Chemosphere 349 (2024) 140863

Contents lists available at ScienceDirect

Chemosphere

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

# Unraveling the chemical fingerprint of the Miami River sources: Insights from high-resolution mass spectrometry and nontarget analysis

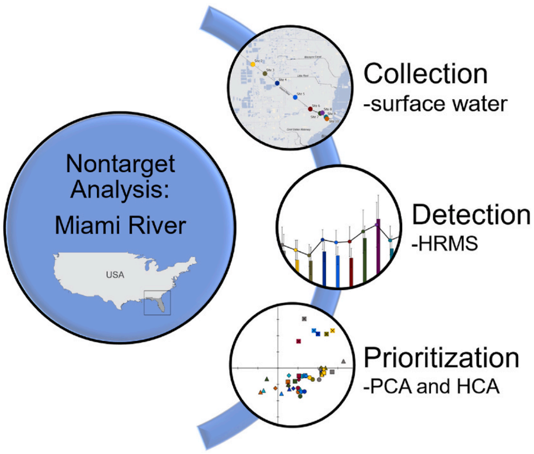
Kassidy Troxell[a,b,c], Milena Ceccopieri[a,b,c], Piero Gardinali[a,b,c,*]

[a] Institute of Environment, Florida International University, 11200 SW 8th Street, Modesto A. Maidique Campus, Miami, FL 33199, USA
[b] Department of Chemistry and Biochemistry, Florida International University, 300 NE 151th Street, Marine Science Building 230/232, Biscayne Bay Campus, North Miami, FL 33181, USA
[c] Environmental Analysis Research Laboratory, Florida International University, 300 NE 151th Street, Marine Science Building 232, Biscayne Bay Campus, North Miami, FL 33181, USA

## HIGHLIGHTS

- Investigated water sources and anthropogenic impact in a highly urbanized river using Nontarget Analysis (NTA).
- Achieved 71.2% average reproducibility for NTA and post-processed compounds detected in 3 out of 5 sampling events.
- Observed a rise in organic contaminants' presence and intensity along the river.
- Consistently differentiated water sources with NTA.
- Prioritized and suggested specific compounds as source tracers.

## GRAPHICAL ABSTRACT



## ARTICLE INFO

Handling Editor: Derek Muir

*Keywords:*
Non-target analysis
High resolution mass spectrometry
Organic contaminants
Sourcing
Tracing

## ABSTRACT

This study investigated the chemical composition of waters in the heavily urbanized and historically polluted Miami River, Florida. The goal was to assess the different water sources and anthropogenic influence in this managed area using nontarget analysis (NTA) combined with high-resolution mass spectrometry (HRMS). Surface water samples were collected from 10 sites during five sampling events in the wet season of 2021 and 2022. The HRMS data was processed using Compound Discoverer™ version 3.3, and the results were analyzed using techniques including Principal Component Analysis and Kendrick Mass Defect plots. Results showed a gradient change in the chemical composition from the Everglades to Biscayne Bay endmembers. The Seybold Canal, an adjacent canal, was consistently identified as a unique source of contaminants, contributing its own specific set of tracers. Increased presence and intensity of organic contaminants along the waterway was observed, indicating significant anthropogenic influence in the area. The NTA and post-processing were evaluated for reproducibility,

---

* Corresponding author. Institute of Environment, Florida International University, 11200 SW 8th Street, Modesto A. Maidique Campus, Miami, FL 33199, USA.
*E-mail address:* gardinal@fiu.edu (P. Gardinali).

https://doi.org/10.1016/j.chemosphere.2023.140863
Received 24 July 2023; Received in revised form 23 October 2023; Accepted 28 November 2023

demonstrating robustness with a 71.2% average reproducibility for compounds detected in 3 out of 5 sampling trips. A detection frequency of 80% (4 out of 5) was the set criterion for detected compounds suggested as tracers. To prioritize samples, hierarchical cluster analysis was employed, and potential tracers for each water source were determined. Tracers included natural products and agricultural herbicides and insecticides for the Everglades, anthropogenic contaminants for the Seybold Canal, and a lack of unique tracers for Biscayne Bay. Additionally, urban-influenced contaminants such as flame retardants, insect repellents, pharmaceuticals, and non-agricultural herbicides were identified along the river. This study highlights the impact of human activities on the Miami River and demonstrates the effectiveness of NTA in differentiating and tracking water sources. The results emphasize the importance of reproducibility in NTA and provide guidance on implementing monitoring strategies by prioritizing samples based on chemical compositions.

## 1. Introduction

The Miami River of Miami-Dade County, Florida was once a completely natural system connecting the Florida Everglades with Biscayne Bay but has been significantly altered throughout time by urbanization. A steady increase of pollution has been a persistent concern for the river as it is a major tributary of Biscayne Bay (Cantillo et al., 2000). Biscayne Bay, also located in Miami-Dade County, Florida, is a large (~700 km$^2$), shallow (~2-m avg. Water depth), subtropical, lagoonal estuary that supports an exquisite and diverse ecosystem with a wide range of plant and animal species (Carey et al., 2011a). The bay is a crucial habitat for numerous endangered species, such as manatees and sea turtles (Lee and Bwenge, 2008). Additionally, Biscayne Bay is a major economic driver for the region supporting a thriving commercial and recreational fishing industry, as well as tourism activities such as boating, kayaking, and recreational fishing. Given its ecological, economic, and environmental importance, it is essential to maintain the health of Biscayne Bay through sustainable practices and responsible management to ensure that the bay continues to provide important benefits for both the environment and local communities.

Efforts to restore and protect the health of Biscayne Bay have shifted to the impaired watersheds that discharge into various bay sections. With the large number of tourists and the already densely populated city, anthropogenic influence on the urban water system is evident. The urban water system contains all the water in and around a major city that is exposed to multiple pathways of anthropogenic influence and pollution from industrial discharges, mobile sources (e.g., cars/trucks), residential and commercial wastewater, trash, and polluted stormwater runoff (Bertrand-Krajewski et al., 2000; Hawkins et al., 2023; Schwarzenbach et al., 2006). Multiple studies have shown elevated concentrations of a variety of organic contaminants and nutrients in surface waters from tributaries of Biscayne Bay including the Miami River (Cui et al., 2022; Millette et al., 2019; Ng et al., 2021; R. Gardinali et al., 2002; Singh et al., 2010). There are many ways to assess the composition of an urban water system such as target analysis and suspect screening, but this leads to gaps in the chemical space and scope of what compounds may or may not be identified (Guardian et al., 2021; Hernández et al., 2019; Joerss et al., 2022; Madikizela et al., 2022). Nontarget analysis (NTA) was chosen to assess the organic contaminants realm of water quality for the Miami River as a holistic approach.

NTA is an analytical technique that uses chemometrics to discover previously undetected, unexpected and/or unknown compounds or substances in a variety of matrices. NTA has a wide variety of environmental applications such as being a tool for assessing anthropogenic and natural inputs throughout the urban water system (Lara-Martín et al., 2020; Ng et al., 2022; Troxell et al., 2022). Unlike target analysis, where only a limited number of specific pollutants are analyzed, NTA allows for the simultaneous detection and identification of a wide range of pollutants, including those that may not be well understood or regulated, such as specific emerging contaminants of concern (Pourchet et al., 2020). In environmental monitoring programs, NTA is often used to complement traditional target analysis methods by providing a more comprehensive picture of the contaminants present in the aquatic environment (da Costa Filho et al., 2022).

NTA is a complex process that involves various steps such as sample preparation, analysis, data acquisition, and post-processing that each can influence the types of chemicals that are detected (Manz et al., 2023). Tools have been developed to identify the chemical space that can be extracted and detected by NTA methods (Black et al., 2023). Due to the time-consuming nature of the process of once compounds are identified, smart prioritization strategies are also needed for samples and identifications to ensure accuracy and reproducibility of the results (Chiaia-Hernández et al., 2017; Menger et al., 2020). There have been guidelines developed in efforts to establish a consensus on NTA-related terms and concepts, as well as consistency in reporting practices (Hollender et al., 2017; Peter et al., 2021; Place et al., 2021). Overall, the use of NTA in environmental surface water samples when implemented appropriately and consistently has been proven a valuable tool for understanding the complex mixture of pollutants present in a watershed to unveil what alterations occur from existing in proximity to anthropogenic sources.

The Miami River, one of over 40 interconnected canal systems that discharge into Biscayne Bay, is a good example of an impaired watershed (RER, 2022) that could be explored by the application of NTA to aid in the evaluation of the composition and fate of pollutants from varying sources reaching Biscayne Bay. The objectives of this study were to characterize the chemical fingerprint of the endmembers and evaluate how the chemical space changes along the urban water system with a standardized NTA and processing workflow. This study utilizes the chemical fingerprint from the NTA and statistical techniques to differentiate sources by probable and tentative identified compounds and their relative intensity. Sources were to be differentiated by applying principal component analysis (PCA) and hierarchical cluster analysis (HCA) to the NTA post-processing results. Probable identified compounds were then prioritized based on their frequency of detection and source specific compounds were suggested as tracers or indicators of individual water sources. By understanding what water sources influence the Miami River and where, guided management efforts can be taken to protect Biscayne Bay and maintain the health of this vital ecosystem that supports the economic, recreational, and ecological needs of Miami-Dade County and beyond.

