategies to (tentatively) identify compounds that would otherwise have remained undetected, and the greater amount of chemical information that can be gathered using NTA strategies. Despite this, many potential organic micropollutants can remain unnoticed due to the large amount of data generated in LC-HRMS and NTA, which makes data processing one of the main bottlenecks in environmental analysis. Future developments in more efficient and reliable data processing tools are crucial to fully exploit the potential of NTA for comprehensive environmental monitoring.

By examining the total concentration loads along downstream courses, it was observed that the accumulation of CECs was primarily due to urban impact, as indicated by the higher presence of pharmaceuticals, their TPs and PCPs. However, some streams near small urban populations showed contamination levels similar to those of the largest river in terms of both number of CECs determined and total concentrations. This could be attributed to the lack of WWTPs or their poor performance in small towns. Therefore, the lack of a WWTP could have a greater impact on pollution than the number of inhabitants. These results can contribute to improving knowledge about potential sources of contamination on aquatic environments, as well as highlight concerns about the impact on surface water quality due to the absence of WWTPs in villages with small populations.

CRediT authorship contribution statement

Alfonso Fernández-García: Writing – original draft, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Ana B. Martínez-Piernas:** Writing – review & editing, Supervision, Investigation, Conceptualization. **David Moreno-González:** Writing – review & editing, Supervision, Conceptualization. **Bienvenida Gilbert-López:** Writing – review & editing, Supervision, Investigation, Funding acquisition. **Juan F. García-Reyes:** Writing – review & editing, Supervision, Resources.

Declaration of competing interest

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

Data availability

Data will be made available on request.

Acknowledgements

A.F.G. acknowledges funding from University of Jaén for the predoctoral contract (Acción 8a, ref. RYC_2_2022). A.B.M.P. acknowledges to the Regional Government of Andalucía and European Social Fund for the postdoctoral research fellowship (ref. DOC_01319). D.M.G. and. B.G. L. acknowledge funding support from the MCIN/AEI/10.13039/501100011033 through the Ramón y Cajal program (RYC2022-035915-I and RYC2019-026581-I, respectively). The authors also thank Servicios Centrales de Apoyo a la Investigación from the University of Jaén (SCAI-UJAEN) and INUO for research support.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2024.176587.

References

Bletsou, A.A., Jeon, J., Hollender, J., Archontaki, E., Thomaidis, N.S., 2015. Targeted and non-targeted liquid chromatography-mass spectrometric workflows for identification of transformation products of emerging pollutants in the aquatic environment. TrAC Trends Anal. Chem. 66, 32–44. https://doi.org/10.1016/j.trac.2014.11.009.

Bolton, E., Schymanski, E., Kondic, T., Thiessen, P., Zhang, J., 2024. PubChemLite for Exposomics (1.35.0). Zenodo. https://doi.org/10.5281/zenodo.12578095.

Bonnefille, B., Karlsson, O., Rian, M.B., Raqib, R., Parvez, F., Papazian, S., Martin, J.W., 2023. Nontarget analysis of polluted surface waters in Bangladesh using Open Science workflows. Environ. Sci. Technol. 57 (17), 6808–6824. https://doi.org/10.1021/acs.est.2c08200.

Cao, M., Fan, J., Guo, C., Chen, M., Lv, J., Sun, W., Xi, B., Xu, J., 2023. Comprehensive investigation and risk assessment of organic contaminants in Yellow River Estuary using suspect and nontarget screening strategies. Environ. Int. 173, 107843. https://doi.org/10.1016/j.envint.2023.107843.

Casado, J., Brigden, K., Santillo, D., Johnston, P., 2019. Screening of pesticides and veterinary drugs in small streams in the European Union by liquid chromatography high resolution mass spectrometry. Sci. Total Environ. 670, 1204–1225. https://doi.org/10.1016/j.scitotenv.2019.03.207.

Čelić, M., Jaén-Gil, A., Briceño-Guevara, S., Rodríguez-Mozaz, S., Gros, M., Petrović, M., 2021. Extended suspect screening to identify contaminants of emerging concern in riverine and coastal ecosystems and assessment of environmental risks. J. Hazard. Mater. 404, 124102. https://doi.org/10.1016/j.jhazmat.2020.124102.

