Data processing strategies for LC-HRMS based non-target analysis of organic micropollutants in aqueous matrices

Dissertation

zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

- Dr.rer.nat. -

vorgelegt von

Lotta Laura Hohrenk-Danzouma

geboren in Hamburg

Lehrstuhl für Instrumentelle Analytische Chemie der
Universität Duisburg-Essen

2022

DOI: 10.17185/duepublico/77351
URN: urn:nbn:de:hbz:465-20230329-121733-4

Die vorliegende Arbeit wurde im Zeitraum von November 2017 bis Mai 2022 im Arbeitskreis von Prof. Dr. Torsten C. Schmidt am Institut für Instrumentelle Analytische Chemie der Universität Duisburg-Essen durchgeführt.

Tag der Disputation: 17.11.2022

Gutachter:	Prof. Dr. Torsten C. Schmidt
	Prof. Dr. Maryam Vosough
	Assoc. Prof. Dr. Emma Schymanski
Vorsitzender:	Prof. Dr. Jochen Niemeyer

Summary

A large variety of organic micropollutants (OMPs) are introduced into the aquatic environment and raise concerns due to their potential impact on ecosystems and human health. The high sensitivity and selectivity of liquid chromatography coupled with high-resolution mass spectrometry (LC-HRMS) enable the screening of a broad range of OMPs at trace concentrations without restriction on predefined analytes. Thus, LC-HRMS based non-target screening (NTS) approaches are of increasing importance in water analysis as they provide the potential to identify formerly unknown compounds and obtain a more comprehensive overview of pollution loads. However, with NTS large amounts of data are recorded within each measurement making sophisticated data processing strategies necessary. The first task of a data processing workflow is a reliable extraction of analyte signals, so-called features, from raw data. After this step complex datasets with thousands of features are obtained. Subsequently, it is essential to reduce and prioritize features that are relevant to the studied research question. This thesis addresses several aspects of data processing strategies, focusing on both the feature extraction step and feature prioritization step based on multivariate chemometric methods.

Nevertheless, high-quality measurement data are essential as a basis for the following data processing. A generic qualitative screening method was developed for an LC-HRMS analytical system. The sensitivity and selectivity to detect a broad range of OMPs at environmentally relevant concentrations and the stability of peak areas and retention times, enabling the comparison of several samples, were confirmed. On this basis, the importance of the feature extraction step was emphasized by first identifying weaknesses in the consistency of results obtained from different programs and secondly presenting an alternative chemometric-based approach.

The comparability of feature extraction with four different commonly used open-source and commercial software tools was investigated by examining the overlap of feature lists obtained from processing the same raw data set with MZmine2, enviMass, Compound Discoverer, and XCMS online. Results show a low coherence between different processing tools, as the overlap of features between all four programs was around 10%, and for each software between 40% and 55% of features did not match with any other program. The deviating implementation of filtering steps such as replicate- and blank filter was identified as one source of observed discrepancies. This comparison showed the necessity for higher robustness of data processing tools, a better understanding of algorithms as well as the influence of different parameter settings for each approach. Even though a general standardization of feature extraction is not

feasible, a higher awareness of the impact of this step and a transparent and detailed reporting of the entire data processing workflow were encouraged with this work.

In addition, an alternative chemometric feature extraction procedure based on regions of interest (ROI) and multivariate curve resolution alternating least squares (MCR-ALS) was employed on an NTS dataset of water samples for the first time. This approach circumvents several error-prone processing steps as there is no need for chromatographic alignment or grouping of multiple features of the same analyte. The approach was employed on samples with different complexity including a target data set of spiked drinking water samples and a NTS dataset obtained from different steps of a waste-water treatment plant and upstream of the receiving river. For all samples, chromatographic peaks and the corresponding mass spectra of OMPs were fully resolved in the presence of highly co-eluting irrelevant and interfering peaks. In the next step, features discriminating between several sample classes of the NTS data set were prioritized based on several multivariate and univariate chemometric methods. By that, from 101 resolved features (MCR-ALS components), 24 were selected and tentatively identified.

Subsequently, the benefits of feature prioritization based on complementary unsupervised and supervised pattern recognition methods were further examined. A comprehensive chemometric-based data processing workflow was employed on a dataset influenced by seasonal and spatial trends to reveal hidden data patterns and to find a subset of discriminating features between samples. Samples were obtained from a passive sampler monitoring campaign of three small streams and one major river over four sampling periods. After employing unsupervised explorative chemometric tools to obtain a general overview of samples, ANOVA simultaneous component analysis (ASCA) was used to disentangle the influence of spatial and seasonal effects as well as their interaction. A target and non-target dataset were compared and showed both a dominant influence of different sampling locations and individual temporal pollution patterns for each river. With the limited set of target analytes, general seasonal pollution patterns were apparent, but NTS data provide a more holistic view of site-specific pollutant loads. With a complementary partial least squares-discriminant analysis (PLS-DA) and Volcano-based prioritization strategy, 223 site and 45 season-specific features were selected and tentatively identified.

Overall, this thesis demonstrates that data processing is crucial in NTS to obtain meaningful results for a comprehensive environmental monitoring. For both feature extraction as well as prioritization remaining challenges and the capabilities of the implementation of advanced multivariate chemometric tools to enable the full exploitation of the potential of NTS were highlighted.

Zusammenfassung

Aufgrund ihrer möglichen Schadwirkung auf Ökosysteme und die menschliche Gesundheit ist es wichtig, den zunehmenden Eintrag von organischen Mikroschadstoffen (engl. OMPs) in die Umwelt zu überwachen. Dazu kann die Flüssigkeitschromatographie gekoppelt mit hochauflösender Massenspektrometrie (engl. LC-HRMS) genutzt werden, welche eine sensitive und selektive Erfassung eines breiten Substanzspektrum ermöglicht, ohne sich auf bekannte OMPs zu beschränken. Daher werden qualitative Verfahren, wie das sogenannte „Non-target Screening“ (NTS) zunehmend in der Wasseranalytik genutzt. Durch NTS können vorher unbekannte OMPs identifiziert und ein umfassenderer Überblick über die Spurenstoffbelastung eines Gewässers ermöglicht werden. Hierbei werden allerdings große Datenmengen generiert, die aufwendige Datenverarbeitungsmethoden nötig machen. Zuerst müssen Analyt-Signale, auch als „Feature“ bezeichnet, aus den Rohdaten extrahiert werden. Dabei entstehen große und komplexe Datensätze mit tausenden Features, die in einem nächsten Schritt reduziert werden müssen, um je nach Forschungsfrage relevante Features auszuwählen. In dieser Doktorarbeit werden verschiedene Aspekte der NTS Datenprozessierung, sowohl mit Fokus auf den Schritt der Featureextraktion als auch der Featurepriorisierung basierend auf multivariaten chemometrischen Verfahren, adressiert. Die Grundlage für die Datenprozessierung sind hochwertige Messdaten. Zu diesem Zweck wurde eine generische LC-HRMS Analysenmethode entwickelt und ihre Sensitivität und Selektivität für ein breites OMP-Spektrum in umweltrelevanten Konzentrationen, sowie die Stabilität von Messsignalen, geprüft. Darauf basierend wurde die Relevanz der Featureextraktion dargelegt, indem Schwächen in der Vergleichbarkeit verschiedener Prozessierungsprogramme aufgezeigt, sowie eine alternative chemometrische Methode präsentiert wurden.

Vier kommerzielle und frei verfügbare Programme zur Featureextraktion (MZmine2, enviMass, Compound Discoverer, and XCMS online) wurden verglichen, indem die gleichen Rohdaten prozessiert und die Überschneidung der erstellten Datensätze überprüft wurde. Hierbei zeigte sich eine geringe Übereinstimmung, da die Überschneidung aller vier Programme bei 10% lag und 40 bis 55% der Feature mit jeweils nur einem Programm detektiert wurden. Als eine mögliche Quelle für diese geringe Übereinstimmung wurde die abweichende Implementierung verschiedener Filterungsschritte (Replikatfilter und Blankabzug) identifiziert. Dieser Vergleich zeigt die Notwendigkeit für robustere Datenextraktionsprogramme, ein besseres Verständnis von Algorithmen und den Einfluss verschiedener Parameter für jedes Programm. Mit dieser Arbeit soll der mögliche Einfluss der Featureextraktion auf das Endergebnis hervorgehoben, und eine transparente und umfassende Veröffentlichung aller Prozessierungsschritte angeregt werden.

In dieser Arbeit wurde ein alternativer Ansatz zur Featureextraktion basierend auf den Methoden „regions of interest“ (ROI) und „multivariate curve resolution alternating least squares“ (MCR-ALS) zum ersten Mal auf NTS Daten von Wasserproben angewandt. Diese Methode umgeht mögliche Fehlerquellen der Datenprozessierung, da keine Anpassung von Retentionszeiten oder Gruppierung von verschiedenen Features, die zu einer Verbindung gehören, nötig ist. Die Methode wurde auf Proben mit verschiedener Komplexität angewendet, darunter dotierte Trinkwasserproben sowie unbekannte Proben entnommen aus verschiedenen Stufen einer Kläranlage sowie des Vorfluters. Chromatographische Peaks und Massenspektren konnten auch in komplexen Proben mit stark überlappenden und störenden Hintergrundsignalen identifiziert werden. In einem nächsten Schritt wurden verschiedene uni- und multivariate chemometrische Verfahren angewendet, um probenspezifische Features zu selektieren. Dadurch konnte die Anzahl von insgesamt 101 extrahierten Featuren auf 24 relevante reduziert und diese vorläufig identifiziert werden.

Anschließend wurden die Vorteile einer Featurepriorisierung basierend auf komplementären chemometrischen Methoden weiter untersucht. Hierfür wurde ein umfassender Datenverarbeitungsprozess auf einen von saisonalen und räumlichen Trends beeinflussten Datensatz angewandt, um verborgene Datenstrukturen aufzudecken und probenspezifische Features zur selektieren. Die untersuchten Proben stammten aus einer Überwachungskampagne mit Passivsammlern von drei Kleingewässern und einem größeren Fluss über vier Probenahmezyklen. Nachdem mit explorativen Verfahren ein allgemeiner Überblick über die Datenstruktur erhalten wurde, konnte mittels „ANOVA simultaneous component analysis“ (ASCA) der Einfluss von räumlichen und saisonalen Effekten und deren Wechselwirkung separat untersucht werden. ASCA wurde auf Target und Non-target Messungen der Proben angewandt und verglichen. Auch mit dem limitierten Subtanzspektrum der Target Messung konnten generelle saisonale Muster abgebildet werden, die NTS Daten zeigten jedoch ein umfassenderes Bild der Spurenstoffbelastung für jeden einzelnen Fluss. Nachfolgend wurden anhand einer komplementären Priorisierungsstrategie mittels „partial least squares-discriminant analysis“ (PLS-DA) und univariater Statistik 223 gewässerspezifische und 45 saisonspezifische Spurenstoffe selektiert und vorläufig identifiziert.

Insgesamt konnte in dieser Arbeit gezeigt werden, dass die Datenverarbeitung beim NTS von entscheidender Bedeutung ist, um aussagekräftige Ergebnisse zu erhalten. Herausforderungen der Datenprozessierung sowie die Vorteile der Nutzung multivariater chemometrischer Methoden für die Featureextraktion und Priorisierung wurden aufgezeigt, um in Zukunft das volle Potenzial des NTS ausschöpfen zu können.

Table of Content

Chapter 1 Introduction and theoretical background

1.1 LC-HRMS based screening of organic micropollutants in water

1.1.1 Organic micropollutants in the aquatic environment

There is a growing public and scientific concern about the presence of organic micropollutants (OMPs) in the aquatic environment and their potential impact on ecosystems and human health [1]. A variety of substance classes such as pharmaceuticals, pesticides, personal care products, and industrial chemicals, among others, have been detected by several studies in surface waters [2,3], groundwater [4,5], or even drinking water [6]. OMPs can enter the aquatic environment from both point sources, such as municipal wastewater effluent or industrial wastewater discharge [7], and non-point sources, such as agricultural runoff or street runoff [5,8]. OMPs are usually present at trace concentrations in the ng-µg/L range. However, several studies have reported adverse environmental effects such as loss of biodiversity [9], an increase of antibiotic resistance in microorganisms [10], or endocrine disrupting effects [11,12] amongst others related to the presence of OMPs. Global challenges such as climate change, growing populations, changing land use, and ongoing industrialization will potentially aggravate the problem of OMPs in the aquatic environment in the future [13–15].

Once the OMPs are released into the environment, they are subject to several biotic and abiotic degradation processes leading to their elimination or transformation into largely unknown substances, so-called transformation products (TPs) [16]. In addition, TPs can be formed during waste-water treatment, especially involving advanced oxidation processes, and released to the environment [17]. Some TPs are known to be more toxic and/or persistent and/or abundant in the aquatic environment than their parent compounds [16]. Therefore, it is necessary to monitor both, parent compounds and TPs. As stated, many of the TPs are unknown, and analytical reference standards are lacking which complicates the evaluation of their environmental risk.

In summary, the wide variety of different OMP classes and related TPs that are potentially present at trace concentrations creates a demand for highly sensitive and selective analytical methods that are capable to detect analytes of a broad range of physicochemical properties.

1.1.2 LC-HRMS screening methods

Recent developments in high-resolution mass spectrometry coupled with liquid chromatography (LC-HRMS) have initiated new possibilities for the analysis of micropollutants. Modern HRMS instruments provide accurate mass data while combining sufficient selectivity and sensitivity for the determination of trace substances in complex environmental matrices [18]. The information of all detectable compounds, including unknown or unexpected

substances, is available from the LC-HRMS data and can be used for more comprehensive monitoring. In the following, several general aspects of LC-HRMS data acquisition, which are relevant for the development of wide scope screening methods for OMPs in water matrices including data processing, will be discussed.

1.1.2.1 Chromatographic separation and ionization

The first step of a wide-scope screening method is the chromatographic separation of complex samples by high-performance liquid chromatography (HPLC). HPLC analysis is used for moderately polar to polar, low-volatile and thermally labile compounds [19]. It is thus the method of choice for OMPs in water matrices. The separation is based on the interaction of each analyte with a stationary (column) and mobile (eluent) phase. In water analysis, often polar embedded reversed-phase (RP) columns are used [20]. RP columns usually contain alkyl chains (most often C18) chemically bonded on silica. Furthermore, polar embedded RP stationary phases can be used for selective retention of polar compounds. These phases contain functional groups, such as amides or carbamates, embedded in the alkyl chains of RP columns [21]. As eluents, usually water and organic solvents are used with the addition of volatile additives to improve chromatography and/or ionization [22]. Complex matrices require a selective separation to reduce extensive overlapping of signals from analytes and/or matrix constituents [22]. Even though overlapping peaks with different m/z values can be separately detected due to the high resolving power used in HRMS measurements, co-elution of compounds can cause interference during ionization [23] and may hinder recording of MS^2 spectra and data processing. As highly polar OMPs are brought more into the focus of water analysis [24], HRMS screening methods with chromatographic separation based on hydrophilic interaction liquid chromatography (HILIC) phases [25], supercritical fluid chromatography (SFC) [26], or capillary electrophoresis (CE) [27] have been reported with the aim to broaden the analytical window.

1.1.2.2 Electrospray ionization

The interface between liquid chromatography and mass spectrometer must transfer the analytes from liquid to the gaseous phase, completely remove the solvent and ionize the analytes before entering the MS. For the coupling of HPLC and mass spectrometer (MS), ionization is generally performed by atmospheric pressure techniques, such as electrospray ionization (ESI), which is usually the method of choice for polar analytes [22] and is thus the most widely used technique in water analysis. ESI can be applied in either positive or negative mode. In ESI positive mode molecules are protonated $[M+H]^+$ or adducts such as $[M+Na]^+$, $[M+NH_4]^+$ and $[M+K]^+$ (besides other) are formed. In ESI negative mode molecules are deprotonated $[M-H]^-$ and rarely adducts such as $[M+HCOO]^-$, $[M+Cl]^-$ are formed. The analyte signal is further distributed over several different isotopologues, which are molecules that only

differ in their isotopic composition. For example, a molecule that incorporates one ^{13}C isotope at a position where an otherwise identical molecule has a ^{12}C atom, shows a mass difference of plus one Dalton (1.003355). For one molecule the probability of the occurrences and ratios of different isotope variants (isotope pattern) can be calculated based on known elemental isotope abundances [28]. Even though ESI is referred to as a rather soft ionization method, the formation of so-called “in-source fragments” is sometimes observed. These fragments are formed in the ESI source by collision of analyte ions with residual gas molecules leading to a neutral loss or dissociation into smaller ions. In addition, during the ionization process proton-bound dimers with formula $[2M+H]^+$ and their corresponding Na^+ and K^+ adducts can be formed or multiple charged ions, e.g., $[M+2H]^{2+}$ can be detected. Multiple charging is most often observed for large molecules (>1000 Da) but also in small molecules with multiple functional groups that are very prone to ionization [28].

In summary, the signal of the molecule of interest is often distributed over multiple entities in the mass spectrum, leading to a high complexity of ESI-MS spectra which is graphically visualized in figure 1-1. In addition, background signals from numerous sources, electronic noise or artefacts can be present [29].

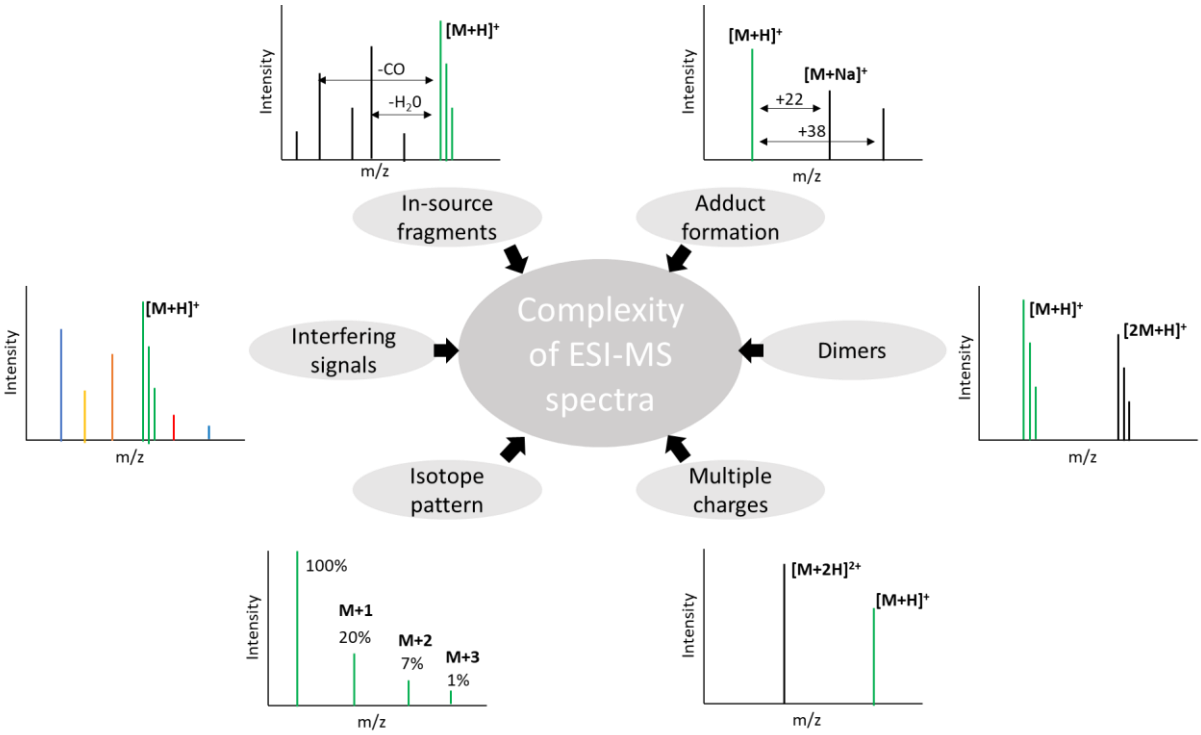


Figure 1-1: Overview of signals produced during electrospray ionisation leading to complex mass spectra.

The extent of adduct formation, in-source fragmentation etc. depends on properties of the molecule but can also vary depending on used instrumentation, measurement and ionization conditions like mobile phase, additives, voltages etc. [28]. Thus, the comparison of ESI-MS spectra is limited and interpretation not always unambiguous.

1.1.2.3 HRMS general terms

A few terms need to be defined when talking about HRMS measurements and data. In mass spectrometry, analytes are measured by their mass-to-charge ratio (m/z). The accurate mass is the experimentally determined mass of an ion of known charge and the exact mass is its theoretically calculated mass, both corresponding to a specific isotopic composition of the molecule. The accurate mass is calculated as sum of the exact masses of the atoms in a molecule using specific isotope masses of each atom instead of the isotope averaged atomic mass. It thus corresponds to the monoisotopic mass (including the mass defect) of a molecule in differentiation to the nominal mass, which is an integer number not including the mass defect. Further on, the average mass of a molecule takes into consideration the natural abundances of different isotopes [30]. For some common elements the nominal and atomic mass of different stable isotopes, their abundances and average masses are given in table 1-1.

Table 1-1: Nominal mass, isotopic mass and average mass of several isotopes of elements H, C, N, O and natural abundances of different isotopes [29,30].

Element	Isotope	Abundance	Nominal Mass	Isotopic Mass	Average Mass
H	^{1}H	99.9885	1	1.007825	1.00794
	^{2}H	0.0115	2	2.014102	
C	^{12}C	98.9300	12	12.000000	12.01100
	^{13}C	1.0800	13	13.003355	
N	^{14}N	99.6320	14	14.003074	14.00674
	^{15}N	0.3680	15	15.000109	
O	^{16}O	99.7570	16	15.994915	15.99940
	^{17}O	0.0380	17	16.999131	
	^{18}O	0.2050	18	17.999160	

For a better illustration of different terms, the nominal-, monoisotopic-, average- and exact mass have been calculated exemplarily for caffeine ($C_8H_{10}N_4O_2$) as shown in table 1-2.

Table 1-2: Exemplary calculation of nominal-, monoisotopic-, average- and exact mass of caffeine.

	Mass	Calculation
Nominal mass	194	= 8x12 + 10x1 + 4x14 + 2x16
Average mass	194.19316	= 8x12.011 + 10x1.00794 + 4x14.00674 + 2x15.99940
Monoisotopic mass	194.080376	= 8x12.0 + 10x1.007825 + 4x14.003074 + 2x15.994915
Exact mass $[M+H]^+$	195.08765	= 8x12.0 + **11**x1.007825 + 4x14.003074 + 2x15.994915 - 0.000549 **(1 electron)**

The mass error is the difference between a measured m/z value of an ion (accurate mass) by the MS and its theoretical m/z (exact mass) shown in equation 1-1. The term mass error is often used as a synonym with mass accuracy, which however describes an average of mass measurement errors as defined in equation 1-2 [31].

$$Equation\ 1\text{-}1:\ \Delta m_i = (m_i - m_a)\ in\ Da\ or\ \Delta m_i = \frac{(m_i - m_a)}{m_a} * 10^6\ in\ ppm$$

$$Equation\ 1\text{-}2:\ \overline{\Delta m_\iota} = \frac{\sum_i |\Delta m_i|}{n}$$

(with: m_i = measured accurate mass, m_a = calculated exact mass, n = number of measurements)

The resolving power describes the capacity of a mass analyzer to distinguish signals for two ions with a small m/z difference which can be expressed as mass resolution according to equation 1-3 [30].

$$Equation\ 1\text{-}3:\ R = \frac{m}{\Delta m}$$

(with m = m/z of ion of interest, Δm = difference of two m/z that can be distinguished)

Depending on the used definition, the mass difference *Δm* of two m/z can be measured at peak maximum, 5%,10%, or 50% of peak height. Since the resolution is always dependent on the mass at which it is determined, the specific mass and underlying definition should always be stated [23].

When complex samples are analyzed, both become highly relevant. With a higher mass resolution, higher selectivity and thus better separation of analytes from matrix constituents is achieved. By that, the error in accurate mass measurement can be minimized as fewer interfering ions are present [23].

1.1.2.4 Orbitrap mass spectrometer

A hybrid quadrupole-Orbitrap mass spectrometer (Q Exactive) operates as an ion trap mass analyser. In figure 1-2 a schematic representation of the Q Exactive model is shown.

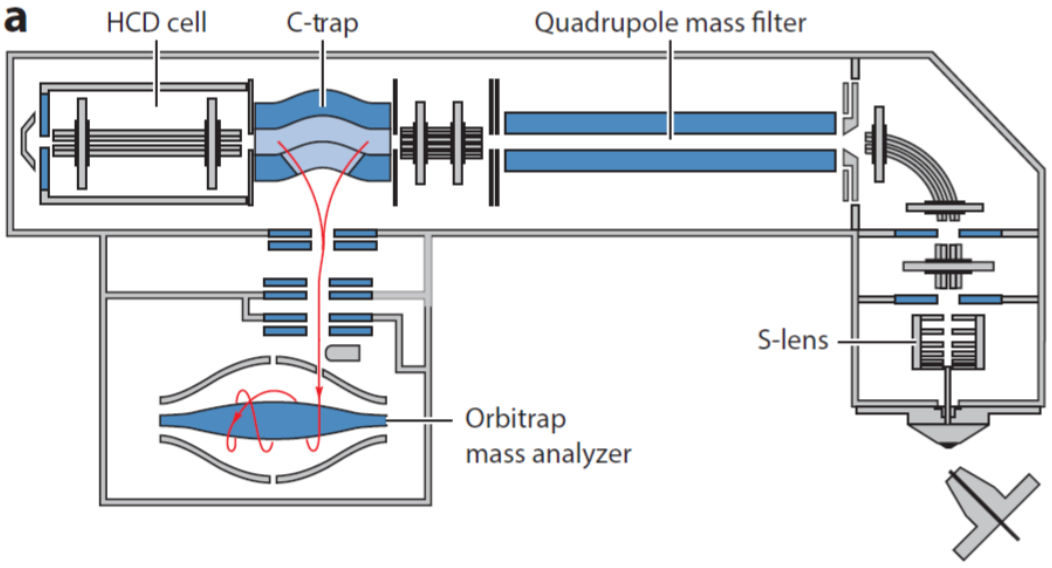


Figure 1-2: Schematic representation of the Q Exactive Orbitrap mass spectrometer [32].

The Orbitrap incorporates a so-called C-trap between quadrupole and Orbitrap analyzer. The C-trap is filled with ions until a certain limit is reached to prevent detector overloading and space charging effects. This limit is defined by the so-called automated gain control (AGC)

target value and/or a defined maximum injection time. After the accumulation of ions is stopped by a lens, the ion package is transferred to the mass analyzer [33]. Here, an axial oscillation is initiated by a strong electrical field. The ions oscillate harmonically around the inner electrode with a period proportional to (m/z)½ [26]. The obtained signal is subsequently converted from a time-domain into a mass spectrum by Fourier transformation [27]. The longer the transient signal is recorded, the higher is the resolution of the obtained mass spectrum [28]. For this reason, the Orbitrap mass resolution is inversely related to the scanning speed. Further, hybrid instruments, combining a high-resolution Orbitrap analyzer with a quadrupole mass filter or a linear ion trap provide the possibility to include MS^2 measurements. This is a technique where mass-selected ions (precursor) are subject to a fragmentation forming product ions which are also transferred to the Orbitrap mass analyzer [30]. Fragmentation is induced by collisions of ions with neutral gas atoms in a collision cell. In Q Exactive Orbitrap systems higher energy collisional dissociation (HCD) cells, located at the far side of the C-trap (see figure 1-2), are used. The precursor ions pass through the C-trap into the HCD cell, where dissociation takes place and the resulting fragment ions are then returned to the C-trap before injection into the Orbitrap for mass analysis.

The acquisition of MS^2 spectra can either be done in a data-dependent acquisition (DDA) mode or data-independent acquisition (DIA) mode. In the DDA mode precursor ions are selected data-dependent by the software based on their intensity or can be defined before measurement by an inclusion list. However, with this approach, MS^2 spectra are recorded only for a limited number of ions, thus for many low-abundance ions, no MS^2 are recorded. During DIA measurements, rather than selecting a precursor, sequential m/z windows are selected and all the ions within the window are fragmented [34]. Indeed, theoretically, MS^2 spectra for all precursor ions can be recorded with this method, opening the possibility of identifying more compounds at low concentrations. However, as the link between precursors and their fragment ions is disconnected, the deconvolution of resulting increasingly complex spectra can be a problem. Guo et al compared MS^2 acquisition methods and found higher spectra quality in DDA mode and a better MS^2 spectral coverage in DIA mode for a metabolomic case study [35].

Further instrumental developments and research on software tools and algorithms for enhanced deconvolution of DIA spectra are expected in the future as it provides an unbiased MS^2 acquisition and better reflects the holistic principle of NTS.

1.1.3 Non-target analysis

LC-HRMS measurements for quantitative and qualitative applications in water analyses can be classified into three different categories as shown in figure 1-3: Target analysis, suspect screening, and non-target screening/analysis [3,18,36].

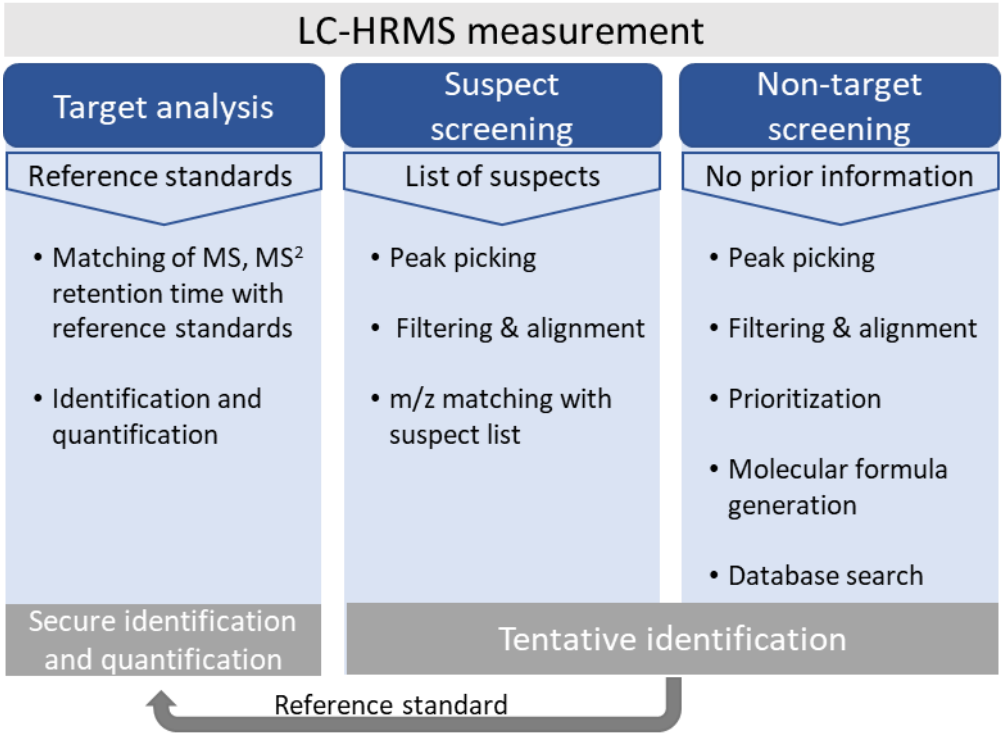


Figure 1-3: Schematic representation of data evaluation strategies used for LC-HRMS measurements including target screening, suspect screening, and non-target screening.