## 2. Materials and methods

### 2.1. Study area descriptions

Miami River, one of the oldest landmarks in Southeast Florida, started out as a tidal channel and over thousands of years converted into a heavily managed freshwater stream that carries fresh water from the Florida Everglades to its saltwater endmember Biscayne Bay. Over time, the river was canalized and expanded to accommodate shipping, resulting in a 5.5-mile long and 2819 square mile wide waterway that discharges about 14.9 m$^3$/s into Biscayne Bay. With the rise of urban development and industrialization, pollution has become a major concern for the river, which is particularly vulnerable to contaminants due to its proximity to urban surroundings. The Miami River faces a variety of pollution sources such as medical and human wastewater, oil spills, airport runoff, and trash disposal (Cantillo et al., 2000). As the largest contributing tributary to Biscayne Bay watershed, it is imperative to understand the dynamic mixing of sources throughout the Miami River and prioritize areas of concern to help improve the overall bay health by directing restoration efforts to where they will have the largest impacts.

Five sampling events were conducted during the wet seasons of 2021 (July and August) and 2022 (August, September, and October) in the Miami River in Southeastern Florida. The wet season was prioritized based on the complex mixing of water sources that occurs in this period between the surface water, stormwater runoff, and groundwater. A total of 10 sampling sites were monitored to represent surface water derived by the two endmembers of the river, Florida Everglades (Site 1) and Biscayne Bay (Site 10), and encompassed sample sites along the interconnected system to assess the extent of urban influence throughout (Fig. 1). An extreme event condition (Trip 4) was assessed in the study during Hurricane Ian Rainfall rates between 6.00 and 8.00 inches were observed over a three-day period in September 2022 (NOAA, 2022). All other trips were collected during days that there was no active precipitation, thus normal behavior of the managed watershed was explored.

### 2.2. *Sample collection and handling*

All samples were collected about 30 cm below the surface level to avoid surface microlayers by a grab sample technique using a sampling pole equipped for holding 500 mL single-use PET sample bottles. The bottles were conditioned by rinsing 3 times with water from the site and preserved on ice until arrival at the laboratory. Samples not processed immediately had a holding time of 30 days and were kept frozen at −20 °C in the dark until time of analysis. On the day of analysis, samples were brought to room temperature, a sub-sample was placed in 50 mL PTFE tubes, filtered through a 0.45 μm membrane with a detachable plunger, and capped for storage until analysis. At each site the basic water quality parameters pH, temperature, conductivity (salinity), dissolved oxygen (DO), turbidity, chlorophyll-a (Chl-a), and fluorescent dissolved organic matter (FDOM) were measured using a YSI EXO2 probe and handheld to understand the tidal influence on the specific section of the river.

### 2.3. *Chemicals and reagents*

All solvents were Optima Liquid Chromatography Mass Spectrometry (LCMS) grade and were purchased from Fisher Scientific including the acetonitrile (ACN), formic acid (FA), methanol (MeOH), and water ($H_2O$). All standards were purchased warranting high purity. The standards used in this study solely for QC purposes were as follows: atrazine (>97.1% purity, Sigma-Aldrich), caffeine (>98.5% purity, Sigma-Aldrich), carbamazepine (>99%, Sigma-Aldrich), (+)-cis-diltiazem hydrochloride (>99%, Sigma-Aldrich), clotrimazole (>98%, Sigma-Aldrich), diclofenac sodium salt (>99%, MP Biomedicals), diphenhydramine hydrochloride (>98%, Sigma-Aldrich), fluoxetine hydrochloride (100%, Sigma-Aldrich), gemfibrozil (>99%, Sigma-Aldrich), hydrochlorothiazide (98.4%, MP Biomedicals), lincomycin (>90%, Sigma-Aldrich), mefenamic acid (>98%, Sigma-Aldrich), norcocaine (>99%, Cerilliant), sertraline hydrochloride (>99%, Sigma-Aldrich), sucralose (99%, AK Scientific), sulfamethoxazole (>99%, Sigma-Aldrich), and trimethoprim (>98%, Sigma-Aldrich). These compounds were used to create a set of QA/QC mixtures that cover a wide range of Log $K_{ow}$ (−1.00–6.26) aimed to implement batch QC checks (intensity thresholds, chromatographic resolution, and mass accuracy) into the analysis of environmental samples to create a quality assurance framework (Ng et al., 2020).

### 2.4. *Nontargeted high resolution mass spectrometry analysis*

The NTA Study Reporting Tool was used in the preparation of this manuscript (Phillips et al., 2023). The method used in this present study was implemented from the published method by Ng et al. (2020). NTA was performed by automated online-solid phase extraction (SPE) coupled to high pressure liquid chromatography (HPLC) on a Q-Exactive Orbitrap (Thermo Scientific) with a mass accuracy of fewer than 5 parts per million (ppm). The SPE column used was a Hypersil GOLD aQ (20 × 2.1 mm, 12 μm, Thermo Scientific) and the analytical separation column used was a Hypersil GOLD aQ C18 polar endcapped (100 × 2.1, 1.9 μm, Thermo Scientific). Samples and optima LCMS water blanks (positive and negative) were run in full scan MS1 positive and negative mode with a scan range from 100.0 to 800.0 *m/z* at a resolution of 140,000 followed by data-dependent MS2 (dd-MS/MS) with a normalized collision energy of 30 at a resolution of 35,000 (4 injections per sample). Additional details on instrument parameters and HPLC can be found in the supporting information.

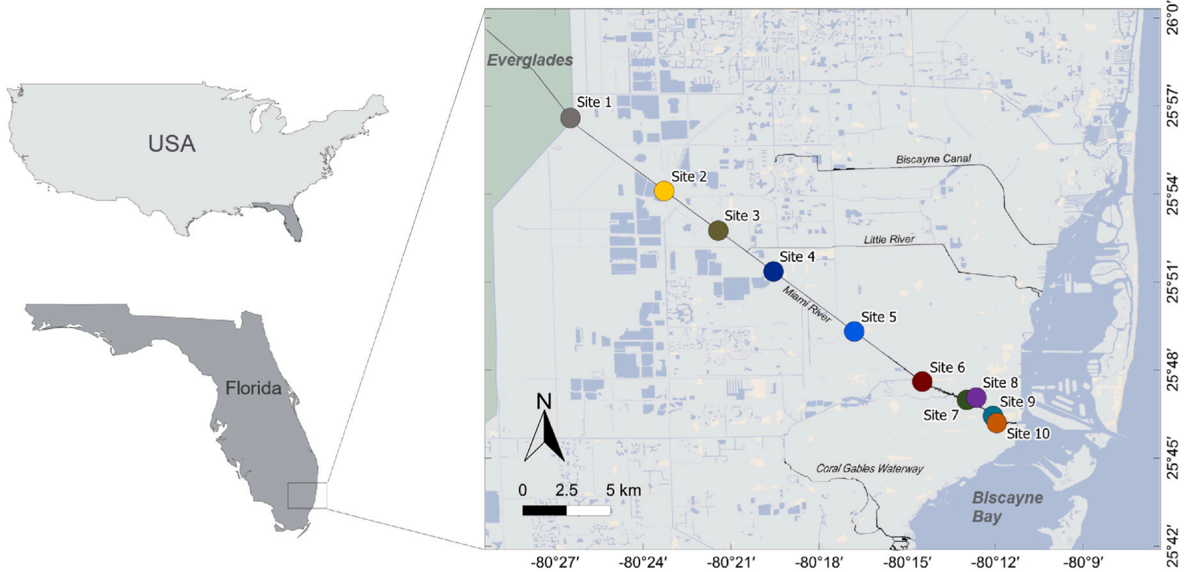


**Fig. 1.** The map of the Miami River in Southeastern Florida and the 10 sample locations starting from the Florida Everglades that drains a total of 2800 sq miles or fresh water into Biscayne Bay.

### 2.5. Quality control

Batch quality control samples (QCs) and blank LCMS water were processed with each set to assess instrument performance over the complete analysis sequence. The quality control checks are adapted from the work of Ng et al. (2020). Consecutive runs of the QCs mixtures (QC+ and QC-) are acquired before and after 8 analytical runs and only the data in between the two sets of QCs meeting the quality control requirements are further processed. QCs were analyzed in full scan mode for both positive and negative NTA method parameters. Stability in mass error, peak shape, peak intensity, and retention time were monitored throughout the analysis sequence to implement a basic quality system for the QCs. Compounds found within the QCs mixtures along with their concentrations and precision figures of merit acceptance can be found in the supporting information. The QCs were additionally used to generate a Log Kow and retention time model to reduce false positives as described and proven effective in Ng et al. (2020). The retention time model was generated for each individual analysis.

### 2.6. NTA Data processing

Raw files for NTA samples were processed using Compound Discoverer™ (CD) version 3.3.0.550 (Thermo Scientific), a small molecule structure identification software. The data processing workflow was adapted from Ng et al. (2020) on CD 3.0 and furthered by Troxell et al. (2022) on CD 3.1. The peak picking conducted by CD 3.3 in this present study consisted of a mass and intensity tolerance of 5 ppm and 30%, signal to noise ratio of 3, and a minimum peak intensity of 500,000 counts. The elemental composition prediction performed by CD stipulations were set at a mass tolerance limit of 5 ppm, element counts of less than C90H190 Br3 Cl4 F6 K2 N10 O18 P3 S5, maximum ring double bond equivalents (RDBE) of 40, and H/C ratio less than 3.5. Pattern matching completed by the software was also done with an intensity tolerance of 30%, threshold for intensity of 0.1%, S/N ratio of 3, spectral fit at least 50%, and pattern coverage of at least 80%. Fragment matching was done with a mass tolerance of 5 ppm and S/N of 3. HPLC solvents blanks were used to distinguish any possible contaminants from the instrument, columns, or solvents and were included in the NTA workflow for background subtraction. The CD 3.3 software matches the information obtained by using online databases such as MzCloud™, Aggregated Computational Toxicology Resource (ACToR), DrugBank, EAWAG Biocatalysis/Biodegradation Database, EPA DSSTox, EPA Toxcast, and FDA UNII-NLM. Additional filters were applied in CD 3.3 to the features detected after the initial data processing; feature must be present in the selected databases, peak areas should be greater than 500, 000 counts, retention time greater than 8.0, full or partial match to predicted compositions, and a peak shape rating of 5 (Fig. S2). Detailed data processing parameters and workflows conducted in CD 3.3 are provided in the supporting information.