Chibwe, L., Parrott, J.L., Shires, K., Khan, H., Clarence, S., Lavalle, C., Sullivan, C., O'Brien, A.M., De Silva, A.O., Muir, D.C.G., Rochman, C.M., 2022. A deep dive into the complex chemical mixture and toxicity of tire wear particle leachate in fathead minnow. Environ. Toxicol. Chem. 41 (5), 1144–1153. https://doi.org/10.1002/etc.5140.

Coelho, S.D., Maricoto, T., Taborda-Barata, L., Annesi-Maesano, I., Isobe, T., Sousa, A.C., 2024. Relationship between flame retardants and respiratory health–a systematic review and meta-analysis of observational studies. Environ. Pollut., 123733 https://doi.org/10.1016/j.envpol.2024.123733.

Confederación Hidrográfica del Guadalquivir, 2024. Plan Hidrológico de la Demarcación Hidrográfica del Guadalquivir (revisión para el tercer ciclo: 2022–2027). Memoria. Ministerio para la Transición Ecológica y el Reto Demográfico, Gobierno de España. https://www.chguadalquivir.es/documents/10182/2564531/PHGuadalquivir_Memoria.pdf/25f1430c-0952-0045-af17-9224edba53e9.

EC, 2013. Directive 2013/39/EU of the European Parliament and of the Council of 12 August 2013 amending Directives 2000/60/EC and 2008/105/EC as regards priority substances in the field of water policy. https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=celex%3A32013L0039.

EC, 2022. Commission Implementing Decision (EU) 2022/1307 of 22 July 2022. Establishing a watch list of substances for Union-wide monitoring in the field of water policy pursuant to Directive 2008/105/EC of the European Parliament and of the Council (notified under document number C(2022) 5098). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022D1307.

Escher, B.I., Fenner, K., 2011. Recent advances in environmental risk assessment of transformation products. Environ. Sci. Technol. 45 (9), 3835–3847. https://doi.org/10.1021/es1030799.

Escher, B.I., Stapleton, H.M., Schymanski, E.L., 2020. Tracking complex mixtures of chemicals in our changing environment. Science 367 (6476), 388–392. https://doi.org/10.1126/science.aay6636.

Fernández-García, A., Martínez-Piernas, A.B., Moreno-González, D., Gilbert-López, B., Molina-Díaz, A., García-Reyes, J.F., 2024. Occurrence and risk assessment of pesticides and their transformation products related to olive groves in surface waters of the Guadalquivir river basin. Chemosphere 357, 142075. https://doi.org/10.1016/j.chemosphere.2024.142075.

Finckh, S., Carmona, E., Borchardt, D., Büttner, O., Krauss, M., Schulze, T., Yang, S., Brack, W., 2024. Mapping chemical footprints of organic micropollutants in European streams. Environ. Int. 183, 108371. https://doi.org/10.1016/j.envint.2023.108371.

Fonseca, E., Renau-Pruñonosa, A., Ibáñez, M., Gracia-Lor, E., Estrela, T., Jiménez, S., Pérez-Martín, M.A., González, F., Hernández, F., Morell, I., 2019. Investigation of pesticides and their transformation products in the Júcar River Hydrographical Basin (Spain) by wide-scope high-resolution mass spectrometry screening. Environ. Res. 177, 108570. https://doi.org/10.1016/j.envres.2019.108570.

Fonseca, E., Hernández, F., Ibáñez, M., Rico, A., Pitarch, E., Bijlsma, L., 2020. Occurrence and ecological risks of pharmaceuticals in a Mediterranean river in Eastern Spain. Environ. Int. 144, 106004. https://doi.org/10.1016/j.envint.2020.106004.

Gouveia, T.I., Alves, A., Santos, M.S., 2019. New insights on cytostatic drug risk assessment in aquatic environments based on measured concentrations in surface waters. Environ. Int. 133, 105236. https://doi.org/10.1016/j.envint.2019.105236.

Grabicová, K., Duchet, C., Švecová, H., Randák, T., Boukal, D.S., Grabic, R., 2024. The effect of warming and seasonality on bioaccumulation of selected pharmaceuticals in freshwater invertebrates. Water Res., 121360 https://doi.org/10.1016/j.watres.2024.121360.