Conventional target analysis is used for the quantification of known (target) analytes with the help of reference standards. However, following this approach alone can result in overlooking relevant OMPs due to the preselection of the substances that are monitored. For the two qualitative approaches suspect screening and non-target screening (NTS) no reference standards are needed beforehand. During suspect screening, lists of compounds are searched by their exact mass, derived from the molecular formulas of known or predicted compounds of interest. This gives the possibility to search for a large number of compounds without the need for reference standards [36]. Suspect lists can be assembled based on different research questions and include e.g. less considered compound classes or predicted compounds as TPs. However, matches from measurement data and suspect lists need to be confirmed with additional information such as MS^2 spectra, fragment ions, isotope patterns etc [37]. NTS approaches do not use any prior information and are therefore not restricted by any preconceived selection criteria other than the technical restrictions imposed by the selected chromatographic and/or mass spectrometric system. Thus, NTS provides a more comprehensive overview of the compounds present in a sample. Previously unknown OMPs such as TPs or unexpected compounds e.g. introduced by spillings etc. can be potentially detected. In addition, NTS provides the possibility to compare samples or e.g. monitor water treatment processes, etc. without identification of all measured signals in the first place. As mining of HRMS data can be performed retrospectively as well, NTS can be used as a digital archive, by that newly discovered or previously not considered substances can be evaluated

after years [38]. It has to be noted, that besides the term “non-target screening” or “non targeted screening” additional expressions such as “non-target analysis” or “untargeted screening/analysis” are used in literature. In some cases these terms are used as synonyms to the approach described here or are used with a broader understanding including untargeted data acquisition methods like DIA [39].

NTS and suspects screening are often used in a complementary way. For both methods, reference standards can be used retrospectively to confirm the identification of a compound. In addition, both NTS and suspect screening have been extended to quantitative screening approaches where concentration ranges are reported instead of precise values and matrix effects are not considered. By that a large number of substances can be analyzed in a single run, important quantitative estimates are generated and in selected cases a fully validated target quantification can be done subsequently [40].

However, an extensive data processing workflow is necessary to extract the information of interest which will be discussed in detail in the second part of the introduction. In addition, several reported prioritization strategies for both suspect screening and NTS are summarized in part 1.2.2 to highlight possible fields of application.

1.2. Data processing strategies in NTS

The most time-consuming and challenging step of a NTS workflow is data processing and interpretation. The data processing workflow comprises several steps that vary depending on the research question, used software tool, etc. The data processing workflow can be distinguished into two parts: data *pre-processing* or *feature extraction* and data *post-processing* or *data analysis*. While the former includes extraction and filtering techniques to generate a feature list or data matrix from raw data, the latter mainly consists of data interpretation, trend or pattern detection, feature selection, and identification. Different automatic data processing methods, software tools, and chemometric methods can be used for all these steps, introduced in the following. An overview of different data processing steps is given in figure 1-4, and subsections of part 2 of the introduction are oriented on this structure.

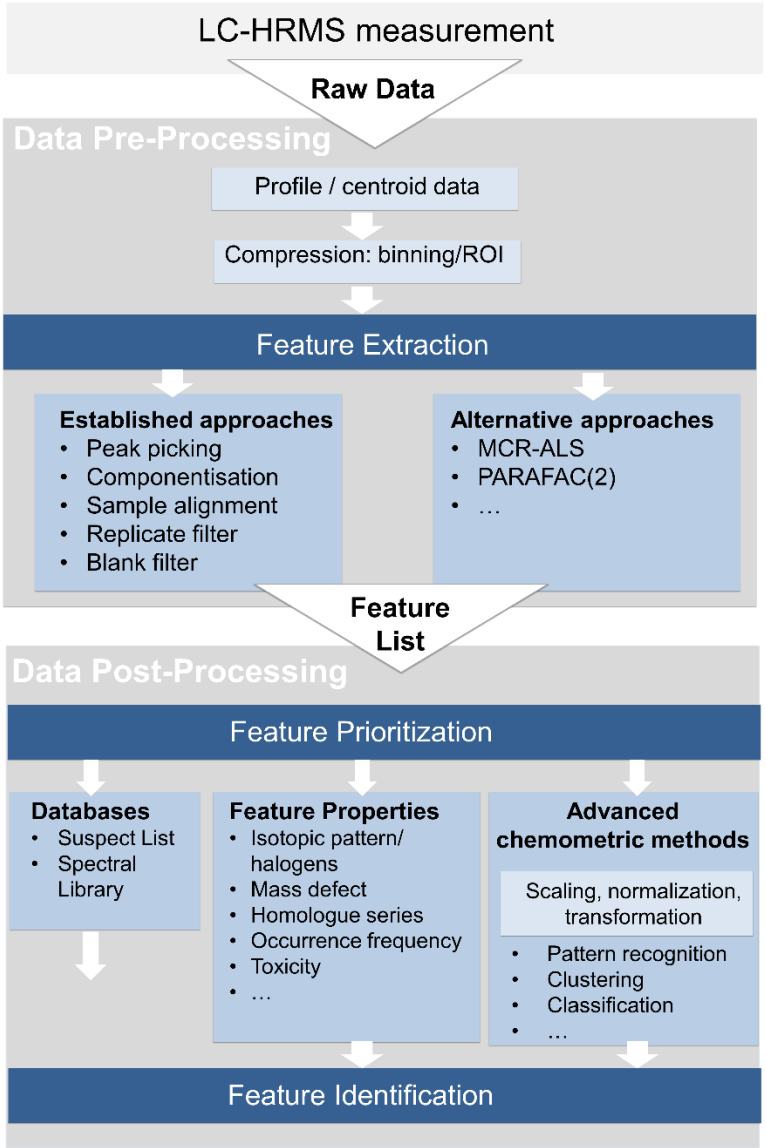


Figure 1-4: Overview of NTS data processing workflow: from data acquisition to compression, feature extraction with conventional software-based or alternative chemometric approaches, feature prioritization based on databases, feature properties, advanced chemometric strategies, and subsequent feature identification.

1.2.1 Data pre-processing – generation of feature lists

In each sample, a variety of ionizable substances is present. These can further form adducts or fragments during the ionization process, leading to thousands of detectable signals. A large amount of information is recorded within each measurement, making sophisticated automatic data processing strategies necessary. Some general strategies will be introduced shortly from the wealth of different approaches.

1.2.1.1 Profile- and centroided data

The ions associated with one compound reach the mass spectrometer ideally in a gaussian-like distribution, the mass spectrometer can record these signals either in continuous mode as profile data, retaining its structure, or convert them during acquisition into centroided data. During centroidization, the different data points of one profile are summarized as a location-intensity pair, e.g., based on mean/median or local maximum [41]. However, next to the reduction of data complexity, centroiding also leads to the potential loss of information or even in erroneous m/z value attribution. In contrast, all available information related to the distribution of a certain mass is kept for profile data, but data files are larger and more complex than compressed centroided ones. Therefore, often data are recorded in profile mode and are converted into centroid data in a subsequent step. This complexity reduction step is necessary to transform data into open-access formats such as mzXML, a standardized interface to most data extraction algorithms [42].

1.2.1.2 Data compression

The next step of data processing usually incorporates a data compression step by binning or searching for regions of interest (ROI).

Binning: This method merges individual mass spectra (vector of m/z – intensity pairs) into a two-dimensional matrix with RT and m/z as column and row headers. For this purpose, the m/z dimension is divided into pre-defined sections, and signals within are summarized. This technique facilitates later feature detection but has the drawback of potential loss of spectral resolution in the m/z domain. Further, if the bin size is not chosen adequately, chromatographic peaks can alternate among bins and might not be detected, or coelutions might be missed [42].

Regions of interest (ROI): The search for ROI is another approach for data compression/merging. First introduced by Stolt et al. [43] this method is based on observing that analyte signals occur in regions with highly dense data points surrounded by void regions. For the search of ROI, thresholds regarding signal intensity, mass error tolerance, and a minimum number of data points need to be defined. Stolt et al. employed the ROI approach on LC-HRMS data of urine samples and extracted ~1200 peaks, including only ~4 ‰ of the

total data but corresponded to 94% of the total variance [43]. The ROI approach was incorporated in the *centWave* peak picking algorithm used in different software tools such as *XCMS*. This principle of data compression has also been used in combination with multivariate curve resolution-alternating least squares (MCR-ALS), which will be discussed in section 1.2.1.4 [44].

1.2.1.3 Established feature detection approaches

Several open-source and commercial software tools are available for data processing, which are constantly improved and updated, and new algorithms and tools are published. In general, the information from MS detection (m/z and intensity) and chromatography (RT=retention time) is combined into a three-dimensional entity – a so-called feature. Instrument vendors usually provide commercial tools restricted to their specific data formats, e.g., *Compound Discoverer* for Thermo Scientific, *UNIFI* for Waters, or *SCIEX OS* for Sciex instruments. However, used algorithms are not published, and data sharing and transparency are limited. Open-source tools include *MZmine2* [45] and several R-based programs such as *enviMass* [46], *XCMS* [47], and its online version [48], or R-based in-house solutions [49]. Recently, *patRoon* has been introduced as an open-source software platform where several algorithms are incorporated for individual NTS workflows [50].

Different software packages use different algorithms with various underlying approaches for feature extraction, e.g., based on iterative clustering of signals, using a decision tree, peak-shape model fit, etc. However, many thresholds have to be chosen regarding intensity, deviations of m/z and RT, number of consecutive data points/scans, peak shape (usually assuming a gaussian shape), etc. In addition, a variety of clean-up and smoothing steps can be included [41,51].

1.2.1.3.1 Componentization

Each analyte can produce various signals during electrospray ionization as visualized in figure 1-1. These groups of m/z signals, all related to one unique molecular structure, can be summarized to reduce the amount of data and collect all available information, which can be helpful for subsequent identification. In the case of adducts and isotope signals, peaks with theoretical mass differences based on knowledge of commonly occurring adducts or isotopes and with identical RT can be systematically searched [41]. However, this procedure is not possible in the case of in-source fragments or less commonly observed adducts. Here in some cases, the similarity of elution profiles of the chromatography is used. For example, the *CAMERA* package has been introduced, integrating multiple methods for grouping related features and using a dynamic rule table to annotate ion species [52]. However, signal grouping sometimes referred to as componentization, is still a developing field and has not yet been implemented in all available programs. Without componentization, however, a high proportion

of redundant signals can remain in the feature list, leading to extremely large and hard to handle data sets. At worst, identification efforts are put into several signals obtained from the same analyte.

1.2.1.3.2 Alignment

Due to natural variations, comparing datasets from different samples, replicate measurements, or blanks requires feature lists aligned by retention times and m/z values. For this reason, the stability of chromatography and MS detection should be as high as possible and evaluated, e.g., with quality control samples or internal isotopically labelled standards. However, instrumental deviations cannot be entirely avoided, and different alignment algorithms have been introduced, usually using defined thresholds for m/z and retention time (RT) deviations. For example, an alignment method was introduced based on the random sample consensus (RANSAC), e.g., used in *Mzmine2*. RANSAC is an iterative method used to determine the optimal parameter of a mathematical model for peak alignment [45]. In some cases, a recalibration of m/z or RT has been employed, e.g., based on internal standards. However, as differences may be non-linear, this step poses a high risk of false data manipulation. Other, more sophisticated alignment algorithms have been introduced. For example, *obiwarp* is an approach (e.g., included in *XCMS*) based on a non-linear retention time alignment, independent from internal standards. It is based on identifying groups present in all samples and evenly distributed over the chromatographic profile, calculating retention time deviations of these groups, and aligning all features using a non-linear warping [47].

1.2.1.3.3 False positive and false negative rates

Robust and accurate detection of features is a big challenge in data processing. Each automated peak detection has to distinguish between signals and noise. Especially for low abundant signals, this can be challenging. Thus, many false positive (FP) detected features are inevitably caused by incorrect chemical or random noise integration. Chemical noise is caused by background signals present in eluents, buffers, or laboratory air or can be introduced during sample preparation and measurement [51]. Random noise is produced during MS detection, e.g., as residuals of the *Fourier Transformation* function in the Orbitrap mass analyzer [53]. The number of FP features can be minimized by optimizing thresholds of different settings during feature detection. However, if too strict settings of thresholds are applied, true positive signals can be missed, increasing the proportion of false negatives (FN). Due to this contrary behavior of FP and FN, parameter optimization is always a trade-off that needs to be evaluated. However, data quality evaluation is not trivial because the number of real (positive) signals in a sample is unknown in NTS screening. Usually, the absence/ presence of spiked analytes or internal standards is evaluated to determine FN rates, and a

subset of features is visually inspected for FP rates [41,54]. Several filtering steps can be implemented to reduce FP detected peaks, summarized in the following part.

1.2.1.3.4 Replicate filtering and blank subtraction

Measuring each sample in technical replicates allows filtering out instrumental noise. If a signal is recorded twice or three times in a row, chances are higher that it is a true positive (TP) signal. Bader et al. 2016 recommend using two or three technical replicates to enhance data quality [55]. However, for low intensity signals with a low S/N ratio, a TP signal might fall just below a certain threshold for one replicate measurement and thus be wrongly filtered out (FN) by a replicate filter.

In addition, the correction of data with a blank measurement is often used to reduce the number of FP signals introduced during sampling, transport, sample preparation, and measurement. For this purpose, a system, field or extraction blank should be measured, depending on the study, type of sample, and pre-treatment [54]. Signals found in these blank samples can then be subtracted from the feature list or filtered out with a specific ratio compared to the corresponding sample signal.

1.2.1.4 Alternative feature extraction approaches

1.2.1.4.1 Multivariate curve resolution by alternating least squares

Multivariate curve resolution by alternating least squares (MCR-ALS) is based on the premise that the measured variation in all samples can be described by a set of so-called "components" (note that the term "component" is used with different meanings in different areas, it will be further on described as MCR-ALS component). These MCR-ALS components combine resolved pure mass spectra and elution profiles of full scan measurements for LC-MS data. MCR-ALS is based on equation 1-4 [56]

$$\textit{Equation 1-4: } \boldsymbol{D} = \boldsymbol{C}\boldsymbol{S}^{\boldsymbol{T}} + \boldsymbol{E}$$

D is the original data matrix (I x J) (compressed by binning or ROI approach) with MS spectra at all retention times (i=1, …I) in rows and m/z channels (j=1, …J) in columns. This matrix is decomposed into matrix **C** (I x N) containing the column vectors of the elution profiles of N components and matrix $\mathbf{S}^{\mathbf{T}}$ (N x J) containing the row vectors of pure mass spectra of N components. The part that the model does not explain remains as residual matrix **E** (I x J) [56].

MCR-ALS can also be applied to decompose several data matrices, also called an augmented or multiset data matrix. MCR-ALS can be applied to data that has been compressed by binning. However, the combination with compression by search of ROI has shown to be advantageous. For this combination of methodologies, also called ROIMCR (in the following referred to as ROI/MCR-ALS), no reduction of spectral resolution occurs since there is no fixed bin size. For column-wise augmentation arrangement, different data matrices with common m/z ROI are

stacked on top of each other. First matrices D_1 and D_2 are augmented to D_{aug1}, again augmented with D_3 to form D_{aug2} and so on. Through this arrangement chromatographic profiles of common components in different samples do not need to be aligned. This property which is called “non-trilinearity condition”, makes ROI/MCR-ALS the most flexible and robust compared to other decomposition methods [44]. Different steps of the data processing workflow ROI/MCR-ALS on augmented matrices are shown in figure 1-5.

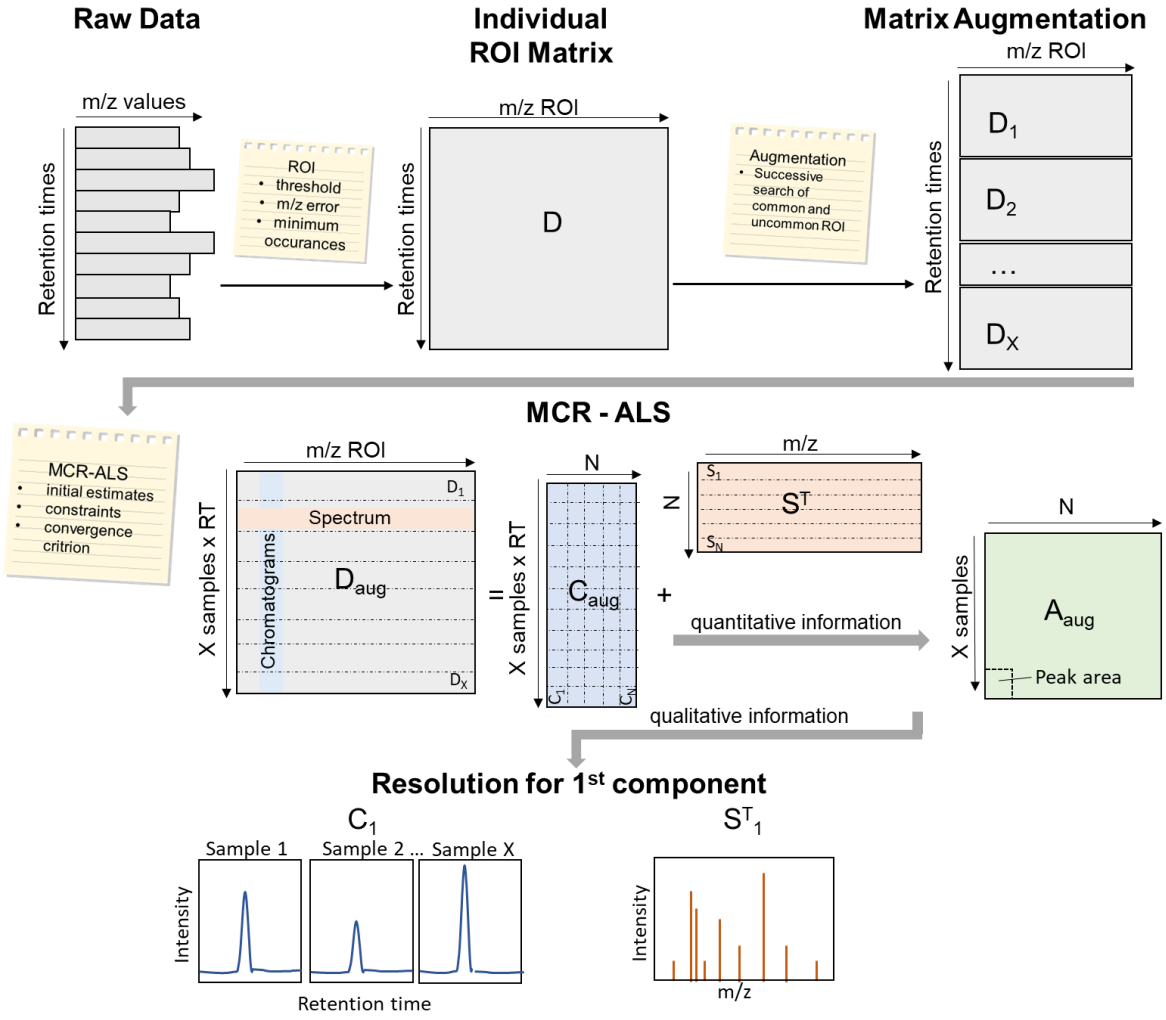


Figure 1-5: Overview of different steps of ROI/MCR-ALS processing pipeline: Raw data compression via ROI procedure and matrix augmentation followed by MCR-ALS resolution of components. Figure modified after [44,57]. Abbreviations: ***Aaug****: matrix of peak areas of N components and X samples;* ***Caug:*** *augmented column vectors of the elution profiles of N components* ***D****: data matrix,* ***Daug:*** *augmented data matrices;* ***N:*** *number of components;* ***S^T****: row vectors of pure spectra of N components;* ***x:*** *number of samples.*

After these data compression and augmentation steps, a set of MCR-ALS components of both chromatographic profiles and pure spectra is determined by alternating least-squares optimization approach starting by a set of initial estimates in an iterative process.

Initial estimates should not be made randomly but can be derived from previous knowledge on the data or with the help of different methods such as, e.g., Simple-to-use Interactive Self-

modeling Mixture Analysis (SIMPLISMA) or Evolving Factor Analysis (EFA): SIMPLISMA aims to select the most dissimilar rows or columns in a single data matrix or a multiset structure. EFA investigates the rise and decay of the components (used for chromatographic mode) [58].

Iterative optimization cycles are continued under the action of suitable constraints until a convergence criterion is fulfilled. Different constraints can be applied to the measurement data's natural properties or mathematical conditions or models [59]. The most known natural constraints are:

- *non-negativity:* forces profiles to be formed by positive values,
- *unimodality*: the presence of a single maximum per profile is allowed (peak shape),
- *local rank/selectivity:* selectivity is given if subregions of the profile exist where only a single component is present (rank one) beside windows where it is absent. For subregions where only some components are present, the information on local rank can be used [56].

However, several coeluting and overlapping peaks are usually present for LC-HRMS data obtained from generic NTS measurements. In addition, peak shapes are not always ideal, e.g., shoulder peaks are present. Thus, *local rank/selectivity* and *unimodality* cannot be applied without restrictions on LC-HRMS data.

The convergence criterion can be a pre-defined number of iterations or a threshold value for the difference in fit improvement between consecutive iterations. Quality parameters related to the final model fit are explained variance (R^2) or the lack of fit (LOF), which relates to the residues. Nevertheless, different ambiguities can be observed when different combinations of resolved MCR-ALS components reproduce the original data set with the same fit quality. Thus, selecting initial estimates and appropriate constraints is crucial to obtaining meaningful and reliable solutions [59].

The outputs of this procedure are MCR-ALS components, each component associated with an elution profile and its corresponding pure MS spectrum. In addition, a matrix containing peak areas or intensities of all MCR-ALS components (columns) for each sample (rows) is obtained, which can be used for further data analysis.

1.2.1.4.2 MCR-ALS in the field of NTS

The resolved MCR-ALS components can be understood as an analog of a feature obtained by conventional feature extraction strategies as both correspond to the signal profile of a chemical compound. However, they differ as one feature is associated with a unique m/z while an MCR-ALS component can contain various m/z values. By that, later filtering steps for adducts or isotope peaks are not necessary. Another advantage is that peak alignment is not required,

which can be helpful if retention time among samples varies highly. Furthermore, complex matrix samples including many overlapping signals can be mathematically resolved by this approach exploiting the second-order advantage [60]. However, depending on the size of the data matrix, it might be necessary to apply several MCR-ALS models to particular time windows of the whole chromatogram to be able to resolve all peak signals of samples.

Navarro-Reig et al. compared the data processing with *XCMS* and MCR-ALS for a metabolomic dataset obtained from LC-MS analysis of rice tissues exposed to Cd and Cu with the help of different chemometric methods. They found very similar results using both approaches. However, they concluded that *XCMS* was more straightforward for non-experienced users and required less processing time, but MCR-ALS gave more robust results [61]. Moreover, the output of a ROI/MCR-ALS approach contains different matrices for elution profiles, MS spectra, and intensities which is different from a feature list with single RT and m/z values. This can complicate compatibility to subsequent data evaluation steps such as identifying with databases or suspect lists.

1.2.1.4.3 Additional LC-HRMS data resolving strategies

Apart from MCR-ALS, other methods such as parallel factor analysis (PARAFAC) and PARAFAC2 or trilinear decomposition (TLD) have been introduced. However, in contrast to MCR-ALS, these approaches require so-called data trilinearity (unique elution and spectral profiles of each compound in all samples apart from a scale factor). This requirement is usually not given for LC-HRMS data, where minor deviations are always present. Thus, the data must be preprocessed and aligned to restore trilinearity [42]. Nevertheless, PARAFAC2 has been applied to LC-MS metabolomics data of complex plant extracts, including overlapped peaks, retention time shifts, and low S/N ratios [62].

1.2.1.5 Research gaps and relevant aspects of feature prioritization for this thesis

It can be concluded that the feature extraction step is a challenging part of any NTS workflow due to the complexity of LC-HRMS data, including overlapping peaks, non-ideal peak shapes, and low abundant signals close to noise. Any presented approach must deal with these circumstances and is prone to false results. As errors during each process propagate through the further workflow and can compromise final data interpretation, quality assurance is essential. Further, numerous different approaches exist with high methodological discrepancies. In addition, many thresholds must usually be adequately selected for each method. However, quality assurance in data processing, including a transparent and comprehensive reporting of parameter settings, has not been the focus of NTS in the past. The consistency of results obtained with different approaches has not been extensively studied.

The ROI/MCR-ALS strategy has shown several advantages for LC-HRMS data resolution in metabolomic studies, such as omitting peak alignment, componentization, or modelling chromatographic peak shapes [44]. However, it has not yet been used in the NTS of OMPs in water samples. Furthermore, although all algorithms and a Matlab-based MCR-ALS toolbox are available with open access, the approach has not yet been incorporated into broadly used software tools or data processing platforms.

1.2.2 Data post-processing – data evaluation

1.2.2.1 Prioritization strategies in NTS

Environmental samples are complex mixtures of chemicals that produce datasets with thousands of features even after several data filtering steps. As a complete identification process is extremely labour-intensive, reducing and prioritizing compounds of interest is essential. One strategy is to systematically search for "suspected" compounds by matching m/z values with a list containing a collection of compounds of interest. These lists can be composed based on different research questions or considering different compound classes of interest. For example, Kiefer et al. performed a suspect screening of many pesticides and TPs in groundwater [63], and Günthardt et al. focused on a less-studied compound group, phytotoxins [64]. Besides suspect screening, for NTS a variety of approaches for feature prioritization have been introduced based on 1) feature properties like intensity [36], occurrence frequency [65], presence of halogens, or distinctive isotopic pattern [3,66], 2) compound concentration profiles like spatial variation [3] and time series analysis [67,68], 3) correlation with further data or experiments, e.g., effect-directed analysis [69], formation of transformation products [70,71] or process monitoring and evaluation [72,73]. In addition, due to the multivariate nature of data originating from HRMS, different state-of-the-art unsupervised and supervised multivariate chemometric methods can be employed to extract information of interest [74–78].

Different approaches can be applied directly and in combinations, e.g., starting with a suspect screening and continuing with time series analysis [39]. A great strength of the NTS approach is its flexible to many different research questions which is demonstrated by the diversity of different prioritization approaches that have been proposed so far.

However, the implementation of advanced chemometric approaches has only recently started to receive more attention, and its high potential to comprehensively analyze complex NTS datasets has not yet been fully exploited. Therefore, several relevant aspects of treating multivariate data are discussed, and unsupervised and supervised chemometric methods are introduced in the following.

1.2.2.2 Multivariate chemometric based prioritization

Different unsupervised and supervised pattern recognition or classification methods can be used for data mining in NTS studies. From the wide range of available methods, only a few relevant for this thesis will be shortly introduced.

Before applying chemometric methods, several data arrangement and preparation steps are necessary. Hereinafter, as the first step of data evaluation, it is recommended to perform exploratory analysis to evaluate data structure with the aim to 1) detect the presence of outliers, 2) recognize patterns in sample distribution, and 3) evaluate relationships between variables and classes [79]. Common methods are principal component analysis (PCA) and hierarchical cluster analysis (HCA). Subsequently, supervised multivariate classification methods, such as partial least squares discriminant analysis (PLS-DA), can be applied to understand different patterns and relations of samples and variables and select a subset of relevant variables. In addition, multivariate methods that take into account the experimental multi-factors and the associated variables to these factors or their interactions, like ANOVA-simultaneous component analysis (ASCA), can further provide an in-depth understanding of the dataset and reveal hidden relationships.

1.2.2.2.1 General steps of data pre-treatment

Chemometric data evaluation is started with arranging the data into a matrix with features in columns aligned over all samples in rows. As matrix entries, intensities, peak areas or, if replicate measurements are used, mean intensities/peak areas are displayed. (Note that the output of many feature extraction software tools is a transpose matrix with features in rows and samples in columns.) Secondly, before any chemometric method can be applied to the NTS data, confounding variations of ion intensities attributed to experimental, or measurement sources must be removed. Simultaneously, the relevant variation containing information on environmental differences among samples needs to be preserved. For this purpose, different data pre-treatment strategies can be applied either column-wise (variables/features) or row-wise (samples). The methods are often summarized in chemometric literature under the term “pre-processing”, however in the context of this thesis the starting point of chemometric data evaluation is meant.

Comparability of variables

If variables have different scales or the distribution of variables is skewed, several scaling or transformation pre-treatment methods can be applied to make them comparable [80]. As an example, for LC-HRMS data the intensities of different features and their variances across different samples can vary highly (heteroscedasticity), leading to the underestimation of lower intensity signals. The most common methods for data transformation are log and power transformation of each element [42].

In addition. variable scaling methods involve calculating statistics on each column/variable, usually either a measure of data dispersion (e.g. standard deviation (SD)) or a size measure (e.g. the mean) is used. A commonly used initial step is column centering (e.g. *mean centering*: dividing each column by its mean). In addition, different variable scaling methods can be applied, often after mean-centering, to obtain a matrix where all the columns have zero mean and unit variance. Autoscaling is an exceptionally common preprocessing method which uses mean-centering followed by division of each column (variable) by the standard deviation of that column. However, when dealing with noisy data, care must be taken using autoscaling; otherwise, the same weight is given to the noisy parts of the signal and the informative features. Commonly used methods that give less weight to noisy signals are *Pareto scaling* (uses the square root of SD), and *Poisson scaling* (uses the square root of the mean) [80]. Other scaling methods, besides others, that are applied to LC-HRMS are *Range scaling* (uses difference between maximum and minimum of each column) and *Vast scaling* (uses coefficient of variation) [81].

Comparability of samples

Different data pre-treatment methods can be used sample-wise to compare different sample matrices and/or reduce systematic or random variation introduced by sample treatment and measurement. These methods (sometimes referred to as normalization) can either be based on the intensity of internal standards as a reference or be data-driven, using mathematical approaches. The latter usually apply the sum, mean, or median of all features across a sample as a normalizing factor. However, data normalization should be used with care to prevent the introduction of artifacts into the dataset [81].

Different element-, column- and sample-wise data pre-treatment steps can be executed one after the other, and trial-and-error approaches are often common practice for deciding which method should be applied to eliminate or reduce unwanted variation influence [82].

Another issue that must be accounted for is usually a high number of missing values in NTS datasets. These gaps in the data matrix occur in case of unique or non-replicable features. Alternatively, the compound falls below the detection limit in part or is not integrated during feature extraction for some samples. However, most chemometric methods only operate robustly if the dataset does not contain missing values. Therefore, missing value imputation is used to fill these gaps while maintaining the data structure. Different approaches can be used for this purpose, including small value replacement and mean or median replacement [81]. Gap-filling is another strategy that intervenes in an earlier data pre-processing stage during feature extraction and alignment. For example, if a feature is detected in one sample, the algorithm searches retrospectively for a signal at the given RT and m/z in all other samples regardless of intensity thresholds. If no feature is found, noise is integrated instead, and gaps are replaced with this value [81].

1.2.2.2.2 Principal component analysis

The extensive data volume and its complexity impede recognizing patterns and relationships. Thus, many multivariate analysis methods aim to reduce data complexity. Often there is some correlation between variables, so some of the information is redundant. When nontrivial correlations between some, most, or all of the variables are present, principal component analysis (PCA) can be applied to reduce the amount of data [83]. It is often used to visualize latent data structures using graphical plots, so data interpretation based on all variables simultaneously is possible. The idea behind PCA is to transform a set of data with correlated variables into a set of uncorrelated principal components (PCs), which are a linear combination of the original variables. The PCA model is presented by an orthogonal decomposition of a data matrix **X** shown in equation 1-5 [84].

$$\textit{Equation 1-5: } \boldsymbol{X} = \boldsymbol{T}\boldsymbol{P}^{\boldsymbol{T}} + \boldsymbol{E}$$

Where **T** is the scores matrix and represents the objects (or samples) in a new reduced space, **P** is the loadings matrix and reflects the importance of each variable (or feature) in the projection. **E** is the residual matrix; it is to be made small and contains measurement and sampling noise [85]. For visual analysis, two-dimensional or three-dimensional projections of samples are usually constructed with the axes (PCs) as the abstract factors. The scores plot shows the inter-location and inter-connection of samples under investigation; it can be used for unsupervised data clustering. The loadings plot shows the influence and inter-connections of the variables in the data set. Scores and loadings can also be visualized together in a biplot. In figure 1-6 results for an exemplary PCA obtained from a model dataset of 30 samples and 42 variables are displayed. In figure 1–6A the scores plot of PC1 and PC2 is given, showing four different sample classes. As can be seen, e.g., sample class 3 is separated on the first PC and sample class 4 on the second PC from other sample classes. Figure 1-6B shows the loadings plot for the second PC, by that the discriminating properties of each variable for this PC can be evaluated based on positive and negative loading values. In this example, variables number 13 and 30 with high loading values for PC2 are discriminative (e.g., show high concentration values) for sample class 2 and variables number 11 and 19 showing negative loading values for PC2 are discriminative for sample class 4. This can also be visualized by a biplot as shown in figure 1-6C indicating the closeness of samples and variables in a two-dimensional space. Variables 13 and 30 are located in the upper part of figure 1-6C, close to sample class 2 and variables 11 and 19 are located close to sample class 4.