### 2.7. Data analysis

For NTA the compounds that passed the post-processing filter were plotted as a Kendrick Mass Defect (KMD) plots for visualization purposes. This method has been previously used in environmental applications and has been proven as a useful tool to help data analysis (Merel, 2023). KMD plots were created by first calculating the Kendrick mass by 1.) $Kendick\ mass = IUPAC\ mass \times \frac{14.00000}{14.01565}$ and then using 2.) $Kendrick\ mass\ defect = nominal\ mass - Kendrick\ exact\ mass$ to calculate the KMD (Hughey et al., 2001). To understand the correlations between the sampling sites, PCA, was also performed based on the compounds that passed the post-processing filter and their corresponding intensities. In CD 3.3, the PCA plot was generated, transformed, and scaled by the mean centering the data and dividing by standard deviation. HCA was performed on prioritized probable and tentative identified compounds with the use of R.

### 2.8. Reproducibility evaluation

Although the study's structure implements the opportunity to assess reproducibility between multiple sampling events at the same locations, it is important to determine the reproducibility of duplicates of an individual sample. Reproducibility of the NTA was assessed in this study by implementing a basic replicate evaluation to determine method robustness. Two environmental samples from the study set were selected and analyzed utilizing the high-resolution mass spectrometry method and CD 3.3 data processing workflow. The samples were run intraday, inter day, two weeks apart, and over one month in both positive and negative mode. The set of features identified as probable and tentative identified compounds (equivalent to Schymanski Level 2 and 3 confidence) from the samples analyzed were compared in each of the replicates.

## 3. Results and discussion

### 3.1. Chemical fingerprints

Urban environments have long been known to introduce anthropogenic chemicals to adjacent water sources (Caccia and Boyer, 2005; Carey et al., 2011b) making an objective of this study to map the constituents present in the Miami River's chemical fingerprint and how those constituents change in the urban water system with the use of KMD plots as a visualization tool. KMD plots can provide a detailed visual comparison of the chemical space of a water sample making it a useful tool for sample prioritization. The features used for the plots were Level 2 and 3 compounds of the Schymanski scale classified as probable and tentative compounds having a mass spectrum database match in positive or negative mode that passed the applied filter in CD 3.3. Total compounds before and after filtering can be found for each site in Table S3. Not only was the KMD calculated for each sample and plotted against the Kendrick Nominal Mass (KNM), but the intensity of the features in the plot are represented by intensity percentiles to indicate where in the chemical space variations in compound abundances were present.

After subjecting each sampling site to the KMD interpretation, three distinct water fingerprints revealed themselves in Site 1 (Everglades based), Site 8 (Seybold Canal), and Site 10 (Biscayne Bay) (Fig. 2). When examining the chemical space fingerprinting for each of these sites the number of features and the location of those features stood out among the rest based on the sites chemical space spread and compound intensities (Figs. S3 and S4). From the KMD plot it was evident that Site 3, the Seybold Canal location, had the most densely populated chemical space in the range of KNM 100–400 and KMD of 0–0.3. Site 1 and Site 10 had a similar chemical fingerprint ranging from KNM of 100–400 and KMD of 0–0.2. The difference between the two fingerprints was that Site 10, the location closest in proximity to Biscayne Bay, had more features and a tightly bound chemical space where Site 1 extends past 0.2 KMD providing evidence of the chemical space being altered throughout the urban system. In this study, KMD plots have shown differences in the chemical fingerprint of the endmembers of the Miami River along with any unique sources present throughout the river warranting deeper investigation into the data set.

Previous studies have shown similar trends of increasing compound density and intensity when analyzing water from an urban watershed (Bonnefille et al., 2023; Nanusha et al., 2022; Yang et al., 2022). After assessing the chemical space fingerprinting of the ten sampling locations and concluding that Site 1, Site 8, and Site 10 had unique features defining their chemical space additional effort was needed to confirm that the visualization tool helped prioritize the samples to define water sources of this system. In Fig. 3, each sampling location was plotted with the average number of probable and tentative compounds and their

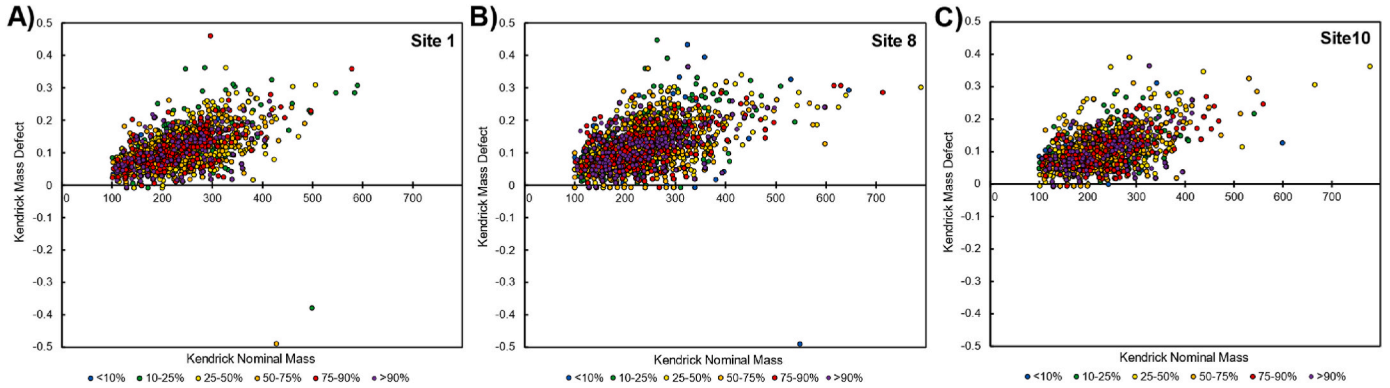


**Fig. 2.** Kendrick Mass Defect plots of probable and tentative compounds at three distinct sources in the Miami River Site 1 (A, n = 2093), 8 (B, n = 2935), and 10 (C, n = 2515) with percentiles based on intensity of peaks.

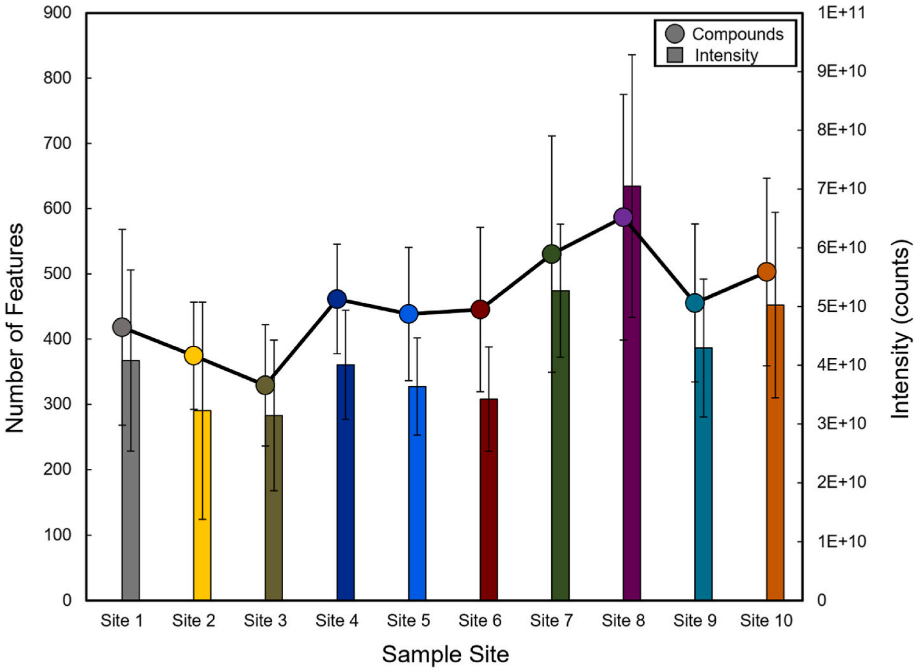


**Fig. 3.** The influence of the surrounding urban environment on the Miami River shown by the increasing average number of Level 2 and 3 compounds and intensity (in count) in positive and negative mode at each sampling location from the Everglades (Site 1) to Biscayne Bay (Site 10).

collective intensities in positive and negative mode from the five sampling events consistent with the compounds from the KMD plots. The results show a noticeable rise in both the number of compounds and their total intensity from the Everglades (Site 1) to the urban environment, which begins at Site 3 and extends to the point where the river discharges into Biscayne Bay (Site 10). Site 1 in the Everglades has a higher average intensity and number of compounds compared to the following two sites (Site 2 and 3). However, at Site 4, situated within the heavily industrialized surroundings on both sides of the river, a distinct increase in both the quantity and intensity of compounds was recorded. Supporting the KMD plots, the Miami River endmembers', Site 1 and Site 10, along with the Seybold Canal showed distinct differences in intensity and compounds number. Throughout this study the relationship between the sites were assessed using PCA, HCA, and compound prioritization to further define sources.

