Journal of Chromatography A 1783 (2026) 467150

Contents lists available at ScienceDirect

Journal of Chromatography A

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

A proof-of-concept evaluation of an ROI-MCR and HILIC-based method for screening environmental contaminants in surface and groundwater

Samuele Pellacani [a], Francesco Ronchetti [a], Andrea Cerrato [b], Anna Laura Capriotti [b], Marina Cocchi [a], Luísa Collischonn [a], Caterina Durante [a,*], Guido Perra [a], Lorenzo Strani [a]

[a] Department of Chemical and Geological Science, University of Modena and Reggio Emilia, via Campi 103, Modena, Italy
[b] Department of Chemistry, Sapienza University of Rome, Piazzale Aldo Moro 5, Rome, Italy

ARTICLE INFO

Keywords:
Emerging contaminants
ROI-MCR
Chemometrics
Untargeted approach
Groundwater
Po alluvial plain

ABSTRACT

This study presents a preliminary evaluation of an untargeted approach for the detection of persistent and mobile organic contaminants, PMOCs, in surface and groundwater samples from northern Italy. The main objective was to assess the performance and applicability of an untargeted screening workflow based on the Regions of Interest-Multivariate Curve Resolution (ROI-MCR) approach on a data set collected by liquid chromatography coupled with high-resolution mass spectrometry (LC—HRMS). Particular attention was given to ensure the reliability and robustness of the extracted spectral features. A targeted analysis was also performed on thirteen PMOCs, eight pharmaceuticals, three central nervous system (CNS) stimulants, one pesticide, and one artificial sweetener, which were selected for their physicochemical properties and their high mobility in aquatic environments. The untargeted screening led to the tentative identification of twenty-four chemical compounds belonging to different classes such as pharmaceuticals, cosmetics, pesticides and industrial products. Their distribution was analysed in relation to the geological and hydrological features of the study area, highlighting the influence of local environmental conditions on the occurrence and mobility of contaminants.

1. Introduction

Emerging Organic Contaminants (EOCs) comprise several groups of organic substances that pose potential risks to both ecosystem integrity and human health [1,2]. Although the term "emerging" implies novelty, many of these compounds have actually been present in the environment for some time. Their recent detection is primarily due to the advances in sampling and analytical techniques [3]. EOCs encompass a broad spectrum of chemicals, originating from both natural and synthetic sources, including plasticizers, pesticides, fertilizers, antibiotics, hormones, parabens, per- and polyfluoroalkyl substances (PFAS) and their metabolites [4]. Over the past few decades, the release of EOCs into the environment has increased due to the expansion of industrial and agricultural activities, alongside rapid urbanization and population growth. As a result, these substances are now widely reported not only in wastewater, but also in surface water and, increasingly, in groundwater [5,6]. Among them, Persistent and Mobile Organic Contaminants (PMOCs), which are investigated in the present work, are rather polar EOCs [7] and owing to their high polarity, water solubility and low sorption to soils sediments, they are particularly challenging to analyse in water samples and are prone to diffuse into the environment due to their physio-chemical characteristics.

As regards emerging contaminants' presence in surface water, the European Union (EU) has introduced several legislative initiatives aimed at safeguarding water quality and ecosystem integrity. In particular, the last EU directive [8], entered into force on 2026, 11 May, revises the lists of pollutants in surface and groundwaters ensuring that the lists are aligned with the latest scientific advice and that new substances will be monitored more closely and subject to stricter controls.

The monitoring of emerging organic contaminants in groundwater has progressively shifted from traditional targeted analyses toward non-target screening (NTS) approaches based on liquid chromatography coupled with high-resolution mass spectrometry (LC—HRMS). This transition has been driven by the need to detect a wider range of contaminants, including transformation products (TPs), which are frequently overlooked by conventional targeted protocols [9–12].

* Corresponding author.
E-mail address: caterina.durante@unimore.it (C. Durante).

https://doi.org/10.1016/j.chroma.2026.467150
Received 8 April 2026; Received in revised form 15 May 2026; Accepted 28 May 2026
Available online 2 June 2026

Regarding the analysis and presence of organic micropollutants in Italian surface and groundwater, some studies on endocrine disrupting compounds have been conducted in the Lamone, Montone, Ronco, Savio, Marecchia river basins (eastern part of the region) [13]. The possibility to carry out a successfully broad investigation of EOCs, to uncover also unexpected or less-studied compounds, through an untargeted approach has still provided useful results in rivers monitoring, as shown in the study of Montone et al. [14]. Additional analyses of EOCs have been carried out in the sparse surface and groundwaters of the Emilia Romagna Region [15]. The Emilia-Romagna Region, located in northern Italy, accounts for the largest share of irrigated land in Italy, and its agricultural sector ranks among the country's most productive [16]. In this region, despite the widespread physio-chemical monitoring of surface water and groundwater carried out by the Regional Environmental Protection Agency (ARPAe), there is still a limited knowledge about the distribution of EOCs, and only a small number of scientific publications/reports exist on the subject. In particular, targeted investigations on endocrine-disrupting compounds have been carried out in the Lamone, Montone, Ronco, Savio and Marecchia river basins in the eastern part of the region [13], and additional analyses of EOCs have been reported for sparse surface and groundwater sites [15]. However, in large parts of Emilia–Romagna, including the Modena district, information on other classes of EOCs, such as pharmaceuticals, estrogens and illicit drugs, remains scarce, particularly for groundwater.

Despite the considerable analytical potential of NTS, the processing and interpretation of the highly complex datasets generated by LC—HRMS remain a major challenge. The use of hydrophilic interaction liquid chromatography (HILIC) columns further enhances the separation and detection of highly polar and zwitterionic compounds. Recent studies have proposed different strategies for the treatment of untargeted LC—HRMS data. Han et al. (2022), for example, employed the R-based XCMS package using the "centWave" algorithm for peak picking after an initial Regions of Interest (ROI) selection aimed at reducing data dimensionality [17]. Their workflow combined compound identification through in-house databases and the GNPS empirical library, while MetFrag (https://msbi.ipb-halle.de/MetFrag/) was used for the tentative annotation of candidate compounds at different confidence levels. Similarly, Zhao et al. (2023) and Aurich et al. (2023) adopted open-source platforms such as patRoon and metID, which enable systematic feature extraction and prioritization [12,18]. In particular, Aurich et al. highlighted the importance of IPO (Isotopologue Parameter Optimization) for refining feature extraction parameters, including *m/z* tolerance and peak width, demonstrating how these settings strongly influence the reliability of the resulting feature lists [18].

Other researchers have increasingly turned toward specialized Python-based workflows. Liu et al. (2025) and Xia et al. (2024) used the PyHRMS package to extract ion features and developed an innovative HRMS fingerprinting strategy based on dilution curves for source apportionment in groundwater and urban streams [10,11]. Their approach enabled semi-quantification through the use of Tanimoto coefficients to evaluate structural similarity between unknown compounds and reference standards. In contrast, some studies rely on vendor-specific software solutions. Jin et al. (2025), for instance, utilized Compound Discoverer 3.3, which prioritizes toxic contaminants in secondary effluents through automated scoring systems based on spectral library matching, detection frequency, and relative peak areas [19].

Although these approaches represent significant progress, several important limitations persist. Conventional peak-picking algorithms, including those implemented in XCMS or commercial software, often struggle to accurately resolve overlapping or co-eluting peaks, which are particularly common in groundwater matrices because of the presence of humic substances and other matrix interferences [11,19]. Furthermore, many componentization procedures currently implemented in platforms such as patRoon, especially those aimed at grouping adducts and isotopes, remain computationally demanding and are frequently affected by false associations when applied to large environmental datasets [18]. Another critical issue concerns quantification uncertainty. As demonstrated by Kruve et al. (2021), electrospray ionization (ESI) response factors may vary across several orders of magnitude, making quantification without analytical standards highly unreliable [20]. Their work also showed that approaches based on the "closest eluting standard" can fail when the acid-base properties of the analyte differ substantially from those of the reference compound [20].