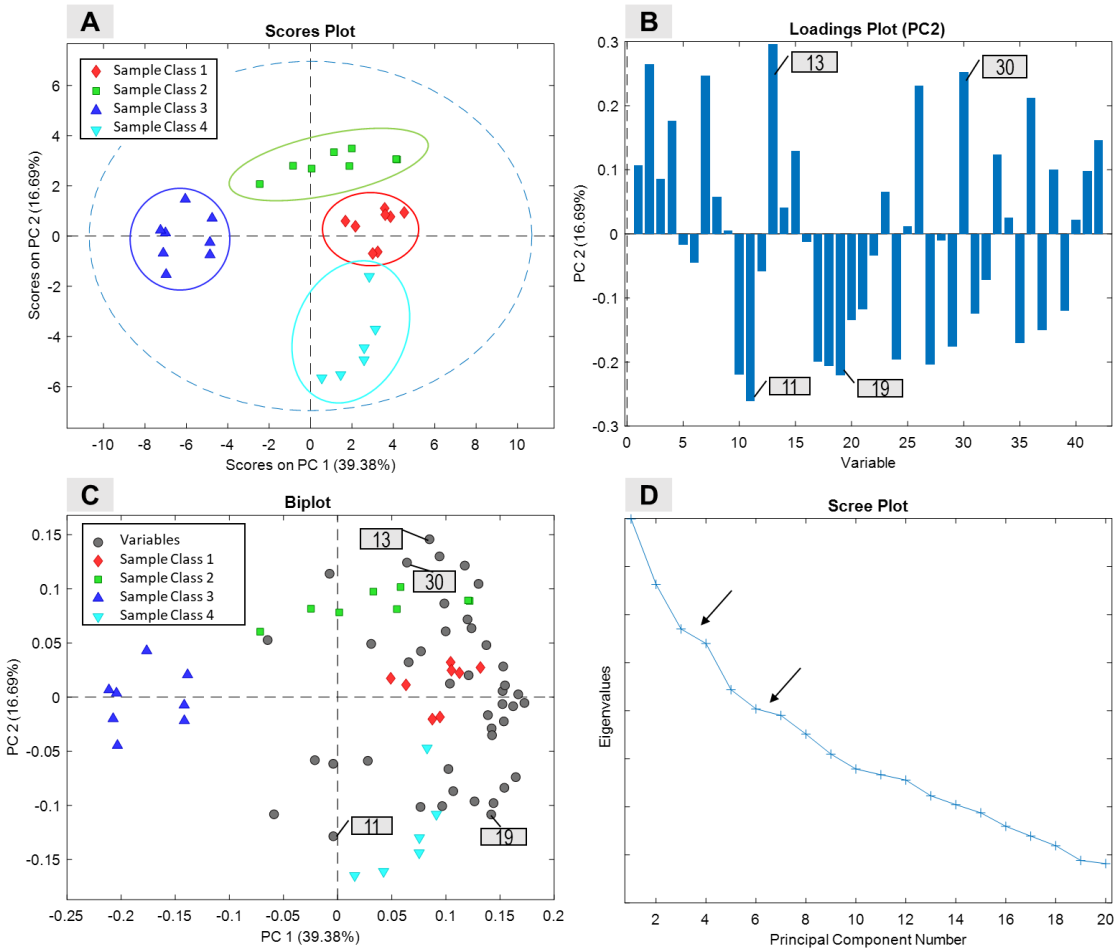


Figure 1-6:Examplarly results of PCA analysis of model data obtained from own measurements. A: Scores Plot (PC1xPC2); B: Loadings Plot (PC2); C: Biplot (PC1xPC2); D: Scree Plot

PCs are calculated so that the first PC accounts for most of the variation in the data set, the second PC for the second most variation, and so on. The primary challenge for PCA is the selection of appropriate PCs, securing that only non-relevant variation ends up in the residual matrix [84]. Several approaches exist, which are often related to the PCs' *eigenvalues*, which show the sum of squares of the scores of the corresponding PC. Successive *eigenvalues* of all PCs can be converted to percentages of the overall sum of squares of the data. A simple approach for having a rough estimation of PCs could be to retain PCs until the cumulative *eigenvalue* accounts for a certain percentage (e.g., 95%). A more straightforward method for determining the number of PCs to be retained is using a graphical representation known as a *scree plot* (e.g. see figure 1-6D) showing eigenvalues plotted against the corresponding PC number in descending order. By identifying “knees” where the curve flattens (indicated by arrows in the figure) the significant number of PCs to be extracted can be selected. The curve shows that an additional PC would add relatively little to the information already extracted. Cross-validation or bootstrap techniques can be applied for a statistically based estimation of the optimum number of PCA components, as well [85].

1.2.2.2.3 Hierarchical cluster analysis

Hierarchical cluster analysis (HCA) is an unsupervised method that visualizes the organization of samples in groups (clusters) and shows its hierarchy. The samples and their relationship are presented in tree form (dendrogram) that can be either agglomerative or divisive. The agglomerative (bottom-up) approach continuously merges the most similar cluster pairs until all points belong to the same cluster. Divisive clustering (top-down) works in reverse order, starting from a single cluster. The algorithm begins with searching for the best possible partition into two clusters, followed by another splitting, etc. Clustering is achieved with an appropriate metric of sample distance (e.g., Euclidean, Mahalanobis, or Manhattan distance) and linkage criterion among groups (e.g., complete, single, average, or Ward's linkage). HCA does not necessitate the number of clusters as an input argument. However, where to stop partitioning or merging must be specified instead.

HCA can be used to reveal inherent hierarchical structures hidden in the data set. It is, thus, similar to PCA, a good starting point for the exploration of a data set. A further advantage of this method is that the similarity or distance of clusters can be evaluated qualitatively and quantitatively. However, dendrograms become very extensive with large data sets, and their interpretation becomes exceptionally complex and challenging [86].

1.2.2.2.4 Partial least square discriminant analysis

Partial least square discriminant analysis (PLS-DA) is a supervised multivariate classification method that aims to recognize each sample's membership to an appropriate pre-defined class. This is achieved by a multivariate regression model between a matrix of independent variables (**X**, predictor variables) and an array of dependent variables (**y**, predicted variables). **X** contains the measurement data, and **y** contains binary dummy variables (0 and 1), indicating the class each sample belongs to. After a classification model has been calibrated, the membership of unknown samples to one of the defined classes can be predicted. In addition, in PLS-DA, the relevant sources of data variability are modeled by Latent Variables (LVs), which are linear combinations of the original variables. Thus, it allows visualization and understanding of the different data patterns and relations by LV scores and loadings. By that, interpretation on sample relationships (scores) and influence of each variable to the model (loadings) can be made similarly to PCA [79].

1.2.2.2.5 ANOVA-simultaneous component analysis

Factors like sampling location or time often influence environmental data sets. Different levels or sample classes can be distinguished for each factor, e.g., different rivers for the factor 'sampling location'. The resulting datasets contain groups of related samples, e.g., taken at the same river, etc. The dataset is of crossed nature if every level of one factor occurs at least once for every level of another factor. Analysis of variance (ANOVA) is commonly applied for

exploring the relationships between controlled experimental factors in biological or chemical studies (e.g., temperature or dose) and a single response (e.g., analyte concentration). It can separate the variability of this response across different samples into the different contributions of the experimental factors, i.e., if the concentration of analyte A changes significantly with chosen temperature or with applied dose, etc. [87]. However, a conventional ANOVA can be only applied to a single variable and is thus not useful for NTS data. ANOVA-simultaneous component analysis (ASCA) was presented by Smilde et al. 2005 as a multivariate extension to ANOVA [88]. It combines the power of ANOVA to separate variance sources with the advantages of simultaneous component analysis (SCA). In the first step, the whole data matrix (**X**) is decomposed according to the different sample factors (for experimental designs with two factors A & B) and the interaction of them (AB) as shown in equation 1-6 [88]:

$$\textit{Equation 1-6: } \boldsymbol{X} = \boldsymbol{X_m} + \boldsymbol{X_A} + \boldsymbol{X_B} + \boldsymbol{X_{(AB)}} + \boldsymbol{E}$$

where $\mathbf{X_m}$ is the matrix of means, $\mathbf{X_A}$, $\mathbf{X_B}$, and $\mathbf{X_{AB}}$ are the effect matrices of the different factors A and B and their interactions. **E** is the residual matrix and contains the variation that the model cannot describe [87]. In addition, the elements of **X** can be patinated in factor-specific sum-of-squares as shown in equation 1-7:

$$\textit{Equation 1-7: } \|X\|^2 = \|X_m\|^2 + \|X_A\|^2 + \|X_B\|^2 + \left\|X_{(AB)}\right\|^2 + \|E\|^2$$

These sums of squares can then quantify the percentage of the total variation in **X** that is explained by each factor and interaction. They are compared to the corresponding null hypothesis-based distribution value, estimated through a randomization test, to obtain p-values [88]. The obtained information is helpful to estimate which factor contributes to overall variation to what extent, if factors are independent of each other or if the interaction of them plays a role, and if these effects are significant.

The second major step of ASCA consists in examining the estimated effects for all variables simultaneously by SCA to each sub-matrix. In other words, one SCA model describing the overall effect of factor A, one SCA model for factor B, and another one describing the interaction factor A and B is built. SCA is a generalization of PCA for several related datasets sharing a set of measured variables. For example, $\mathbf{X_A}$, the effect matrix of factor A (equation 1-6), is decomposed according to equation 1-8 [87].

$$\textit{Equation 1-8 = } \boldsymbol{X_A} = \boldsymbol{T_A P_A^T} + \boldsymbol{E_A}$$

Thus, similar to PCA for each sub-model, scores $\mathbf{T_A}$ and loadings $\mathbf{P^T_A}$ matrices are obtained, and different patterns and relations of samples and variables can be visualized for decomposed data. For unbalanced data sets, where an equal number of samples for each factor of a crossed design are not possible, ASCA+ has been proposed as an extension [89].

1.2.2.2.6 Model validation

Obtaining reliable and robust results requires validating the performance of chemometric models. Initially, outliers in the model can be evaluated by indices, such as residuals of Q statistics and Hotelling's T^2. Q statistics are used to determine the size of the part of each sample not explained by the model. Thus, the smaller Q residuals are, the better the model fits the data. The Hotelling's T^2 values represent the variation in each sample within the model. It indicates how far each sample differs from the model's center (scores = 0). Confidence limits can also be established for Q and T^2 to visualize outliers [79].

Tuning and validating a model is highly important for supervised classification methods like PLS-DA. However, especially for high-dimensional data, over-fitting can become a problem.
The most common strategies for model validation are independent test set, cross-validation, and permutation test. Ideally, an independent (external) representative test set is available. In addition, many algorithms can be used to divide samples into training and test sets. However, this requires a high sample size, which is not always given [90].
Cross-validation is usually performed by dividing samples into different sets of training and testing groups, e.g., by leave-one-out, leave-group-out (by k-fold Venetian blind, for example), and Monte Carlo cross-validation. So, each cross-validation group is removed from the whole data set, one at a time. Finally, the model is calibrated based on the remaining (k-1) groups (selected as training samples) and then used to predict the testing groups and validating the performance of the model[79].
The permutation test is another powerful approach for validating regression and classification models. For this purpose, the class labels of samples are permutated randomly. A group of "wrong" models is built by repeating the permutation test numerous times. Subsequently, it can be tested if the results of the original model differ significantly from the perturbed model [90].

1.2.2.7 Variable selection

As a complete identification process is highly labour-intensive, reducing and prioritizing compounds of interest is essential. For this purpose, various approaches have been introduced, as discussed in part 1.2.2. In addition, an efficient selection or ranking of a subset of features that reveal similarities and/or differences between samples can be performed with multivariate chemometrics-based strategies [42]. After e.g., pattern recognition, uncovering relationships between samples, or classification of samples has been successfully performed, the primary question is usually which variables/features are responsible for these observations. As described above, the loadings matrix, e.g., of a PCA, PLS-DA, or ASCA analysis, gives information on the influence of each variable on the model. However, for large datasets, interpretation of loading plots can get rather complicated. Therefore, several variable selection methods have been proposed.

The variable importance on projection (VIP) score, first proposed by Wold et al. 1993 [91], can be calculated for each variable as a weighted sum of squared PLS variable weights. The VIP score for the j^{th} variable is calculated according to equation 1-9 [92].

$$Equation\ 1\text{-}9:\ VIP_j = \sqrt{\frac{\sum_{f=1}^{F} w_{jf}^2 * SSY_f * J}{SSY_{total} * F}}$$

(with w_{jf} = weight value for j variable and f component; SSY_f : sum of squares of explained variance for the f^{th} component; J : number of X variables; SSY_{total} : total sum of squares explained of the dependent variable; F : total number of components)

However, finding a suitable threshold of VIP for variable selection should be done with care to avoid missing relevant variables and to prevent the selection of false-positive variables. Usually, the average of the squared values of VIPs equals 1. Thus, VIP values greater than one are often considered to be relevant. Depending on the type of data structure, other threshold values can be used, e.g., concerning the dispersion of all obtained VIP scores like the upper quartile can be considered [93].

Another commonly used approach for variable selection is the selectivity ratio (SR). It is the ratio between each variable's explained and residual variance in a target projection. VIP and SR have been compared in various studies, with contradicting conclusions on the superiority of one of the methods depending on the type of data. Farrés et al. compared both methods on different data sets and observed a higher number of false-positive selected variables with VIP and false-negative selected variables with the SR approach [92]. Both approaches search for a minimal set of variables to discriminate between groups and are thus also called "sparse methods." Sparse methods have been critically discussed lately, as redundant or correlated variables can be ignored by these methods but might as well be relevant [94].

Besides multivariate variable selection methods, univariate statistical analysis can also be implemented on, e.g., intensity changes. Several commonly used methods include the coefficient of variation analysis, t-test, ANOVA, and volcano plots. With volcano plots, features are selected based on fold changes and p-values (significance level). This is especially useful when only two groups (e.g., with effect/without effect) must be compared. However, because univariate statistics necessitate significance testing of hundreds of features for non-targeted data sets, multiple test corrections should be considered as an integral part of these methods to protect against the increasing risk of false positives [95].

In conclusion, the selection of variables based on univariate or multivariate methods has a high potential to reduce data and prioritize relevant information in NTS studies, and only a small part of available methods has been discussed here. Nevertheless, as for any data processing step, methods and thresholds must be carefully chosen to avoid FP or FN findings.

1.2.3 Identification

The final step of a NTS workflow is identifying which OMPs correspond to the accurate mass and RT of the selected relevant features. For a secure identification a comparison with a reference standard is necessary, which is not always available. However, the high mass accuracy of HRMS can enable the determination of sum formulas, and by including further information available from MS and MS^2 spectra (molecular ion, isotope pattern, and fragments) and compound databases as well as meta information, molecular structure elucidation can succeed [38].

The first step during compound identification can be searching for m/z or derived molecular formulas in broad compound databases like *PubChem* or *ChemSpider* or specific databases including water-relevant OMPs such as *FORIDENT*. If MS^2 spectra are available, spectral libraries such as *mzCloud, Metlin* or *MassBank* can also be included.

An open-source platform including several tools for identification is *MetFrag* [96,97]. Potential candidates are retrieved by matching mass with compound databases. Subsequently, candidates can be scored according to how well the experimental spectrum matches with in-silico predicted fragments using a bond dissociation modelling approach. In addition, further information on retention time, availability of literature, patent information, or presence in suspect lists can be used for scoring candidates [98].

However, in some cases, all identification efforts remain without success. Problems arise, e.g., when large databases are used that do not only contain environmentally relevant compounds, resulting in too many or not meaningful candidates. In other cases, not enough information is available to distinguish several possible candidates. Ultimately, there exist compounds that are not listed in any chemical databases, e.g. most TPs, which have been referred to as “unknown unknowns” or “true unknowns” [39].

Several stages of identification confidence of unknown compounds can be derived from the available information and should be communicated, such as the confidence level proposed by Schymanski et al. [99] summarized in table 1-3.

Table 1-3: Level of identification confidence as proposed by [99].

Level	Identification confidence	Minimum data requirements
1	Confirmed structure (by reference standard)	MS, MS^2, reference standard
2a	Probable structure (by library spectrum match)	MS, MS^2, library MS^2
2b	Probable structure (by diagnostic evidence)	MS, MS^2, experimental data
3	Tentative candidate(s) (structure, substituents, class)	MS, MS^2, experimental data
4	Unequivocal molecular formula	MS isotope/adduct
5	Exact mass of interest	MS

The "experimental data" mentioned for levels 2b and 3 include e.g., diagnostic MS^2 fragments or ionization behaviour, knowledge on parent compounds, etc. When only one structure fits this information, but no standard or library MS^2 is available, it is categorized as level 2a, if e.g., several isomers exist it is categorized as level 3. The proposed categorization of identification confidence has been widely accepted in the NTS field and was, e.g., included with adaptions in a guideline for NTS in water [54].

Nevertheless, final identification of pollutants remains a bottleneck in non-targeted workflows and relevant potentially harmful features remain unknown in many cases. Thus, prioritized features that could not be elucidated should be reported, to preserve their information until better structure elucidation tools become available. Improved automated software solutions are necessary to make this processing step more time-efficient and reliable in the future. Furthermore, collaborations with researchers working in synthetic chemistry and incorporation of orthogonal techniques such as magnetic resonance spectroscopy (NMR) should be extended [39].

.

Chapter 2 Scope and aims of the thesis

Nontarget screening (NTS) based on high-resolution mass spectrometry (HRMS) coupled with liquid chromatography (LC) offers the potential for selective and sensitive detection of a broad range of organic micropollutants (OMPs) within one measurement. By that, a more holistic picture of pollution loads in the aquatic environment can be drawn, previously unknown compounds can be discovered, and several data mining strategies can be employed on NTS data, even retrospectively, to evaluate differentiation/similarity between samples without initially identifying each signal. Thus, NTS and suspect screening are becoming increasingly important approaches for environmental monitoring and evaluation of wastewater or drinking water treatment processes.

Due to the recording of large amounts of data and the generation of complex datasets with HRMS-based screening methods, sophisticated data processing strategies are necessary. As illustrated in the introduction in **Chapter 1**, data processing and mining is often not only the most time-consuming part but also a crucial step to obtain meaningful results to be able to exploit the full potential of NTS. The development of feature extraction algorithms, prioritization strategies, and identification approaches has become a major development field in the NTS area.

The scope of this work is thus focused on several aspects of the NTS data processing workflow as shown in figure 2-1. A NTS workflow usually consists of the following general steps 1) data acquisition with LC-HRMS 2) extraction of features from raw data 3) prioritization of relevant features and subsequent identification. This work aims to improve feature extraction strategies by highlighting the importance of this step, identifying weaknesses in the consistency of results obtained from different methods, and presenting an alternative chemometric-based approach. For the step of feature prioritization, the benefits of implementing multiple advanced chemometric approaches are illustrated. The contribution of the individual chapters to this thesis is highlighted in the workflow shown in figure 2-1.

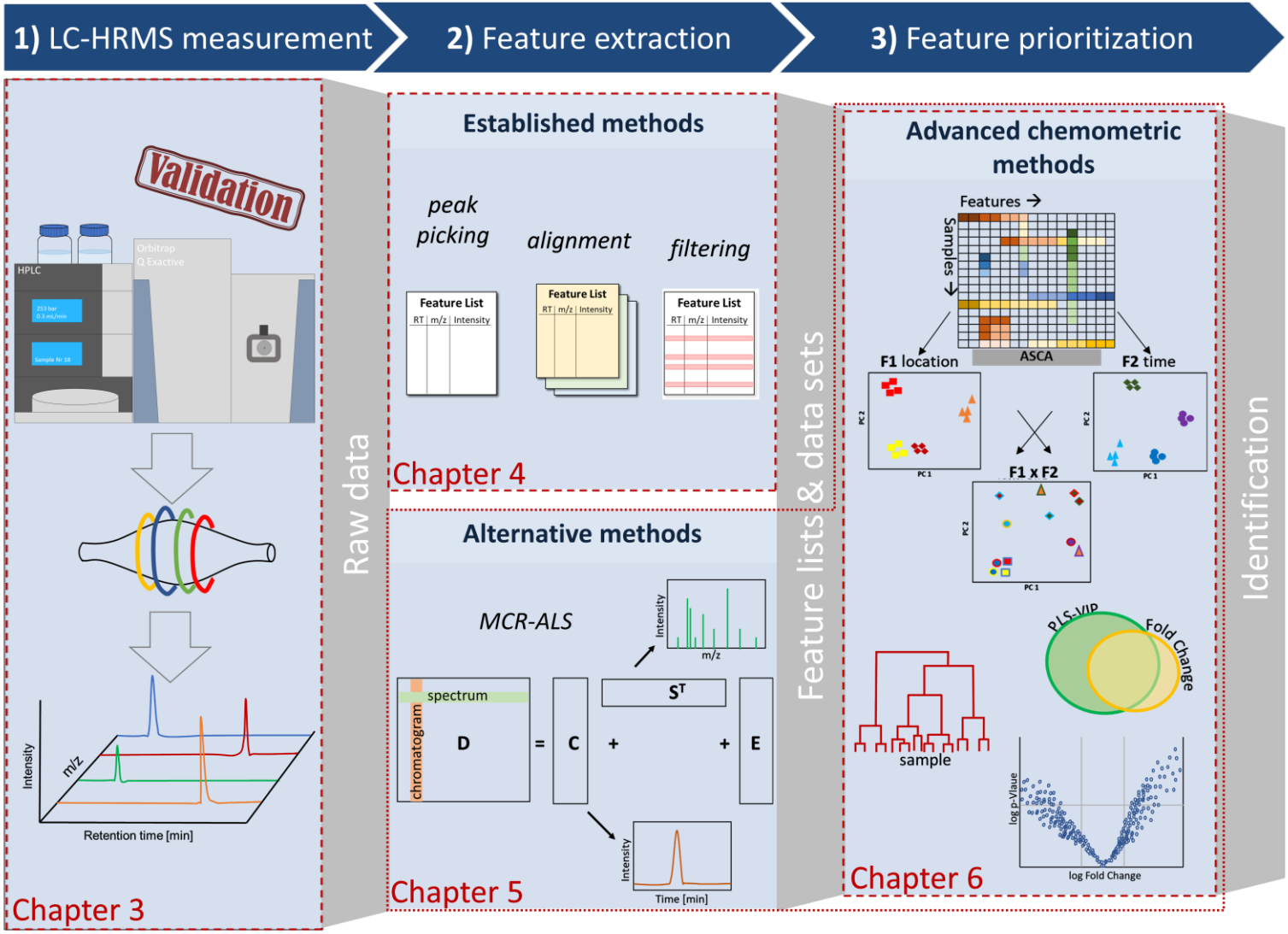


Figure 2-1: Graphical summary of the presented thesis. Chapters 3 to 6 discuss different aspects of NTS workflow from data acquisition to feature extraction and prioritization.

Chapter 3 deals with the development and validation of a generic LC-HRMS screening method (part 1 in figure 2-1) as the basis for data acquisition for the following work. Parameters of chromatography, electrospray ionization, and detection with an Orbitrap QExactive mass spectrometer were optimized for a wide scope of model OMPs. In the next step, the sensitivity and stability of measurements were validated to ensure an adequate method performance as required for NTS screening studies and meaningful data processing. In the following two chapters, several aspects of feature extraction (part 2 of figure 2-1) are highlighted. In **Chapter 4** feature extraction with different open source and commercial software tools is compared to investigate the consistency of resulting feature lists. For this purpose, the overlap of feature lists obtained from the processing of the same raw data set with four different commonly used tools is evaluated. In addition, the implementation of filtering steps such as replicate and blank filter was investigated as a source of observed discrepancies. In **Chapter 5** an alternative data processing procedure based on regions of interest (ROI) and multivariate curve resolution alternating least squares (MCR-ALS) is presented. This approach was employed on an NTS dataset of water samples for the first time. With this approach raw LC-HRMS data can be processed to simultaneously resolve chromatographic profiles and pure mass spectra. In a subsequent step, relevant features are prioritized based on univariate and multivariate chemometric methods and tentatively identified. The next chapter focuses on the feature prioritization part of the data processing workflow (part 3 of figure 2-1). In **Chapter 6** chemometric-based feature prioritization strategies for a comprehensive investigation of complex NTS datasets are presented. Multiple advanced chemometric methods were used in a complementary way on a set of target and non-target data to investigate both spatial and temporal trends. The dataset was obtained from a passive sampler monitoring campaign of small agriculturally influenced streams. By that, the influence of the diffusive introduction of OMPs by nonpoint source run-off into small streams was studied. Finally, **Chapter 7** provides an overall conclusion of this thesis and an outlook and discussion on future tasks and challenges.

Chapter 3 Development and evaluation of a generic HPLC-ESI-HRMS wide-scope screening method for non-target analysis

3.1 Introduction

For environmental monitoring and evaluation of wastewater or drinking water treatment processes suspect and non-target analysis is becoming an increasingly important tool [100]. These approaches require fast, sensitive and selective analytical methods for the detection of a broad range of organic micropollutants (OMPs). The combination of liquid chromatography with high-resolution mass spectrometry (LC-HRMS) has shown great potential for this purpose as it covers the detection of a wide range of OMPs [101]. As for any other measurement workflow, for non-target screening (NTS) various workflow steps from sampling, sample treatment and storage to chromatographic separation, ionization and detection can be combined. For the development of a generic qualitative screening method, it needs to be kept in mind that within each step, some analytes can be lost, leading to a limitation of the scope of the method. This can happen by various processes e.g. when analytes adsorb to sampling material or are not stable during transportation and storage, when they are not retained by the chromatographic column or not ionized or are not captured during sampling or sample preparation etc. In addition, contaminations, or background signals (e.g. plasticizer, cross-contamination etc.) can be introduced, which can complicate final data evaluation and should be minimized [102].

Commonly used methods for sampling and sample treatment include grab sampling or composite samples followed by direct large injection or solid-phase extraction, and passive sampling. The different approaches show advantages and drawbacks regarding enrichment of analytes, removal of matrix constituents, discrimination of analytes, representativeness of sample etc. [101]. A suitable procedure has to be chosen depending on the aim of the study. To limit potential loss of analytes or introduction of contamination into the samples, several NTS studies aim to keep sample treatment as limited as possible and use direct large injection after filtration [103], centrifugation [40,55] or vacuum-assisted evaporation [63]. The next part of a measurement workflow is a suitable chromatographic method that is capable to separate a wide range of analytes of different polarities. Usually reversed phase C18 columns with modifications to enhance retention of very polar OMPs are used [20]. As in NTS the analytes of interest are not known beforehand, method development has to be based on a range of model compounds and parameter settings suitable for the majority of compounds need to be selected. To evaluate the performance of chromatographic separation, several well-established parameters such as peak resolution R or retention factor k can be evaluated. To

asses peak shape quality measures like tailing factor T can be used [104]. In addition, for HRMS multicomponent methods an efficient exploration of the chromatographic space is desirable to keep measurement runs as short as possible and simultaneously prevent extensive overlapping of analyte signals and/or signals from matrix constituents [105]. Even though with the high resolving power used in HRMS measurements overlapping peaks with different m/z values can be separately detected, co-elution of compounds can cause interference during ionization [23] and may complicate data processing.

After chromatographic separation, ionization of analytes is necessary before mass spectrometric (MS) detection. For the coupling of LC with HRMS usually electrospray ionization (ESI) in both positive and negative mode is employed. Here again settings suitable for a range of compounds must be selected. Especially for the optimization of ionization voltages and temperatures a tradeoff between ion yield and ion fragmentation must be found [105]. HRMS screening methods require a MS detection of ions with a mass accuracy of at least 5 ppm [102] [23] to be able to identify unknown substances. Further, high quality MS^2 spectra are required for compound identification with a high confidence. In table 3-1 an overview is given of different ESI-HRMS detection methods employed on Orbitrap MS instruments reported in literature. For several laboratories or research institutes with frequent application of NTS only one or two representative studies are listed.

*Table 3-1 Overview of Orbitrap detection methods used in NTS studies. **AIF**: all ion fragmentation; **AGC target**: automated gain control target; **BWB:** Berliner Wasserbetriebe; **ddMS^2** data dependent MS2; **Eawag**: Swiss federal institute of aquatic science and technology; **loop count:** defines number of MS^2 cycle; **max IT**: maximum injection time; **min AGC target:** minimum number of ions required to trigger MS^2**; NCE:** normalized collision energy; **UFZ**: Helmholtz center for environmental research **WWU:** Westfälische Wasser und Umweltanalytik GmbH*

Institute	**NTS Guideline*[1]**	**Eawag**	**UFZ**	**WWU**	**BWB**
Reference	[102]	[63,64]	[68]	*[2]	[106]
Instrument	*n.i.*	QExactive Plus	QExactive Plus	QExactive	*n.i.*
Electrospray Ionisation					
		pos/neg	pos/neg	pos/neg	pos/neg
spray voltage [kV]	3.5	4.0/3.0	3.8/4.5	3.0	3.5
spray temperature [°C]	350	320	300/280	360	320
sheat gas [a.u]	15	15	45/25	40	40/40
aux gas [a.u]	50	40	1	15	20/20
s-lens	*n.i.*	50	70	55	50
Full scan					
mass range [m/z]	120-1200	100-1200	100-1000	80-1100	103-900
resolution	120,000	140,000	70,000	70,000	70,000
AGC target	3E+06	1E+06	*n.i.*	1E+06	1E+06
max IT [ms]	100	100	*n.i.*	100	200
MS^2 acquisition					
	Full MS/ddMS^2	Full MS/ddMS^2	Full MS/ddMS^2	Full MS/ddMS^2	Full MS -/AIF
mass range (only for AIF)					60-840
resolution	15,000	17,500	35,000	17,500	35,000
AGC target	1E+05	2E+05	*n.i.*	1E+05	5E+05
max IT [ms]	50	50	*n.i*	50	*n.i*
loop count	5	4	*n.i*	5	*n.i*
isolation window [m/z]	1.3			1.5	
NCE	80	15; 60; 105	35; 55	CE 20;50	40
min AGC target	8E+03	*n.i*	*n.i*	8E+02	*n.i*
apex trigger	3-10 s	*n.i*	*n.i*	*n.i*	*n.i*
exclude isotopes	on	*n.i*	*n.i*	off	*n.i*
dynamic exclusion	15 sec	4 sec	*n.i*	3 sec	*n.i*

**1: An exemplary method for Orbitraps is given in the supporting information of this guideline, the laboratory using this method is not stated *2: Parameter settings were obtained directly from WWU; n.i.= value/information not indicated*

As shown, by most studies a full MS/ddMS^2 TopN method is used. This method uses a full scan measurement from which the N most intense precursor ions are selected, isolated, fragmented and sequenced by a product ion scanning. Several parameters for MS detection and ionization can be varied. One crucial Orbitrap-specific parameter is the so-called

automated gain control (AGC) which is implemented to avoid detector overloading and space charging effects. The ion current within the mass range of interest is determined by a short pre-scan, by that the number of ions (the AGC target value) that should be stored in the C-trap for the subsequent analytical scan can be defined [32]. This parameter can enhance mass accuracy of the measurement by avoiding space charging effects but also influences the sensitivity of the detection especially for trace components in the presence of a complex matrix [23]. In addition, the mass resolving power is an essential parameter as it influences the scan rate of the measurement. With a higher resolution, a higher selectivity and thus better separation of analytes from matrix constitutes is achieved, however the acquisition time in the orbitrap mass analyser is longer which results in longer cycle times of the method. A higher cycle time compromises the number of data points that are recorded and thus can lead to a poor temporal resolution of chromatographic peaks [107].