### 3.2. Principal component analysis

PCA is a widely used method for examining extensive datasets with numerous variables or features. It enhances data interpretability while retaining the maximum amount of information, making it possible to visualize multidimensional data effectively (Wold et al., 1987). The aim is to understand how the variables of the input data set vary from the mean with respect to each other, to assess if there is any relationship between them. PCA has been shown to differentiate water sources and shared pollution pathways among surface waters in previous studies (Beckers et al., 2020; Hu et al., 2021; Troxell et al., 2022).

The PCA was applied to the NTA post-processing results of the Level 2 probable and 3 tentative compounds. Two PCA's for each sampling events were created with the probable identified compounds from positive and negative ion modes along with their corresponding intensities. An example of the positive and negative PCA from the first sampling event is shown in Fig. 4, where there is a gradient change between Site 1 to Site 10 with an observable outlier among the gradient, Site 8 the Seybold Canal. The positive and negative mode PCA's described the total variance between data of 44.9% and 49.8%. Although there was a gradient change between Site 10 to Site 1, there were three distinct quadrantes of the PCA that house sites that resemble the Everglades (upper left), Biscayne Bay (lower right), and Seybold Canal (upper right) in both the positive and negative mode. Each trip (1–5) resembles this

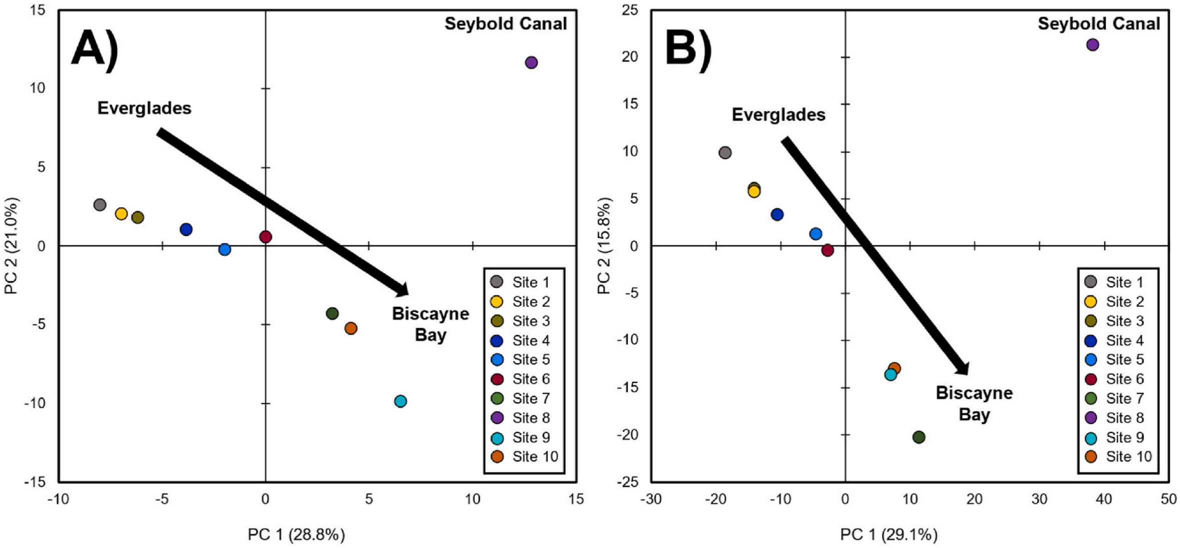


**Fig. 4.** Principal Component Analysis comparing all samples based on Level 2 probable and 3 tentative compounds and their intensities throughout the Miami River in positive (n = 2045) and negative (n = 1067) mode. Positive (A) and negative (B) mode for Trip 1 both showing three distinctive water sources from site 1, 8, and 10 (see Figs. S3 and S4 for all individual sampling events).

trend, apart from the heavy rain event the separation between the Everglades and Biscayne Bay sampling locations was exacerbated and can be found in the supporting information along with the other individual sampling events' PCAs (Figs. S3 and S4).

PCA has provided evidence that the observed trends in the Miami River are consistent across individual events. However, it is crucial to evaluate whether these trends persist when combining all events and comparing them using a single PCA for both positive and negative modes. In Fig. 5 (A and B), positive mode results demonstrate that the Seybold Canal remains a visible outlier, but the trend from Everglades to Biscayne Bay (Site 1–10) is not as readily evident. As anticipated, the relationship is not as linear as previously described, nevertheless, there are still some consistencies with the Everglades sites from various sampling events being grouped together, as well as samples from Biscayne Bay. Negative mode PCA's can be found in the supplementary information showing similar results to the positive mode PCA described here. The PCA results of the heavy rain event differ from all others, suggesting a change in how the sources influence the river. For example, an increased presence of perception and stormwater runoff, which was not a factor in the other sampling events. The individual and collective PCAs show reproducibility in the results for surface water in an interconnected system and its suitability for distinguishing event conditions, but further compound prioritization is needed for a higher resolving power of sources and to identify unique source tracers.

### 3.3. Reproducibility

Based on previous in-house studies to assess method performance and metrics using the same set of QC samples as specified in this study, a precision of RSD ≤50% and accuracy of ≥70% was suggested as an acceptable benchmark for non-target analysis (Ng et al., 2020). In this study, reproducibility of compound identification of the NTA results were examined with the goal of not only understanding the method reproducibility, but to establish a threshold for frequency of detection

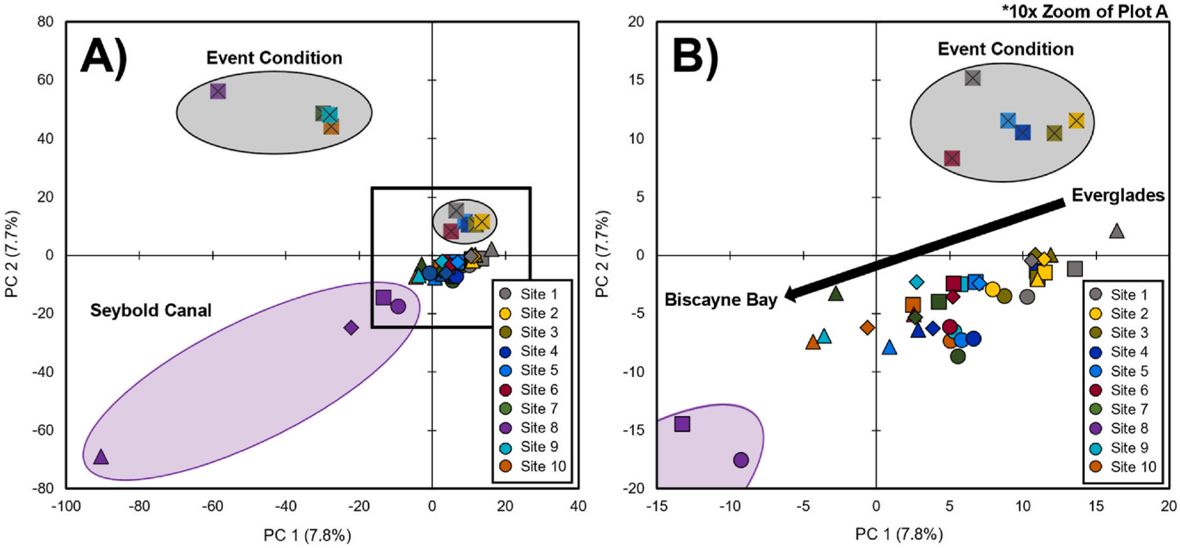


**Fig. 5.** Positive mode Principal Component Analysis for the comparison of all trips Level 2 and 3 probable and tentative compounds (n = 4578) and intensities (A) and the scaled plot (B) (Circle, Trip 1; Triangle, Trip 2; Square, Trip 3; Square X, Trip 4; Diamond, Trip 5). The negative mode PCA can be found in the supplementary information (Fig. S7).

when prioritizing probable identified compounds. The reproducibility of the method was tested by subjecting two of the environmental samples from the Miami River to a five aliquot quintuplicate analysis. The two samples were analyzed intraday, interday one day and week later, and one month later. All results were subjected to the same post-process filtering using CD 3.3. The findings revealed that among the duplicates, 39.8% and 41.4% of the total number of compounds were present in at least three out of the five samples.

A significant proportion of the overall compounds detected (42.2% and 43.8%) consisted of probable identified compounds in one of the five replicates. By removing probable identified compounds only detected once and considering those with higher detection frequencies, 71.6% and 70.8% of the compounds were identified in at least three out of the five or 60% of replicates. The increased weight of the single detections in the reproducibility emphasizes the importance of replication in NTA. This reproducibility study supports the establishment of a frequency of detection threshold. Setting such a threshold is crucial in non-target analysis, and conducting replicate studies using different aliquots of the same samples or different samples from the same sampling location is essential to reduce the number of single identifications when aiming to identify source tracers. This study further suggests that a strict frequency of detection threshold of 80% is beneficial in enhancing confidence and prioritizing compounds for potential tracers.

### 3.4. *Prioritizing frequency of detection*

Considering features with at least a Level 2 probable identification with a frequency of detection of 80% per site, there were a total of 470 different compounds prioritized from the three identified sources the Everglades (Site 1), Seybold Canal (Site 8), and Biscayne Bay (Site 10). Of these detected compounds, the Everglades based sampling location (Site 1) had 160 compounds detected among the five sampling events with 11% (51) unique to the Everglades, the Seybold Canal had 339 compounds detected with 29% (136) unique to the Seybold Canal, and the Biscayne Bay location had 272 with 12% (58) unique to the Biscayne Bay location (Fig. 6). There were 16% (76) compounds commonly detected among the three sources. When comparing the compounds detected at each site, the Seybold Canal and Biscayne Bay samples had the most compounds in common at 41% (192). Biscayne Bay and the Everglades had 20% (98) compounds in common followed by the Everglades and Seybold Canal site having the least number of compounds in common at 18% (87). The results highlight the Seybold Canal again as an outlier and an area of concern due to the number of unique compounds originating from the canal and the evident influence it has on the site closest to Biscayne Bay.