To address these challenges, workflows based on the integration of Regions of Interest with Multivariate Curve Resolution (ROI-MCR) [21] have emerged as particularly promising alternatives [22,23]. While Han et al. introduced ROI mainly as a preliminary data reduction strategy [17], the complete integration of Multivariate Curve Resolution–Alternating Least Squares (MCR-ALS) enables the mathematical separation of co-eluting compounds into pure elution profiles and corresponding mass spectra [21]. The use of LC—HRMS combined with the Region of Interest Multivariate Curve Resolution (ROI-MCR) approach provides significant advantages in the identification of analytes in untargeted studies, particularly when dealing with complex matrices such as environmental water samples [22]. This method could be especially valuable when analytes are present at trace levels, resulting in low-intensity signals often hidden by overlapping peaks in the chromatogram or by low signal-to-noise values. The ROI approach requires the definition of three input parameters, namely a threshold for signal-to-noise ratio, an *m/z* tolerance deviation and the minimum number of consecutive signals at a particular *m/z* value. Although this automated feature extraction is designed to improve signal clarity and reduce data complexity, it is not devoid of critical limitations. Specifically, the ROI selection process may include irrelevant features related to column bleed (e.g., compounds naturally released by the chromatographic column), baseline drift, or sample pretreatment processes, where reagents or procedural steps may introduce contaminants. Therefore, rigorous evaluation of selected ROIs is essential to obtain trustworthy results. To address this challenge, two key control steps are recommended [24]. First the inclusion of blanks in the analysis is mandatory. In particular, both solvent blanks (containing only the dilution medium, identical to the mobile phase) and method blanks (subjected to the same SPE treatment as the samples) signals help differentiate between true sample-related features and those introduced by the analytical process or solvent. Selected ROIs must be critically assessed by comparing their presence or absence in the blanks. Additionally, the elution profile associated to each ROI should be examined for chromatographic peak consistency. Peaks lacking defined Gaussian-like profiles may indicate artefacts or noise.

Within this context, recent developments in ROI-based workflows have focused not only on improving feature extraction efficiency, but also on enhancing the reliability and reproducibility of untargeted analyses. Recent work in data-independent acquisition (DIA) [22] has emphasized the necessity of establishing standardized protocols for ROI selection, which greatly improves the reproducibility and interpretability of untargeted studies. In our study, we adapted and rigorously applied these principles within a data-dependent acquisition (DDA) framework. The challenge is more pronounced in the present study due to the inherently lower concentrations of emerging contaminants in surface and subsurface water samples compared to matrices like wastewater, where contaminant levels are typically higher.

For these reasons, this work aims to evaluate the performance of an HILIC—HRMS and ROI-MCR based untargeted analytical approach for investigation of EOCs in surface and subsurface waters of the Modena province (Emilia-Romagna Region). This study area is in the central sector of the Po alluvial plain, one of the most densely populated and industrialised areas of Europe (Fig. 1). The retrieved components were putatively identified by means of the MSident tool [25] and validated using Compound Discoverer software and targeted analysis for a selected number of contaminants.

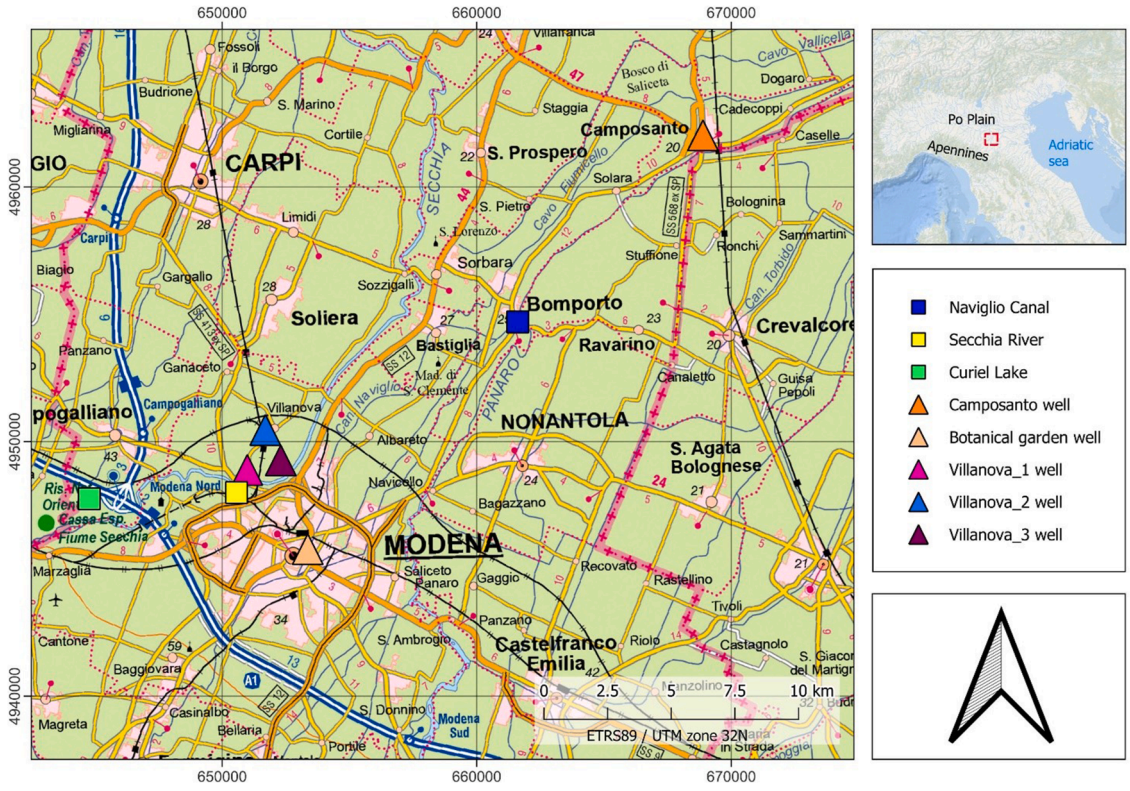


Fig. 1. Geographical setting and detailed map of the water sampling points. In the legend, the symbol of the surface (river, lake, ditch) and groundwater (well) sampling points are reported with the corresponding names. Square symbols indicate position of the surface water sampling point; triangle symbols indicate position of the groundwater sampling point.

2. Materials and methods

2.1. Sampling and preprocessing

Eight water samples were collected from different locations within the Modena district (Fig. 1).

The samples were categorized as follows:

- Groundwater: five samples were obtained from five distinct groundwater wells with varying depths (Table 1). These samples were labelled Villanova_1, Villanova_2, Villanova_3, Botanical Garden, and Camposanto.
- Surface water: three samples were collected from three different sampling points: an artificial water basin; a river; an artificial ditch (Table 1). These were labelled Curiel_Lake, Secchia_river, and Naviglio_canal.

The sampled wells are shallow and draw water from unconfined aquifers. These aquifers are directly recharged by precipitation through one to three meters of unsaturated zone, or they are connected to surface waters. The Curiel Lake is an artificial water basin (an abandoned gravel quarry), and the lake's water results from the outcropping of the local water table. The unconfined and shallow aquifer surrounding Curiel Lake is fed by the Secchia River.

The geochemical and hydrochemical information of the investigated water basins are reported in the Supplementary Materials (**Table S1** and **Figure S1**).

All samples were collected in one-liter amber glass bottles and stored at 5 °C under refrigeration until analysis. Methanol was added to the samples to obtain a final methanol content of 5 % (v/v) to inhibit microbial activity after sampling, as reported in the literature [26]. Analyses were performed within one week of collection. No pretreatment was applied to the subsurface water samples, as they were already suitable for solid-phase extraction. The three surface water samples underwent vacuum filtration with 1.6 µm glass microfiber filters to

Table 1
Sampling points, depths, and geological notes of the five groundwater samples.

Groundwater wells	Sampling date	Sampling point (city/ inhabited)	Well depth (m)	Hydrogeological notes
Villanova_1	*Nov-24*	*Villanova*	*3,8*	*Shallow groundwater connected to the surface water body (Secchia River)*
Villanova_2	*Dic-24*	*Villanova*	*20*	*Groundwater connected to the surface water body (Secchia River)*
Villanova_3	*Dic-24*	*Villanova*	*40*	*Groundwater connected to the surface water body (Secchia River)*
Botanical Garden	*Nov-24*	*Modena*	*30*	*Shallow groundwater connected to the surface water body (Secchia River)*
Camposanto	*Apr-25*	*Camposanto*	*6,6*	*Groundwater connected to the surface water body (Panaro River and Naviglio Canal)*
Curiel_Lake	*Jan-25*	*Campogalliano*	*0*	*Surface water; a water mixing between groundwater (aquifer water) and Secchia River (surface water)*
Secchia_river	*Jan-25*	*Ponte Alto*	*0*	*Surface water*
Naviglio_canal	*Jan-25*	*Bomporto*	*0*	*Surface water*

remove any sediment.