During MS^2 recording Orbitrap QExactive uses the so-called normalized collision energy (NCE) for fragmentation. NCE is a dimensionless number that is approximately equivalent to the collision energy (in eV) for a reference ion of mass 500 and charge 1 used in the high collision HCD. The actual HCD energy is calculated on basis of mass and charge ratio (m/z) of the selected precursor ion [108]. An adequate collision energy is necessary to obtain MS^2 of high quality for identification. If the collision energy is too high the precursor ion is completely fragmented, in some cases fragments further break down into smaller parts and no meaningful fragments are left. If the collision energy is too low the precursor ion remains intact and no fragmentation pattern can be evaluated [109]. Thus, a balance has to be found, which is dependent on molecule properties such as size or charge, and therefore for unknown screening suitable generic settings must be selected.

After method development and optimization for a range of model compounds an evaluation of method performance is of interest for later NTS data processing and evaluation of results. The precision of a method is a measure on how close results of several measurements are to each other. It is usually assessed by repeatability, short-term precision of multiple replicate measurements under similar conditions and reproducibility, the differences of measurements made under different conditions e.g. different operators, instruments and large time frames [31]. In the context of this study, for the alignment of signals during data processing from different technical replicate measurements the repeatability and for measurements of different samples or batches the reproducibility of retention times and mass errors is of high interest to choose adequate thresholds. If peak areas of different samples should be compared, the repeatability/reproducibility of them needs to be known as well. For compound identification and database searches a general assessment of mass accuracy, which describes the average mass error, of the instrument is necessary.

In qualitative screening methods, there are no required detection limits that must be reached, but sensitivity has to be sufficient to detect OMPs at environmentally relevant concentrations. Only few substances are regulated by authorities on a national or European level. Nevertheless, for water samples detection limits in the lower ng/L range are required, as some compounds have shown adverse effects even at trace concentrations especially as they are most often present in mixtures [110]. Schulz et al. 2019 recommend a sensitivity of approximately 10 pg on column [102] .

Several broad range screening methods employed on Orbitrap or quadrupole time of flight (QTOF) MS instruments have been reported before [40,55,63,64,68,103,106], however parameter settings differ and depend on the analytical system. In this chapter, a generic LC-ESI-HRMS screening method is developed based on reported methods and adjusted for the given analytical system. Its performance is evaluated to ensure suitability for a broad range unknown OMP screening. The gradient elution program of the HPLC separation method is optimized and different chromatographic columns are compared based on a set of model compounds. Conditions of ESI and parameter settings of the Orbitrap detection method are optimized based on methods reported in literature. Subsequently, different settings of the NCE are compared. The reproducibility of peak areas, retention times and mass accuracy are evaluated to assess the precision of the method. The sensitivity of the method with an injection volume of 20 and 100 µL is evaluated based on limit of detection (LOD) calculated according to DIN 32456. The calculation based on DIN 3245 is further compared to (i) an LOD estimation from a signal-to-noise ratio (S/N) of 3 and (ii) the < (MDL) as defined by United States Environmental Protection Agency [111].

3.2 Methodology

3.2.1 Chemicals and reagents

Different parts of method development and evaluation were performed with different mixtures of model compounds (Mix A-D). Model compounds were selected to cover a broad range of OMPs, including pharmaceuticals, pesticides and TPs with a range of polarities and molecular weight sizes ionizable in positive and negative mode. Table 3-2 summarizes information of all used compounds and indicates which compound was part of each mixture. In addition, a standard mixture D containing 289 compounds was used in part 3.1.4 and 3.2.2 with all compounds listed in table S3-1. Standard solutions of each compound were prepared in methanol and stored at -20°C. Standard mixtures were freshly prepared before each measurement in ultra-pure water (UPW) or matrix.

Table 3-2: Overview of model compounds in different standard mixtures.

Compound	Formula	log D*	Ionization mode	Adduct	m/z	Mix A	Mix B	Mix C
17-α-Methyltestosterone	$C_{20}H_{30}O_2$	3.6	pos	$[M+H]^+$	303.2319	X	x	
1H-Benzotriazole	$C_6H_5N_3$	1.3	pos	$[M+H]^+$	120.0556	x	x	x
			neg	$[M-H]^-$	118.0411			
5,6-Dimethyl-1H-Benzotriazole	$C_8H_9N_3$	2.3	pos	$[M+H]^+$	148.0869	x	x	x
5-Chloro-1H-benzotriazole	$C_6H_4ClN_3$	1.9	pos	$[M+H]^+$	154.0167	x	x	x
			neg	$[M-H]^-$	152.0021			
Acesulfame	$C_4H_4KNO_4S$	-0.8	neg	$[M-H]^-$	161.9867		x	x
Acetanilide	C_8H_9NO	1.2	pos	$[M+H]^+$	136.0757	x	x	x
Acetylsalicylic acid	$C_9H_8O_4$	1.1	pos	$[M+H]^+$	203.0315	x	x	
Anthranilic acid	$C_7H_7NO_2$	1.3	neg	$[M-H]^-$	136.0404	x		
Atenolol	$C_{14}H_{22}N_2O_3$	0.4	pos	$[M+H]^+$	267.1703	x	x	x
Atrazine-desethyl	$C_6H_{10}ClN_5$	1.5	pos	$[M+H]^+$	188.0698	x	x	x
Atrazine-desethyl-desisopropyl		0.5	pos	$[M+H]^+$	174.0541	x	x	x
Atrazine-desisopropyl	$C_5H_8ClN_5$	-0.2	pos	$[M+H]^+$	146.0228	x	x	x
Atrazine	$C_8H_{14}ClN_5$	1.8	pos	$[M+H]^+$	216.1011			x
Bicalutamide	$C_{18}H_{14}F_4N_2O_4S$	2.7	neg	$[M-H]^-$	429.0538	x	x	x
Bisphenol S	$C_{12}H_{10}O_4S$	5.4	pos	$[M+H]^+$	251.0373	x	x	
Caffein	$C_8H_{10}N_4O_2$	-0.6	pos	$[M+H]^+$	195.0877		x	x
Carbamazepine	$C_{15}H_{12}N_2O$	2.8	pos	$[M+H]^+$	237.1022	x	x	x
Carbamazepine – D_8	$C_{15}D_8H_4N_2O$		pos	$[M+H]^+$	245.1525	x		
Catechol	$C_6H_6O_2$	1.4	neg	$[M-H]^-$	109.0295	x		
Chloridazon-desphenyl	$C_4H_4ClN_3O$	-0.8	pos	$[M+H]^+$	146.0116		x	x
Clofibric acid	$C_{10}H_{11}ClO_3$	2.8	neg	$[M-H]^-$	213.032	x	x	x
Diclofenac	$C_{14}H_{11}Cl_2NO_2$	4.2	pos	$[M+H]^+$	296.024	x	x	x
Diclofenac - $^{13}C_6$	$^{13}C_6C_8H_{10}Cl_2NNaO_2$		pos	$[M+H]^+$	324.026	x		
Ibuprofen	$C_{13}H_{18}O_2$	3.8	neg	$[M+Na]^+$	229.1199	x	x	
Ketoprofen	C16H14O3	3.6	pos	$[M+H]^+$	255.1016			x
Malic acid	$C_4H_6O_5$	-1.3	neg	$[M-H]^-$	133.0143	x		
Melamine	$C_3H_6N_6$	-2.6	pos	$[M+H]^+$	127.0727		x	x
Metformin	$C_4H_{11}N_5$	-3.6	pos	$[M+H]^+$	130.1087		x	x

Compound	Formula	log D*	Ionization mode	Adduct	m/z	Mix A	Mix B	Mix C
Metoprolol - D_7	$C_{30}D_{14}H_{36}N_2O_6 * C_4H_6O_8$		pos	$[M+2H]^{2+}$	275.2352	x		
Metoprolol	$C_{15}H_{25}NO_3$	-1.5	pos	$[M+H]^+$	268.1907	x	x	x
Paracetamol	$C_8H_9NO_2$	0.9	pos	$[M+H]^+$	152.0706	x	x	x
p-Nitrophenol	$C_6H_5NO_3$	1.6	neg	$[M-H]^-$	138.0184	x	x	x
Propazine	$C_9H_{16}ClN_5$	2.2	pos	$[M+H]^+$	230.1167	x	x	x
Saccharin	$C_7H5NO3S$	0.1	neg	$[M-H]^-$	181.9917		x	x
Simazine	$C_7H_{12}ClN_5$	1.3	pos	$[M+H]^+$	202.0854	x	x	x
Succinic acid	$C_7H_{12}ClN_5$	-0.5	neg	$[M-H]^-$	117.0193	x		
Sulfamethoxazole	$C_{10}H_{11}N_3O_3S$	0.8	pos	$[M+H]^+$	254.0594	x	x	x
Sulfamethoxazole - D_4	$C_{10}D_4H_7N_3O_3S$		pos	$[M+H]^+$	258.0845	x		
Tamoxifen	$C_{26}H_{29}NO$	2.9	pos	$[M+H]^+$	372.2322	x	x	x
Terbutylazin-desethyl	$C_7H_{12}ClN_5$	1.3	pos	$[M+H]^+$	202.0854	x	x	x
Toremifene	$C_{26}H_{28}ClNO$	2.8	pos	$[M+H]^+$	406.1932	x	x	x
Triclosan	$C_{12}H_7Cl_3O_2$	5.0	neg	$[M-H]^-$	286.9446	x	x	x

*logD values for pH 3 were obtained from ChemAxon (http://www.chemaxon.com)

In table 3-3 all other chemicals that were used for measurements are summarized. Several consumables were obtained from different suppliers over the time but always with LC-MS grade quality.

Table 3-3: List of used chemicals

Chemical	CAS	Supplier	Purity
Methanol	67-56-1	VWR; Fisher Scientific	LC-MS grade
Water	7732-18-5	Merck; Sigma Aldrich	LC-MS grade
Formic acid	64-18-6	VWR; Merck	LC-MS grade
Pierce ESI Negative Ion Calibration Solution	n.s.	Thermo Fisher Scientific Scientific	n.s.
Pierce LTQ Velos ESI Positive Ion Calibration Solution	n.s.	Thermo Fisher Scientific Scientific	n.s.

* n.s.: not specified

3.2.1 Exemplary surface water matrix

As exemplary surface water matrix, water from lake Baldeney was used. Sampling was performed in June 2017 from a sampling location in Essen-Heisingen at 24°C. The water was filtered with mixed cellulose ester filter (0.45 µm, 50 mm diameter) using a vacuum pump. After filtration, the matrix was transferred into 2.5 L amber glass shot-bottles. The shot-bottles were previously rinsed with methanol and UPW and dried in a heat cabinet overnight. The matrix was stored at 4 °C in the dark. For experiments performed for optimization of MS parameter (3.2.1) and precision of the method (3.2.1) mix A was spiked into matrix samples. All measurements including matrix samples were performed between June and September 2017 and the pH (8.0) and non-purgeable organic carbon (NPOC = 3.8 mg/L) were checked regularly (SI 3.5.2).

3.2.2 Instrumentation

3.2.2.1 Chromatographic separation

The HPLC system consisted of an Dionex UltiMate 3000 degaser, Dionex UltiMate 3000 binary pump and a Dionex UltiMate 3000 Autosampler and column compartment from Thermo Scientific (Bremen, Germany). The chromatographic separation was carried out with different gradient methods shown in table 3-4 on different chromatographic columns given in table 3-5. The mobile phase consisted of eluent A: water + 0.1% formic acid, and eluent B: methanol + 0.1% formic acid. The HPLC system was equipped with a 20-µL sample loop, operated with 20 µL full loop injection and for different parts of the study with a 200-µL sample loop, operated with 100 µL partial injection.

Table 3-4: Overview of different gradient elution methods with eluent B: methanol + 0.1% formic acid.

Gradient 1		**Gradient 2**		**Gradient 3**		**Gradient 4**	
min	*%B*	*min*	*%B*	*min*	*%B*	*min*	*%B*
0	5	0	2	0	2	0	0
5	5	3	2	2	2	2	0
10	50	5	70	4	50	4	50
15	95	15	95	16	95	17	98
20	95	20	95	20	95	22	98
20.1	5	20.1	2	20.1	2	22.1	0
30	5	30	2	30	2	30	0
flow rate							
0.35 ml/min		0.35 ml/min		0.3 ml/min		0.3 ml/min	

Table 3-5: Overview of used chromatographic columns.

Name	**Column properties**	**Supplier**
Atlantis	Atlantis T3 2.1 x 150 mm. 3 µm particle size	Waters
YMC	YMC-Triart C18 2.1x150 mm, 3 µm particle size	YMC
XSelect	XSelect HSS T3 2.1x100 mm + Precolumn 2.1x5mm, 2.5 µm particle size	Waters

Measurements of part 3.1.2, 3.1.3 and 3.2.1 were run with the Atlantis chromatographic column with gradient 1 and an injection volume of 20 µL. Measurements of part 3.1.4, 3.2.2 and 3.2.3 were run with XSelect column with gradient 4 and an injection volume of 20 µL or 100 µL as indicated.

3.2.2.2 Mass spectrometry

Mass spectrometric detection was performed with an Orbitrap mass spectrometer from the series Q Exactive by Thermo Scientific (Bremen, Germany). Mass calibration was performed with the calibration solution shown in table 3-3 before a new measurement series was started and thereupon repeated every three days. Measurements in the positive and negative ionization mode were performed separately. The final parameters of the heated electrospray ionization (HESI) source are summarized in table 3-6 and of Full MS/ddMS^2 detection method in table 3-7 and Full MS detection in table 3-8.

Table 3-6: HESI source parameter after optimization.

Parameter	Positive	Negative
Sheath gas flow rate	37	40
Aux gas flow rate	15	15
Sweep gas flow rate	1	0
Spray voltage (k.V)	4.0	4.0
Capillary temperature(°C)	350	350
S-Lens RF level	50	50
Aux gas heater temperature (°C)	50	100

Table 3-7: Parameter settings of the full MS/ddMS2 method after optimization.

Parameter	Positive	Negative
General		
Polarity	Positive	Negative
Runtime	28 min	28 min
Full MS		
Resolution	70,000	70,000
AGC Target	1E+06	5E+05
Maximum injection time	100 ms	100 ms
Scan Range	100 – 1000 m/z	100 – 1000 m/z
dd-MS^2		
Resolution	17,500	17,500
AGC Target	5E+04	5E+04
Maximum IT	50 ms	50 ms
Loop Count	5*	5*
Isolation window	1.4 m/z	1.4 m/z
NCE (stepped mode)	20,50,70	20,50,70
Intensity Threshold	1.6E+04	1.6E+04
Minimum AGC target	8.0E+02	8.0E+02
Dynamic exclusion	3.0 s	3.0 s

Table 3-8: Parameter settings of full MS method

Parameter	Setting
Polarity	Positive/Negative
Runtime	28 min
Resolution	70,000
AGC Target	3E+06
Maximum Injection Time	100 ms
Scan Range	100 - 1000 m/z

Measurements for the development of the chromatographic method in part 3.1.1 were done with the full MS detection method (table 3-8), all other measurements were performed with the full MS/ddMS^2 method of table 3-7.

3.2.3 Data processing

Raw data were processed with *Xcalibur Quan Browser* to obtain peak areas, retention times (RT) and signal to noise ratios (S/N). *Xcalibur Qual Browser* was used to obtain measured accurate masses for calculation of mass error in ppm after equation 3-3. Raw data were imported to *MZmine2* version 2.34 and a target peak picking was performed with a target mass list containing m/z for development of the chromatographic method in part 3.1.1 as it provides information on full width at half maximum (FWHM), tailing factor and data points per peak. For additional calculation of S/N in part 3.2.3 data were processed with a R script [49] after conversion to mzXML format with *MSConvert* with parameter settings in table 3-9.

Table 3-9: Parameter settings for data processing with MSConvert and R-script.

Parameter	**Setting**
MS Convert	
Filter	Peak picking
Algorithm	Vendor
MS Levels	1-2
Min SNR	0.1
Min peak spacing	0.1
R-script	
Mass range [m/z]	100 -1000
RT range [min]	0 - 28
m/z step	0.01
Min. Intensity	1E+04
S/N	3
Noise [scans]	30
Max peaks per peak	10
Peak width [sec]	5-15

3.2.4 Data evaluation and calculations

3.2.4.1 Development of chromatographic method

For evaluation of chromatographic method following criteria were used 1) good peak quality 2) sufficient separation of single analytes and 3) even distribution of peaks over chromatographic space. The peak quality was evaluated based on the fraction of peaks with tailing factor (T) above 1.5, mean FWHM and mean number of datapoints per peak.

Information of FWHM, tailing factor and data points per peak were taken from MZmine2 and median and mean values were calculated for 32 model compounds (Mix B).

Peak separation was evaluated based on the median of peak resolution (R) and retention factor k calculated after equations 3-1 and 3-2 [104].

$$\textit{Equation 3-1: } R = 1.18 * \frac{RT_2 - RT_1}{FWHM_1 + FWHM_2}$$

(with RT_1/RT_2: retention times of one peak and subsequent peak and $FWHM_1$/$FWHM_2$: peak width at half maximum for corresponding peaks.)

Mean and median values were built for R of all model compounds.

$$Equation\ 3\text{-}2:\ k = \frac{RTi - RT_0}{RT_0}$$

(with RT_i: retention time of model compound, RT_0: dead volume of the column determined by injection peak, which was 0.95 min for Atlantis column, 0.75 min for XSelect column and 1.3 min for YMC column.)

The chromatographic space Δk was calculated by subtracting k of the first eluting peak from k of the last eluting peak. The distribution of peaks over the chromatogram was determined by first calculation of ideal equidistant distribution of peaks and subsequent determination of deviations from ideal distribution. For this purpose, for each peak X^2 from optimum value was calculated and all values were summarized into $\sum X^2$. Each gradient was evaluated based on $\sum X^2$, with minimal values representing best correspondence to optimal distribution of peaks.

3.2.4.2 Development of ESI-HRMS detection method

The ionization and detection conditions were optimized with compound mix A. The number of scans per second were exemplarily calculated with the total number of scans over measurement time obtained from *Xcalibur Qual Browser* divided by 28 min measurement time and 60 seconds per minute. As the applied full MS/ddMS^2 method uses one scan for full scan measurement and subsequent 5 scans for MS^2 measurements the number of scan/s was divided by six to obtain the number of data points of the full scan measurement.

The mass error in ppm, as difference between individual measurement and true value, was calculated according to equation 3-3 [31].

$$Equation\ 3\text{-}3:\ \Delta mi = \frac{mi - ma}{ma}\ x\ 10^6\ \ ppm$$

(with mi: measured accurate mass in Da and ma: calculated exact mass in Da)

The term mass accuracy is used to describe the average of mass errors of different compounds [31]. Peak areas for comparison of resolution and AGC target settings were obtained from *Xcalibur Quan Browser*. The rate of obtained MS^2 spectra of spiked model compounds in part 3.1.3 and 3.1.4 was calculated according to equation 3-4. Presence of MS^2 spectra and total intensities were evaluated in *Xcalibur Qual Browser*.

$$Equation\ 3\text{-}4:\ MS^2\ rate = \frac{number\ of\ spiked\ analytes}{number\ of\ MS^2 spectra\ of\ spiked\ analytes}$$

3.2.4.3 Precision of the method

The intraday or within run repeatability and the interday or within laboratory reproducibility for peak areas, retention times and mass errors were assessed to estimate the precision of the method. A low, middle, and high concentration level (1000 ng/L, 500 ng/L and 100 ng/L) of Mix A were measured in triplicates over three different measurement days within three weeks. For each measurement day, freshly spiked matrix samples were prepared from standard Mix A.

The repeatability (SD/RSD) and reproducibility (SDI/RSDI) were determined as within column and between column deviation of an ANOVA analysis according to equation 3-5 and equation 3-6 [112]. The ANOVA was performed with Microsoft Excel using the one-factor variance analysis with α=0.05.

$$Equation\ 3\text{-}5:\ Repeatability = SD: \sqrt{MS_w}$$

$$Equation\ 3\text{-}6:\ Reproducibility\ =\ SDI\sqrt{SD^2 + SD_R^2}$$

$$with\ SD_R: \sqrt{\frac{MS_B - MS_w}{n}}$$

(with MS_W: mean square within groups, MS_B: mean square between groups and n: number of observations in each group)

In addition, in part 3.2.3 the repeatability of peak areas and retention times was evaluated after changing of injection volume based on 10 injections of Mix C at 1 µg/L and calculation of relative standard deviations (RSD) of peak areas and standard deviation (SD) of retention times.

3.2.4.4 Method Sensitivity

The limits of detection (LODs) were determined for each analyte of Mix C individually with three different approaches.

- LOD calculation according to the German standard DIN 32645:2008-11 [113] using the equations given for the calibration method, with α=0.01 and k= 3. An 8-point calibration with concentrations of 10, 50, 100, 200, 400, 600, 800 and 1000 ng/L was measured in triplicates.
- LOD extrapolation was done for a S/N of 3 for each analyte based on triplicate measurement of a 10 ng/L standard. S/N was determined using *Xcalibur Quan Brower* and an R script as described in part 2.3.
- MDL calculation was based on standard deviation of 10 injections of 100 ng/L standard after equation 3-7 given for spiked samples in guideline from United States Environmental Protection Agency (EPA) [111].

$$Equation\ 3\text{-}7:\ MDL = t_{(n-1,1-\alpha=0.99)} * S_s$$

(with $t_{(n-1,\ 1-\alpha=0.99)}$: the Student´s t-value appropriate for a single-sided 99th percentile t statistic and a standard deviation estimate with n-1 degrees of freedom and S_s: sample standard deviation of the replicate spiked sample analyses)

The given $t_{(n-1,\ 1-\alpha=0.99)}$ for 10 injections and 9 degrees of freedom of 2.821 according to [111] was used.

3.2.4.5 Test of significant difference

To examine if the results of two different measurement settings were significantly different the two-variable t-test was used. As this test is only applicable for data groups with variance homogeneity, a Fischer's F-test was conducted beforehand according to equation 3-8 [114]. If the calculated F-value was lower than the tabulated one, the samples differed randomly and not significantly and the two-sided t-test according to equation 3-9 [114] was applied. If the F-value was higher, the general t-test, after Welch, was applied according to equation 3-10 [114]. The calculated t-value was then compared to the tabulated t-value for 95% significance. If the calculated t-value was smaller than the tabulated one, the samples varied only randomly, and not significantly.

$$\textit{Equation 3-8: Fischer´s } F-Test\text{: } F = \frac{s_1^2}{s_2^2}$$

$$\textit{Equation 3-9: two-sided } t-test\text{: } t = \frac{|\bar{x}_1 - \bar{x}_2|}{SD}\sqrt{\frac{n_1 n_2}{n_1+n_2}}$$

$$\textit{Equation 3-10: general } t-test\text{: } = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}$$

with: weighted averaged standard deviation according $SD = \sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}}$

(with n_1& n_2: number of parallel measurements; $\overline{x_1}$ & $\overline{x_2}$: mean value of measurements; s_1^2& s_2^2 variances of measurements)

The statistical significance testing was performed with *Microsoft Excel* for the individual compounds for the comparison of spray voltages and spray temperatures during the method development part 3.1.2. All other significance tests were not carried out for individual compounds but for the whole data set and were performed with *GraphPad Prism 5*. If the F-test indicated a significant difference of variances, the Welsh correction was performed as well.

Graphical representation of results is done in several figures with so called box and whiskers plots which show distribution of data. The box shows the 25th and 75th percentiles of data, the line in the middle of the box is plotted at the median and the mean is indicated by a cross. In the chosen format, the whiskers are drawn down to the 10th percentile and up to the 90th percentile and values below or above the whiskers are shown as individual points.

3.3 Results

3.3.1 Method development

3.3.1.1 Chromatography

The aim of the development of a generic screening method is to cover a wide range of OMPs with different polarities with a sufficient separation of single analytes, even distribution of peaks over chromatographic space and good peak quality. Method development was based on 32 model compounds (Mix B) with m/z in the range of 120 to 429 and logD (at pH 3 as eluents) values from -3.6 (metformin) to 5.0 (triclosan). The gradient elution program with water and methanol both containing 0.1% formic acid was first optimized. Figure 3-1 shows for each tested gradient the fraction of methanol over time in part A and distribution of peaks over chromatogram in part B.

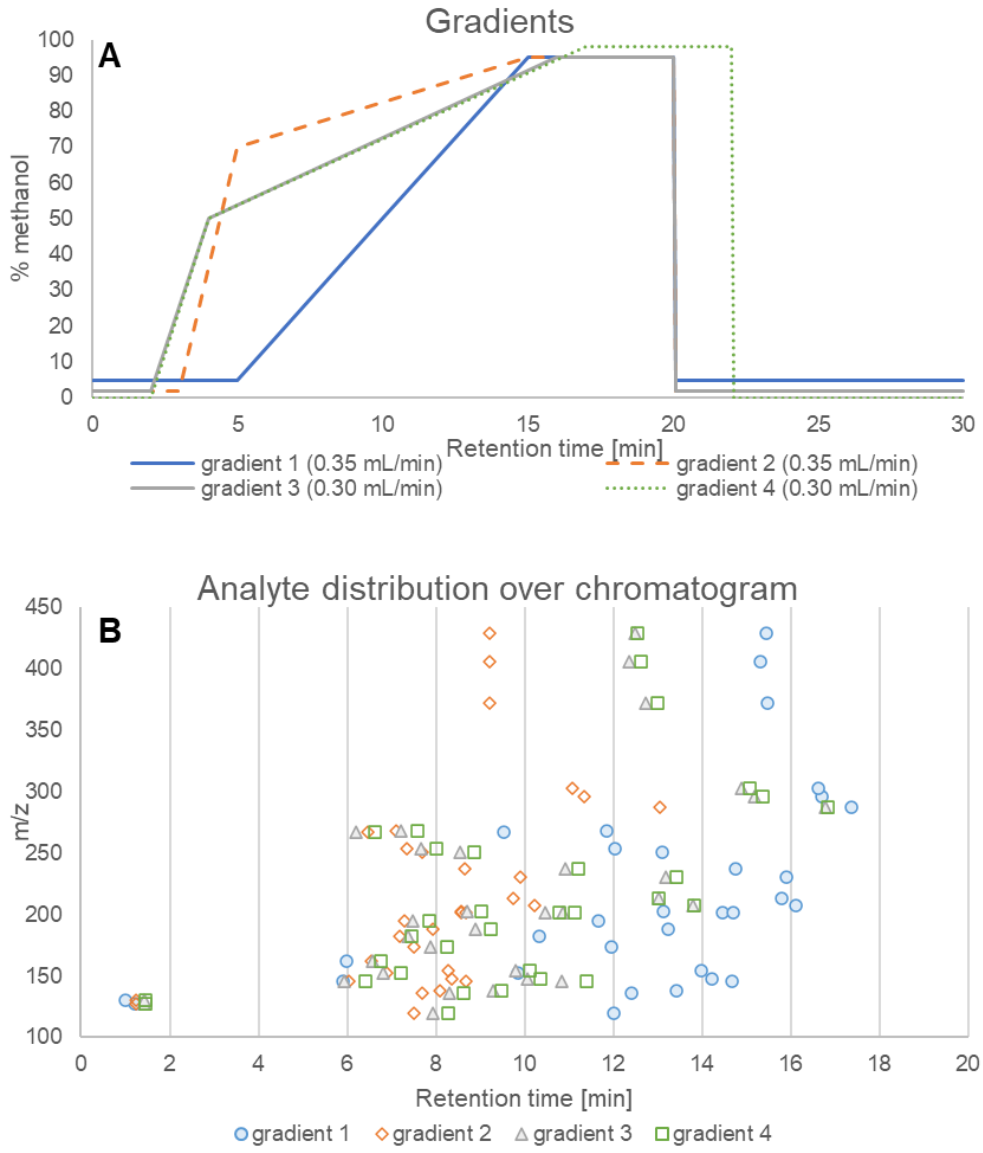


Figure 3-1: Fraction of methanol over time and analyte distribution (Mix B, 1 µg/L in UPW) over chromatogram of gradients 1-4.

Gradient 1 shows a very late elution of most analytes and a large gap between RT min 1.3 and 6 and another gap until min 10. However, peak resolution was high indicated by high median of R and peak shapes were of high quality indicated by low number of peaks with T>1.5 as shown in table 3-10. For gradient 2 the proportion of methanol was increased more rapidly for a faster elution of analytes. However, as can be seen a gap between RT min 1.3 and 6 remains and peak resolution declines, indicated by lower median for R. After further adjustments of the gradient, the distribution of peaks over the chromatogram was improved, as shown for gradient 3. In addition, for gradient 3 the flow rate of the mobile phase was lowered to 0.3 mL/min which has the advantage of reduced eluent consumption, lower backpressure on the column and a slightly retarded elution of very polar compounds such as melamin and metformin and thus better separation from injection peak. A further effect is broadening of peaks and therefore a higher number of data points per peak (see table 3-10). As for later MS detection a method that constantly switches from full scan to MS^2 mode is used, slower scan rates are obtained which can become problematic for extremely narrow peaks leading to an insufficient number of data points per peak as discussed in more detail in part 3.1.3. For this reason, even though usually narrow peaks are favorable due to better peak resolution and peak shapes, in case of Orbitrap MS detection, broader peaks are required sometimes. However, a higher proportion of peaks with T>1.5 was observed for lower flow rates as well. Finally, the total range of methanol was increased, starting at 0% and rising to 98% at the end of the gradient elution method (gradient 4). As can be seen in figure 3-1 this did not highly influence the RTs of compounds. The reduction to 0% methanol did not increase RT of polar compounds as expected. However, this extreme range was further on used to make sure all analytes and possible contaminations are eluted from the column when real water samples are used.

Table 3-10: Overview of peak quality measures, peak resolution and distribution for gradients 1-4.

	N peaks with T > 1.5	**mean FWHM [min]**	**mean data points/peak**	**median R**	$\sum X^2$
Gradient 1	5	0.077	35	2.2	99
Gradient 2	5	0.076	33	1.0	40
Gradient 3	9	0.080	39	1.9	20
Gradient 4	7	0.078	38	2.1	24

For a better visualization of the distribution of peaks over the chromatographic space, peaks were sorted by retention time and plotted in elution order as shown in figure 3-2. A plateau of the curve visualizes an accumulation of peaks in a short RT range and a void RT range is indicated by stepwise curve shape. For all four gradients a gap between ~min 1.3 and 6 is observed. For gradient 1 an additional gap between min 6 and 10 is present. In addition, gradient 1 shows several plateaus of the curve, e.g., around min 12, that indicate accumulation of peaks in this range. Gradient 2 shows a very flat slope of the curve which is caused by lower

separation of peaks. The most even distribution of peaks is observed for gradient 3 and 4 showing best approximation to linear behavior.