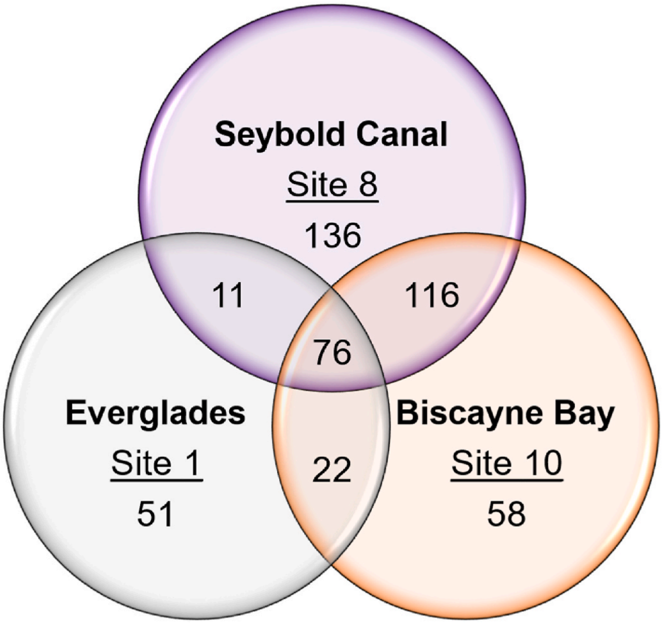


**Fig. 6.** Level 2 and 3 compounds found in the three distinct water sources determined by PCA from samples in the Everglades (grey), Seybold Canal (purple), and Biscayne Bay (orange).

HCA was used to identify groups or clusters within the NTA data. HCA has been previously used with NTA applications (Schollée et al., 2018). When applied to surface water analysis to determine sources, HCA can help identify similarities and differences among different water samples based on probable and tentative identified compounds and their intensities. The data was scaled to ensure that all variables have the same influence on the analysis, regardless of the compound's intensity scale, with z-score normalization chosen for this cluster analysis. Ward's linkage was selected as a popular agglomerative HCA algorithm (Kimes et al., 2017; Roux, 2018). The method is based on the concept of minimizing the sum of squared differences within clusters and is calculated by considering the increase in the total sum of squares that would result from merging the clusters and is commonly computed using the squared Euclidean distance between cluster centroids or means (Ward, 1963). Euclidean distance is used as the metric for continuous variables and was chosen to exhibit the correlation distance to determine the similarity between each sampling sites and events.

The cluster analysis generated from analyzing the compounds with at least an 80% detection frequency in any sampling site over five events and their intensities resulted in four clusters (Fig. 7). The clusters showed a clear group of the Seybold Canal samples and the only Seybold Canal sample that was not clustered among the rest was sampling event for the heavy rain event. The rain event strongly clustered together in the dendrogram indicating that this was a true event condition altering the compounds present and their intensities to be drastically different among the other standard wet season events. While the Seybold Canal and heavy rain event were distinctly defined clusters, the last two clusters generated represented samples that were close in proximity to the Everglades and Biscayne Bay sources. By using compounds that were detected with an 80% frequency, the sources were further and more clearly differentiated providing support that improving compound confidence by having a detection frequency criteria improves the discriminatory power NTA has at defining sources and prioritizing sample locations for further water quality monitoring. By utilizing a standardized NTA approach, it is possible to identify competing water sources in an urban water system and determine their correspondent source-specific tracers.

### 3.5. *Identification of potential tracers*

By utilizing a standardized and reproducible NTA approach to identify competing water sources in an urban water system, three distinctive water sources were defined with KMD, PCA plots, and further refined by HCA. Understanding the probable identified compounds that define the water sources are of high priority to understand which compounds may be used as tracers of these sources. Tracers are compounds that are easily detectable, unique to the source, and relatively stable in the environment. With the addition of identifying source specific tracers' further source discriminatory power and water quality monitoring prioritization can be conducted. Tracers were defined in three identified sources Everglades, Seybold Canal, and Biscayne Bay from the probable compounds (Level 2a and 2b) that had an 80% detection frequency originating from the identified sources. Compounds were also identified as being contributed from the surrounding urban environment along the Miami River that didn't originate from the defined sources. For these probable compounds to qualify as tracers, as indicated in Table 1, they had to satisfy distinct criteria. This entailed a Level 2a requirement, necessitating a match within an MS2 database (mzCloud). The fragmentation pattern match must be more than 60% to

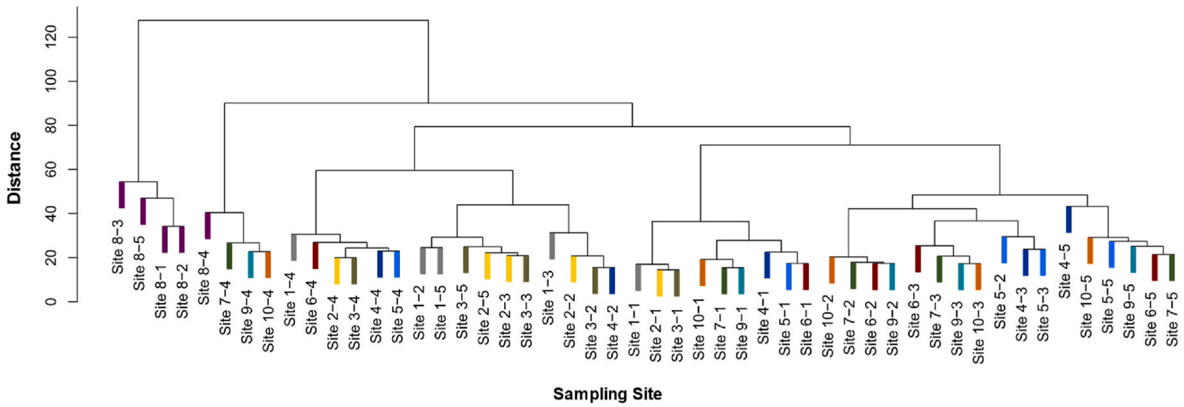


**Fig. 7.** Hierarchal Cluster Analysis of the intensities of compounds probable and tentative identified at an 80% detection frequency over five sampling events for positive (n = 529) and negative mode (n = 63).

place compounds in this category. If an MS2 spectrum was unavailable, the evaluation relied on diagnostic support (Level 2b), including Log Kow, to predict the compound's retention time, allowing for a 1-min deviation of the actual retention time. Lastly, the compounds were required to achieve the detection frequency threshold of 80%.

Probable compounds that defined the Everglades based source were mostly natural products, herbicides, and insecticides. The Florida Everglades is a unique and diverse ecosystem that is influenced by numerous factors, including its proximity to the large agricultural industry directly North of the protected wetland. The heavy influence of agriculture on the Everglades can result in the presence of anthropogenic compounds to the receiving into the surface waters. One such compound detected regularly in the Everglades is atrazine, which is a commonly used herbicide in the agricultural industry. Atraton, a metabolite of atrazine, was a probable identified compounds selected as an Everglades tracer where the intensity is highest in the Everglades and decreases down the Miami River (Fig. 8). Atrazine is widely used to control weeds in crops, particularly in corn and sugarcane cultivation, both of which are prevalent in the agricultural regions north of the Everglades. Due to the extensive use of atrazine in these areas, it is not uncommon for the compound and its metabolites to find their way into the surrounding environment in the Everglades and have been monitored along with other herbicides over the years (Harman-Fetcho et al., 2005). Atraton, Bufencarb, Difenoxuron, Picardin, and Simetone are examples of the prioritized herbicides, insecticides, and insect repellents selected as source tracers of the Everglade's. Monitoring the presence of these tracers in the Everglades and connecting water system allows the ability to track this source throughout the urban water system.

The tracers identified for the Seybold Canal source primarily consisted of anthropogenic compounds such as pharmaceuticals, personal care products, and surfactants. These tracers were selected based on their unique origin in the Seybold Canal and their influence on the subsequent sites along the Miami River. An example, Lidocaine, a commonly used anesthetic, was detected in the Seybold Canal and in the following two sampling locations (Site 1 and 2) in the direction of Biscayne Bay (Fig. 8). Other compounds originating from the Seybold Canal included perfumes such as Androstadienone and Hedione, as well as Cladosporin and 4-hydroxychlorothalonil, which are antifungal and fungicide metabolites. Pharmaceuticals such as Phenyltoloxamine and Trimelamol, along with industrial additives like Sudan 2 and Thiabendazole, were also found in the canal. The presence of these compounds can be attributed to the surrounding land use patterns near the canal, which include residential areas, industrial zones, and several hospitals. To track the Seybold Canal source throughout the urban water system, it is essential to monitor the presence of tracers in the canal to establish the extent of influence it has on the Miami River and ultimately Biscayne Bay. Based on the number of unique compounds present and the common use tracers identified, the Seybold Canal is a location chosen for further water quality monitoring.