Hollender, J., Van Bavel, B., Dulio, V., Farmen, E., Furtmann, K., Koschorreck, J., Kunkel, U., Krauss, M., Munthe, J., Schlabach, M., Slobodnik, J., Stroomberg, G., Ternes, T., Thomaidis, N.T., Togola, A., 2019. High resolution mass spectrometry-based non-target screening can support regulatory environmental monitoring and chemicals management. Environ. Sci. Eur. 31 (1), 1–11. https://doi.org/10.1186/s12302-019-0225-x.

Reifycs Abf (Analysis Base File) Converter. Available online: http://www.reifycs.com/AbfConverter/index.html. (Accessed May 2023).

Hu, J., Lyu, Y., Chen, H., Li, S., Sun, W., 2023. Suspect and nontarget screening reveal the underestimated risks of antibiotic transformation products in wastewater treatment plant effluents. Environ. Sci. Technol. 57 (45), 17439–17451. https://doi.org/10.1021/acs.est.3c05008.

Kandie, F.J., Krauss, M., Beckers, L.M., Massei, R., Fillinger, U., Becker, J., Liess, M., Torto, B., Brack, W., 2020. Occurrence and risk assessment of organic micropollutants in freshwater systems within the Lake Victoria South Basin. Kenya. Sci. Total Environ. 714, 136748. https://doi.org/10.1016/j.scitotenv.2020.136748.

Katajamaa, M., Miettinen, J., Orešič, M., 2006. MZmine: toolbox for processing and visualization of mass spectrometry based molecular profile data. Bioinformatics 22 (5), 634–636. https://doi.org/10.1093/bioinformatics/btk039.

Khan, H.K., Rehman, M.Y.A., Malik, R.N., 2020. Fate and toxicity of pharmaceuticals in water environment: an insight on their occurrence in South Asia. J. Environ. Manag. 271, 111030. https://doi.org/10.1016/j.jenvman.2020.111030.

Kiefer, K., Du, L., Singer, H., Hollender, J., 2021. Identification of LC-HRMS nontarget signals in groundwater after source related prioritization. Water Res. 196, 116994. https://doi.org/10.1016/j.watres.2021.116994.

Lahens, L., Cabana, H., Huot, Y., Segura, P.A., 2024. Trace organic contaminants in lake waters: occurrence and environmental risk assessment at the national scale in Canada. Environ. Pollut. 347, 123764. https://doi.org/10.1016/j.envpol.2024.123764.

Lentz, M.P., Graham, D.J., van Vliet, M.T., 2024. Drought impact on pharmaceuticals in surface waters in Europe: case study for the Rhine and Elbe basins. Sci. Total Environ., 171186 https://doi.org/10.1016/j.scitotenv.2024.171186.

Llamas-Dios, M.I., Vadillo, I., Jiménez-Gavilán, P., Candela, L., Corada-Fernández, C., 2021. Assessment of a wide array of contaminants of emerging concern in a Mediterranean water basin (Guadalhorce river, Spain): motivations for an improvement of water management and pollutants surveillance. Sci. Total Environ. 788, 147822. https://doi.org/10.1016/j.scitotenv.2021.147822.

López-Serna, R., Kasprzyk-Hordern, B., Petrović, M., Barceló, D., 2013. Multi-residue enantiomeric analysis of pharmaceuticals and their active metabolites in the Guadalquivir River basin (South Spain) by chiral liquid chromatography coupled with tandem mass spectrometry. Anal. Bioanal. Chem. 405, 5859–5873. https://doi.org/10.1007/s00216-013-6900-7.

Malm, L., Palm, E., Souihi, A., Plassmann, M., Liigand, J., Kruve, A., 2021. Guide to semi-quantitative non-targeted screening using LC/ESI/HRMS. Molecules 26 (12), 3524. https://doi.org/10.3390/molecules26123524.

Malnes, D., Ahrens, L., Köhler, S., Forsberg, M., Golovko, O., 2022. Occurrence and mass flows of contaminants of emerging concern (CECs) in Sweden's three largest lakes and associated rivers. Chemosphere 294, 133825. https://doi.org/10.1016/j.chemosphere.2022.133825.

Mandaric, L., Mor, J.R., Sabater, S., Petrovic, M., 2018. Impact of urban chemical pollution on water quality in small, rural and effluent-dominated Mediterranean streams and rivers. Sci. Total Environ. 613, 763–772. https://doi.org/10.1016/j.scitotenv.2017.09.128.