2.2. Chemicals and reagents

All solvents and additives used were LC-MS grade. Acetonitrile (ACN), methanol (MeOH), acetic acid and ammonium acetate (ANH4) were purchased from Incofar Srl (Modena, Italy). Ultra-pure water was obtained in the lab using a Milli-Q Millipore system (Millipore, Bedford MD). Analytical standards (all above 98 % of purity) were purchased from different suppliers: 4-dichlorophenoxyacetic acid, atenolol, caffeine, clenbuterol, hydrochlorothiazide, furosemide, metformin, metoprolol, nicotine, taurine and terbutaline were from Sigma- Aldrich; salbutamol from Alfa Aesar (Haverhill, MA, USA). Stock solutions of single standards were prepared at a concentration ranging from 1000 to 2000 μg L^{-1} in MeOH or MeOH:D.I. water (1:1 v/v) and stored at −18 °C. A mixture of the standard solutions at 50 μg L^{-1} was obtained diluting single solutions and used to prepare daily working solutions by further dilutions. The StrataX-CW cartridge (WCX, 200 mg / 6 mL, 33 μm polymeric weak cation mixed-mode solid phase adsorbent), Strata-X-AW cartridge (WAX, 200 mg/ 6 mL, 33 μm polymeric weak anion mixed-mode solid phase adsorbent), Strata-X cartridge (NEU, 200 mg/ 6 mL, 33 μm polymeric reversed solid phase adsorbent) were both from Phenomenex Srl (Phenomenex Srl, Castelmaggiore, BO, Italy). A vacuum manifold for 12 cartridges (Phenomenex Srl, Castelmaggiore, BO, Italy) was used for the solid phase extraction (SPE).

2.3. Sample treatment

All SPE cartridges were conditioned and equilibrated with 3 mL of methanol and 3 mL of deionized water, respectively, following SPE vendor suggested method. An aliquot of 300 mL of sample was loaded onto the SPE cartridge. The flow was no faster than 1 drop/s. Next, 3 mL of water was used for washing and finally, 3 mL of 2 % acid formic, 3 mL of 5 % NH_4OH and 3 mL of 2 % acid formic in methanol, were used for Strata-X-CW, Strata-X-AW and Strata-X, respectively, to elute the analytes. The extractants were evaporated to dryness under a stream of nitrogen. The residues were dissolved in 1000 μL of an acetonitrile/ deionized water mixture (95:5 v/v, eluent B used for HPLC analysis) and transferred into a 2 mL vial for HPLC autosampler.

2.4. Instrumental analysis

2.4.1. Untargeted analysis

Samples were analysed using the Thermo Fisher Scientific Vanquish Core (Thermo-Fisher Scientific, Waltham, MA, USA) as the UHPLC instrument coupled to a heated electrospray ionization system and a mass spectrometer, the Exploris 120 Orbitrap (UHPLC—HESI-Orbitrap). The chromatographic column and separation were the same used in the targeted analysis. The following HESI source characteristics were used: sheath gas 70 arbitrary units (au), auxiliary gas 5 au, sweep gas 0.5 au, ion transfer tube temperature 390 °C, evaporator temperature 150 °C and electrospray voltage 4.2 kV (positive mode). Both positive and negative mode ions were used for the extraction and putative identification of features. Analyses were acquired by using the Xcalibur software version 4.4 (Thermo-Fisher Scientific, Waltham, MA, USA) in full-scan (FS), data-dependent (dd-MS2) modes and a resolving power of 60,000 full width at half maximum (FWHM) at *m/z* values of 200 for FS mode and 30,000 for dd-MS2 mode. The isolation window for filtering the precursor ions was set to *m/z* values of 1.2, and a progressive step collision energy was used to fragment the precursor ions.

Extracts were analyzed on two different days and loaded into the autosampler in a random order (the measurement sequence was reported in **Table S2**, Supplementary Materials).

2.4.2. Targeted analysis

Thirteen different persistent and mobile organic contaminants (2,4-dichlorophenoxyacetic acid, atenolol, caffeine, clenbuterol, hydrochlorothiazide, furosemide, metformin, metoprolol, nicotine, salbutamol, taurine, and terbutaline) were quantified in the investigated samples by an analytical procedure optimized in a previous work [27]. The procedure involved the use of a 1200 series HPLC coupled to a 6410B triple quadrupole mass spectrometer by Agilent Technologies (Santa Clara, CA, USA) to achieve chromatographic separation and analyte detection. Electrospray ionization (ESI) in polarity switching mode was used as ion source for the MS system. The following values were set for the ESI source for both positive and negative polarities: drying gas (N_2) set at a temperature and flow of 300 °C and 9 L min^{-1}, respectively, and nebulizer pressure 32 psi. Capillary voltage was set at 4000 V and 3500 V for positive and negative ionization, respectively. Data related to the MRM transitions for each analyte are reported in **Table S3**, Supplementary Materials [27].

The column used for the separation was a SeQuant® ZIC®-cHILIC column (100 mm x 2.1 mm i.d., 3 μm particle size) by Merck (Darmstadt, Germany). The ZIC®-cHILIC column stationary phase is constituted of porous silica (100 Å pore size) modified with phosphorylcholine zwitterionic functional groups. The stationary phase was considered interesting because it contains both positively and negatively charged functional groups, thus probably providing interaction with the selected analytes characterized by different chemical properties [27]. The column was maintained at 25 °C, and the injection volume was 5 μL. Eluent A was ultra-pure water with 0.5 % of acetic acid and 1 mM of ammonium acetate, while eluent B was ACN: H_2O 95:5 v/v with the same amount of modifiers. The mobile phase flow was set at 0.1 mL $min^{-1,}$ and the gradient program was as follows: initially 100 % of phase B was kept constant for 2 min; then the percentage of phase B was gradually decreased to 75 % in 30 min (kept constant for 10 min); finally, the column was restored to the initial condition at minute 40. The column was re-equilibrated for a further 12 min (total time run: 52 min). The Agilent MassHunter workstation software (version B.04.01) was used for data acquisition, qualitative and quantitative analysis. The parameters measured for monitoring the performance of the used analytical method, namely LODs and LOQs, linear dynamic range, precision, and accuracy, are reported in **Table S3**, Supplementary Materials. Moreover, specificity was evaluated by verifying the analytes' retention time and by calculating the ratios between the quantifier and qualifier MRM transitions of each compound. The ratios of the MRM signal transition (qualifier and quantifier) were calculated in all samples and confirmed to deviate by no >30 % from the reference standard. Some examples are reported in **Figure S2**, Supplementary Materials.

2.5. Data analysis

2.5.1. Data acquisition planning

To account for potential background signals and to control for procedural artefacts, a total of four different sample blanks were included in the analysis. Three of these were method blanks, i.e. water samples pre-processed with each of the three SPE cartridges employed in the analysis (i.e. WAX, WCX and NEU). Finally, a solvent blank, consisting of the dilution medium used for sample preparation, was also analysed. The latter was injected four times. All these blanks served as critical references for identifying and eliminating features arising from the extraction procedure, the instrumental system or the mobile phase.

To evaluate instrumental reproducibility and provide a benchmark for comparison with putatively identified compounds, a standard mixture used in targeted analysis was injected five times. This provided a reference matrix for further assessing the consistency of feature identification made by MSident protocol. Each environmental sample was subjected to all three SPE procedures (WAX, WCX, and NEU), in order to broaden the spectrum of retained analytes and maximize compound coverage.

This comprehensive pretreatment yielded a total of 24 distinct water samples (8 samples x 3 different SPE processing). Furthermore, a quality

control (QC) sample was prepared by pooling 100 μL aliquots from each of the 24 samples. This pooled QC was analysed five times to monitor the analytical performance and detect any potential drift across the batch.

In total, considering blanks, standards and waters, 41 samples were analysed by LC—HRMS (**Table S3**, Supplementary Materials).

2.5.2. ROI-MCR protocol

Region of Interest (ROI) search combined with Multivariate Curve Resolution – Alternating Least Square (MCR-ALS) [28,29], i.e. ROI-MCR, were applied to achieve data compression, chromatographic resolution and features extractions.