As predicted, it was difficult to determine unique tracers of the Biscayne Bay site because the bay is the discharge point of several canals including the Miami River. However, this specific location within Biscayne Bay did exhibit a substantial presence of compounds classified as urban influence compounds. These compounds were selected based on their addition throughout the Miami River and their persistence until reaching the discharge point of Biscayne Bay. Among them, Irgarol 1051, an algaecide commonly used in antifouling paints, was detected throughout the middle section of the Miami River, where significant boating industries are concentrated (Fig. 8). Irgarol 1051 continued to be probable identified until the end of the river location and was therefore selected as an urban influence tracer. The use of Irgarol 1051 in boat paint has been canceled by the EPA in 2021 due to its detrimental effects on freshwater and marine plants, aiming to safeguard aquatic ecosystems (EPA, 2023). Previous studies in Southeastern Florida, including in the Miami River, have targeted Irgarol 1051 at levels below the effects benchmark for the plant 10th centile (251 ng/L), but was highest in the Miami River at locations similar to the probable detection in his study (Hall and Gardinali, 2004; Hall et al., 2009). Other suggested urban influence tracers include phosphorus flame retardants (Tributyl Phosphate), insect repellent (DEET), stimulant (Caffeine), pharmaceutical (Carbamazepine), nonagricultural herbicide (Prometon), algaecide (Irgarol 1051), and detergent (Xylenesulfonate) compounds. To assess their feasibility as tracers, targeted quantification is necessary for the suggested tracers in the Everglades, Seybold Canal, and urban influence along the Miami River.

### 3.6. *Study limitations*

A limitation of this study is the inability to assess matrix effects within the samples from the freshwater everglades and the saltwater bay. The work with the QC samples does indicate that suppression may be important for saltwater samples, but the use of a few selected internal standards will not likely resolve the issue over the whole NTA chemical space. Great progress has been made towards the standardization of NTA work, but the assessment of bias is still a challenge in any study using complex environmental samples. It is possible that compounds detected in freshwater samples may be below the instrumental detection thresholds for saltwater and vice versa. We have partially addressed the issue by introducing S/N thresholds and intensity post-filtering algorithms in the NTA workflow, but even the introduction of internal standards could create additional bias as not all internal standards will be representative of the data set as a whole. LC-HRMS is an accepted technology for identifying low concentrations of compounds in the environment, but at trace-levels additional knowledge, techniques, or a

**Table 1**
Probable compounds proposed as source indicators throughout the Miami River.

| Compound Name | Uses | Ionization Mode | Monoisotopic Mass | Log Kow |
|---|---|---|---|---|
| *Everglades* | | | | |
| Atratone[a] | herbicide | (+) | 211.14331 | 2.56 |
| Bufencarb | Insecticide | (+) | 442.28317 | 3.70 |
| Difenoxuron | herbicide | (+) | 286.13174 | 2.73 |
| Picaridan | insect repellant | (+) | 229.16780 | 1.56 |
| Simetone | herbicide | (+) | 197.12765 | 2.21 |
| 2-[(8-Methyl-2-oxo-4-propyl-2H-chromen-7-yl)oxy] propanoic acid | fungicide and bactericide | (−) | 290.11541 | 3.45 |
| (±)-Equol | natural estrogenic metabolite from soy isoflavones | (+) | 242.09429 | 2.98 |
| *Seybold Canal* | | | | |
| Androstadienone | perfumes | (+) | 270.19836 | 5.03 |
| Cladosporin | antimalarial, antifungal, antibacterial | (+) | 292.13107 | 2.48 |
| Flavanone | antimicrobial and antifungal | (+) | 224.08372 | 3.62 |
| Hedione | perfumes | (−) | 226.15689 | 2.50 |
| Lidocaine | anesthetic | (+) | 234.17321 | 3.63 |
| Methyl undec-10-enoate | cosmetics and personal care products | (−) | 198.16197 | 4.45 |
| Phenyltoloxamine | pharmaceutical | (+) | 255.16230 | 3.98 |
| Sudan 2 | coloring agent in oils, plastics, textiles, shoes, printing inks, floor polishes and cosmetics | (+) | 276.12625 | 5.16 |
| Thiabendazole[a] | preservative in paints, carpets, adhesives, textiles, and waxing fruits and vegetables | (+) | 201.03607 | 2.47 |
| Trans-Azobenzene | acaricide, fumigant, and chemical intermediate | (+) | 182.08439 | 3.82 |
| Trimelamol | pharmaceutical | (+) | 258.14404 | |
| 4-hydroxychlorothalonil | metabolite of widely used fungicide | (−) | 245.91544 | 3.24 |
| *Urban Influence* | | | | |
| Acetyl norfentanyl[a] | major metabolite of acetyl fentanyl an opioid analgesic drug | (+) | 218.14203 | 1.21 |
| Butyl diphenyl phosphate | fire retardants, plasticizers and oils | (+) | 306.10208 | 4.69 |
| Caffeine | stimulant | (+) | 203.13687 | −0.13 |
| Carbamazepine[a] | pharmaceutical | (+) | 236.09495 | 2.67 |
| Cycluron | herbicide | (+) | 198.17321 | 2.98 |
| DEET[a] | insect repellant | (+) | 191.13101 | 1.96 |
| Desmetryn | herbicide | (+) | 213.10481 | 2.56 |
| Dibutyl 3-hydroxybutyl phosphate | metabolite of tributyl phosphate | (+) | 282.15960 | 2.61 |
| Diethyl phosphate | metabolite of organophosphorus pesticides | (+) | 154.03949 | 0.08 |
| Diuron | broad-spectrum residual herbicide and algaecide | (+) | 232.01701 | 2.88 |
| Estriol | weak estrogen hormone | (+) | 288.17254 | 2.94 |
| Fludrocortisone | corticosteroid | (−) | 380.19990 | 1.70 |
| Gemfibrozil | pharmaceutical | (−) | 250.15689 | 4.39 |
| Hexadecanedioic acid | insecticide and acaricide | (−) | 286.21440 | 5.40 |
| Irgarol 1051[a] | algaecide commonly used in antifouling paints | (+) | 253.13610 | 3.27 |
| Isoproturon | herbicide | (+) | 206.14190 | 2.32 |
| Methyl cinnamate[a] | flavor and perfume industries | (+) | 162.06808 | 2.18 |
| Molindone | pharmaceutical | (+) | 276.18377 | 1.96 |
| Prilocaine | local anesthetics | (+) | 220.15756 | 1.74 |
| Prometon | nonagricultural herbicide | (+) | 225.15896 | 2.91 |
| Tebuthiuron[a] | non-cropland areas herbicide | (+) | 228.10447 | 1.79 |
| Terbucarb | insecticide and pesticide | (+) | 277.20419 | 5.00 |
| Traumatic Acid | antifibrinolytic drug | (−) | 228.13615 | 2.95 |
| Triisobutyl Phosphate[a] | organophosphorus flame retardant | (−) (+) | 266.16470 | 4.26 |
| Xylenesulfonate[a] | household detergents surfactant and shampoos | (−) | 185.02778 | 1.39 |
| 1,2,3-Benzotriazole | anti-rust and corrosion inhibitor | (+) | 119.04834 | 1.34 |
| 2,6-di-*tert*-Butylphenol | antioxidant in polymers, lubes, jet fuels, metal cutting fluids | (−) | 206.16706 | 4.90 |
| 2-Nonylphenyl hydrogen sulfate | detergents, lubricants, and antistatic agents | (−) | 300.13953 | 5.60 |
| 5-Methylbenzotriazole tolytriazole[a] | antirust and corrosion inhibitor for metals | (+) | 133.06399 | 1.80 |
| 17beta-Trenbolone[a] | metabolite of synthetic anabolic steroid | (+) | 270.16198 | 2.27 |

[a] Compounds with MS2 database match.

reference standard are usually needed for conclusive identification (Krauss et al., 2010). Acknowledging these biases, this study does not aim for any quantification or semi quantification, but solely for the purpose of compound and sample prioritization.

## 4. Conclusion

This study provides valuable insights into the water quality of the Miami River, shedding light on the influence of the urban environment and human activities on its chemical composition. By employing HRMS and NTA, a distinct gradient in the river's chemical fingerprint from the Florida Everglades to Biscayne Bay was observed by the average compounds present at each sampling location and their intensities. Through post-processing using CD 3.3 and subsequent statistical analysis utilizing techniques such as KMD plots PCA, significant variations in the chemical makeup of the river along its course were observed. Reproducibility of the analytical method were also addressed in this study by comparing the results of the five sampling events and conducting a sample duplicate study. The robustness of the method and compound prioritization was demonstrated, and a detection frequency of 80% was set as a criterion to improve confidence of probable compounds identification. The compound prioritization techniques employed in this study, based on the frequency of detection, provided further source discrimination shown by HCA and allowed for the identification of potential tracers for each water source along the Miami River. Tracers were established for the Everglades including natural products, agricultural herbicides, and insecticides such as Atraton, Bufencarb, Difenoxuron, Picardin, and Simetone; Seybold canal including anthropogenic contaminants such as