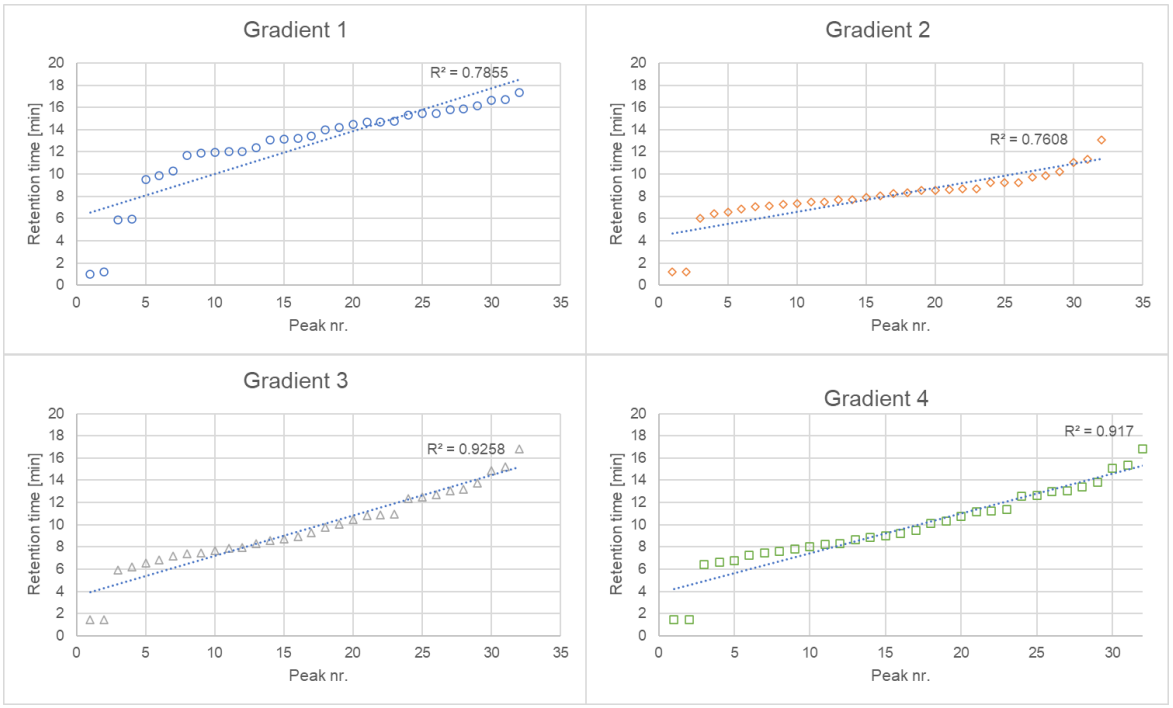


Figure 3-2: Retention times of peaks with ascending order for gradients 1-4.

The approximation to ideal distribution of peaks was further assessed by sum X^2 (see table 3-10). This confirms that gradients 3 and 4 show best exploration of the chromatographic space as they had lowest values for X^2.

In a next step, two further chromatographic columns were tested during method development. A mixture of 32 standards (Mix B) was measured with gradient 4 as described above. For evaluation the number of compounds with tailing factor (T) above 1.5, the mean FWHM, mean number of data points per peak, median peak resolution (R) were compared. The chromatographic space was determined by subtracting of the the lowest retention factor from the highest retention factor (Δk). As can be seen in table 3-11 the YMC column showed the narrowest peak widths with a mean FWHM of 3 sec. This also leads to a lower number of recorded data points per peak and a higher mean resolution of peaks. However, as already discussed above, too narrow chromatographic peak widths are inconvenient for Orbitrap MS detection. In addition, the YMC column showed an elevated back pressure compared to the other two columns with a maximum at 380 bar which is very close to the upper pressure limit of the HPLC system of 400 bar. Furthermore, with the YMC column and the given gradient it was not possible to separate the two isomers Simazin and Terbuthylazin-desethyl (both with m/z 202). Thus, the YMC column was considered less suitable for the purpose of this study.

The other two columns, Atlantis and XSelect, have a very similar stationary phase and also showed very similar performance. The chromatographic space Δk and mean k was slightly higher for the XSelect column, which is also the successor model of the Atlantis column and was therefore finally selected.

Table 3-11: Overview of peak quality measures, peak resolution and distribution for different chromatographic columns.

Column	T > 1.5	mean FWHM [s]	mean data points / peak	median R	Δk
YMC	8	3.22	11	4.4	12.5
Atlantis	5	5.24	30	2.1	16.4
XSelect	5	5.19	25	2.0	19.9

3.3.1.2 Ionization conditions

To optimize the ionization efficiency three different spray temperatures and three different spray voltages were tested for each ionization mode. For this purpose, the mean peak areas of a triplicate injection for each compound of Mix A were determined. Spray voltages used in literature range from 3.5 to 4.5 kV for positive mode and 3.0 to 3.8 kV for negative mode (see table 3-1). The general comparison of spray voltages 3.5, 4.0 and 4.5 kV for positive ionization in figure 3-3 (left) shows no significant difference. However, in figure S3-1 peak areas of individual compounds are compared and show higher mean peak areas with 4.0 kV for some compounds and for others with 4.5 kV. As several compounds, which showed a low ionization efficiency, were better ionizable at 4.0 kV, this spray voltage was selected for the positive mode. For negative ionization mode in general smaller peak areas were obtained compared to positive ionization mode. The comparison of spray voltages 3.0, 3.5 and 4.0 kV in figure 3-3 (right) showed slightly higher mean peak areas for 4.0 kV, but differences were not significant. Nevertheless, 4.0 kV was selected as spray voltage for negative mode.

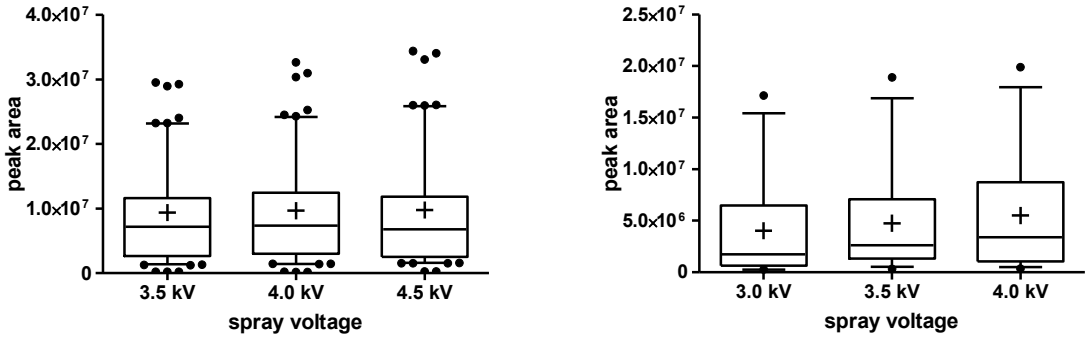


Figure 3-3: Comparison of spray voltages for positive mode (left) and negative mode (right), evaluated for triplicate measurement of Mix A at 1 µg/L in UPW at 300 °C ionization temperature.

In a next step, spray temperatures were varied to find optimal settings for most compounds. As shown in table 3-1, spray temperatures in literature were reported between 280 and 360°C, in some cases with lower temperatures selected for negative ionization mode. The general comparison of ionization efficiencies for all compounds with spray temperatures 320, 350 and 380°C for positive ionization mode given in figure 3-4 (left) shows no significant differences between settings. However, a comparison for individual model compounds in figure S3-3 showed significantly smaller peak areas for 17 compounds with 380°C and 12 with 320°C compared to 350°C. Thus, 350 °C was selected. The comparison between spray temperatures 300, 320 and 350 °C for negative ionization mode given in figure 3-4 (right) showed very small differences in ionization efficiency, which was also true for individual compounds (figure S3-4). However, spray formation was more stable at 350°C during visual inspection. In addition, changes from positive to negative ionization mode in large sample sequences can be faster implemented if the same spray temperature is applied, as cooling of the ion source takes some time. Thus, finally 350 °C was selected for negative ionization mode as well.

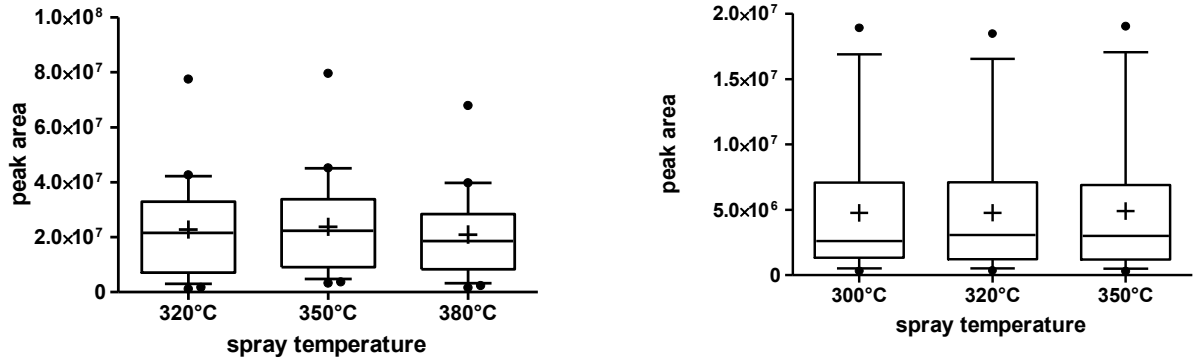


Figure 3-4: Comparison of spray temperatures for positive mode (left) and negative mode (right), evaluated for triplicate measurement of Mix A at 10 µg/L in UPW at 3.5 kV spray voltage. Represented as boxplots 10-90 quartile.

As already well known from literature, optimal ionization conditions vary between compounds [105]. However, observed differences were small and compromises that are expected in the case of generic screening methods seem acceptable. The observed lower general lower intensities in the negative ionization mode were also reported in other studies [55] and can be partly caused by addition of formic acid to eluents in negative mode.

Settings for sheath and auxiliary gas in a.u. vary in literature and can be adjusted depending on HPLC conditions such as flow rate etc. or sample matrix. For this study parameters were selected as shown in table 3-6 based on visual inspection stability of electrospray.

3.3.1.3 MS detection parameter

Most NTS studies with Orbitrap mass spectrometers use a combination of full MS and ddMS^2 as detection method. For both full MS^1 and MS^2 the mass resolution and AGC target can be adjusted.

The mass resolution has an influence on the selectivity and mass accuracy of the measurement but also impacts the cycle time of the instrument and thus the number of scans/sec and ultimately the number of data points/peak. Table 3-12 gives an overview of average scans/sec for different settings of mass resolution for both full scan and ddMS^2 based on calculation of three replicate measurements of Mix A spiked into matrix at 1 µg/L. Within each measurement cycle one MS^1 and five MS^2 scans are recorded, thus the number of data points per second in full scan corresponds to scans/sec divided by six. As can be seen with decreasing resolution of both full scan and ddMS^2 the number of scans per sec increases as acquisition time of lower resolution mass spectra is shorter.

Table 3-12:Overview of average number of scans per second and data points per sec of full MS measurement at different combination of settings for resolution and AGC target of full MS/ddMS^2 method.

	Resolution		
Full Scan	140,000	70,000	70,000
ddM^2	17,500	35,000	17,500
~scans/sec	6.1	5.9	8.1
~data points/sec MS^1	1.0	1.0	1.4

**Measurements with different resolution settings were performed with AGC target 1E+06 for full scan and 1E+05 for ddMS^2*

Figure 3-5 shows an example of an extracted ion chromatogram of 1H-benzotriazole recorded with ~6.1 and ~8.1 scans/sec. As can be seen the number of data points for a rather narrow peak can be comprimised with a long cycle time. According to literature a cylce time that allows at least 15 but not less than 8-10 data points per peak is recommended for quantification [105] but is also required for an adequate and reproducable definition of a chromatographic peak for screening methods [115].

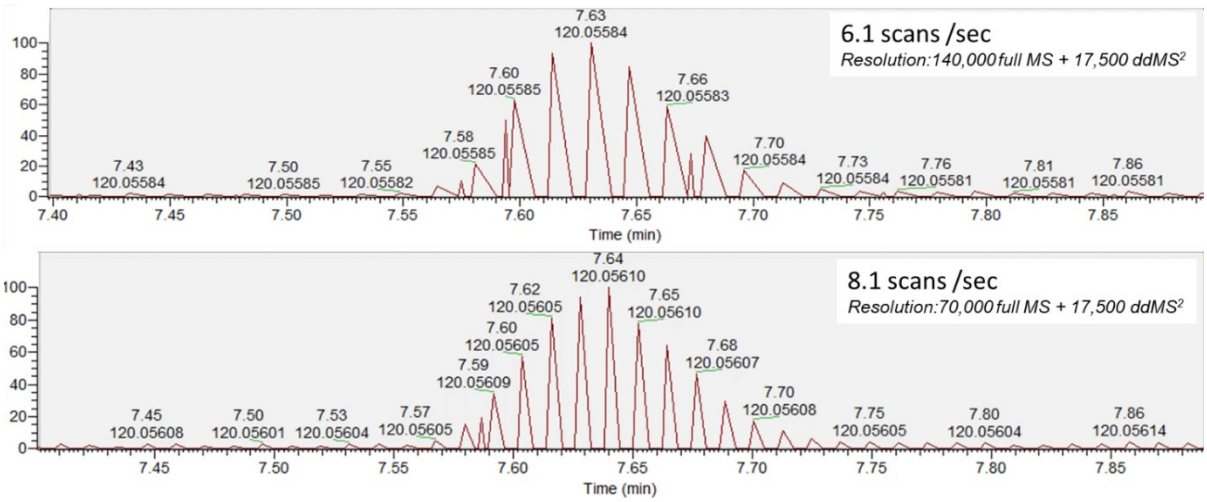


Figure 3-5: Exemplary EIC of 1H-Benzotriazole measured in positive mode with different settings for resolution. Jagged baseline due to acquisition mode changes from MS^1 to MS^2

The use of higher resolving power can result in an enhanced selectivity and accuracy of mass assignment [115]. For this reason, for the evaluation of the mass resolution of the full scan the mass accuracy was evaluated by determination of mass errors of individual analytes. However, figure 3-6 shows no improved mass accuracy for the measurements with mass resolution of 140,000. Therefore, the combination of mass resolution for full scan of 70,000 and for $ddMS^2$ of 17,500 was selected as this provides the highest scan speed and is also in accordance with the majority of methods from literature.

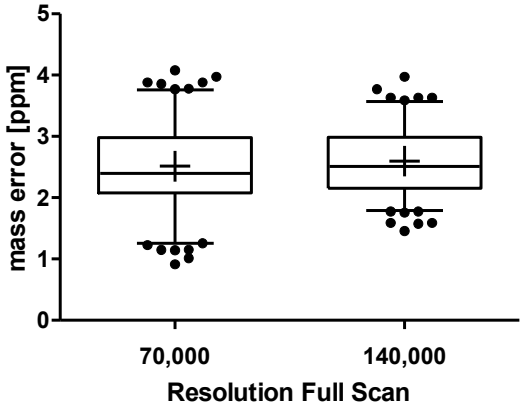


Figure 3-6: Comparison of mass error in ppm evaluated for triplicate measurement of Mix A at 1 µg/L in matrix in positive ionization with different resolution of the full scan.

For the $ddMS^2$ acquisition a resolution of 17,500 was selected as this was mainly used in literature. A higher resolution of 30,000 would additionally increase cycle time and reduce number of data points/peak.

For the selection of optimal settings for the AGC target value the sensitivity based on peak areas and mass accuracy based on mass errors of model compounds of a triplicate measurement of Mix A in matrix was evaluated. Table 3-1 shows that for AGC targets in literature for full scan mainly 1E+06 and in one case 3E+06, for the ddMS^2 1E+05 or 2E+05 were used. Figure 3-7 shows decreasing peak areas with decreasing AGC targets for the full scan in positive mode. This can be explained by the lower number of ions that are collected. For the negative mode, generally lower intensities were observed and thus lower AGC target values 1E+06 and 5E+05 were tested. As shown in figure 3-7 differences of both settings were minimal.

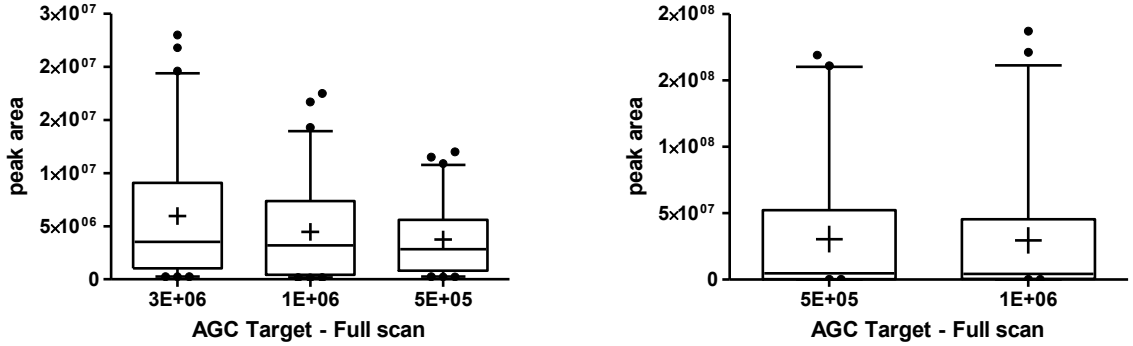


Figure 3-7: Comparison of peak areas for different AGC target values, evaluated for triplicate measurement of Mix A at 1 µg/L in matrix for positive mode (left) and negative mode (right).

Figure 3-8 shows that for positive mode the mass error differed significantly with different AGC target values. The general lowest mass errors and thus highest mass accuracy was obtained with an AGC target of 1E+06. A decreasing mass accuracy for too high AGC target values due to space charge effects is known from literature, e.g., Kalli et al. 2013 observed increasing mass deviations for AGC target values >2E+06 for a peptide mixture measured on a LTQ Orbitrap [107]. An explanation for the higher mass deviation at AGC target 5E+05 could be that for this measurement no sufficient number of ions for an accurate mass detection was reached. The mass error for the measurement in negative ionization was not evaluated because of an erroneous mass calibration in negative mode, which will be discussed in part 3.3.2.2.

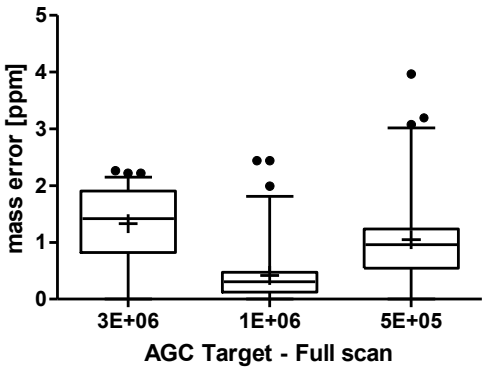


Figure 3-8: Comparison of mass errors for different AGC target values for the full MS, evaluated for triplicate measurement of Mix A at 1 µg/L in matrix in positive mode.

For the positive ionization mode an AGC target of 1E+06 was selected as tradeoff between sensitivity and mass accuracy, as differences of mass accuracy for different settings were much more pronounced compared to differences in sensitivity. For the negative ionization mode an AGC target of 5E+05 was selected, even though differences between settings were not significant.

For the selection of AGC target of the ddMS^2 part it was determined if MS^2 spectra were recorded for the model compounds. Further, the mean intensity of all recorded spectra was calculated. For the positive ionization mode, the target values 2E+05, 1E+05 and 5E+04 were tested and the AGC target values 1E+05, 5E+04 and 2E+04 were compared for the negative ionization mode, respectively. The rate of recorded MS^2 spectra of the spiked analytes in matrix in percent and the average intensity for both ionization modes is shown in figure 3-9. In both ionization modes, the average intensity of the MS^2 spectra increased with decreasing AGC target value. In the positive ionization mode, the rate of recorded MS^2 spectra of the spiked analytes was 94% for all three settings. In the negative ionization mode, less MS^2 spectra were recorded for the model compounds. The highest rate was achieved for the target value 5E+04 with 47%. This is probably due to the low intensity of several analytes in the negative ionization mode. As a ddMS^2 scan of the five most intense precursor ions is generated, no ddMS^2 is triggered if the analyte intensities are too low. For higher AGC target values, also a higher amount of matrix ions are collected which can also compromise triggering of MS^2 spectra of analytes.

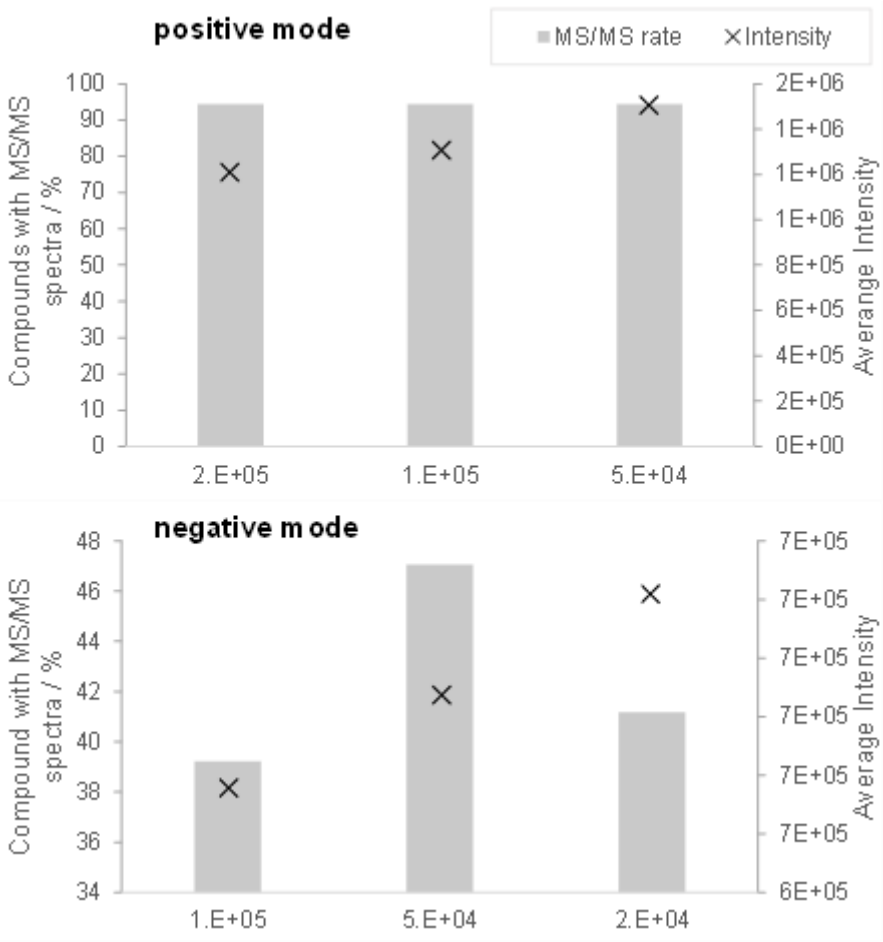


Figure 3-9:Rate of MS² spectra of target analytes (grey columns) and average intensity of MS² spectra (black X) for different AGC target values for the ddMS² measurement; positive ionization mode (top), negative ionization mode (bottom). For triplicate measurement

Finally, the AGC target value 5E+04 was selected for the positive ionization mode as it showed the highest average sensitivity of MS^2 spectra. In the negative ionization mode, the AGC target 5E+04 was also selected, as it showed the highest rate of recorded MS^2 spectra of the spiked compounds, even though a higher average intensity of MS^2 was obtained at 2E+04.

It should be considered, that the parameter AGC target further interacts with the setting of 'maximum injection time'. This parameter limits the maximum time of ion collection in the C-trap and comes into action when the AGC target is not reached within this time frame. However, this parameter was not variated in this work but was kept constant at 100 ms according to literature values. Further, the so-called space-charge effects in the C-Trap do not only compromise the mass accuracy but also resolution of peaks, which was not evaluated in this work. The optimal AGC target is dependent on the ion flux reaching the instrument and needs adaption if samples and/or matrices with extremely high or low content of ionizable substances are used. It can be concluded that different parameter settings correlate with each other and influence the measurement in a complex manner. Nevertheless, with generic

settings, as used in literature for small molecules in water matrices, no high differences were observed.

3.3.1.4 NCE for MS^2 spectra

The NCE can be used in fixed or stepwise mode. For the stepwise NCE mode, up to three different values are entered, the mass spectrometer will perform a stepwise fragmentation on the precursor ion. All fragments created in the steps are collected and sent to the Orbitrap analyser for one scan detection. An adequate normalized collision energy (NCE) is necessary to obtain MS^2 of high quality for identification. In figure S3-5 exemplary the MS^2 spectra of 1H-benzotriazole and sulfamethoxazole at NCE 35, 50 and 70 are given to show dependency of optimal conditions with molecule properties. As shown, with NCE 35 Benzotriazole at m/z 120 is almost not fragmented, but for the precursor of Sulfamethoxazole at m/z 254 only a very low intensity is left at this collision energy. With increasing NCE settings fragmentation increases for both compounds. Nevertheless, all of the shown MS^2 spectra were correctly identified with mzCloud database search.

As in NTS the compounds of interest are not known beforehand, and molecule properties can vary, a tradeoff with a collision energy appropriate for most compounds must be determined. To that end, a mixture of 289 compounds (Mix D), of which 206 were ionized in positive and 83 in negative ionization mode, was used. It was manually evaluated if MS^2 spectra were recorded, the m/z of the precursor ion was still present in the spectra and if besides the precursor ion further fragments were present. Figure 3-10 gives a summary of the results of a measurement with stepped NCE of 20, 50 and 70 as well as 30, 60 and 80 eV.

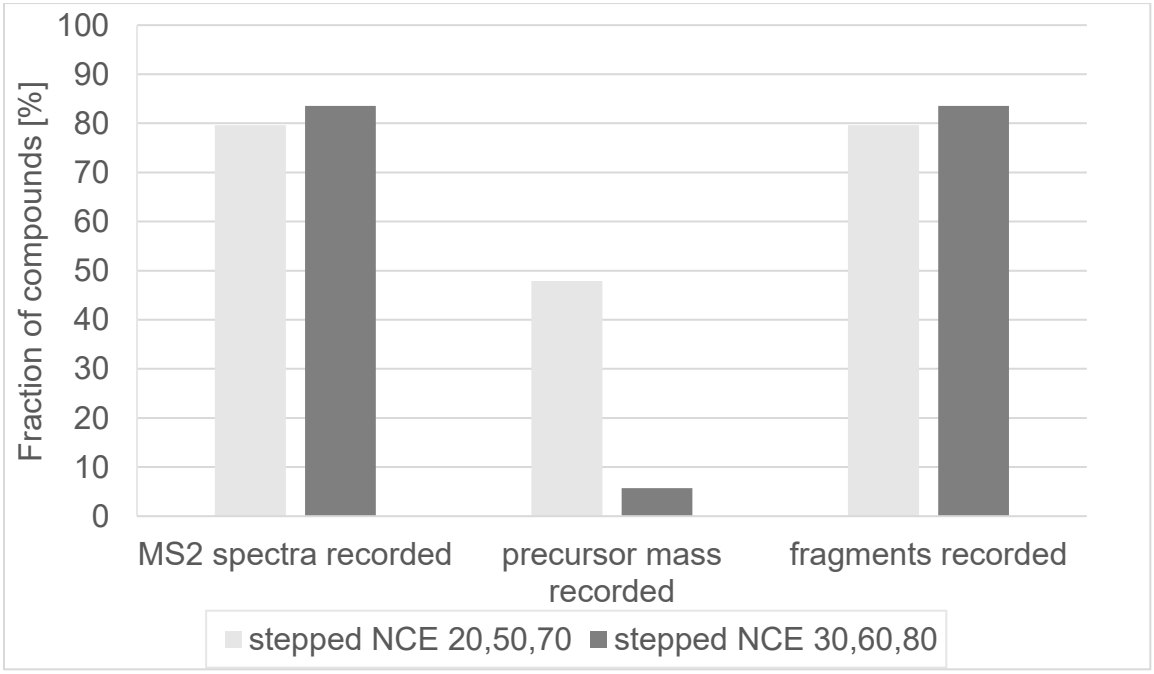


Figure 3-10: Comparison of different settings for NCE during MS^2 recording, Mix D 1 µg/L in UPW.

As can be seen, not for all the compounds MS^2 spectra were recorded with either method, which is probably due to low intensities of the precursors. For the NCE setting with lower energies (light grey) the recorded MS^2 spectra showed a higher proportion of spectra with the precursor ion mass still present. Thus, the method setting with lower energies was selected. Yet it is not assured that for all NTS signals, that retrospectively appear as relevant during data evaluation, high quality MS^2 spectra are obtained. A comparison of further NCE values and a more detailed evaluation of MS^2 spectra quality would be interesting topics for further research.

3.3.2 Method validation

3.3.2.1 Precision of peak areas, retention times and mass error

In figure 3-11 the repeatability (left) of 10 injections and reproducibility (right) over 3 weeks of the peak areas are given as relative deviations for both ionization modes at three spiking level in matrix. With decreasing concentration, repeatability decreased. The median RSD of peak areas increased from 2.7% for a concentration level 1000 ng/L to 3.9% for 100 ng/L, but differences were not significant. Nevertheless, RSD were below 10% in all cases and below 5% with some compound specific exceptions (tamoxifen and toremifene). For 65% of the compounds RSDI were < 10% at a spiking level of 1000 ng/L. However, several compounds show a very low reproducibility with RSDI around 30% or higher. For the reproducibility, no significant differences were observed between concentration level.

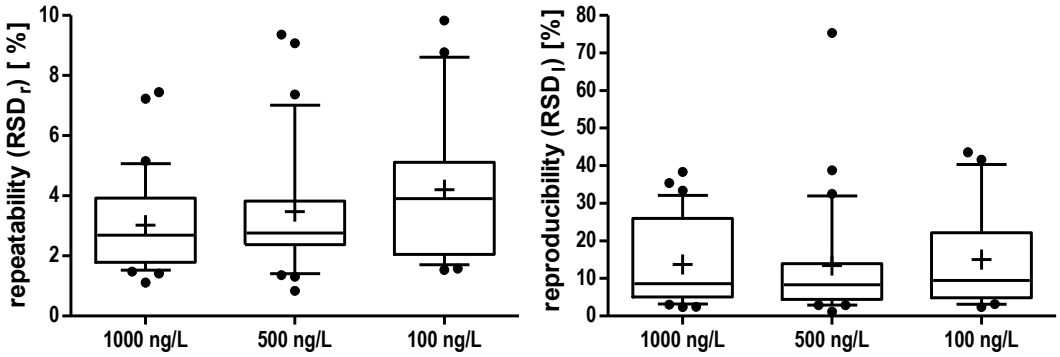


Figure 3-11: Repeatability (left) and reproducibility (right) of peak areas at different concentration levels of Mix A in matrix.

It was observed that mean peak areas of several compounds on the second measurement day were different compared to day one and three. This indicates a variation of the instrument performance or a mistake during preparation of samples as solutions were freshly prepared for each measurement day. Furthermore, several compounds showed decreasing peak areas over time which indicates a low stability of these compounds in standard solution. Moschet et al. (2013) obtained RSDs of peak areas in a range of 1 to 25% for a set of 45 target compounds [116]. This shows that compound dependent variations are usual. The results emphasize the

importance of including internal standards to monitor instrument performance and sample stability. Furthermore, a variability of peak areas caused by measurement uncertainties has to be considered during data processing and evaluation of long time series.

Figure 3-12 shows repeatability (left) and reproducibility (right) of retention times in seconds at three different concentration levels.

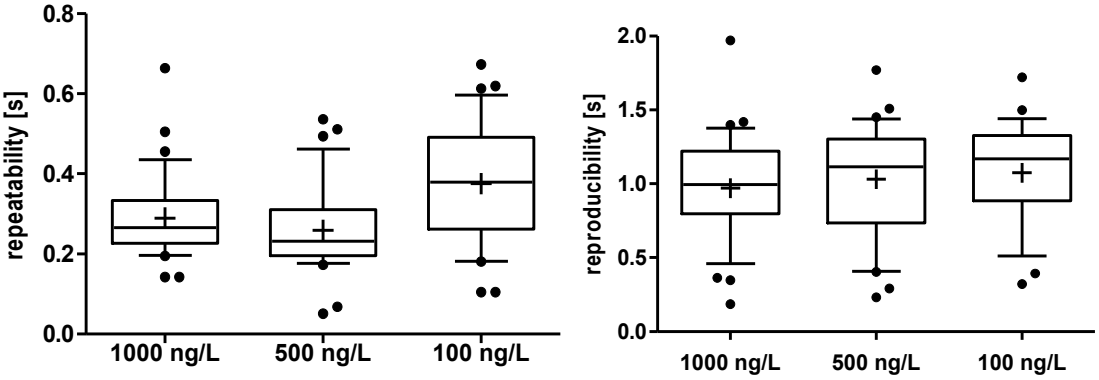


Figure 3-12: Repeatability (left) and reproducibility (right) of retention times at different concentration levels of Mix A in matrix.