The ROI selection is driven by three key input parameters: (i) a signal-to-noise threshold to filter out background noise, (ii) an *m/z* tolerance that determines whether signals are attributed to the same *m/z* feature, and (iii) a minimum number of consecutive scans at a given *m/z* to define a chromatographic peak. In this study, the parameter values were established through a comprehensive analysis of literature [24,30] and conducting several trials on the dataset. following initial tests and based on the literature [21] for both positive and negative ions dataset, a threshold of 0.1 % of the maximum ion peak intensity (automatically detected by the software) and 20 minimum number of occurrences proved to be the optimal choice. The *m/z* error was set, as usually for Orbitrap analysis, to 0.005 Da. Every ROI was then examined carefully (see *Section 3.1.1*) to make sure that it was not present in any of the blanks and it depicted a chromatographic peak.

Scheme 1 outlines the protocol developed and implemented in this study for the robust and systematic selection of valid ROIs from the positive dataset. The procedure used for negative dataset is reported in the Supplementary Materials.

2.5.3. Multivariate curve resolution-alternating least squares (MCR-ALS)

MCR-ALS [28] resolves elution profiles and spectra of distinct sample constituents by decomposing the data matrix (**D**) according to Eq. (1)

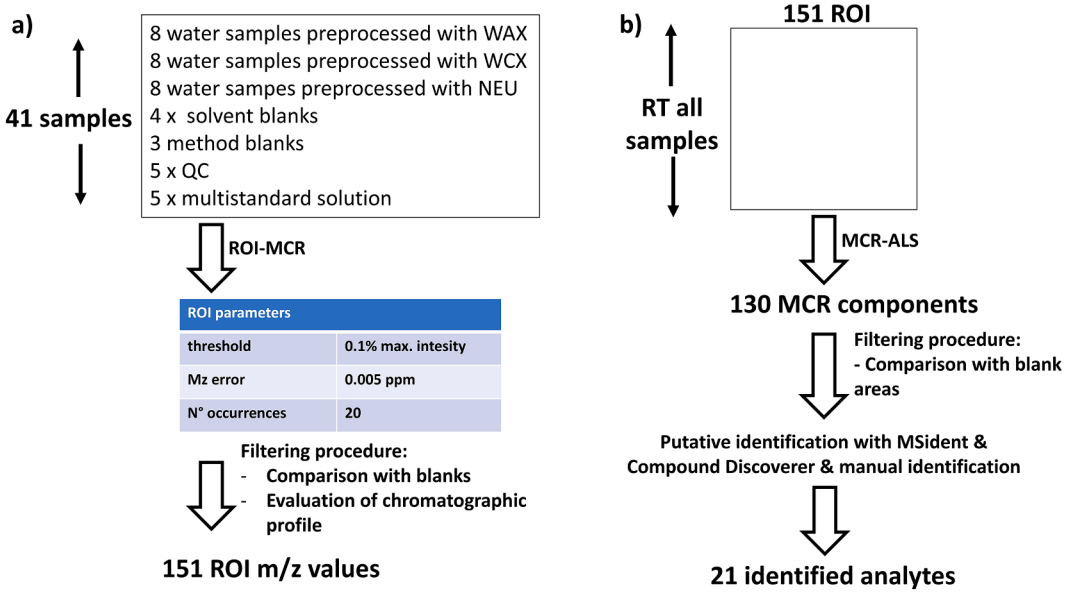
$$\mathbf{D} = \mathbf{C}\mathbf{S}^{T} + \mathbf{E} \quad (1)$$

where **D** is the final MSroi matrix obtained from the ROI analysis (as described in Section 2.4.2), and the two factor matrices, **C** and **S**, contain the elution profiles of the resolved components (single chemical constituents) for all chromatograms and their related pure mass spectra, respectively. Matrix **E** holds the residuals, i.e. the unmodelled part of **D**. To reduce the rotational ambiguity, non-negativity constraints on both **C** and **S** were applied. Initial pure mass spectra profile estimates (**S**) were derived by a SIMPLISMA-based method [31].

2.5.4. MSident

MSident is a recently developed software application introduced by Pérez et al. in 2024 [25] and designed to automate the identification of chemical compounds using MS spectra resolved through the ROI-MCR method. The input for MSident consists of two key variables obtained from the ROI-MCR analysis: the **S** matrix and the mzROI vector, which contains the relevant *m/z* values for the regions of interest. This identification process is carried out by directly comparing the spectra of ROI-MCR resolved components with reference spectra from existing spectral libraries. The accuracy and reliability of this spectral comparison are influenced by various factors, including experimental conditions and the presence or absence of MS2 data. When MS2 information is available, MSident first applies a precursor ion filter, taking into account the mass accuracy tolerance of the mass spectrometry instrument used during data acquisition. From the filtered pool of candidates, similarity and correlation scores are calculated to facilitate compound identification. Alternatively, putative identification is still possible using solely MS1 signals, as in the present study. In this case, three main parameters must be defined. First, an intensity threshold is established to filter out low-intensity signals. For normalized data (maximum intensity scaled to 1), a threshold of 0.1 was used to exclude signals below this cutoff. The remaining two parameters include the *m/z* tolerance in parts per million (ppm), which restricts matches to those within a specified error margin, and the intensity threshold for identifying potential precursor ions within the MS1 spectrum. In this case, an *m/z* tolerance of 5 ppm and a precursor ion intensity threshold of 0.9 were applied.

Once these parameters are configured, MSident initiates the identification routine by loading the selected spectral reference libraries (MS1 spectra). The libraries used in this study included MassBank, MassBank of North America, and MS-DIAL. To enhance the reliability of the tentative identifications, the identity of the compounds has been further confirmed through MS2 fragment ion matching. Indeed, an additional identification step was performed using the Compound Discoverer software [32]. This involved querying the results of Compound Discoverer by entering both the *m/z* values and retention times of each feature resolved by ROI-MCR, exploiting the results obtained from the MZcloud library ('full score match', MS2 spectra) to support putative identification [33,34] . The resulting identifications fall under Level 3



Scheme 1. ROI-MCR (a) and MCR-ALS (b) procedures for positive ionization dataset.

("Tentative candidate") according to the classification proposed by Schymanski et al. [35].

2.5.5. Compound discoverer

The same dataset was analysed by means of Compound Discoverer software [32], version 3.3 by Thermo Fisher Scientific in order to validate the ROI-MCR results. The software consists of several tuneable nodes. The details regarding the functioning of each node can be found in the literature [34]. The detailed workflow is reported in **Figure S3**, Supplementary Materials, while the parameters for each node are reported in Table SX, Supplementary Materials.

2.5.6. Software

Xcalibur 3.0 was used as LC-MS spectra acquisition software, while raw data were processed using Compound Discoverer, both from Thermo Fisher (Thermo Fisher Scientific, Waltham, MA, USA).

The ROI-MCR and MCR-ALS (2.0) software were downloaded from the official developers' website (https://mcrals.wordpress.com/theory/mcr-als/) and are implemented in the MATLAB environment (Mathworks, Natick, Massachusetts, USA). They have been used in their GUI (Graphical User Interface) version, which allows us to work with a user-friendly interface: command-line versions can also be downloaded for both [28,36].

3. Results

3.1. Untargeted analysis

3.1.1. Preprocessing of extracted features

In particular, the resulting raw LC—HRMS data were processed using the ROI-MCR procedure (Scheme 1) with parameter settings as described in *Section 2.4.2*. This initial processing step led to the selection of 231 ROIs. Each ROI was then subjected to a two-step filtering process to limit the presence of false positives. First, the presence or absence of the ROI in the blank samples was examined. Only those ROIs not detected in any of the blank runs were retained for further consideration. Second, the chromatographic elution profile corresponding to each selected ROI was visually inspected to assess its conformity to a Gaussian-like shape, which is indicative of a genuine chromatographic peak rather than background noise or artefact. ROIs exhibiting irregular or non-Gaussian elution profiles were excluded (**Figure S4A, S4B**, and **Figure S5A, S5B** Supplementary Materials). Finally, the presence of each ROI was verified in the QC samples to ensure signal consistency. Following this inspection, 151 ROIs were retained, representing 65 % of the initial ROIs, forming an MSroi matrix of 70,642 × 151 dimensions. MSroi matrix is a column-wise augmented matrix (Scheme 1), where the sample chromatograms, for each ROI, are placed one on top of each other, i.e. the columns hold the elution profiles at a given *m/z* channel for all samples, and the rows hold the ROIs signals at each chromatographic retention point. MCR-ALS has then been applied considering all the samples in the column-wise augmented data set. From this analysis, 130 components were extracted, based on explained variance, convergence criteria, and the minimization of lack-of-fit [21] (lof = 3.9917, R^2 = 99.84). After an inspection of elution and *m/z* profiles of each component resolved components, the peak areas were calculated for each component and for each sample by integrating the elution profiles in each **C** column (a single component). Thus a 41 × 130 components matrix was obtained.