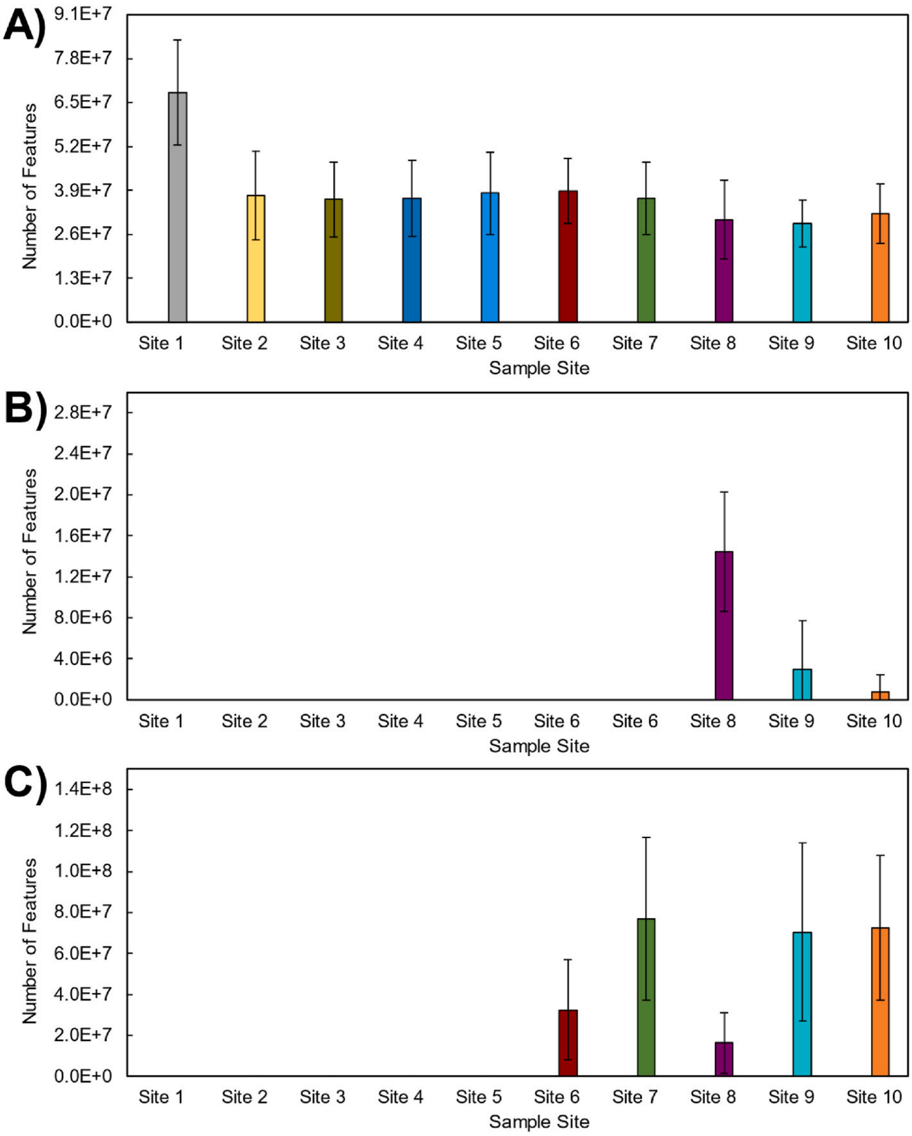


**Fig. 8.** Examples of the probable compounds that were selected as tracers for Everglades based location Site 10 and the Seybold Canal Site 1. Atraton was an example of a tracer from the Everglades (A). An example of a tracer from the Seybold Canal was Lidocaine (B). An example of a compound that was selected as an urban influence tracer due to its clear inputs into the Miami River from the urban surrounds was Irgarol 1051(C).

Androstadienone, Hedione, Cladosporin, and 4-hydroxychlorothalonil; and a lack of unique tracers for Biscayne Bay. Flame retardants, insect repellents, pharmaceuticals, and nonagricultural herbicides, such as Tributyl Phosphate, DEET, Caffeine, Carbamazepine, Prometon, Irgarol 1051, and Xylenesulfonate, were established as urban influenced contaminants along the river. Overall, this study emphasizes the complex interplay between land use and urban development on the chemical fingerprint in the Miami River. It highlights the effectiveness of HRMS and NTA in tracking specific water sources and understanding the effects of human activities through the urban environment. Ongoing research is vital to inform decision-making when it comes to prioritizing locations for sustainable management actions aimed at protecting and restoring the water resources of South Florida, particularly in relation to Biscayne Bay.

## CRediT authorship contribution statement

**Kassidy Troxell:** Conceptualization, Methodology, Data curation, Formal analysis, Writing - original draft. **Milena Ceccopieri:** Conceptualization, Data Curation, Writing - review & editing, Visualization. **Piero Gardinali:** Conceptualization, Writing - review & editing, Supervision, Project Administration, Funding Acquisition.

## 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.

## Acknowledgement

This is publication #1666 from the Institute of Environment at Florida International University. This material is based upon work supported by the National Science Foundation under Grant No. HRD-1547798 and Grant No. HRD-2111661. These NSF Grants were awarded to Florida International University as part of the Centers of Research

Excellence in Science and Technology (CREST) Program. Thank you to the two anonymous reviewers whose comments and suggestions helped improve the manuscript.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.chemosphere.2023.140863.

## References

Beckers, L.-M., Brack, W., Dann, J.P., Krauss, M., Müller, E., Schulze, T., 2020. Unraveling longitudinal pollution patterns of organic micropollutants in a river by non-target screening and cluster analysis. Sci. Total Environ. 727, 138388 https://doi.org/10.1016/j.scitotenv.2020.138388.

Bertrand-Krajewski, J.-L., Barraud, S., Chocat, B., 2000. Need for improved methodologies and measurements for sustainable management of urban water systems. Environ. Impact Assess. Rev. 20 (3), 323–331. https://doi.org/10.1016/S0195-9255(00)00044-5.

Black, G., Lowe, C., Anumol, T., Bade, J., Favela, K., Feng, Y.-L., Young, T., 2023. Exploring chemical space in non-targeted analysis: a proposed ChemSpace tool. Anal. Bioanal. Chem. 415 (1), 35–44. https://doi.org/10.1007/s00216-022-04434-4.

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. Environmental Science & Technology. https://doi.org/10.1021/acs.est.2c08200.

Caccia, V.G., Boyer, J.N., 2005. Spatial patterning of water quality in Biscayne Bay, Florida as a function of land use and water management. Mar. Pollut. Bull. 50 (11), 1416–1429. https://doi.org/10.1016/j.marpolbul.2005.08.002.

Cantillo, A.Y., Pikula, L., Hale, K.K., Collins, E.V., Caballero, R., 2000. Biscayne Bay Environmental History and Annotated Bibliography [Technical Memorandum], vol. 145. NOAA technical memorandum NOS NCCOS CCMA.

Carey, R.O., Migliaccio, K.W., Brown, M.T., 2011a. Nutrient discharges to Biscayne Bay, Florida: trends, loads, and a pollutant index. Sci. Total Environ. 409 (3), 530–539. https://doi.org/10.1016/j.scitotenv.2010.10.029.

Carey, R.O., Migliaccio, K.W., Li, Y., Schaffer, B., Kiker, G.A., Brown, M.T., 2011b. Land use disturbance indicators and water quality variability in the Biscayne Bay Watershed, Florida. Ecol. Indicat. 11 (5), 1093–1104. https://doi.org/10.1016/j.ecolind.2010.12.009.

Chiaia-Hernández, A.C., Günthardt, B.F., Frey, M.P., Hollender, J., 2017. Unravelling contaminants in the anthropocene using statistical analysis of liquid chromatography–high-resolution mass spectrometry nontarget screening data recorded in lake sediments. Environ. Sci. Technol. 51 (21), 12547–12556. https://doi.org/10.1021/acs.est.7b03357.

da Costa Filho, B.M., Duarte, A.C., Rocha-Santos, T.A.P., 2022. Environmental monitoring approaches for the detection of organic contaminants in marine environments: a critical review. Trends in Environmental Analytical Chemistry 33, e00154. https://doi.org/10.1016/j.teac.2022.e00154.

Cui, D., Ricardo, M., Quinete, N., 2022. A novel report on phthalates levels in Biscayne Bay surface waters and drinking water from South Florida. Mar. Pollut. Bull. 180, 113802 https://doi.org/10.1016/j.marpolbul.2022.113802.

EPA, 2023. EPA protects aquatic ecosystems by finalizing Irgarol antifoulant paint cancellation. https://www.epa.gov/pesticides/epa-protects-aquatic-ecosystems-finalizing-irgarol-antifoulant-paint-cancellation.

Gardinali, R., P Plasencia, M., Mack, S., Poppell, C., 2002. Occurrence of IRGAROL 1051 in coastal waters from Biscayne bay, Florida, USA. Mar. Pollut. Bull. 44 (8), 781–788. https://doi.org/10.1016/S0025-326X(02)00058-9.

Guardian, M.G.E., He, P., Bermudez, A., Duan, S., Kaushal, S.S., Rosenfeldt, E., Aga, D.S., 2021. Optimized suspect screening approach for a comprehensive assessment of the impact of best management practices in reducing micropollutants transport in the Potomac River watershed. Water Res. X 11, 100088. https://doi.org/10.1016/j.wroa.2021.100088.

Hall, L.W., Gardinali, P., 2004. Ecological risk assessment for Irgarol 1051 and its major metabolite in United States surface waters. Hum. Ecol. Risk Assess. 10 (3), 525–542. https://doi.org/10.1080/10807030490452188.

Hall Jr., L.W., Killen, W.D., Anderson, R.D., Balcomb, R., Gardinali, P., 2009. Ecological risk of Irgarol 1051 and its major metabolite in coastal California marinas and reference areas. Mar. Pollut. Bull. 58 (5), 702–710. https://doi.org/10.1016/j.marpolbul.2008.12.019.

Harman-Fetcho, J.A., Hapeman, C.J., McConnell, L.L., Potter, T.L., Rice, C.P., Sadeghi, A.M., Curry, R., 2005. Pesticide occurrence in selected South Florida canals and Biscayne Bay during high agricultural activity. J. Agric. Food Chem. 53 (15), 6040–6048. https://doi.org/10.1021/jf047803g.