Marrocco, A., Meade, B.J., Long, C.M., Lukomska, E., Marshall, N.B., Anderson, S.E., 2015. Investigations into the immunotoxicity and allergic potential induced by topical application of N-butylbenzenesulfonamide (NBBS) in a murine model. J. Toxicol. Env. Heal. A. 78 (17), 1122–1132. https://doi.org/10.1080/15287394.2015.1056898.

Martin, J., Buchberger, W., Santos, J.L., Alonso, E., Aparicio, I., 2012. High-performance liquid chromatography quadrupole time-of-flight mass spectrometry method for the analysis of antidiabetic drugs in aqueous environmental samples. J. Chromatogr. B 895, 94–101. https://doi.org/10.1016/j.jchromb.2012.03.023.

Masiá, A., Campo, J., Vázquez-Roig, P., Blasco, C., Picó, Y., 2013. Screening of currently used pesticides in water, sediments and biota of the Guadalquivir River Basin (Spain). J. Hazard. Mater. 263, 95–104. https://doi.org/10.1016/j.jhazmat.2013.09.035.

MassBank of North America, 2024. https://mona.fiehnlab.ucdavis.edu/. (Accessed May 2024).

Menger, F., Gago-Ferrero, P., Wiberg, K., Ahrens, L., 2020. Wide-scope screening of polar contaminants of concern in water: a critical review of liquid chromatography-high resolution mass spectrometry-based strategies. Tren. Environ. Anal. Chem. 28, e00102. https://doi.org/10.1016/j.teac.2020.e00102.

Mofijur, M., Hasan, M.M., Ahmed, S.F., Djavanroodi, F., Fattah, I.M.R., Silitonga, A.S., Kalam, M.A., Zhou, J.L., Khan, T.Y., 2023. Advances in identifying and managing emerging contaminants in aquatic ecosystems: analytical approaches, toxicity assessment, transformation pathways, environmental fate, and remediation strategies. Environ. Pollut., 122889 https://doi.org/10.1016/j.envpol.2023.122889.

National Institute of Health, 2024. PubChem, National Institute of Health. https://pubchem.ncbi.nlm.nih.gov/. (Accessed May 2024).

Ng, K., Alygizakis, N., Nika, M.-C., Galani, A., Oswald, P., Oswaldova, M., Cirka, Ľ., Kunkel, U., Macherius, A., Sengl, M., Mariani, G., Tavazzi, S., Skejo, H., Gawlik, B. M., Thomaidis, N.S., Slobodnik, J., 2023. Wide-scope target screening characterization of legacy and emerging contaminants in the Danube River Basin by liquid and gas chromatography coupled with high-resolution mass spectrometry. Water Res. 230, 119539. https://doi.org/10.1016/j.watres.2022.119539.

Oertel, R., Baldauf, J., Rossmann, J., 2018. Development and validation of a hydrophilic interaction liquid chromatography-tandem mass spectrometry method for the quantification of the antidiabetic drug metformin and six others pharmaceuticals in wastewater. J. Chromatogr. A 1556, 73–80. https://doi.org/10.1016/j.chroma.2018.04.068.

Palmiotto, M., Castiglioni, S., Zuccato, E., Manenti, A., Riva, F., Davoli, E., 2018. Personal care products in surface, ground and wastewater of a complex aquifer system, a potential planning tool for contemporary urban settings. J. Environ. Manag. 214, 76–85. https://doi.org/10.1016/j.jenvman.2017.10.069.

Riva, F., Zuccato, E., Davoli, E., Fattore, E., Castiglioni, S., 2019. Risk assessment of a mixture of emerging contaminants in surface water in a highly urbanized area in Italy. J. Hazard. Mater. 361, 103–110. https://doi.org/10.1016/j.jhazmat.2018.07.099.

Robles-Molina, J., Gilbert-López, B., García-Reyes, J.F., Molina-Díaz, A., 2014. Monitoring of selected priority and emerging contaminants in the Guadalquivir River and other related surface waters in the province of Jaén. South East Spain. Sci. Total Environ. 479, 247–257. https://doi.org/10.1016/j.scitotenv.2014.01.121.