The signal areas obtained from the MCR output were used as an additional level of verification. For each of the 130 components, the maximum signal intensity observed across all blank samples—referred to as the "max blank"—was calculated. Only those features exhibiting a signal area in the environmental samples at least five times greater than the corresponding value in the blanks were considered for putative identification. The 34 MCR components were searched in the Compound Discoverer output and were putatively identified using MSident and Compound Discoverer, which incorporate several spectral reference libraries, as explained in *Section 3.1.2*.

Furthermore, to ensure the robustness of the selected features, a final verification step was conducted using the raw LC—HRMS signal. For each candidate feature, putatively identified by MSident, the chromatographic profile was re-evaluated and were confirmed only those features that displayed a well-defined elution profile in the raw signal. The same procedures were followed for the extraction and the putative identification of negative mode datasets where 25 ROIs were retained (*Supplementary materials*). Finally, 21 and 3 features were putatively identified for positive and negative datasets, respectively.

3.1.2. Untargeted results

24 features in positive mode and three in negative were finally putatively identified based on their MS and MS/MS data. Their identifications and some assessment parameters are displayed in Table 3 and **Table S5a** and **Table S5b**. Chemical classes and assigned compound category information was obtained by searching in Food Database (FOODB), Human Metabolome Database (HMDB) or PubChem. The filtered MCR components were subject to careful manual annotation of the MS and MS/MS spectra and annotated based on match of their spectra to those in databases (e.g., mzCloud) or their diagnostic fragmentation mechanisms. Most compounds were annotated based on the match with MS/MS spectra reported in the mzCloud database in the same conditions as our experimental spectra. As an example, component 11 was annotated as 4-methylquinoline, a small heterocyclic aromatic organic compound that is commonly employed as a chemical intermediate for the synthesis of pharmaceuticals, agrochemicals, and pigments. The annotation of this compound was based on the diagnostic product ion at *m/z* 103.0551, that distinguishes it from its more common isomers and 2- and 6-methylquinoline. In some other cases, careful manual inspection of the MS/MS spectra allowed tentative identifications. Component 4 was annotated as N,N-Dibutylethanolamine, an irritating and corrosive industrial chemical, based on its neutral loss of 56 Da (a butyl group) that generate an intense ion at *m/z* 118.1239 and its other main diagnostic product ion at *m/z* 57.0708, that corresponds to the butyl cation. Ten more peaks were reported as unknowns in **Table S5b**. Despite not being thoroughly elucidated, several compounds were at least partially annotated. MCR components 9 and 77 were both listed as benzylamine derivatives based on their base peak at *m/z* 91.0550 that is typical of the stable benzyl carbocation. Moreover, component 9 shared with component 7 (annotated as dibenzylamine) an ion at *m/z* 181.1019, that corresponds to a dibenzyl cation, thus component 9 is likely a dibenzylamine derivative. On the other hand, components 15 and 74 were both butylamine derivatives, sharing with component 4 the neutral loss of 56 Da and the butyl carbocation at *m/z* 57.0708.

One of the most frequently individuated categories was drugs, mainly pharmaceuticals, comprising 7 analytes. The widespread presence of pharmaceuticals in surface waters has been well documented across Europe, often prompting investigations into their ecotoxicity, persistence, and removal efficiency [37,38]. The individuated pharmaceuticals include a variety of therapeutic classes such as antihypertensive (e.g., atenolol), beta-blockers (e.g., metoprolol), antidiabetic agents (e.g., metformin), and opioid analgesics (e.g., tramadol, tapentadol). In addition to pharmaceuticals, two cosmetic ingredients and surfactants (cetyl sulfate and diethanolamine) were identified. Notably, some of them can be found in a wide range of consumer products such as shampoos, soaps, cleaners, and cutting oils. Furthermore, the analysis also revealed the presence of compounds commonly used as food additives or natural metabolites (e.g., 4-methylquinoline, 1-stearoylglycerol). 4-methylquinoline has been detected but not quantified in brassicas. Finally, the screening identified eight compounds related to industries (N,N-dibutylethanolamine, dibenzylamine, methyldicyclohexylamine, N-benzyl-1-tetradecanamine, hydroxyoctadecenoic acid methyl ester, N,N'-diphenylguanidine, propyl p-toluenesulfonate and butyl p-toluenesulfonate). The presence and distribution of the

identified compounds across the eight investigated water samples are summarized in **Table S6, Supplementary materials**, where each green cell indicates the detection of a specific analyte in each sample. Information on the best performing SPE cartridge (in terms of maximum final area obtained, hence maximum extraction) was also included in the respective cell. Notably, two compounds were consistently detected across most of the samples: 4-methylquinoline, and 1-stearoylglycerol. Their widespread occurrence could suggest a high environmental persistence and/or extensive usage in anthropogenic activities. When comparing sample types, groundwater samples generally showed a higher frequency of detected analytes compared to surface waters. Samples from Villanova 1, and Camposanto exhibited the highest detection frequency. This could reflect local anthropogenic activities, local geological factors and hydrological factors.

Concerning geological factors, for example, layers rich in clay or organic matter can accumulate contaminants, reduce groundwater velocity, or divert underground water flows (RAO, 1998). As for hydrological factors, such as rainfall intensity and duration and the alternation of wet and dry periods, these have a strong short-term effect on surface water runoff, even more so than on groundwater. This leads to frequent changes in pollutant concentrations in surface waters, as the mixing ratios between polluted and unpolluted runoff water often change. This specific phenomenon has also been documented by other authors in various plains and rivers around the world [39,40], and it is the main reason for the differing concentrations of EOCs between surface water and groundwater in our area during the sampling period. The lower frequency of analytes in the sampled surface waters is due to dilution phenomena affecting surface waters, resulting from local meteorological events that occurred prior to sampling.

Regarding the groundwater sampled, the presence of analytes in the shallow well Villanova_1 is mainly due to the infiltration of water from the river into the adjacent aquifer, since the well is the closest to the Secchia River and located in an agricultural area. By contrast, the presence of analytes in the shallow Camposanto well is likely due to small, localized and scattered pollution phenomena, as well as to exchanges between surface waters and adjacent phreatic aquifers. Near this well there are small industries and a dense network of channels for stormwater drainage.

3.2. Targeted analysis

In this study, given the lack of information on PMOCs in the Emilia Romagna region, the targeted investigation was focused on the detection of twelve different PMOCs in surface and subsurface water samples using a method previously optimized through experimental design techniques [27]. The selected target compounds included pharmaceuticals (atenolol, clenbuterol, hydrochlorothiazide, furosemide, metformin, metoprolol, terbutaline, salbutamol), stimulants and lifestyle markers (caffeine, nicotine, taurine), and the herbicide 2,4-dichlorophenoxyacetic acid. These analytes were selected not to exhaustively represent the entire PMOC chemical space, but because they belong to highly polar and mobile contaminants that are still relatively underexplored in environmental monitoring studies, especially in reclaimed and surface waters, despite their high mobility and potential environmental relevance. Moreover, the targeted analysis was used to support and validate the untargeted workflow through the confirmation of selected compounds with reference standards and optimized MS/MS transitions. The concentrations of all detected analytes in the real samples were corrected for their respective process efficiencies (**Table S1**, Supplementary Materials) and reported in Table 2.

Most of the analytes were above the limit of detection (LOD) in the samples analysed, with the exception of 2,4-dichlorophenoxyacetic acid (pesticide), furosemide (pharmaceutical) and salbutamol (pharmaceutical), which were not present in either sample and were therefore not listed in the table.