70% of the compounds showed a repeatability of retention times of < 0.3 s. The repeatability slightly increased with decreasing concentration but was still < 0.8 s. The reproducibility (figure 3-12 right) was around one second and showed a generally higher spread of data compared to repeatability. A general shift of the retention times of 0.5 s from day one to two and another 0.5 s from day two to three was observed. Nürenberg et al. (2015) specified the deviations of RT of their measurements as < 0.2 min or 12 s [103]. In conclusion, the repeatability and reproducibility of retention times were highly satisfying compared to literature values, allowing a peak alignment during NTS data processing. The obtained information can be used to find adequate thresholds for peak alignment during data processing

The repeatability and reproducibility of the mass errors in the positive ionization mode are shown in figure 3-13 for three different concentration levels. As can be seen, the deviation within and between measurements is very low (< 1 ppm)

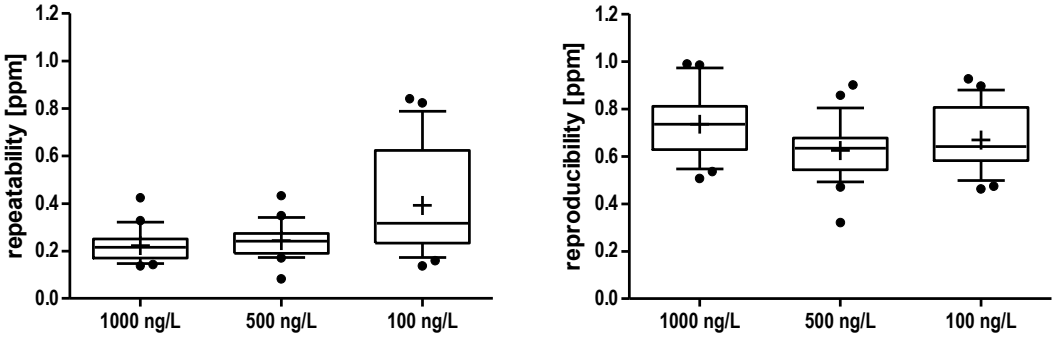


Figure 3-13: Repeatability (left) and reproducibility (right) of mass error in positive ionization mode at different concentration levels of Mix A in matrix.

Further, for the repeatability lower deviations (median 0.2 ppm at 1000 ng/L) were observed compared to the reproducibility (median 0.7 ppm at 1000 ng/L). This can be explained by the fact that between measurement days the instrument was calibrated. At the lowest spiking level (100 ng/L) the repeatability decreased significantly compared to both higher concentration levels. However, the lower concentration did not compromise the within laboratory reproducibility. The mass errors for the measurements in negative ionization were again not evaluated because of an erroneous mass calibration in negative mode as mentioned above. However, this aspect will be discussed in part 3.3.2.2.

Nürenberg et al. (2015) obtained deviations of around 1 ppm for both intraday and interday precision of the mass error on an TOF MS at 1 µg/L. In comparison the performance of Orbitrap MS is highly satisfying as even much lower mass errors were obtained for repeatability in positive mode compared to reported intraday precision. The stability of measured masses over several runs can be helpful information during data processing steps like peak alignment of several samples for settings of ppm window thresholds.

3.3.2.2 Mass accuracy

The mass accuracy, in terms of general deviation of theoretical and measured masses, of the instrument is of high interest during unknown compound identification or database queries. Therefore, in addition to the repeatability and reproducibility of measured masses discussed above, the mass errors of a broad range of compounds (Mix D) was evaluated to assess the general mass accuracy of the instrument.

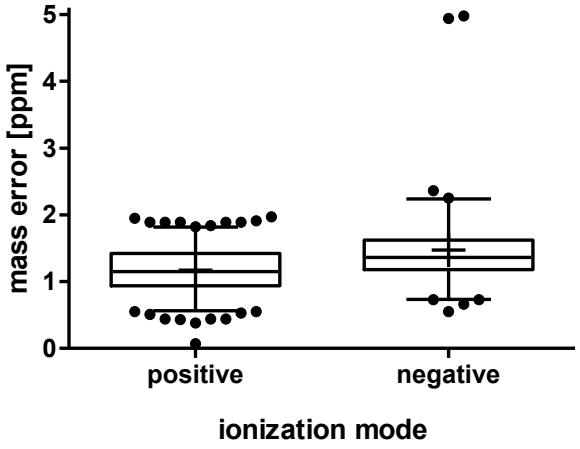


Figure 3-14: Mass errors for positive ionization mode (206 compounds) and negative ionization mode (83 compounds) at 1 µg/L in UPW.

As shown in figure 3-14 for positive ionization mode mass errors below 2 ppm were obtained for all compounds with a mean of 1.2 ppm. For negative ionization a few more outlier above 2 ppm were observed but mean mass error of 1.5 ppm was comparable to positive ionization mode. The slightly higher deviations in negative mode can be due to lower signal intensities. Nevertheless, the mass error of both ionization modes measurement was far below the required 5 ppm for NTS screening methods [23,54]. However, for previous measurements, e.g., in part 3.3.1.3 and 3.3.2.1 extremely high mass errors (up to 8 ppm) were observed for some compounds in negative ionization mode. The reason for this was later identified: the calibration solution for negative mode provided for Orbitrap MS by the vendor (table 3-2) does not contain a small mass. This problem was solved by adjustment of the calibration routine by adding the m/z of an in-source fragment of one standard contained in the calibration solution. As shown in figure 3-14 the mass accuracy in negative mode was comparable to positive mode after this adjustment.

3.3.2.3 Sensitivity

The sensitivity of the method was evaluated based on limits of detection for 30 model compounds of Mix C. As described in the introduction, it can be favourable for NTS studies to use a direct large injection of 100 µL to minimize sample preparation and discrimination of compounds. For this reason, the instrumental setting of the HPLC system was changed from a 20 µL sample loop and full loop injection volume of 20 µL to a sample loop of 200 µL with the option of partial injection of 100 µL. In this context, it was of interest to evaluate if the sensitivity of the method was enhanced with larger injection volume and if repeatability of peak areas and retention times remained constant.

For 30 model compounds limits of detection (LOD) were calculated according to DIN 32645 for both injection volumes. Detection limits were between 11 and 108 ng/L with a mean of 30

ng/L (see table 3-13) for the 20 µL injection volume. LODs for the 100 µL injection volume were between 7 and 64 ng/L with a mean of 17 ng/L. The general comparison of LODs obtained for both injection volumes in figure 3-15 shows that LODs with 100 µL injection volume were significantly smaller.

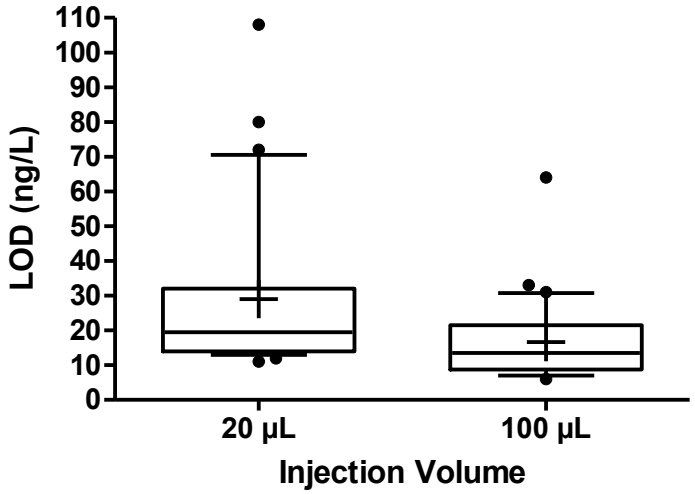


Figure 3-15: Detection limits after DIN 32456 for 20 and 100 µL injection volume for Mix C in UPW.

However, differences did not show a factor of 5. In addition, in a few cases even lower LODs with the 20 µL injection volume were reached (see table 3-13). Calibration curves of these compounds in figure S3-6 show that peak areas obtained with higher injection volume were higher. However, for these compounds the deviations of peak areas during triplicate measurement for the 100 µL injection volume were higher compromising LOD calculation after DIN. Furthermore, it has to be taken into account that a concentration range of 10 ng/L to 1000 ng/L was used, which includes two orders of magnitude and is not optimal for the calculation of LOD after DIN. In addition, this approach is rather conservative and for most screening methods without quantification purposes other approaches were used in literature. Several studies in the NTS field use a LOD estimation based on the S/N of a low concentration standard with the factor 3 or LOQ with a factor of 10, respectively [40,64,103]. Figure 3-16 shows the obtained LODs for a S/N of 3, estimated from the lowest calibration standard 10 ng/L with two different data processing programs.

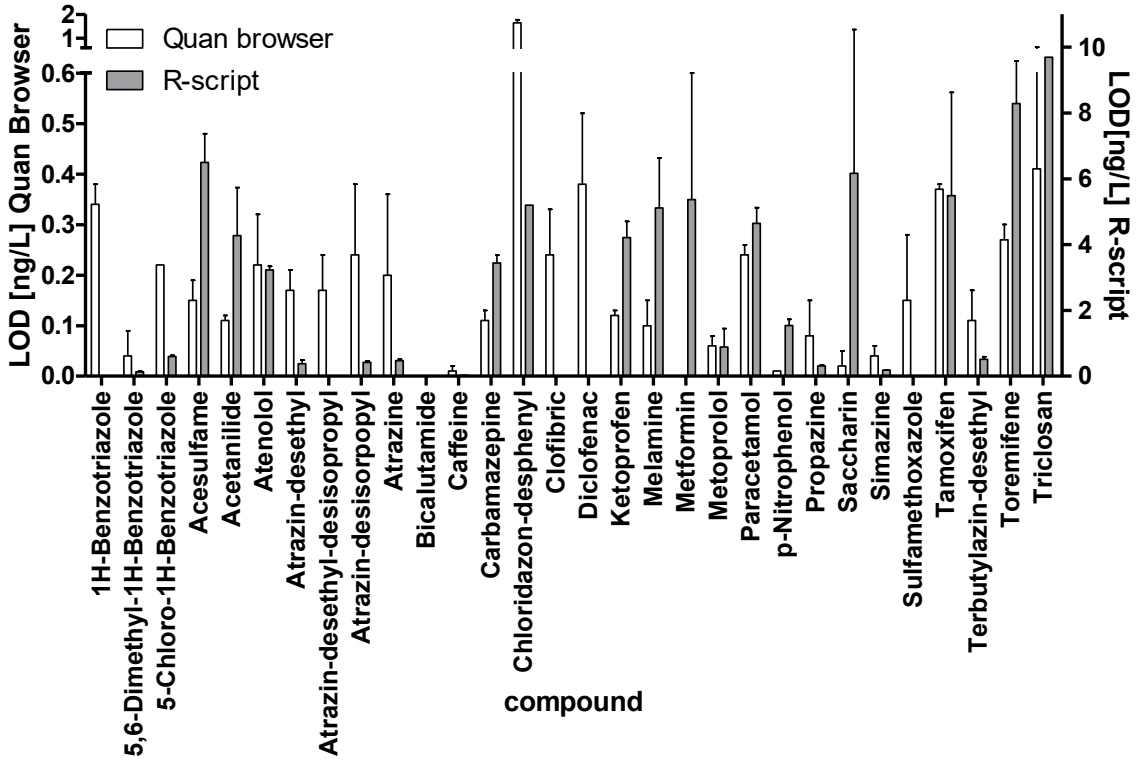


Figure 3-16: Detection limits based on a S/N of 3, estimated for a triplicate measurement of 10 ng/L of Mix C in UPW with 100 µL injection volume and evaluated with Quan Browser and R script. Error bars represent SD of triplicate measurement.

Xcalibur Quan Browser is the standard program for the Orbitrap provided by the vendor *Thermo Scientific* for automated peak integration of targets. Further, a R script for non-target peak picking was used, which better reflects the conditions of NTS data processing. Overall, very low LODs were reached with both data processing methods. LODs calculated from S/N in *Quan Browser* software with a mean of 0.2 ng/L were significantly smaller compared to LODs calculated from S/N with the R script with a mean of 3.4 ng/L. Hinnenkamp et al. 2019 reported LOD < 5 ng/L for 59% of compounds based on a S/N 3 [40]. In a study by Nürenberg 2015 50% of compounds showed a LOQ < 22 ng/L based on a S/N 10 [103]. Günthardt et al obtained LODs based on a S/N of 3 for the lowest concentration standards with a median of 0.4 ng/L and LODs < 5 ng/L for 81% of the analysed phytotoxins [64]. The obtained LODs based on S/N of 3 are comparable to literature values. However, large error bars indicate high variation of obtained S/N values of triplicate injection. Further, for some substances no noise was recorded and thus no LOD could be calculated, this was especially observed for the R script. In addition, figure 3-16 shows that this approach highly depends on the method for the determination of noise, that can vary depending on the applied data processing method. The difficulties for HRMS instruments for LOD calculation based on S/N 3 due to very low background noise in the chromatogram was already discussed in literature. Especially for Orbitrap MS instruments often a baseline noise cut-off is applied by the instrument control

software by default to reduce data-file size. The use of S/N based approaches is thus not recommend for HRMS instruments, yet it is widely applied [18].

Therefore, as third approach a determination of the method detection limit (MDL) based on standard deviations of 10 injections of 100 ng/L standard was included. MDLs of all compounds are in a range of 3 to 18.5 ng/L with a mean of 9 ng/L (see table 3-3). Figure 3-17 shows a comparison of all different approaches used for determination of sensitivity. Lowest values were obtained with S/N 3 method with a mean of 3.2 ng/L (R-script) (right y-axis) and most conservative estimation of the sensitivity with the calculation after DIN 3246 with a mean of 17 ng/L.

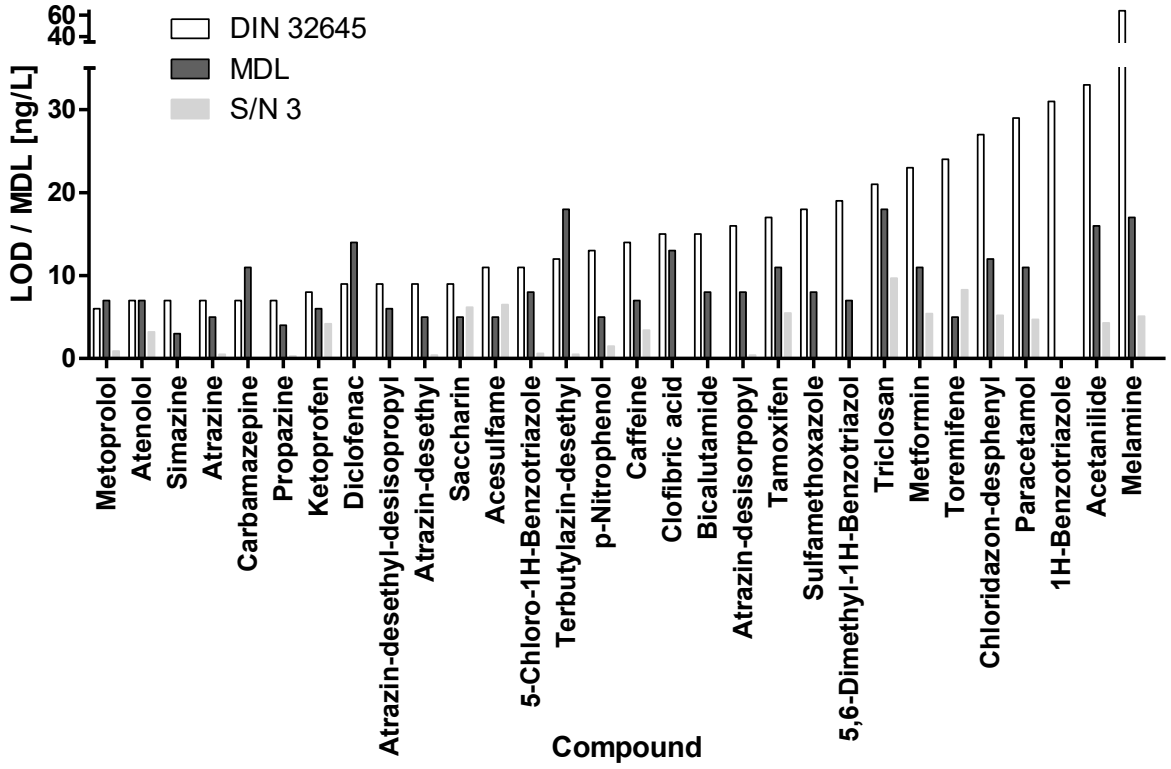


Figure 3-17: Comparison of LODs of Mix C in UPW after DIN 32345, MDL calculated from 10 injections of 100 ng/L and S/N for triplicate injection of 10 ng/L evaluated with R-script, all measurements with 100 µL injection volume.

The comparison of different approaches for the individual compounds in figure 3-17 demonstrates compound-dependent differences among the sensitivity estimates from the three approaches. In addition, this comparison highlights that LOD calculation is highly depending on used calculation approach and comparison with literature values must be done with care. Nevertheless, in comparison to extremely low LODs calculated from S/N 3 approach, the differences between MDL and LOD after DIN 32645 are in the same range in most of the cases. Table 3-13 gives an overview of all LODs calculated based on different approaches and with different injection volumes.

Table 3-13: Overview of LODs for Mix C in UPW calculated based on S/N of 3 for different data processing methods, after MDL and after DIN 32456 with different injection volumes.

Compound	S/N 3 *Quan Browser*	S/N 3 *R script*	MDL	DIN 32456 *20 µL*	DIN 32456 *100 µL*
1H-Benzotriazole	0.3	*	*	24	31
5,6-Dimethyl-1H-Benzotriazole	0.0	0.1	7	11	19
5-Chloro-1H-Benzotriazole	0.2	0.6	8	16	11
Acesulfame	0.1	6.5	5	108	11
Acetanilide	0.1	4.3	16	21	33
Atenolol	0.2	3.2	7	21	7
Atrazine-desethyl	0.2	0.4	5	12	9
Atrazine-desethyl-desisopropyl	0.2	**	6	15	9
Atrazine-desisorpopyl	0.2	0.4	8	16	16
Atrazine	0.2	0.5	5	13	7
Bicalutamide	0.0	**	8	14	15
Caffeine	0.0	3.4	7	30	14
Carbamazepine	0.1	**	11	13	7
Chloridazon-desphenyl	1.6	5.2	12	20	27
Clofibric acid	0.2	*	13	38	15
Diclofenac	0.4	**	14	24	9
Ketoprofen	0.1	4.2	6	19	8
Melamine	0.1	5.1	17	57	64
Metformin	*	5.4	11	51	23
Metoprolol	0.1	0.9	7	16	6
Paracetamol	0.2	4.7	11	24	29
p-Nitrophenol	0.0	1.5	5	16	13
Propazine	0.1	0.3	4	13	7
Saccharin	0.0	6.2	5	19	9
Simazine	0.0	0.2	3	14	7
Sulfamethoxazole	0.1	**	8	14	18
Tamoxifen	0.4	5.5	11	80	17
Terbutylazin-desethyl	0.1	0.5	18	30	12
Toremifene	0.3	8.3	5	72	24
Triclosan	0.4	9.7	18	48	21

All values are given in ng/L. Determination based on S/N 3 and MDL with 100 µL injection volume

** not detected at concentration 10 ng/L (S/N 3) or 100 ng/L (MDL)*

*** no noise detected*

It can be concluded that a higher sensitivity was reached with the higher injection volume of 100 µL as expected. All applied methods for determination of LODs showed different results and had limitations. Nevertheless, with all approaches, even with the most conservative determination after DIN, a high sensitivity and thus suitability for the detection of environmental concentrations in the lower ng/L range without prior enrichment was verified. Nevertheless, the determination was performed in ultrapure water and effects from sample matrix were not

considered. However, matrix effects vary with different sample matrices and are individual for each compound and a general evaluation of the matrix effect is thus not feasible. Nevertheless, the matrix effect should be examined, e.g., with internal standards for environmental samples for each study to ensure sensitivity.

Finally, as a higher variation of peak areas was observed during LOD calculation after DIN 32456 for some compounds with the 100 µL injection volume the repeatability of peak areas was tested after changing the instrument setup. For this purpose, 10 injections of 1 µg/L Mix C in UPW were performed with 20 µL sample loop and full loop injection and with 200 µL sample loop with partial injection of 20 and 100 µL. Relative standard deviation (RSD%) of peak areas and standard deviations (SD) of RTs for each setting were calculated. The measurements with 200 µL sample loop with partial injection of 20 µL were included, to assess if lower injection volumes are still possible with this setup. Results displayed in figure 3-18 show a RSD below 10 or even 5% for most compounds with different loops and injection volume. For the 20 µL injection volume with the 200 µL loop, two compounds showed a high deviation of peak areas (triclosan and clofibric acid).

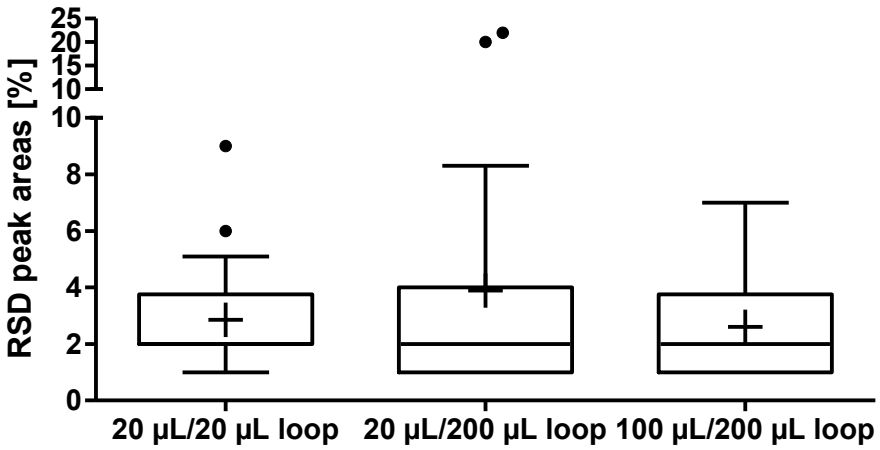


Figure 3-18: Repeatability of peak areas with different sample loops and injection volume. Evaluated for 10 replicate measurements of 1µg/L Mix C in UPW for each setting.

Results, displayed in figure 3-19 show that repeatability of retention times was high with mean values below 1 second for all settings.

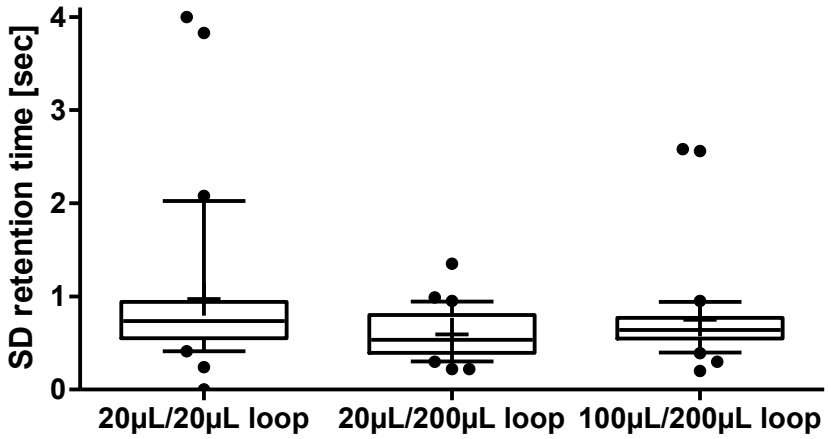


Figure 3-19: Repeatability of retention times with different sample loops and injection volume. Evaluated for 10 replicate measurements of 1 µg/L Mix C for each setting.

Thus, it can be concluded that overall repeatability was not compromised by changing the setup of the HPLC injection system.

3.4. Conclusion and outlook

The method shows sufficient sensitivity and selectivity to detect a broad range of OMPs at environmentally relevant concentrations. In general, different parameter settings during ionization and MS detection did not highly influence sensitivity or mass accuracy of the measurement. Thus, the range of parameter settings found in literature can be applied for a generic OMP screening with satisfying results. However, as the optimal setting for the AGC target is dependent on the ion current should be adjusted if extremely different sample matrices or highly enriched extracts are measured. As for Orbitrap MS instruments, the mass resolution is inversely related to scanning speed, tradeoffs must be found. This is also the case for chromatographic peak separation, as broader peaks are necessary for Orbitrap detection. The stability of peak areas and retention times is satisfying and enables the data evaluation and comparison of samples of large sampling campaigns. However, quality measures need to be implemented to monitor sample and measurement stability. The findings on detection limits highly variated depending on applied determination approach but were satisfactory in all cases. However, for each sample matrix the sensitivity should be examined with internal standards, even though only qualitative measurements are performed. Further on, the method development and evaluation were based on a limited number of model compounds. As the method is designed for a broad unknown screening, it would be of high interest to evaluate different parameter settings during method development based on the general quality of all measured signals. However, such a quality measure is not available to date but future research on it would be beneficial for method development in NTS studies. The Orbitrap HRMS detection shows a high mass accuracy and recording of MS^2 spectra for a high proportion of analytes. Nevertheless. improvement of MS^2 detection parameters (e.g. different NCE settings) or a different acquisition mode of MS^2 detection is a field of development for future research. The drawback of the data dependent acquisition modus, which was also used in this study, is that not for all analytes MS^2 are recorded. As feature prioritization is done retrospectively, recording of MS^2 spectra for selected compounds is not ensured. An alternative is the MS^2 acquisition in data independent mode (DIA) which reflects more the objectives of NTS. However, DIA MS^2 spectra have to date the main drawback that spectra interpretation is complicated as precursor and their fragment ions are not connected with each other. This makes MS^2 spectra deconvolution necessary to enable database search [35].

As for any LC-HRMS method the scope of analytes that can be detected with this method has limitations. Recently, more scientific focus has been put on very polar compounds as these compounds can be potentially problematic for drinking water resources [24]. New developments in the NTS field regarding instrumentation have thus a focus on broadening the analytical window. Chromatographic separation on hydrophilic interaction liquid

chromatography (HILIC) phases [25], a supercritical fluid chromatography (SFC) [26] or capillary electrophoresis (CE) [27] have been proposed. But also a coupling of HRMS with gas chromatography [117] or additional ionization methods such as atmospheric pressure chemical ionization (APCI) or atmospheric pressure photoionization (APPI) [101] can complete the detectable substance spectrum. Furthermore, two-dimensional separation during chromatography with orthogonality of separation mechanisms [118] or the implementation of an additional ion mobility (IM) separation step [119] before MS detection have been proposed in literature and are interesting instrumental fields of development.

3.5 Supporting information

3.5.1 Overview of compounds contained in Mix D

Table S 3-1: List of sum formula, ionization polarity, measured adduct and m/z of all compounds of Mix D.

Compound Name	Sum formula	Polarity	m/z	Adduct
Clothianidin	C6H8ClN5O2S	Positive	250.016	M+H
(+)-Methamphetamine	C10H15N	Positive	150.12773	M+H
2,6-Dichlorbenzamid	C7H5Cl2NO	Positive	189.9821	M+H
2-Hydroxybenzothiazol	C7H5NOS	Positive	152.01646	M+H
2-Methylmercaptobenzothiazol	C8H7NS2	Positive	182.00927	M+H
4-Acetamidoantipyrin (AAA)	C13H15N3O2	Positive	246.1237	M+H
4-Formylaminoantipyrin (FAA)	C12H13N3O2	Positive	232.10805	M+H
Acephate	C4H10NO3PS	Positive	184.01918	M+H
Aclonifen	C12H9ClN2O3	Positive	265.03745	M+H
Alprenolol	C15H23NO2	Positive	250.18016	M+H
Amidotrizic acid	C11H9I3N2O4	Positive	631.80346	$M+NH_4$
Amisulpride	C17H27N3O4S	Positive	370.1795	M+H
Amoxilline	C16H19N3O5S	Positive	366.11182	M+H
Amphetamine	C9H13N	Positive	136.11208	M+H
Anatoxin-a	C10H15NO	Positive	166.12264	M+H
Asulame	C8H10N2O4S	Positive	231.0434	M+H
Atenolol	C14H22N2O3	Positive	267.17032	M+H
Atrazine	C8H14ClN5	Positive	216.10105	M+H
Atrazine-desethyl	C6H10ClN5	Positive	188.06975	M+H
Atrazine-desisopropyl	C5H8ClN5	Positive	174.0541	M+H
Azilsartan	C25H20N4O5	Positive	457.15065	M+H
Bentazon-N-methyl	C11H14N2O3S	Positive	255.07979	M+H
Benzotriazole	C6H5N3	Positive	120.05562	M+H
Benzotriazol-5,6-dimethyl	C8H9N3	Positive	148.08692	M+H
Bisoprolol	C18H31NO4	Positive	326.23258	M+H
Bupirimate	C13H24N4O3S	Positive	317.16419	M+H
Candesartan	C24H20N6O3	Positive	441.16697	M+H
Candesartan-cilexetil	C33H34N6O6	Positive	611.26126	M+H
Carbamazepine	C15H12N2O	Positive	237.10224	M+H
Carbamazepine	C15H12N2O	Positive	220.07529	M+H
Carbamazepine	C15H12N2O	Positive	194.09612	M+H
Carbamazepine 10,11-dihydro-10,11-dihydroxy	C15H14N2O3	Positive	271.10772	M+H
Carbamazepin-10,11-epoxide	C15H12N2O2	Positive	253.09715	M+H
Carbendazim	C9H9N3O2	Positive	192.07675	M+H
Carbetamid	C12H16N2O3	Positive	237.12337	M+H
Carfentrazon-ethyl	C15H14Cl2F3N3O3	Positive	412.04371	M+H
Cetirizin_x_2HCl	C21H25ClN2O3	Positive	389.16265	M+H
Chinolin	C9H7N	Positive	130.06513	M+H
Chloridazon	C10H8ClN3O	Positive	222.04287	M+H
Chloridazon	C10H8ClN3O	Positive	146.01158	M+H

Compound Name	Sum formula	Polarity	m/z	Adduct
Chloridazon-desphenyl	C4H4ClN3O	Positive	146.01157	M+H
Chloridazon-methyldesphenyl	C5H6ClN3O	Positive	160.02722	M+H
Chlorthalonil-M05	C8H4Cl3NO3	Positive	267.93295	M-H
Chlortoluron	C10H13ClN2O	Positive	213.07892	M+H
Ciprofloxacin	C17H18FN3O3	Positive	332.1405	M+H
Citalopram	C20H21FN2O	Positive	325.17107	M+H
Clarithromycin	C38H69NO13	Positive	748.48417	M+H
Clindamycin	C18H33ClN2O5S	Positive	425.18715	M+H
Clodinafop-propargyl	C17H13ClFNO4	Positive	350.05899	M+H
Clomazone	C12H14ClNO2	Positive	240.07858	M+H
Codein	C18H21NO3	Positive	300.15942	M+H
Caffeine	C8H10N4O2	Positive	195.08765	M+H
Danofloxacin	C19H20FN3O3	Positive	358.15615	M+H
Desmedipham/Phenmedipham	C16H16N2O4	Positive	301.11828	M+H
Dexamethason	C22H29FO5	Positive	393.20718	M+H
Diazepam	C16H13ClN2O	Positive	285.07892	M+H
Diclofenac	C14H11Cl2NO2	Positive	296.02396	M+H
Diethyltoluamid (DEET)	C12H17NO	Positive	192.13829	M+H
Diflufenican	C19H11F5N2O2	Positive	395.08135	M+H
Dimefuron	C15H19ClN4O3	Positive	339.12184	M+H
Dimethachlor	C13H18ClNO2	Positive	256.10988	M+H
Dimethenamid	C12H18ClNO2S	Positive	276.08195	M+H
Diuron	C9H10Cl2N2O	Positive	233.02429	M+H
Diuron-D6_IS	C9H4 2H6Cl2N2O	Positive	239.06196	M+
DMSA	C8H12N2O2S	Positive	201.06922	M+H
DMST	C9H14N2O2S	Positive	215.08487	M+H
Enrofloxacin	C19H22FN3O3	Positive	360.1718	M+H
Ephedrin	C10H15NO	Positive	166.12264	M+H
Eprosartan	C23H24N2O4S	Positive	425.15295	M+H
Erythromycin	C37H67NO13	Positive	734.46852	M+H
Ethidimuron	C7H12N4O3S2	Positive	265.04236	M+H
Ethofumesat	C13H18O5S	Positive	287.09477	M+H
Fenoxaprop-P-ethyl	C18H16ClNO5	Positive	362.07898	M+H
Fenpropidin	C19H31N	Positive	274.25293	M+H
Flecainide	C17H20F6N2O3	Positive	415.14509	M+H
Flufenacet	C14H13F4N3O2S	Positive	364.07374	M+H
Flumioxazin	C19H15FN2O4	Positive	355.10886	M+H
Fluorochloridon	C12H10Cl2F3NO	Positive	312.01643	M+H
Flurtamone	C18H14F3NO2	Positive	334.10494	M+H
Foramsulfuron	C17H20N6O7S	Positive	453.11869	M+H
Fuberidazol	C11H8N2O	Positive	185.07094	M+H
Gabapentin	C9H17NO2	Positive	172.13321	M+H
Gabapentin-lactam	C9H15NO	Positive	154.12264	M+H
Guanylharnstoff	C2H6N4O	Positive	103.06144	M+H
Hexazinon	C12H20N4O2	Positive	253.1659	M+H
Imidacloprid	C9H10ClN5O2	Positive	256.05958	M+H