Royano, S., de la Torre, A., Navarro, I., Martínez, M.Á., 2023. Pharmaceutically active compounds (PhACs) in surface water: occurrence, trends and risk assessment in the Tagus River Basin (Spain). Sci. Total Environ. 905, 167422. https://doi.org/10.1016/j.scitotenv.2023.167422.

Ruttkies, C., Schymanski, E.L., Wolf, S., Hollender, J., Neumann, S., 2016. MetFrag relaunched: incorporating strategies beyond in silico fragmentation. J. Chemother. 8, 1–16. https://doi.org/10.1186/s13321-016-0115-9.

Schymanski, E.L., Jeon, J., Gulde, R., Fenner, K., Ruff, M., Singer, H.P., Hollender, J., 2014. Identifying small molecules via high resolution mass spectrometry:

communicating confidence. Environ. Sci. Technol. 48 (4), 2097–2098. https://doi.org/10.1021/es5002105.

Smith, C.A., Want, E.J., O'Maille, G., Abagyan, R., Siuzdak, G., 2006. XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. Anal. Chem. 78 (3), 779–787. https://doi.org/10.1021/ac051437y.

Song, H., Fan, X., Liu, G., Xu, J., Li, X., Tan, Y., Qian, H., 2016. Inhibitory effects of tributyl phosphate on algal growth, photosynthesis, and fatty acid synthesis in the marine diatom Phaeodactylum tricornutum. Environ. Sci. Pollut. R. 23, 24009–24018. https://doi.org/10.1007/s11356-016-7531-2.

Tian, Z., Peter, K.T., Gipe, A.D., Zhao, H., Hou, F., Wark, D.A., Khangaonkar, T., Kolodziej, E.P., James, C.A., 2019. Suspect and nontarget screening for contaminants of emerging concern in an urban estuary. Environ. Sci. Technol. 54 (2), 889–901. https://doi.org/10.1021/acs.est.9b06126.

Toth, J., Fugère, V., Yargeau, V., 2024. Relationship between stream size, watershed land use, and pesticide concentrations in headwater streams. Environ. Pollut. 349, 123940. https://doi.org/10.1016/j.envpol.2024.123940.

Tsugawa, H., Cajka, T., Kind, T., Ma, Y., Higgins, B., Ikeda, K., Kanazawa, M., VanderGheynst, J., Fiehn, O., Arita, M., 2015. MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis. Nat. Methods 12 (6), 523–526. https://doi.org/10.1038/nmeth.3393.

Valcárcel, Y., Alonso, S.G., Rodríguez-Gil, J.L., Maroto, R.R., Gil, A., Catalá, M., 2011. Analysis of the presence of cardiovascular and analgesic/anti-inflammatory/antipyretic pharmaceuticals in river-and drinking-water of the Madrid Region in Spain. Chemosphere 82 (7), 1062–1071. https://doi.org/10.1016/j.chemosphere.2010.10.041.

Vosough, M., Schmidt, T.C., Renner, G., 2024. Non-target screening in water analysis: recent trends of data evaluation, quality assurance, and their future perspectives. Anal. Bioanal. Chem. 416 (9), 2125–2136. https://doi.org/10.1007/s00216-024-05153-8.

Wu, J.-L., Liu, Z.H., Ma, Q.G., Dai, L., Dang, Z., 2023. Occurrence, removal and risk evaluation of ibuprofen and acetaminophen in municipal wastewater treatment plants: a critical review. Sci. Total Environ., 164600 https://doi.org/10.1016/j.scitotenv.2023.164600.

Zahn, D., Arp, H.P.H., Fenner, K., Georgi, A., Hafner, J., Hale, S.E., Hollender, J., Letzel, T., Schymanski, E.L., Reemtsma, T., 2024. Should transformation products change the way we manage chemicals? Environ. Sci. Technol. 58 (18), 7710–7718. https://doi.org/10.1021/acs.est.4c00125.

Zhu, K., Sarvajayakesavalu, S., Han, Y., Zhang, H., Gao, J., Li, X., Ma, M., 2022. Occurrence, distribution and risk assessment of organophosphate esters (OPEs) in water sources from Northeast to Southeast China. Environ. Pollut. 307, 119461. https://doi.org/10.1016/j.envpol.2022.119461.