In the subsurface water samples, eight analytes showed concentrations above the LOD, even if they were not detected in all samples. Conversely, in surface water samples, seven contaminants were present in all samples, while clenbuterol, hydrochlorothiazide, and terbutaline were individuated specifically in the Secchia River and the Naviglio Canal.

A pharmaceutical, metformin, commonly used for diabetes treatment [41], was also detected all samples. Other pharmaceuticals, such as atenolol (primarily used to prevent heart attacks and myocardial problems [42], metoprolol (used to treat hypertension by reducing heart rate [43]), and terbutaline (mainly used for bronchopulmonary issues [44]), were detected and quantified in some samples, particularly in river waters. It is important to note that some of the detected analytes with the triple quadrupole, namely caffeine, clenbuterol, hydrochlorothiazide, nicotine, taurine, terbutaline were not detected by the

Table 2
Heatmap of the analytes investigated in the subsurface and surface water samples. Concentrations are reported in ng L^{-1} and the reported standard deviations were calculated from triplicate analysis.

	Type of water	Atenolol	Caffeine	Clenbuterol	Hydrochlorothiazide	Metformin	Metoprolol	Nicotine	Taurine	Terbutaline
Villanova_1	Groundwater									
Villanova_2	Groundwater									
Villanova_3	Groundwater									
Botanical garden	Groundwater									
Camposanto	Groundwater					48±7	15,0±0,3			
Curiel lakes	surface									
Secchia river	surface	10±1								28±4
Naviglio canal	surface	40±4			19±2		40,7±0,8			

Legend. Grey: analyte with a concentration at least lower than LOD; yellow: analyte detected but below LOQ; green: analyte quantified (scales of green vary according to the concentration reported in the box). LOD and LOQ values are reported in Table S1, Supplementary Material.

Table 3
Chemical compounds identified in the investigated samples.

MCR component	Chemical compound	RT	Molecular Formula	Experimental *m/z*	Mass Error	Diagnostic Product Ions	C. L.
2	Octodrine	7.3	C8H19N	130.1592	1.3	130.1594; 74.0973; 57.0708	2
4	N,N-Dibutylethanolamine	6.9	C10H23NO	174.1854	0.9	174.1855; 118.1230; 57.0708	3
7	Dibenzylamine	5.0	C14H15N	198.1278	0.5	198.1286; 181.1019; 106.0659; 91.0551	2
10	Bis(4-ethylbenzylidene)sorbitol	3.2	C24H30O6	415.2122	0.5	181.0628; 133.0652; 119.0859	2
11	4-Methylquinoline	3.7	C10H9N	144.0809	0.6	144.0814; 103.0551	2
20	methyldicyclohexylamine	6.7	C13H25N	196.2064	3.9	196.2073; 114.1288; 83.0866	2
25	Metoprolol	19.5	C15H25NO3	268.1907	0.0	268.1925; 191.1079; 159.0816; 121.0658; 116.1080; 98.0975; 74.0613; 72.0820	2
27	Memantine	20.2	C12H21N	180.1746	0.4	180.1750; 163.1483; 107.0860	2
31	N-benzyl-1-tetradecanamine	3.9	C21H37N	304.2999	−0.6	304.3001; 212.2376; 91.0548; 58.0661	3
34	Tapentadol	16.5	C14H23NO	222.1860	3.4	222.1863; 135.0812; 121.0656; 107.0500	2
44	Hydroxyoctadecenoic acid methyl ester	3.2	C19H36O3	313.2742	1.4	313.2748; 257.2483; 109.1020; 95.0864; 85.1021; 71.0866; 57.0710	3
47	Tramadol	18.3	C16H25NO2	264.1967	0.9	264.1973; 246.1862; 201.1286; 133.0652; 107.0499; 58.0662	2
52	3-Hydroxycotinine	5.6	C10H12N2O2	193.0972	0.1	198.0980; 134.0607; 118.0658; 80.0504	2
60	N,N'-Diphenylguanidine	5.0	C13H13N3	212.1185	1.5	212.1180; 195.0915; 119.0605; 94.0655	2
61	Natalensin	18.8	C18H15N5	302.1390	0.8	302.1396; 270.1133; 226.0868; 211.0761; 181.0654	2
63	Diethanolamine	35.5	C4H11NO2	106.0867	0.3	106.0868; 88.0763; 70.0658	2
64	Atenolol	27.8	C14H22N2O3	267.1702	0.2	267.1722; 225.1257; 208.0982; 190.0876; 145.0695; 116.1080; 98.0975; 74.0613; 72.0820	2
72	1-Stearoylglycerol	3.8	C21H42O4	359.3163	2.2	341.3069; 109.1019; 95.0864; 85.1021; 71.0865; 57.0710	2
85	Atenolol acid	25.8	C14H21NO4	268.1542	0.5	268.1547; 226.1079; 191.0706; 145.0695; 116.1080; 98.0975; 74.0613; 72.0820	2
102	Valine	28.5	C5H11NO2	118.0865	2.2	118.0867; 72.0815	2
115	Metformin	26.6	C4H11N5	130.1086	0.4	130.1090; 88.0875; 87.0048; 84.0515; 60.0565	2
16 (ESI-)	Propyl p-toluenesulfonate	14.0	C10H14O3S	213.0584	−3.2	213.0582; 79.9555	3
19 (ESI-)	Butyl p-toluenesulfonate	13.7	C11H16O3S	227.0742	−2.0	227.0753; 79.9554	3
46 (ESI-)	Cetyl sulfate	6.5	C16H34O4S	322.2172	−2.1	321.2101; 96.9687	3

RT: retention time (min).
C.L.: confidence limit according to Schymanski et al., 2014.

developed approach (Section 3.1.2). This could be likely due to the different sensitivity of the two instruments [45,46]. In fact, the mass spectrometry conditions for HRMS were not optimized on specific analytes, as it was performed for the targeted analysis [27].

4. Conclusion

In this study, both targeted and untargeted analytical strategies were employed to investigate the presence of emerging contaminants in a range of surface and subsurface water samples collected across the province of Modena. Among the detected compounds through targeted approach, caffeine, nicotine, and taurine emerged as some of the most frequently found compounds (targeted results), which aligns with existing literature. Conversely, the broader set of untargeted findings revealed compounds originating from diverse sources, including pharmaceuticals, personal care products, industrial chemicals, pesticides, and food additives, underscoring the multiplicity of contamination sources impacting both surface and groundwater in the study area. However, it is important to note that the implementation of the ROI-MCR approach in the untargeted strategy required strict control procedures to ensure the reliability of the selected features. Careful evaluation steps were incorporated, including comparison with reagent and process blanks, analysis of elution profiles, and signal-to-blank intensity ratios. These quality control steps were essential to minimize false positives and discriminate genuine features from artifacts or background signals arising from the chromatographic system, SPE pretreatment, or baseline drift. Although this study did not aim to quantify contaminant concentrations or assess toxicological risks, the presence of several emerging contaminants in the water bodies of Modena province emphasizes the need for further research. Considering, the limited number of geographical samples and the absence of temporal replicates, it is acknowledged that it remains difficult to determine whether the detected concentration are occasional or recurrent. Nevertheless, the occurrence of pharmaceutical observed in this study is consistent with the findings reported in previous works conducted in Italian rivers basins [14,15]. Future investigations involving a larger number of samples and extended temporal coverage will be essential to validate these preliminary findings and to better assess the contribution of diverse anthropogenic sources to EOC contamination in the Modena district. Regarding to the hydrogeological aspect, this study confirms the occurrence of emerging contaminants in coarse-grained unconfined aquifers recharged by meteoric water or local surface watercourses. Furthermore, it demonstrates that the spatial distribution and transport of emerging contaminants in the subsurface are strongly influenced by the geometric configuration of subsurface geological layers. This complexity poses significant challenges in accurately delineating contaminant distribution in geologically heterogeneous settings, such as alluvial plains. The analysis results show that the river samples exhibit a lower concentration compared to the groundwater samples. In line with other studies on the spread of emerging contaminants in surface water bodies, the concentration values detected in rivers are due to dilution effects associated with autumn/winter rainfall pattern, rather than to changes in the contaminant source. Current evidence indicates that, within the Po River alluvial plain, emerging contaminants are frequently detected in the shallowest aquifers, particularly within the first 30 m below the ground surface. However, further investigations are necessary to assess the occurrence and migration of these compounds at greater depths. Future research should focus on the potential sorption of emerging contaminants by fine-grained geological deposits, as well as on the degradation processes these compounds may undergo during their transition from surface water bodies to groundwater systems. About this last issue, special attention should be given to the riparian zone, where both biotic and abiotic processes could significantly influence contaminant transformation.