Compound Name	Sum formula	Polarity	m/z	Adduct
Indomethacin	C19H16ClNO4	Positive	358.08406	M+H
Iohexol	C19H26I3N3O9	Positive	821.88758	M+H
Iomeprol	C17H22I3N3O8	Positive	777.86137	M+H
Iopamidol	C17H22I3N3O8	Positive	777.86137	M+H
Iopromid	C18H24I3N3O8	Positive	791.87702	M+H
Ioxitalamin acid	C12H11I3N2O5	Positive	644.78748	M+H
Iprovalicarb	C18H28N2O3	Positive	321.21727	M+H
Irbesartan	C25H28N6O	Positive	429.23974	M+H
Isoproturon	C12H18N2O	Positive	207.14919	M+H
Isoxaflutole	C15H12F3NO4S	Positive	360.05119	M+H
Ketoprofen	C16H14O3	Positive	255.10157	M+H
Lamotrigin	C9H7Cl2N5	Positive	256.01513	M+H
Lidocain	C14H22N2O	Positive	235.18049	M+H
Losartan	C22H23ClN6O	Positive	423.16946	M+H
Marbofloxacin	C17H19FN4O4	Positive	363.14631	M+H
Melamin	C3H6N6	Positive	127.07267	M+H
Mesotrione	C14H13NO7S	Positive	340.04855	M+H
Metalaxyl	C15H21NO4	Positive	280.15433	M+H
Metalaxyl CGA 62826	C14H19NO4	Positive	266.13868	M+H
Metamitron	C10H10N4O	Positive	203.09274	M+H
Metamitron-desamino	C10H9N3O	Positive	188.08184	M+H
Metazachlor	C14H16ClN3O	Positive	278.10547	M+H
Metformin	C4H11N5	Positive	130.10872	M+H
Methabenzthiazuron	C10H11N3OS	Positive	222.06956	M+H
Methadon	C21H27NO	Positive	310.21654	M+H
Methadon-Metabolit EDDP	C20H23N	Positive	278.19033	M+H
Metobromuron	C9H11BrN2O2	Positive	259.00767	M+H
Metolachlor	C15H22ClNO2	Positive	284.14118	M+H
Metolachlor	C15H22ClNO2	Positive	252.11487	M+H
Metolachlor	C15H22ClNO2	Positive	212.08364	M+H
Metoprolol	C15H25NO3	Positive	268.19072	M+H
Metoprolol	C14H21NO4	Positive	268.15433	M+H
Metribuzin	C8H14N4OS	Positive	215.09611	M+H
Metronidazol	C6H9N3O3	Positive	172.07167	M+H
Microcystin-LR	C49H74N10O12	Positive	498.28166	M+H
Microcystin-RR	C49H75N13O12	Positive	519.79018	M+H
Microcystin-YR	C52H72N10O13	Positive	523.27129	M+H
Myclobutanil	C15H17ClN4	Positive	289.12145	M+H
N-Methylbenzenesulfonamide	C7H9NO2S	Positive	172.04268	M+H
Nadolol	C17H27NO4	Positive	310.20128	M+H
Nicosulfuron	C15H18N6O6S	Positive	411.10813	M+H
Nicosulfuron ASDM	C8H11N3O3S	Positive	230.05939	M+H
Nicosulfuron AUSN	C10H14N6O4S	Positive	315.087	M+H
Nicosulfuron HMUD	C14H16N6O6S	Positive	397.09248	M+H
Nicosulfuron UCSN	C10H13N5O5S	Positive	316.07102	M+H
Nicotine	C10H14N2	Positive	163.12298	M+H

Compound Name	Sum formula	Polarity	m/z	Adduct
Olmesartan	C24H26N6O3	Positive	447.21392	M+H
Orbifloxacin	C19H20F3N3O3	Positive	396.15295	M+H
Oxadixyl	C14H18N2O4	Positive	279.13393	M+H
Oxazepam	C15H11ClN2O2	Positive	287.05818	M+H
Oxcarbazepin	C15H12N2O2	Positive	253.09715	M+H
Parbendazol	C13H17N3O2	Positive	248.13935	M+H
Pendimethalin (Stomp)	C13H19N3O4	Positive	282.14483	M+H
Pethoxamid	C16H22ClNO2	Positive	296.14118	M+H
Phenazon (Antipyrin)	C11H12N2O	Positive	189.10224	M+H
Phenylalanine	C9H11NO2	Positive	166.08626	M+H
Picolinafen	C19H12F4N2O2	Positive	377.09077	M+H
Pinoxaden	C23H32N2O4	Positive	401.24348	M+H
Pregabalin	C8H17NO2	Positive	160.13321	M+H
Prilocaine	C13H20N2O	Positive	221.16484	M+H
Primidon	C12H14N2O2	Positive	219.1128	M+H
Propaquizafop	C22H22ClN3O5	Positive	444.13207	M+H
Propranolol	C16H21NO2	Positive	260.16451	M+H
Propyphenazon	C14H18N2O	Positive	231.14919	M+H
Prosulfocarb	C14H21NOS	Positive	252.14166	M+H
Pyroxsulam	C14H13F3N6O5S	Positive	435.0693	M+H
Quinmerac	C11H8ClNO2	Positive	222.03163	M+H
Quinmerac BH 515-2 (CA)	C11H6ClNO4	Positive	252.00581	M+H
Ranitidin	C13H22N4O3S	Positive	315.14854	M+H
Rimsulfuron	C14H17N5O7S2	Positive	432.06422	M+H
Ritalinic acid	C13H17NO2	Positive	220.13321	M+H
Ronidazol	C6H8N4O4	Positive	201.06183	M+H
Roxithromycin	C41H76N2O15	Positive	837.53185	M+H
Simazin / Terbuthylazin-desethyl	C7H12ClN5	Positive	202.0854	M+H
Sitagliptin	C16H15F6N5O	Positive	408.12536	M+H
Sotalol	C12H20N2O3S	Positive	273.12674	M+H
Spice JWH-073	C23H21NO	Positive	328.16959	M+H
Spice JWH-18	C24H23NO	Positive	342.18524	M+H
Sulcotrione	C14H13ClO5S	Positive	329.0245	M+H
Sulcotrione	C14H13ClO5S	Positive	351.0064	M+H
Sulfadiazin	C10H10N4O2S	Positive	251.05972	M+H
Sulfadiazin-n-acethyl	C12H12N4O3S	Positive	293.07029	M+H
Sulfadimethoxin	C12H14N4O4S	Positive	311.08085	M+H
Sulfadoxin	C12H14N4O4S	Positive	311.08085	M+H
Sulfaethoxypyridazin	C12H14N4O3S	Positive	295.08594	M+H
Sulfamerazin	C11H12N4O2S	Positive	265.07537	M+H
Sulfamethazin (Sulfadimidin)	C12H14N4O2S	Positive	279.09102	M+H
Sulfamethizol	C9H10N4O2S2	Positive	271.03179	M+H
Sulfamethoxazol	C10H11N3O3S	Positive	254.05939	M+H
Sulfamethoxazol-acetyl	C12H13N3O4S	Positive	296.06995	M+H
Sulfamethoxypyridazin	C11H12N4O3S	Positive	281.07029	M+H

Compound Name	Sum formula	Polarity	m/z	Adduct
Sulfapyridin	C11H11N3O2S	Positive	250.06447	M+H
Sulfaquinoxalin	C14H12N4O2S	Positive	301.07537	M+H
Sulfathiazol	C9H9N3O2S2	Positive	256.02089	M+H
Tebuconazol	C16H22ClN3O	Positive	308.15242	M+H
Telmisartan	C33H30N4O2	Positive	515.24415	M+H
Temazepam	C16H13ClN2O2	Positive	301.07383	M+H
Terbuthylazin	C9H16ClN5	Positive	230.1167	M+H
Terbuthylazin-2-hydroxy	C9H17N5O	Positive	212.15059	M+H
Terbutryn	C10H19N5S	Positive	242.14339	M+H
Tetracycline	C22H24N2O8	Positive	445.16054	M+H
Tetramethylurea	C5H12N2O	Positive	117.10224	M+H
Thiabendazol	C10H7N3S	Positive	202.04334	M+H
Thiacloprid	C10H9ClN4S	Positive	253.03092	M+H
Thiamethoxam	C8H10ClN5O3S	Positive	292.02656	M+H
Thiencarbazone-methyl	C12H14N4O7S2	Positive	391.03767	M+H
Tiamulin	C28H47NO4S	Positive	494.32986	M+H
Tolyltriazol (Isomers 4- und 5-Methylbenzotriazol)	C7H7N3	Positive	134.07127	M+H
Tramadol	C16H25NO2	Positive	264.19581	M+H
Tramadol-N-desmethyl	C15H23NO2	Positive	250.18016	M+H
Tramadol-N-Oxid	C16H25NO3	Positive	280.19072	M+H
Tramadol-O-desmethyl	C15H23NO2	Positive	250.18016	M+H
Triethylphosphate	C6H15O4P	Positive	183.07807	M+H
Trifloxystrobin CGA 321113	C19H17F3N2O4	Positive	395.12132	M+H
Trifloxystrobin NOA (413161, 413163 Isomere)	C19H15F3N2O6	Positive	425.0955	M+H
Trimethoprim	C14H18N4O3	Positive	291.14517	M+H
Tylosin	C46H77NO17	Positive	916.52643	M+H
Tyrosin	C9H11NO3	Positive	182.08117	M+H
Venlaflaxin	C17H27NO2	Positive	278.21146	M+H
Venlaflaxin-O-desmethyl	C16H25NO2	Positive	264.19581	M+H
2,4-D (2,4-Dichlorphenoxy acetic acid)	C8H6Cl2O3	Negative	218.9621	M-H
2,4-D (2,4-Dichlorphenoxy acetic acid)	C8H6Cl2O3	Negative	160.9565	M-H
2,4-DB	C10H10Cl2O3	Negative	160.9567	M-H
2,4-Dinitrophenol (DNP)	C6H4N2O5	Negative	183.00474	M-H
2-Mercapto-benzothiazole	C7H5NS2	Negative	165.97906	M-H
3-Trifluoromethylbenzoic acid	C8F3H5O2	Negative	189.01689	M-H
Acesulfame	C4H5NO4S	Negative	161.98665	M-H
Bentazon	C10H12N2O3S	Negative	239.04959	M-H
Bezafibrat	C19H20ClNO4	Negative	360.10081	M-H
Bromacil	C9H13BrN2O2	Negative	259.0088	M+H
Bromoxynil	C7H3Br2NO	Negative	273.8509	M-H
Bromoxynil	C7H3Br2NO	Negative	275.8488	M-H
Caprylsaeure	C8H16O2	Negative	143.10775	M-H

Compound Name	Sum formula	Polarity	m/z	Adduct
Chlorthalonil-M12	C8H3Cl3N2O4S	Negative	326.88063	M-H
Clofibric acid	C10H11ClO3	Negative	213.0324	M-H
Clopyralid	C6H3Cl2NO2	Negative	189.94681	M-H
Clothianidin	C6H8ClN5O2S	Negative	248.00145	M-H?
Cyclamat_Na	C6H13NO3S	Negative	178.05434	M-H
Dicamba	C8H6Cl2O3	Negative	218.9621	M-H
Dicamba	C8H6Cl2O3	Negative	174.9725	M-H
Dichlorprop (2,4 DP)	C9H8Cl2O3	Negative	232.9778	M-H
Dimethachlor CGA 354742 (ESA)	C13H19NO5S	Negative	300.09112	M-H
Dimethachlor CGA 50266 (OA)	C13H17NO4	Negative	250.1085	M-H
Dimethenamid M23 (OA)	C12H17NO4S	Negative	270.08055	M-H
Dimethenamid M27 (ESA)	C12H19NO5S2	Negative	320.06319	M-H
Dinoterb	C10H12N2O5	Negative	239.06735	M-
Fenoprop	C9H7Cl3O3	Negative	266.9388	M-H
Fluazinam	C13H4Cl2F6N4O4	Negative	463.9508	M-
Fluazinam	C13H4Cl2F6N4O4	Negative	462.9441	M-
Flufenacet M9 (Thiadone)	C3HF3N2OS	Negative	168.96889	M-H
Flufenacet-ESA	C11H14FNO4S	Negative	274.05548	M-H
Flufenacet-OA	C11H12FNO3	Negative	224.07284	M-H
Fluroxypyr	C7H5Cl2FN2O3	Negative	252.95885	M-H
Gemfibrozil	C15H22O3	Negative	249.1495	M+H
Gemfibrozil	C15H22O3	Negative	250.1529	M+H
H4-PFOS (C8)	C8H5F13O3S	Negative	426.9679	M-H
HPFHA (C7)	C7H2F12O2	Negative	344.97904	M-H
Ibuprofen	C13H18O2	Negative	205.1234	M-H
Ioxynil	C7H3I2NO	Negative	369.82312	M-H
MCPA	C9H9ClO3	Negative	199.0168	M-H
MCPB	C11H13ClO3	Negative	227.0481	M-H
MCPB	C11H13ClO3	Negative	141.0117	M-H
Mecoprop (MCPP)	C10H11ClO3	Negative	213.0324	M-H
Mecoprop (MCPP)	C10H11ClO3	Negative	141.0113	M-H
Metalaxyl CGA 108906 (CA)	C14H17NO6	Negative	294.09831	M-H
Metazachlor BH 479-04 (OA)	C14H15N3O3	Negative	272.10406	M-H
Metazachlor BH 479-08 (ESA)	C14H17N3O4S	Negative	322.0867	M-H
Metolachlor CGA 351916 (OA)	C15H21NO4	Negative	278.13978	M-H
Metolachlor CGA 380168 (ESA)	C15H23NO5S	Negative	328.12242	M-H
Nodularin	C41H60N8O10	Negative	823.43596	M-H
Oxipurinol	C5H4N4O2	Negative	151.02615	M-H
Pentobarbital	C11H18N2O3	Negative	225.12447	M-H
Pethoxamid MET-42	C16H23NO5S	Negative	340.12242	M-H
PFBA (C4-acid)	C4HF7O2	Negative	212.9792	M-H
PFBS (C4-Sulfonat)	C4HO3F9S	Negative	298.94299	M-H
PFDA (C10-acid)	C10HF19O2	Negative	512.96004	M-H

Compound Name	Sum formula	Polarity	m/z	Adduct
PFDS (C10-Sulfonat)	C10HF21O3S	Negative	598.92383	M-H
PFHA (C7-acid)	C7HF13O2	Negative	362.96962	M-H
PFHS (C7-Sulfonat)	C7HF15O3S	Negative	448.93341	M-H
PFHxA (C6-acid)	C6HF11O2	Negative	312.97281	M-H
PFHxS (C6-Sulfonat)	C6HF13O3S	Negative	398.9366	M-H
PFNA (C9-acid)	C9HF17O2	Negative	462.96323	M-H
PFNS (C9-Sulfonat)	C9HF19O3S	Negative	548.92702	M-H
PFOA (C8-acid)	C8HF15O2	Negative	412.96643	M-H
PFOS (C8-Sulfonat)	C8HF17O3S	Negative	498.93022	M-H
PFPA (C5-acid)	C5HF9O2	Negative	262.97601	M-H
PFPS (C5-Sulfonat)	C5HF11O3S	Negative	348.9398	M-H
Phenobarbital	C12H12N2O3	Negative	231.07752	M-H
Pikrins acid	C6H3N3O7	Negative	227.98982	M-H
Saccharin_Na	C7H5NO3S	Negative	181.99174	M-H
Secobarbital	C12H18N2O3	Negative	237.12447	M-H
Sulbactam	C8H11NO5S	Negative	232.02852	M-
Tembotrione	C17H16ClF3O6S	Negative	439.02354	M-H
Terbacil	C9H13ClN2O2	Negative	215.05928	M-H
Topramezon M670H05	C12H13NO5S	Negative	282.04417	M-H
Topramezone	C16H17N3O5S	Negative	362.0841	M-H
Topramezone-M01	C14H13N3O4S	Negative	318.0554	M-H
Triclopyr	C7H4Cl3NO3	Negative	195.9129	M-H
Triclosan	C12H7Cl3O2	Negative	286.94389	M-H
Tritosulfuron	C13H9F6N5O4S	Negative	444.02067	M-H
Valsartan	C24H29N5O3	Negative	434.21976	M-H
Valsartanic acid	C14H10N4O2	Negative	265.0731	M-H
Zearalenol	C18H24O5	Negative	319.1551	M-H

*for some compounds several adducts or fragments were evaluated

3.5.2 Stability of matrix sample

The pH value was measured with a pH electrode (827 pH lab) from Metrohm (Filderstadt, Germany). The UV-Vis spectrum was recorded with an UV-Vis spectrometer (UV-1650 PC) operating with the software UV-Probe from Shimadzu (Duisburg, Germany). The NPOC was measured with a TOC-L analyzer equipped with ASI-L autosampler from Shimadzu (Duisburg, Germany). The parameters of the method are summarized in table S3-2. First, method A was used, but as no acid was added, the determination of NPOC was not correct. Thus, a new method including the addition of acid (Method B) was used. However, in the following measurement with both methods was performed to be able to compare the results of previous measurements to assess the stability of the matrix.

Table S 3-2: Summary of the parameter used for the determination of NPOC in matrix.

Parameter	Method A	Method B
Number injection	3	5
Number wash cycle	2	2
Maximal coefficient of variation	2%	2%
Acid addition	0%	3%
Purge flow rate	80 mL	80 mL
Purge time	1 min	1 min
Automatic dilution	1	1
Injection volume	40 µL	40 µL
Expected concentration	50 mg/L	50 mg/L

Table S3-3 shows the results of the monitoring of the stability of the pH value and NPOC content of the matrix. NPOC Method A is included to show the stability of the NPOC. However, NPOC Method B shows the correctly determined NPOC content.

Table S 3-3:NPOC and pH value of the surface water matrix.

Date	pH	NPOC Method A [mg/L]	NPOC Method B [mg/L]
16.08.2017	7.96	22.50	
31.07.2017	8.02	22.35	
25.08.2017	8.01	23.05	
06.09.2017	8.08	22.52	3.79
20.09.2017	8.01	22.38	3.80

3.5.3 Additional figures of ionization parameter optimization

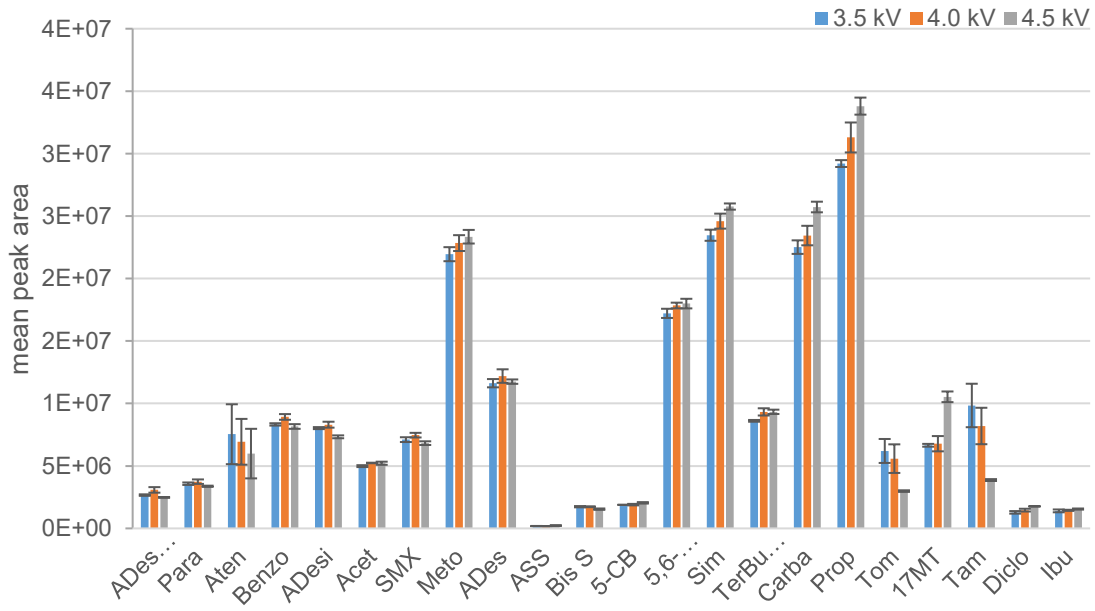


Figure S 3-1: Mean peak areas at different spray voltages for the positive ionization mode. Error bars show SD. Mix A 1 µg/Lin UPW.

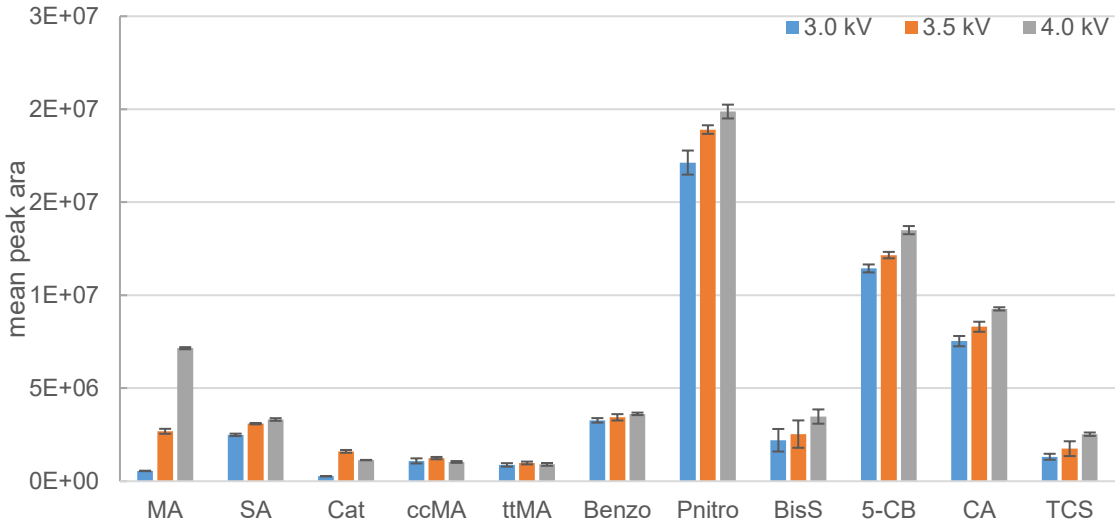


Figure S 3-2: Mean peak areas at different spray voltages for the negative ionization mode. Error bars show SD Mix A 1 µg/L in UPW.

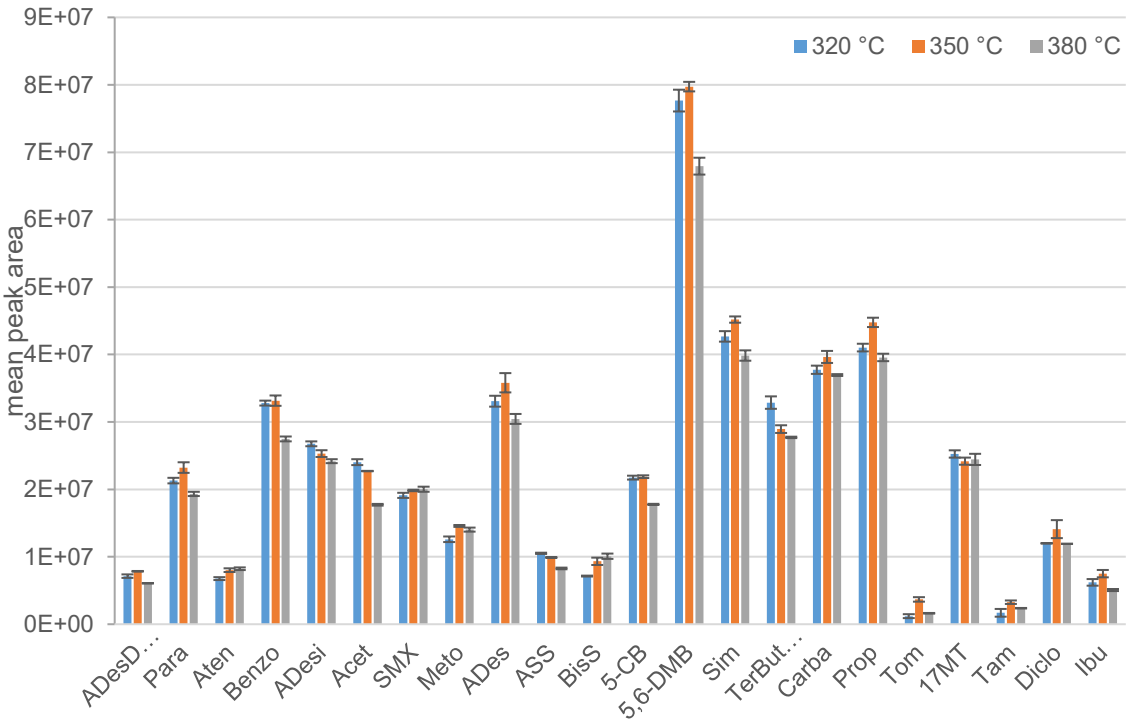


Figure S 3-3: Mean peak areas of triplicate injection at different spray temperatures for the positive ionization mode. Error bars show SD Mix A 10 µg/L in UPW.

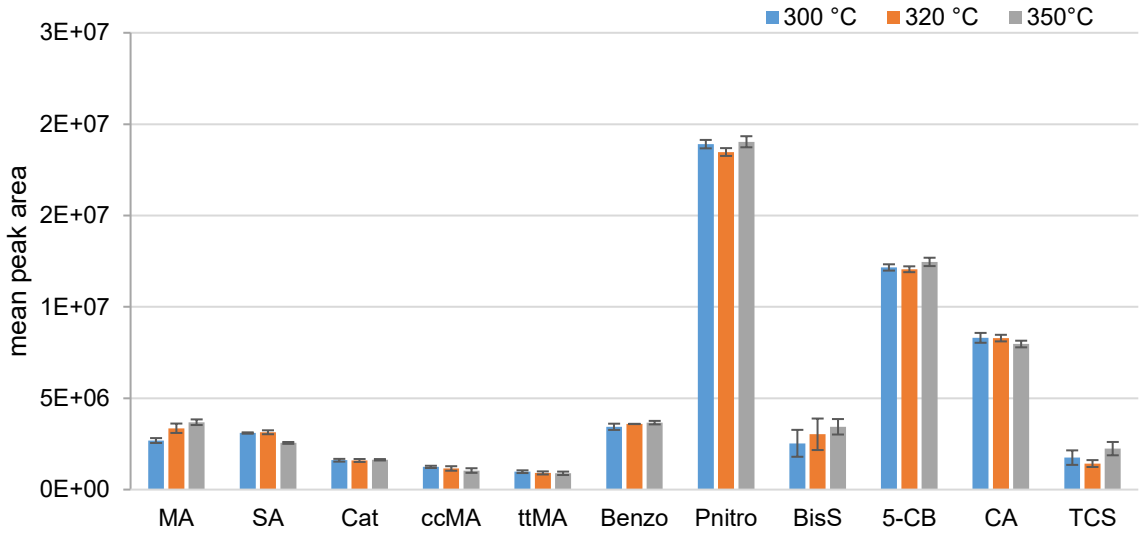


Figure S 3-4: Mean peak areas of triplicate injection at different spray temperatures for the negative ionization mode. Error bars show SD Mix A 10 µg/L in UPW.

3.5.4 Exemplary MS^2 spectra

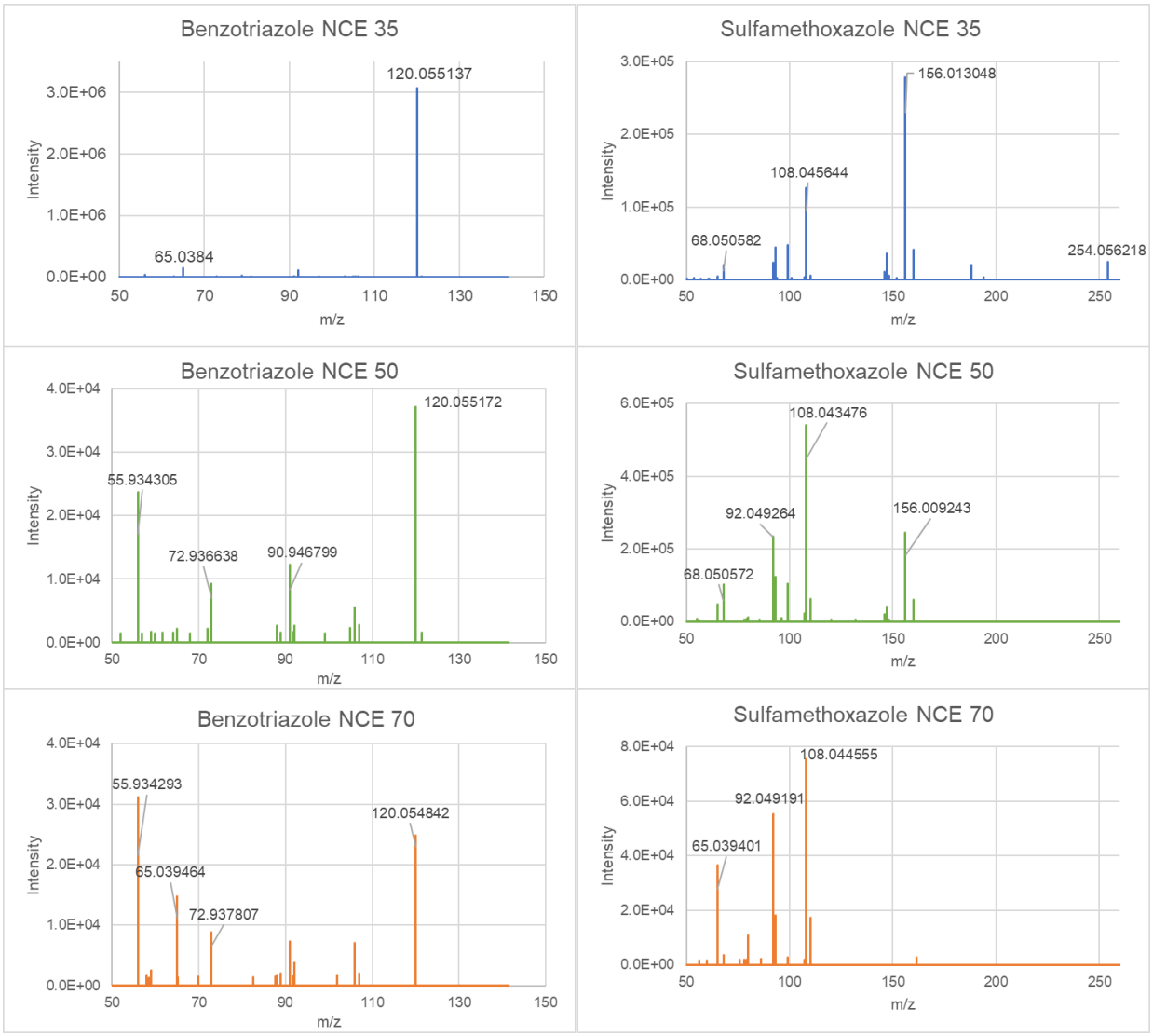


Figure S 3-5: Exemplary MS^2 spectra at different NCE settings for 1H-benzotriazole and sulfamethoxazole at 1 µg/L in UPW.