Hawkins, C., Foster, G., Glaberman, S., 2023. Chemical prioritization of pharmaceuticals and personal care products in an urban tributary of the Potomac River. Sci. Total Environ. 881, 163514 https://doi.org/10.1016/j.scitotenv.2023.163514.

Hernández, F., Bakker, J., Bijlsma, L., de Boer, J., Botero-Coy, A.M., Bruinen de Bruin, Y., Hogendoorn, E.A., 2019. The role of analytical chemistry in exposure science: focus on the aquatic environment. Chemosphere 222, 564–583. https://doi.org/10.1016/j.chemosphere.2019.01.118.

Hollender, J., Schymanski, E.L., Singer, H.P., Ferguson, P.L., 2017. Nontarget screening with high resolution mass spectrometry in the environment: ready to go? Environ. Sci. Technol. 51 (20), 11505–11512. https://doi.org/10.1021/acs.est.7b02184.

Hu, L.-X., Olaitan, O.J., Li, Z., Yang, Y.-Y., Chimezie, A., Adepoju-Bello, A.A., Chen, C.-E., 2021. What is in Nigerian waters? Target and non-target screening analysis for organic chemicals. Chemosphere 284, 131546. https://doi.org/10.1016/j.chemosphere.2021.131546.

Hughey, C.A., Hendrickson, C.L., Rodgers, R.P., Marshall, A.G., Qian, K., 2001. Kendrick mass defect spectrum: a compact visual analysis for ultrahigh-resolution broadband mass spectra. Anal. Chem. 73 (19), 4676–4681. https://doi.org/10.1021/ac010560w.

Joerss, H., Menger, F., Tang, J., Ebinghaus, R., Ahrens, L., 2022. Beyond the tip of the iceberg: suspect screening reveals point source-specific patterns of emerging and novel per- and polyfluoroalkyl substances in German and Chinese rivers. Environ. Sci. Technol. 56 (9), 5456–5465. https://doi.org/10.1021/acs.est.1c07987.

Kimes, P.K., Liu, Y., Neil Hayes, D., Marron, J.S., 2017. Statistical significance for hierarchical clustering. Biometrics 73 (3), 811–821. https://doi.org/10.1111/biom.12647.

Krauss, M., Singer, H., Hollender, J., 2010. LC–high resolution MS in environmental analysis: from target screening to the identification of unknowns. Anal. Bioanal. Chem. 397 (3), 943–951. https://doi.org/10.1007/s00216-010-3608-9.

Lara-Martín, P.A., Chiaia-Hernández, A.C., Biel-Maeso, M., Baena-Nogueras, R.M., Hollender, J., 2020. Tracing urban wastewater contaminants into the atlantic ocean by nontarget screening. Environ. Sci. Technol. 54 (7), 3996–4005. https://doi.org/10.1021/acs.est.9b06114.

Lee, D.J., Bwenge, A., 2008. Estimating the benefits from restoring coastal ecosystems: a case study of Biscayne Bay, Florida. Coastal Watershed Management 13, 283.

Madikizela, L.M., Nuapia, Y.B., Chimuka, L., Ncube, S., Etale, A., 2022. Target and suspect screening of pharmaceuticals and their transformation products in the klip river, South Africa, using ultra-high-performance liquid chromatography-mass spectrometry. Environ. Toxicol. Chem. 41 (2), 437–447. https://doi.org/10.1002/etc.5265.

Manz, K.E., Feerick, A., Braun, J.M., Feng, Y.-L., Hall, A., Koelmel, J., Young, J.A., 2023. Non-targeted analysis (NTA) and suspect screening analysis (SSA): a review of examining the chemical exposome. J. Expo. Sci. Environ. Epidemiol. 33 (4), 524–536. https://doi.org/10.1038/s41370-023-00574-6.

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. Trends in Environmental Analytical Chemistry 28, e00102. https://doi.org/10.1016/j.teac.2020.e00102.

Merel, S., 2023. Critical assessment of the Kendrick mass defect analysis as an innovative approach to process high resolution mass spectrometry data for environmental applications. Chemosphere 313, 137443. https://doi.org/10.1016/j.chemosphere.2022.137443.

Millette, N.C., Kelble, C., Linhoss, A., Ashby, S., Visser, L., 2019. Using spatial variability in the rate of change of chlorophyll a to improve water quality management in a subtropical oligotrophic estuary. Estuar. Coast 42 (7), 1792–1803. https://doi.org/10.1007/s12237-019-00610-5.

Nanusha, M.Y., Frøkjær, E.E., Liigand, J., Christensen, M.R., Hansen, H.R., Hansen, M., 2022. Unravelling the occurrence of trace contaminants in surface waters using semi-quantitative suspected non-target screening analyses. Environ. Pollut. 315, 120346 https://doi.org/10.1016/j.envpol.2022.120346.

Ng, B., Quinete, N., Gardinali, P.R., 2020. Assessing accuracy, precision and selectivity using quality controls for non-targeted analysis. Sci. Total Environ. 713, 136568 https://doi.org/10.1016/j.scitotenv.2020.136568.

Ng, B., Quinete, N., Maldonado, S., Lugo, K., Purrinos, J., Briceño, H., Gardinali, P., 2021. Understanding the occurrence and distribution of emerging pollutants and endocrine disruptors in sensitive coastal South Florida Ecosystems. Sci. Total Environ. 757, 143720 https://doi.org/10.1016/j.scitotenv.2020.143720.

Ng, B., Quinete, N., Gardinali, P., 2022. Differential organic contaminant ionization source detection and identification in environmental waters by nontargeted analysis. Environ. Toxicol. Chem. 41 (5), 1154–1164. https://doi.org/10.1002/etc.5268.

NOAA, 2022. Major and/or recent tropical cyclones: 2022 major Hurricane ian. National Weather Service: https://www.weather.gov/mfl/events_index.

Peter, K.T., Phillips, A.L., Knolhoff, A.M., Gardinali, P.R., Manzano, C.A., Miller, K.E., Sobus, J.R., 2021. Nontargeted analysis study reporting tool: a framework to improve research transparency and reproducibility. Anal. Chem. 93 (41), 13870–13879. https://doi.org/10.1021/acs.analchem.1c02621.

Phillips, A.L., Peter, K.T., Sobus, J.R., Fisher, C.M., Manzano, C.A., McEachran, A.D., Ulrich, E.M., 2023. Standardizing non-targeted analysis reporting to advance exposure science and environmental epidemiology. J. Expo. Sci. Environ. Epidemiol. https://doi.org/10.1038/s41370-022-00490-1.

Place, B.J., Ulrich, E.M., Challis, J.K., Chao, A., Du, B., Favela, K., Williams, A.J., 2021. An introduction to the benchmarking and publications for non-targeted analysis working group. Anal. Chem. 93 (49), 16289–16296. https://doi.org/10.1021/acs.analchem.1c02660.

Pourchet, M., Debrauwer, L., Klanova, J., Price, E.J., Covaci, A., Caballero-Casero, N., Antignac, J.-P., 2020. Suspect and non-targeted screening of chemicals of emerging concern for human biomonitoring, environmental health studies and support to risk assessment: from promises to challenges and harmonisation issues. Environ. Int. 139, 105545 https://doi.org/10.1016/j.envint.2020.105545.

RER, 2022. DERM Biscayne bay report card 2022. In: New Release. Miami-Dade County. https://www.miamidade.gov/releases/2022-04-26-mayor-bay-report.asp.

Roux, M., 2018. A comparative study of divisive and agglomerative hierarchical clustering algorithms. J. Classif. 35 (2), 345–366. https://doi.org/10.1007/s00357-018-9259-9.

Schollée, J.E., Bourgin, M., von Gunten, U., McArdell, C.S., Hollender, J., 2018. Non-target screening to trace ozonation transformation products in a wastewater treatment train including different post-treatments. Water Res. 142, 267–278. https://doi.org/10.1016/j.watres.2018.05.045.

Schwarzenbach, R.P., Escher, B.I., Fenner, K., Hofstetter, T.B., Johnson, C.A., von Gunten, U., Wehrli, B., 2006. The challenge of micropollutants in aquatic systems. Science 313 (5790), 1072–1077. https://doi.org/10.1126/science.1127291.

Singh, S.P., Azua, A., Chaudhary, A., Khan, S., Willett, K.L., Gardinali, P.R., 2010. Occurrence and distribution of steroids, hormones and selected pharmaceuticals in South Florida coastal environments. Ecotoxicology 19 (2), 338–350. https://doi.org/10.1007/s10646-009-0416-0.

Troxell, K., Ng, B., Zamora-Ley, I., Gardinali, P., 2022. Detecting water constituents unique to septic tanks as a wastewater source in the environment by nontarget analysis: South Florida's deering estate rehydration project case study. Environ. Toxicol. Chem. 41 (5), 1165–1178. https://doi.org/10.1002/etc.5309.

Ward, J.H., 1963. Hierarchical grouping to optimize an objective function. J. Am. Stat. Assoc. 58 (301), 236–244. https://doi.org/10.2307/2282967.

Wold, S., Esbensen, K., Geladi, P., 1987. Principal component analysis. Chemometr. Intell. Lab. Syst. 2 (1), 37–52. https://doi.org/10.1016/0169-7439(87)80084-9.

Yang, W., Tang, Y., Jiang, L., Luo, P., Wu, Y., Cao, Y., Xiong, J., 2022. Coupling suspect and non-target analytical methods for screening organic contaminants of concern in agricultural & urban impacted waters: optimization and application. Sci. Total Environ. 809, 151117 https://doi.org/10.1016/j.scitotenv.2021.151117.