CRediT authorship contribution statement

Samuele Pellacani: Writing – original draft, Visualization, Software, Methodology, Formal analysis, Data curation. **Francesco Ronchetti:** Writing – original draft, Resources, Funding acquisition, Conceptualization. **Andrea Cerrato:** Writing – original draft, Visualization, Software, Formal analysis, Data curation. **Anna Laura Capriotti:** Visualization, Supervision, Software, Formal analysis, Data curation. **Marina Cocchi:** Visualization, Supervision, Methodology, Conceptualization. **Luísa Collischonn:** Writing – original draft, Visualization, Formal analysis, Data curation. **Caterina Durante:** Writing – original draft, Supervision, Resources, Project administration, Funding acquisition, Conceptualization. **Guido Perra:** Visualization, Software, Formal analysis, Data curation. **Lorenzo Strani:** Validation, Software, Investigation, Formal analysis, Data curation.

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.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.chroma.2026.467150.

Data availability

Data will be made available on request.

References

[1] M. Stuart, D. Lapworth, E. Crane, A. Hart, Review of risk from potential emerging contaminants in UK groundwater, Sci. Total. Environ. 416 (2012) 1–21, https://doi.org/10.1016/j.scitotenv.2011.11.072.

[2] M.E. Stuart, D.J. Lapworth, Transformation products of emerging organic compounds as future groundwater and drinking water contaminants. Transformation Products of Emerging Contaminants in the Environment, John Wiley & Sons, Ltd, 2014, pp. 65–86, https://doi.org/10.1002/9781118339558.ch02.

[3] S. Khan, M.U. Naushad, M. Govarthanan, J. Iqbal, S.M. Alfadul, Emerging contaminants of high concern for the environment: current trends and future research, Environ. Res. 207 (2022) 112609, https://doi.org/10.1016/j.envres.2021.112609.

[4] V. Pérez-Fernández, L. Mainero Rocca, P. Tomai, S. Fanali, A. Gentili, Recent advancements and future trends in environmental analysis: sample preparation, liquid chromatography and mass spectrometry, Anal. Chim. Acta 983 (2017) 9–41, https://doi.org/10.1016/j.aca.2017.06.029.

[5] S.Y. Bunting, D.J. Lapworth, E.J. Crane, J. Grima-Olmedo, A. Koroša, A. Kuczyńska, N. Mali, L. Rosenqvist, M.E. Van Vliet, A. Togola, B. Lopez, Emerging organic compounds in European groundwater, Environ. Pollut. 269 (2021) 115945, https://doi.org/10.1016/j.envpol.2020.115945.

[6] N. Morin-Crini, E. Lichtfouse, G. Liu, V. Balaram, A.R.L. Ribeiro, Z. Lu, F. Stock, E. Carmona, M.R. Teixeira, L.A. Picos-Corrales, J.C. Moreno-Piraján, L. Giraldo, C. Li, A. Pandey, D. Hocquet, G. Torri, G. Crini, Worldwide cases of water pollution by emerging contaminants: a review, Environ. Chem. Lett. 20 (2022) 2311–2338, https://doi.org/10.1007/s10311-022-01447-4.

[7] T. Reemtsma, U. Berger, H.P.H. Arp, H. Gallard, T.P. Knepper, M. Neumann, J. B. Quintana, P. de Voogt, Mind the gap: persistent and mobile organic compounds—Water contaminants that slip through, Environ. Sci. Technol. 50 (2016) 10308–10315, https://doi.org/10.1021/acs.est.6b03338.

[8] Directive (EU) 2026/805 of the European Parliament and of the Council of 30 March 2026 amending Directive 2000/60/EC establishing a framework for Community action in the field of water policy, Directive 2006/118/EC on the protection of groundwater against pollution and deterioration and Directive 2008/105/EC on environmental quality standards in the field of water policy (Text with EEA relevance), (n.d.).

[9] O.A. Abafe, M.A. Lawal, T.B. Chokwe, Non-targeted screening of emerging contaminants in South African surface and wastewater, Emerg. Contam. 9 (2023) 100246, https://doi.org/10.1016/j.emcon.2023.100246.

[10] K. Liu, J. Qiu, C.-H. Weng, J. Tao, Z. Tang, J. Zeng, N. Liu, X. Wang, Target and non-target screening for the existence, ecological risk and source analysis of contaminants of emerging concern (CECs) in groundwater, Environ. Chem. Ecotoxicol. 7 (2025) 1315–1325, https://doi.org/10.1016/j.enceco.2025.06.013.

[11] D. Xia, L. Liu, B. Zhao, D. Xie, G. Lu, R. Wang, Application of nontarget high-resolution mass spectrometry fingerprints for qualitative and quantitative source apportionment: a real case study, Environ. Sci. Technol. 58 (2024) 727–738, https://doi.org/10.1021/acs.est.3c06688.

[12] J.-H. Zhao, L.-X. Hu, S. Xiao, J.-L. Zhao, Y.-S. Liu, B. Yang, Q.-Q. Zhang, G.-G. Ying, Screening and prioritization of organic chemicals in a large river basin by suspect and non-target analysis, Environ. Pollut. 333 (2023) 122098, https://doi.org/10.1016/j.envpol.2023.122098.

[13] E. Pignotti, E. Dinelli, Distribution and partition of endocrine disrupting compounds in water and sediment: case study of the Romagna area (North Italy), J. Geochem. Explor. 195 (2018) 66–77, https://doi.org/10.1016/j.gexplo.2018.02.008.

[14] C.M. Montone, B. Giannelli Moneta, S.E. Aita, F. Aulenta, C. Cavaliere, A. Cerrato, S. Fazi, A. Laganà, V. Paolini, F. Petracchini, S. Piovesana, A.L. Capriotti, Untargeted analysis of contaminants in river water samples: comparison between two different sorbents for solid-phase extraction followed by liquid chromatography-high-resolution mass spectrometry determination, Microchem. J. 172 (2022) 106979, https://doi.org/10.1016/j.microc.2021.106979.

[15] R. Meffe, I. De Bustamante, Emerging organic contaminants in surface water and groundwater: a first overview of the situation in Italy, Sci. Total. Environ. 481 (2014) 280–295, https://doi.org/10.1016/j.scitotenv.2014.02.053.

[16] C.D. Pérez-Blanco, G. Standardi, J. Mysiak, R. Parrado, C. Gutiérrez-Martín, Incremental water charging in agriculture. A case study of the Regione Emilia Romagna in Italy, Environ. Model. Softw. 78 (2016) 202–215, https://doi.org/10.1016/j.envsoft.2015.12.016.

[17] Y. Han, L.-X. Hu, T. Liu, J. Liu, Y.-Q. Wang, J.-H. Zhao, Y.-S. Liu, J.-L. Zhao, G.-G. Ying, Non-target, suspect and target screening of chemicals of emerging concern in landfill leachates and groundwater in Guangzhou, South China, Sci. Total. Environ. 837 (2022) 155705, https://doi.org/10.1016/j.scitotenv.2022.155705.

[18] D. Aurich, P. Diderich, R. Helmus, E.L. Schymanski, Non-target screening of surface water samples to identify exposome-related pollutants: a case study from Luxembourg, Environ. Sci. Eur. 35 (2023) 94, https://doi.org/10.1186/s12302-023-00805-5.

[19] Y. Jin, B. Ye, W.-M. Wang, M.-F. Cen, W.-L. Wang, Q.-Y. Wu, Non-target screening of toxic contaminants in secondary effluents (from several sources) using SPE-LC-Orbitrap mass spectrometry, J. Environ. Sci. (2025), https://doi.org/10.1016/j.jes.2025.12.044.

[20] A. Kruve, K. Kiefer, J. Hollender, Benchmarking of the quantification approaches for the non-targeted screening of micropollutants and their transformation products in groundwater, Anal. Bioanal. Chem. 413 (2021) 1549–1559, https://doi.org/10.1007/s00216-020-03109-2.

[21] E. Gorrochategui, J. Jaumot, R. Tauler, ROIMCR: a powerful analysis strategy for LC-MS metabolomic datasets, BMC Bioinform. 20 (2019), https://doi.org/10.1186/s12859-019-2848-8.

[22] B. Benedetti, C. Perez-Lopez, H. MacKeown, E. Magi, R. Tauler, Improving regions of interest multivariate curve resolution: development of an empirical metric system through the study of passive sampling extracts of wastewater in Antarctica, Anal. Chem. 97 (2025) 13110–13119, https://doi.org/10.1021/acs.analchem.5c00777.

[23] S. Pellacani, G. Tomasi, G. Renner, S. Tisler, Comparative assessment of feature and component profiling workflow for supercritical fluid chromatography – high resolution mass spectrometry analysis, J. Chromatogr. A 1769 (2026) 466730, https://doi.org/10.1016/j.chroma.2026.466730.

[24] C. Pérez-López, L.H.M.L.M. Santos, A. Ginebreda, L.L. Alonso, G. Buttiglieri, E. Statiris, D. Barcelo, S. Rodriguez-Mozaz, R. Tauler, Changes in the rhizosphere metabolome profile of wastewater-irrigated lettuces (Lactuca sativa L.) highlighted using non-target HRMS and ROIMCR chemometrics method, Green Anal. Chem. 12 (2025) 100203, https://doi.org/10.1016/j.greeac.2024.100203.

[25] C. Perez-Lopez, A. Ginebreda, J. Jaumot, F.Y. Yamamoto, D. Barcelo, R. Tauler, MSident: straightforward identification of chemical compounds from MS-resolved spectra, Chemom. Intell. Lab. Syst. 245 (2024) 105063, https://doi.org/10.1016/j.chemolab.2024.105063.

[26] Y.-Q. Wang, L.-X. Hu, J.-H. Zhao, Y. Han, Y.-S. Liu, J.-L. Zhao, B. Yang, G.-G. Ying, Suspect, non-target and target screening of pharmaceuticals and personal care products (PPCPs) in a drinking water system, Sci. Total. Environ. 808 (2022) 151866, https://doi.org/10.1016/j.scitotenv.2021.151866.

[27] L. Strani, B. Benedetti, M. Cocchi, C. Durante, G. Perra, M. Pietropaolo, S. Pellacani, D. Tanzilli, Optimization of an analytical method based on the use of zwitterionic- phosphorylcholine -HILIC column for the determination of multiple polar emerging contaminants in reclaimed water, J. Chromatogr. A 1741 (2025) 465605, https://doi.org/10.1016/j.chroma.2024.465605.

[28] J. Jaumot, A. de Juan, R. Tauler, MCR-ALS GUI 2.0: new features and applications, Chemom. Intell. Lab. Syst. 140 (2015) 1–12, https://doi.org/10.1016/j.chemolab.2014.10.003.

[29] R. Tauler, E. Gorrochategui, J. Jaumot, R. Tauler, A protocol for LC-MS metabolomic data processing using chemometric tools, Protoc. Exch. (2015), https://doi.org/10.1038/protex.2015.102.

[30] R.L. Khatoonabadi, M. Vosough, L.L. Hohrenk, T.C. Schmidt, Employing complementary multivariate methods for a designed nontarget LC-HRMS screening of a wastewater-influenced river, Microchem. J. 160 (2021) 105641, https://doi.org/10.1016/j.microc.2020.105641.

[31] W. Windig, C.E. Heckler, F.A. Agblevor, R.J. Evans, Self-modeling mixture analysis of categorized pyrolysis mass spectral data with the SIMPLISMA approach, Chemom. Intell. Lab. Syst. 14 (1992) 195–207, https://doi.org/10.1016/0169-7439(92)80104-C.

[32] Compound Discoverer software - IT, (n.d.). https://www.thermofisher.com/uk/en/home/industrial/mass-spectrometry/liquid-chromatography-mass-spectrometry-lc-ms/lc-ms-software/multi-omics-data-analysis/compound-discoverer-software.html (Accessed 8 August 2024).
[33] J. Li, M. Zhou, Z. Chen, J. Guo, F. Fang, A. Schäffer, Y. Shao, Identification of pollutant markers in rural mountainous areas of China by combining non-targeted analysis with zebrafish embryo toxicity tests, Sci. Total. Environ. 965 (2025) 178625, https://doi.org/10.1016/j.scitotenv.2025.178625.
[34] S. Pellacani, C. Citti, L. Strani, B. Benedetti, P.P. Becchi, V. Pizzamiglio, S. Michelini, G. Cannazza, A. De Juan, M. Cocchi, C. Durante, Comparative analysis of features extraction protocols for LC-HRMS untargeted metabolomics in mountain cheese 'identitation,', Microchem. J. 207 (2024) 111863 https://doi.org/10.1016/j.microc.2024.111863.
[35] E.L. Schymanski, J. Jeon, R. Gulde, K. Fenner, M. Ruff, H.P. Singer, J. Hollender, Identifying small molecules via high resolution mass spectrometry: communicating confidence, Environ. Sci. Technol. 48 (2014) 2097–2098, https://doi.org/10.1021/es5002105.
[36] M. Pérez-Cova, C. Bedia, D.R. Stoll, R. Tauler, J. Jaumot, MSroi: a pre-processing tool for mass spectrometry-based studies, Chemom. Intell. Lab. Syst. 215 (2021) 104333, https://doi.org/10.1016/j.chemolab.2021.104333.
[37] J. Wang, L. Chu, Irradiation treatment of pharmaceutical and personal care products (PPCPs) in water and wastewater: an overview, radiation, Phys. Chem. 125 (2016) 56–64, https://doi.org/10.1016/j.radphyschem.2016.03.012.
[38] Y. Yang, Y.S. Ok, K.-H. Kim, E.E. Kwon, Y.F. Tsang, Occurrences and removal of pharmaceuticals and personal care products (PPCPs) in drinking water and water/sewage treatment plants: a review, Sci. Total. Environ. 596-597 (2017) 303–320, https://doi.org/10.1016/j.scitotenv.2017.04.102.
[39] F. Labad, A. Ginebreda, R. Criollo, E. Vázquez-Suñé, S. Pérez, A. Jurado, Occurrence, data-based modelling, and risk assessment of emerging contaminants in an alluvial aquifer polluted by river recharge, Environ. Pollut. 316 (2023) 120504, https://doi.org/10.1016/j.envpol.2022.120504.
[40] C.A. Schwanen, P.M. Kronsbein, B. Balik, J. Schwarzbauer, Dynamic transport and distribution of organic pollutants in water and sediments of the Rur River, Water. Air. Soil. Pollut. 235 (2024) 9, https://doi.org/10.1007/s11270-023-06786-8.
[41] G. Rena, D.G. Hardie, E.R. Pearson, The mechanisms of action of metformin, Diabetologia 60 (2017) 1577–1585, https://doi.org/10.1007/s00125-017-4342-z.
[42] A.N. Wadworth, D. Murdoch, R.N. Brogden, Atenolol. A reappraisal of its pharmacological properties and therapeutic use in cardiovascular disorders, Drugs 42 (1991) 468–510, https://doi.org/10.2165/00003495-199142030-00007.
[43] C.-G. Regårdh, G. Johnsson, Clinical pharmacokinetics of metoprolol, Clin.-Pharmacokinet. 5 (1980) 557–569, https://doi.org/10.2165/00003088-198005060-00004.
[44] G.H. Jones, S.J. Scott, Continuous infusions of terbutaline in asthma - a review, J. Asthma 48 (2011) 753–756, https://doi.org/10.3109/02770903.2011.613509.
[45] J. Aceña, S. Stampachiacchiere, S. Pérez, D. Barceló, Advances in liquid chromatography–high-resolution mass spectrometry for quantitative and qualitative environmental analysis, Anal. Bioanal. Chem. 407 (2015) 6289–6299, https://doi.org/10.1007/s00216-015-8852-6.
[46] W. Chen, H. Huang, C.-E. Chen, S. Qi, O.R. Price, H. Zhang, K.C. Jones, A. J. Sweetman, Simultaneous determination of 20 trace organic chemicals in waters by solid-phase extraction (SPE) with triple-quadrupole mass spectrometer (QqQ-MS) and hybrid quadrupole orbitrap high resolution MS (Q-Orbitrap-HRMS), Chemosphere 163 (2016) 99–107, https://doi.org/10.1016/j.chemosphere.2016.07.080.