3.5.5 Calibration curve comparison with 20 and 100 µL injection volume for selected compounds

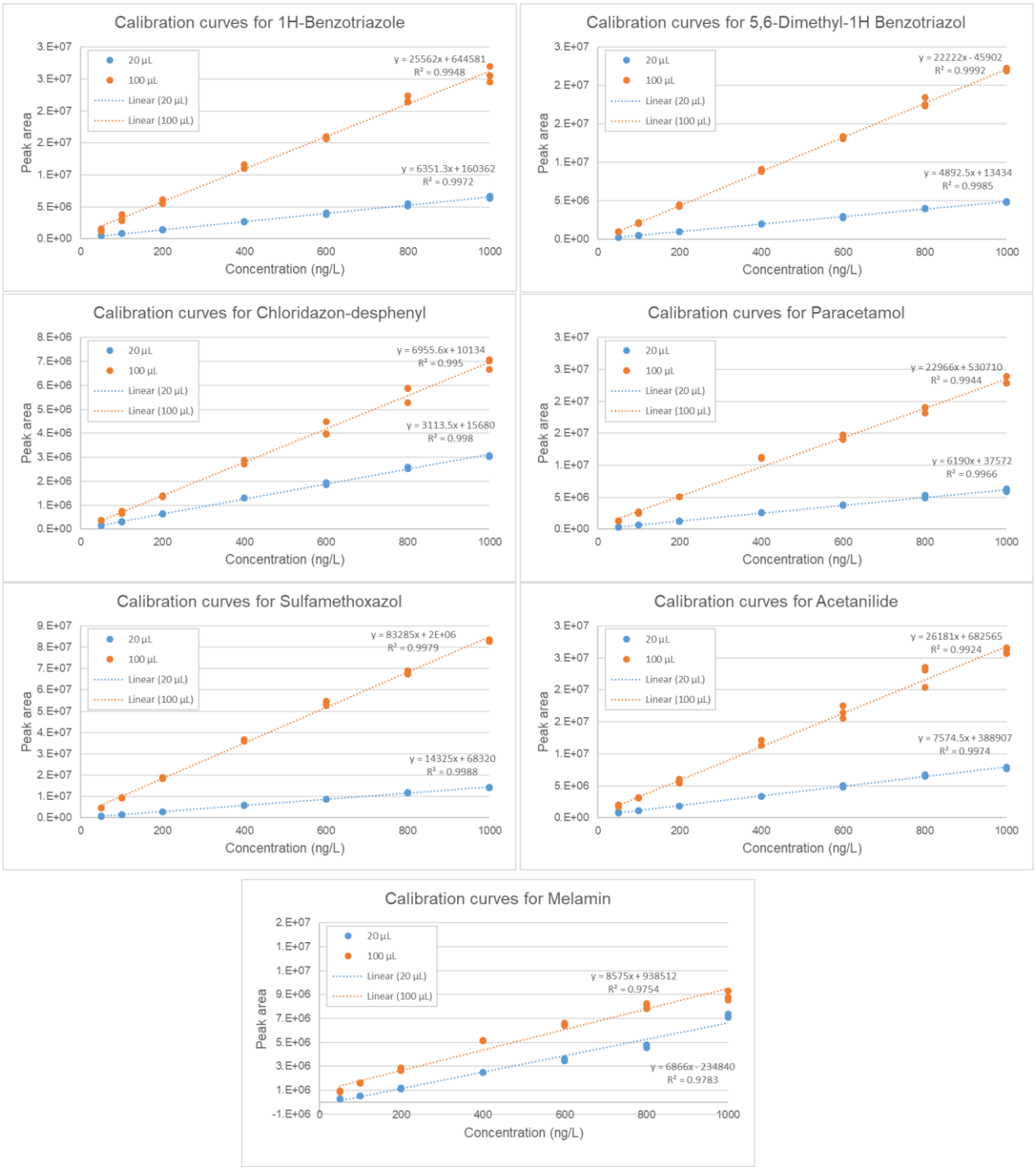


Figure S 3-6: Calibration curves for compounds with lower LODs obtained with 20 µL injection volume

Chapter 4 Comparison of software tools for LC-HRMS data processing in non-target screening of environmental samples

This chapter was adapted from: L.L. Hohrenk, F. Itzel, N. Baetz, J. Tuerk, M. Vosough, T.C. Schmidt, Comparison of Software Tools for Liquid Chromatography-High-Resolution Mass Spectrometry Data Processing in Nontarget Screening of Environmental Samples, Anal. Chem. 92 (2020) 1898–1907. doi.org/10.1021/acs.analchem.9b04095.

Abstract

The field of high resolution mass spectrometry has undergone a rapid progress in the last years due to instrumental improvements leading to a higher sensitivity and selectivity of instruments. A variety of qualitative screening approaches, summarized as non-target screening, have been introduced and have successfully extended the environmental monitoring of organic micropollutants. Several automated data processing workflows have been developed to handle the immense amount of data that are recorded in short time frames by these methods. Most data processing workflows include similar steps, but underlying algorithms and implementation of different processing steps vary. In this study the consistency of data processing with different software tools was investigated. For this purpose, the same raw data files were processed with the software packages MZmine2, enviMass, Compound Discoverer and XCMS online and resulting feature lists were compared. Results show a low coherence between different processing tools, as overlap of features between all four programs was around 10 % and for each software between 40 – 55 % of features did not match with any other program. The implementation of replicate and blank filter was identified as one of the sources of observed divergences. However, there is a need for a better understanding and user-instructions on the influence of different algorithms and settings on feature extraction and following filtering steps. In future studies it would be of interest to investigate how final data interpretation is influenced by different processing software. With this work we want to encourage more awareness on data processing as a crucial step in the workflow of non-target screening.

4.1 Introduction

Non-Target Screening (NTS) based on high resolution mass spectrometry (HRMS) coupled to liquid chromatography (LC) offers the potential of detecting a broad range of micropollutants at environmentally relevant concentration. These qualitative screening approaches give new opportunities for spatial- or time-trend analysis and process evaluation without previous compound identification, as well as retrospective data mining and the potential to identify formerly unknown compounds such as transformation products [100].Therefore, NTS has been a tool of increasing importance for the environmental monitoring of micropollutants in the last years [38]. However, as being a new field of research, standardization of analytical methods and workflows is not yet advanced. This was for example noticeable in a collaborative trial in the NTS field with 18 participating institutes, where a variety of instruments, LC- and HRMS-methods, software and analyte identification strategies were used [37].

The final, but yet often most time-consuming, step of a NTS workflow is data processing and interpretation. Sophisticated automatic data processing strategies are necessary as in each sample up to thousands of signals are detectable leading to the recording of a large amount of information in a short time frame. Several open-source and commercial software tools are available for data processing which are constantly improved and updated and new algorithms and tools are published. All steps of data extraction, filtering and analysis are time consuming and prone to false interpretation. Thus, data processing remains a main challenge in NTS approaches and is, besides the non-harmonized instrumentation and measurement part, an under recognized source of variation between laboratories and research groups.

Most data processing strategies include similar steps such as raw data preprocessing and filtering, peak picking, blank subtraction, replicate filtering and componentization, though their implementation can vary. The starting point of data processing is a set of raw data files consisting of successively recorded MS^1 scans [51]. Depending on the software and instrument used a conversion of the raw data files into a common open-source data format such as mzXML is necessary, by which centroidization of data takes place. During the centroiding process multiple data points of an ion peak in profile mode are combined into a single data point with one *m/z* and intensity value [120]. With either profile or centroid data the next step is the extraction of peaks, the so-called “peak picking”, resulting in a list of features (defined as mass-to-charge-ratio (*m/z*) at a certain retention time (RT) with a certain signal intensity). For this data extraction step, different algorithms based on different strategies are available. The *vectorized peak detection* method searches for data points separately in two directions (*m/z* and RT) and identifies peaks that meet an intensity threshold or peak shape criteria [51,121]. Another strategy is slicing data to extracted ion chromatograms (XIC or EIC) with narrow *m/z* ranges and subsequent peak integration in the time domain for each XIC [51].

Further, model fitting against the raw data can be used, for example by an iterative fitting of a three-dimensional model of a generic isotope pattern on the highest raw signal [51,122]. These first steps are crucial, as errors at this stage propagate throughout the entire data processing workflow and can affect subsequent data interpretation. However, any algorithm produces a number of false positive (FP) detected features caused by either chemical or random noise [51]. Random noise is attributed to the detector for example as residuals of the *Fourier Transformation* function if Orbitrap mass analysers are used [90]. Chemical noise is caused by background signals that are present in eluents, buffers or laboratory air or are introduced during sample preparation and measurement [53]. Therefore, several steps for data cleaning and filtering need to be implemented. A *blank filter* can be applied to remove signals not specific to the sample. It can either be done by subtraction of all features present in a blank sample data set or only of features that exceed a certain intensity ratio compared to the sample [103]. Bader et al. showed that the measurement of at least two or even three technical replicates of a sample and subsequent filtering of features that are pervasive, can be a useful tool as *replicate filter* to reduce FP [55]. When soft ionization methods such as electrospray ionization are used, signals from isotopic peaks, adducts, different charge states and in-source fragment ions of a compound should be grouped [38]. For isotope grouping most algorithms search for features with a lower intensity and a *m/z* difference which is a multiple of the neutron mass with the same retention time to identify isotopologues. The adduct search filters for peaks with a *m/z* difference defined by a certain adduct mass from features with the same retention time. However, this so-called *componentization* is not included in all software packages, sometimes adduct and isotopic peaks are subtracted, sometimes the information is annotated or even summarized into groups. Finally, an alignment of the features over several samples is necessary for further statistical analysis. For the optimization of peak detection, different programs include different input parameters, which can be adjusted by the user. Often these parameters are not well defined, and their influence on the results is not conceivable by the user. As different workflows and software packages are used by different research groups the question arises, how this may influence data evaluation. The influence of data processing programs was investigated by a few studies in the metabolomics/proteomics field [123–127], all showing considerable differences between performance of software tools. Rafei and Sleno found an overlap of feature lists between Peakview®, Markerview™, MetabolitePilot™ and XCMS online of less than 10 % [127]. Myers et al. compared the use of the same peak picking algorithm *centwave* implemented in XCMS and MZmine2 and found an overlap of 60 % [123]. Gürdeniz et al. found 37-46 % common features with MarkerLynx™, MZmine and XCMS [126]. In metabolomic studies often two groups (e.g. healthy and disease state) are compared and for this purpose features that are dysregulated are selected, fold changes are determined and biomarkers identified [48]. Therefore, most comparative studies in the metabolomics field

evaluated the performance of software tools based on the identification of biomarkers and their quantification. In addition, in most of these studies a replicate filter and blank subtraction was not generally implemented during data processing [123,124,126], but manually applied after processing in a few cases [125,127]. Beside different sample types and analyte classes in environmental NTS approaches research aims can differ. Common applications are investigations of time or spatial trends of a pollution [3,128] a process evaluation [72,73,129] or a suspect screening [128,130,131]. As a result of different research foci, data processing strategies differ although certain tools such as a replicate filter, blank subtraction and componentization is widely applied in NTS workflows.

In this work, it was studied how the use of different software tools influences NTS data processing results. The same dataset was processed with four software packages commonly used in NTS approaches. Compound Discoverer is the commercial software from Thermo Scientific for NTS workflows. XCMS online and MZmine2 are open-source programs that were developed in the *omics* field but are also frequently used in the NTS field [78,132,133]. EnviMass is a R based script that was specially developed for LC-MS based NTS data sets [134].

4.2. Material and methods

4.2.1 Sampling and sample preparation

Grab samples (table S4-1) were taken from the effluent of the wastewater treatment plant Warburg (Stadtwerke Warburg GmbH, Warburg, Germany) comprising of an ozonation and biological treatment step as well as of the receiving river Diemel upstream and downstream of the discharge at 3 different days over a period of three weeks (N=9). A solid phase extraction (SPE) was carried out within 48 h after sampling. SPE cartridges (150 mg, 6 mL, Oasis HLB, Waters, Germany) were conditioned (2 x 5 mL methanol, LC-MS grade) and equilibrated (2 x 5 mL water, LC-MS grade). Prior to the enrichment 10 ng of an internal standard mix (see table S4-2) were added. Subsequently cartridges were loaded with 1 L sample. The cartridges were dried under vacuum and stored at -18 °C until further sample preparation. The elution was done using 5 x 5 mL methanol (LC-MS grade) and the solvent was evaporated to complete dryness at 50 °C under a gentle nitrogen stream. Before analysis, the samples were re-dissolved in water (LC-MS grade). Further, SPE blanks were processed the same way as described above with ultra-pure water instead of sample [135].

4.2.2 Instrumentation

Chromatographic separation was performed using a Dionex UltiMate 3000 HPLC system (Thermo Scientific, Bremen, Germany). Gradient elution was carried out on a XSelect HSS T3 (2.1 mm x 75 mm, 3.5 µm particle size) column from Waters (Milford, MA, USA). The mobile

phase consisted of eluent A: ultrapure water + 0.1 % formic acid, and eluent B: methanol + 0.1 % formic acid. After an isocratic step with of 5 % B for 5 min, the concentration of eluent B was raised to 95 % within 10 min and kept constant for 10 min. Following, the column was re-equilibrated with initial conditions. The injection volume was 20 µL and the flow rate 0.35 $mLmin^{-1}$. All samples were measured in technical triplicates. Detection was performed on an Orbitrap mass spectrometer (Q Exactive Thermo Scientific, Bremen, Germany). Ionization was done in positive electrospray ionization mode with a spray voltage of 3.5 kV and a capillary temperature of 350 °C. The full scan HRMS spectra (*m/z* 100-1000) with a resolution of 70,000 was followed by five data dependent MS^2 scans of the most intense ions with a resolution of 17,500 for each scan. However, later on MS^2 spectra were not part of software comparison in this work. Details of electrospray conditions and measuring parameters are given in table S4-3 and S4-4.

4.2.3 Software tools and implemented algorithm

MZmine2

MZmine 2 is a java based open source application programmed by Orešič and Katajamaa [136] and later updated to MZmine2 by Pluskal et al.[45] which enables users to program and add plugins for specified data processing of MS data. The program is widely used in the metabolomic and proteomic context, but is applicable for all other MS based data and was recently also applied in the NTS field [132,133]. In this study, the data were processed with a NTS workflow suggested by Verkh et al.[132] with a few adaptations. The processing steps of the workflow will be shortly explained. First the mass detection step was performed with the *Exact Mass* algorithm, that uses the full width at half maximum paradigm to determine the center of each peak and pick *m/z* values and intensities [45]. Artificial peaks are filtered out by a *Lorentzian* function, this is done with the help of a theoretical peak shape model (Lorentzian or Cauchy) which is built with a given mass resolution around each peak, and all noise peaks below this model are removed [45].The next plugin is the so called “chromatogram builder”, that combines peaks of neighbouring scans using user-defined thresholds for *m/z* variation, intensity and peak width. Following, each chromatogram is deconvoluted into individual chromatographic peaks. This was done with the *Noise Amplitude Algorithm,* which calculates a noise level for each chromatogram individually by separating the chromatogram in small bins, identifying the level where most noise is concentrated and setting a cut off score accordingly [45]. The peak lists were further filtered by a smoothing algorithm and a duplicate filter. The isotope grouping plugin was applied, which deletes features of an isotopic pattern of the main peak and saves the information in the metadata. Unlike that, the adduct peaks found are marked as identified but the information is not saved in the metadata. The adduct peak either remains in the peak list or can be manually deleted. In a next step, features across multiple samples, replicate measurements or blanks can be aligned. In this study, this was done by the

RANSAC aligner. It is based on the RANdom SAmple Consensus (RANSAC) algorithm developed by Fischler and Bolles, which is a non-deterministic iterative algorithm, that estimates parameters of a mathematical model from a set of observed data [137]. As a result, a new feature list is created giving information about the average *m/z* and RT of all aligned features in a row and additionally of the original lists. This aligned peak list was manually filtered to exclude features that were not present in all replicates or features that were also present in a blank sample.

enviMass

enviMass, developed by Loos, is an automatized data mining tool based on the R coding environment including a graphical user interface in a web-browser. enviMass has the aim to provide a workflow for trend and spill detection of known and unknown micropollutants in aquatic systems from LC-HRMS data. In the following the main steps of data processing as described in [138] are shortly summarized, however the software is updated and extended on a regular base. Peak picking is done in three sub steps: 1) The centroided data are partitioned into subsets by linking each data point to its neighbours within large tolerances of *m/z* and RT. 2) Data points in each partition are clustered according to their *m/z* value with an intensity descending routine. 3) From these clusters, single EICs are built, by merging locally connected clusters which fulfil set parameters and saving mean *m/z*, mean RT and maximum signal intensity for each peak. Systematic deviations of *m/z*, RT and intensity can be normalized with spiked internal standards, but this step was not implemented in this study. Following, all features that are connected via isotopologue or adduct linkages are grouped into components [138]. After this step, the lists can also be screened for homologues series, target and suspect compounds. Finally, components are aligned over several samples by the building of so-called profiles. For profiling, all features of the selected samples are pooled together. A first profile is initialized by the most intense peak and by iterating over decreasing intensities other peaks are either added to this profile, if there is overlap in *m/z* and RT with this profile, or otherwise used to initialize new profiles themselves. The resulting profile list can be filtered for features that are absent in a defined blank sample or exceed an intensity ratio compared to the blank.

XCMS Online

XCMS online is a web-based platform, initially developed for the processing data from untargeted metabolomic experiments. The online version is based on the XCMS algorithm published in 2006, however, it does not require familiarity with a command-line interface or programming scripts. Predefined parameter sets for different instrument setups are available. The following automated workflow includes raw data preprocessing, retention time correction, sample alignment and several statistical analyses [48]. Data preprocessing and peak picking in XCMS are based on the *centwave* algorithm, which combines detection of regions of interest

(ROI) in the *m/z* domain, and a *Continuous Wavelet Transform* (CWT) based approach for chromatographic peak resolution [139]. ROIs are detected with an incremental approach starting with the most intense data point of the first scan. Following, each data point is tested if it falls in the user defined *m/z* tolerance of an existing ROI. If this is the case, it is added to the existing ROI initiating an update of its average *m/z*, if not a new ROI is created. Several filter steps are employed to remove improper ROIs, such as the ones that do not have a certain amount of sequential points above the base line, do not exceed an intensity threshold, or do not pass a signal-to-noise check. In a next step chromatographic peaks are built from the ROI with the help of local maxima of calculated CWT coefficients. Peak boundaries, intensities, average *m/z* and S/N are determined for each feature [139]. Further, XCMS includes a nonlinear retention time alignment which is independent of the use of internal standards. The approach called *Obiwarp* is based on the identification of groups that are present in all samples and are evenly distributed over the chromatographic profile, calculation of retention time deviation of these groups and alignment of all features using a non-linear warping [47]. Further, information of adducts and isotopes are annotated in the final peak list [48].

Compound Discoverer

Compound Discoverer is the commercial software solution for NTS workflows from Thermo Scientific. It comprises a customizable untargeted workflow including options for peak picking, RT alignment, formula prediction, background annotation and an automated library and database search for identification purposes and several statistical tools. However, underlying algorithms are not published. Several adducts and isotopologues are summarized to “compounds”. Their formulas are predicted, and molecular weights are given as output parameter. To enable comparison with other software the output list “merged features” including *m/z* values was used. The merged feature list was manually filtered for features that were present in all replicates and absent in the blank.

4.2.4 Data processing

4.2.4.1 Data pre-processing

The same 30 raw data files (9 samples and 1 SPE blank as triplicates) were processed with the different software tools. Data were acquired in profile mode, for processing with enviMass and XCMS online the raw data files were converted to mzXML format with msConvertGUI[140] software. For data processing several parameters (mass error (5 ppm), mass range (100 - 1000 *m/z*), and retention time window (0 - 28 min)) were commonly set for all programs. Different adjustable parameters are implemented by different software for the intensity/height during peak picking, a threshold of 1E+05 was selected. Further software-individual parameters were optimized for each software with the help of the recovery of a set of internal standards and suspects (table S4-2). Following, filtering steps like componentization, replicate

filter and blank subtraction were implemented depending on each software's capabilities as described in the previous part and summarized in the following table 4-1. The m/z tolerance for the replicate filter and blank subtraction was set to 5 ppm for all programs, however for the RT window different optimal settings were determined. Details on all settings are found in the SI part 4.6.2.

Table 4-1 Overview of data processing workflows with used software tools

	enviMass	**MZmine2**	**Compound Discoverer**	**XCMS online**
Software version	4.0	2.34	3.0	03.03
Data input format	mzXML	.raw	.raw	mzXML
Peak picking algorithm	enviPick	Exact Mass	?	centWave
Componentization	Adducts and isotopes were summarized to components	Isotopes were summarized, adducts were deleted	"Merged Features" lists were used as they provide *m/z* values. No subtraction of adducts or isotopes	Adducts and isotopes were annotated but not subtracted
Feature alignment	enviMass profiles	RANSAC	Adaptive curve	Obiwarp
Replicate filter	Replicates were processed with replicate Tag. Profiles containing only peaks present in replicates were evaluated	Alignment of peak lists of replicates followed by removing of rows that were not common in all replicates	Filtering of "merged feature" list for features that were present in all three replicates	Upload of replicate files as group
Blank subtraction	Sample/blank ratio (10) was used to subtract blank peaks	Replicate samples were aligned with blank and rows containing blank peaks were deleted	Filtering "merged feature" list for features that were not present in blanks	Filter criteria sample vs. blank group fold change > 10 UP regulated

4.2.4.2 Validation of data pre-processing

A list of 22 suspects including pharmaceuticals, pesticides and industrial chemicals with a *m/z* range from 120 to 750 and retention time range of 0.6 to 16 min that were present in most of the samples was selected. Their presence or absence in each sample was verified with Thermo Xcalibur Quan Browser as benchmark. The peak lists obtained by the different software packages were manually searched for suspects by their theoretical *m/z* and observed RT. Even though several programs include a target or suspect detection function, the aim was to ensure that peak picking was successful, not to evaluate the target detection functionalities. For each software and sample a recall rate of suspects was calculated with the number of detected suspects by each software (true positives) divided by the total number of suspects detected by the benchmark approach as described in formula 4-1.

$$\textit{Equation 4-1: } recall\ rate = \frac{true\ positives}{total\ number\ of\ true\ positives} * 100\ \%$$

For the determination of the total number of true positives, suspects had to be present in all three replicates, absent in the blank sample and exceed a peak area of 1E+05.

Further, the variation of *m/z* values and RT during peak picking with different programs was explored to determine filter criteria for feature overlap evaluation. For this purpose, for the found 22 suspects in each sample the *m/z* and RTs values were collected without rounding. For each individual suspect in each sample the difference of min and max of *m/z* and RT over all four programs was determined. Thus, only deviations created during feature extraction with different algorithm were evaluated as only the same raw files were compared with each other and deviations caused by measurement variation were not included.

4.2.4.3 Determination of overlapping features

In order to find overlapping features from different processing programs, the final feature lists were exported from each software as either .xls or .tsv file and imported into a custom-build script in MATLAB programming environment (R2016a; Mathworks, Inc., Natick, MA). Here, the information of retention time in minutes, *m/z* and peak areas (MZmine2, Compound Discoverer) or intensities (enviMass, XCMS online) were extracted and combined into individual tables for each software. Following, RTs were rounded to 0.1 min and *m/z* values to 0.01. Duplicates within each software-independent table, which were created during rounding were deleted to avoid double counting during comparison. For calculation of the overlapping regions of the whole feature lists the area/intensity domain was not included. For the calculation of the overlap of the 100 most intensive features of each software, feature lists were sorted by decreasing peak area/intensity. Subsequently, the 100 most intensive features were cut out and a new table without the area column was created for calculation of overlapping regions. Then, the overlaps between the different feature lists (pairs of *m/z* and RT) of each software tool were calculated either for all or for the 100 most intensive features. The overlapping regions of feature lists between two, three and all four software tools were determined. A common feature of all four tools should not be counted again as overlap of three or two tools. Therefore, intersections were further reduced to show only one intersection for each feature. Finally, the remaining number of unique features for each software that showed no overlap with any other program were determined. For a better comparison between results from different samples, the numbers of overlapping features were calculated as fractions of the total feature number identified in each software.

4.2.4.4 Exploration of observed discrepancies between software tools

The filtering steps replicate and blank filter as intermediate feature reduction steps after peak picking, were tested as possible source of variation between feature lists. For this purpose, exemplary for one sample (river Diemel upstream 23.11), all three replicates were individually processed, and the resulting feature lists were compared with each other as described in

section 4.2.4.3. As a next step, replicates were processed together and features, which were not common in all three replicates, were removed as described in table 4-1 (replicate filter) but features contained in the blank sample were not subtracted. For these lists, overlapping features were again determined. The reduction of total feature numbers of the single replicates, after application of the replicate filter and with whole processing workflow after additional blank subtraction, were determined and overlaps between programs were compared. The data analysis and calculations were performed employing MATLAB R2016a.

4.2.4.5 Comparison of identification with FOR-IDENT

To evaluate if further data analysis and interpretation after data processing is affected by different data processing tools a suspect screening based on a FOR-IDENT[141] data base search was conducted. FOR-IDENT incorporates a data base in which relevant organic molecules, their transformation products and metabolites are listed. Exemplarily for one WWTP effluent sample (WWTP final effluent 23.11) and one surface water sample (Diemel upstream 23.11), the 1000 most intensive features for each software were extracted and uploaded to a simple FOR-IDENT search (specification: ppm precursor ion 5, ion species +H, stationary phase C18 (polar endcapped), pH 3). Not all features showed a hit and for some masses several hits were reported. All search hits were exported, and the lists were compared based on SMILES codes of the proposed suspects.

4.2.5. Statistical analysis

To test if findings of overlapping features for different samples follow a similar trend, several statistical tests were implemented. For this purpose, the relative numbers of non-overlapping and overlapping features were combined in a matrix of which each column represented a sample and each row showed all combinations of the relative non-overlapping region and overlapping regions of two, three or four software tools. Data were preliminary checked for normal distribution. To test if findings of overlapping features for different samples follow a similar trend, the non-parametric test of Spearman's rank correlation was conducted due to the non-normal distribution of the data set (Shapiro-Wilk test). Also, Wilcoxon signed-rank test was employed for pair-wise comparison of sample groups with different matrices (surface water and WWTP effluents). In order to analyze the differences between the overlapping features obtained through implementation of intensity filter (section 4.3.3) and intermediate steps (section 4.3.4), Mann–Whitney U-test was used. A p-value of less than 0.05 was considered to be statistically significant. Statistical analyses were performed using SPSS software ver. 21.0 (IBM, Armonk, NY, USA) and in-house MATLAB program.

4.3. Results and discussion

4.3.1 Validation of data processing

To evaluate the data processing with different software packages the recall rates of 22 suspects in all samples were determined. As can be seen in figure 4-1 Compound Discoverer showed the highest recall rate (mean 88 %) followed by enviMass (mean 83 %), MZmine2 (mean 82 %), and XCMS online (mean 64 %).

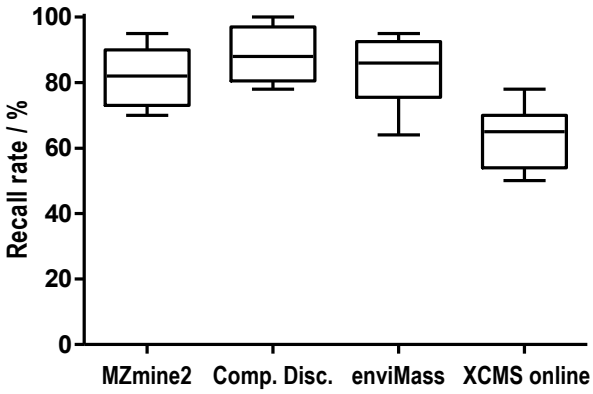


Figure 4-1 Recall rates of 22 suspects in all 9 sample. Depicted as Box-and-Whisker plot (5-95 percentile) for each software. Abbreviation: Comp.Disc = Compound Discoverer.

It has to be taken into account, that suspects were not spiked compounds and their presence in the environmental samples was not constant. Several suspects that were not detected with some programs showed a small peak area during target detection with Xcalibur, which is probably the reason they were filtered out during peak picking with tested programs. Also, this characteristic was differently reported by different software for example as mean or maximal peak area or as intensity, which probably caused unequally strict filtering of peaks. Furthermore, it was noticed that the WWTP effluent samples showed an overall lower recall rate, which is probably due to higher total feature numbers in these samples and thus overlaying signals impeding feature extraction. However, for example for XCMS online a few suspects (e.g. valsartan, bisoprolol, cyanazine) were not detectable in any sample, even though detection with other programs was possible and intensities were comparable.

The description of peak picking algorithms in part 2.3 illustrated that several steps like setting of a peak apex and determination of RTs or combining *m/z* values from several scans can slightly differ between algorithms. Also, alignment of several replicates and calculation of average *m/z* and RTs further promotes differences. Thus, a closer look on the deviations of *m/z* and RTs was taken to find adequate criteria for feature overlap determination in the following steps. Figure 4-2 shows the mean variation of m/z and RT of the same features extracted with different software.

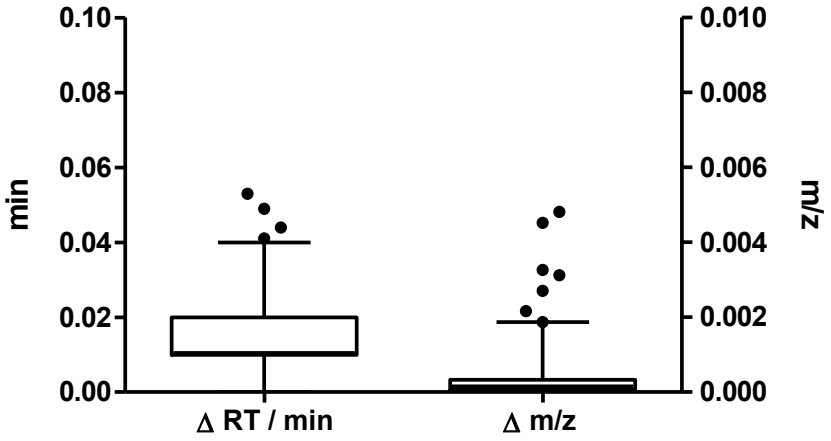


Figure 4-2: Deviation of RT and m/z over all data processing programs for 22 found suspects in 9 sample. Depicted as Box-and-Whisker plot (5-95 percentile).

The *m/z* of one feature varied with a mean of 0.0004 between different software and the RT with a mean of 0.02 min, respectively. The maximum deviation of one feature within different software was 0.6 min and 0.0048 m/z. Thus, a rounding of features to 0.1 min and 0.01 m/z is sufficient to ensure that no mismatches during feature overlap determination are caused by small deviations of algorithm performance during peak picking.

4.3.2 Overlapping features

Figure 4-3 gives an overview of the intersections of feature lists of all processing software for six surface water samples. Figure S4-11 shows the results for the 3 WWTP effluent samples. The results were not combined, as total feature numbers of part A and B were generally higher for WWTP effluent samples. Nevertheless, the distribution of relative intersecting features for all 9 samples showed a significant correlation ($p<0.05$) with the bootstrapped Spearman´s rank correlation test. Furthermore, Wilcoxon signed-rank test did not elicit a statistically significant change in the relative intersections given in part C for two groups ($Z = -0.965$, $p = 0.334$, table S4-11). Part A of figure 4-3 shows that between 3000 and 4000 features remained for the surface water samples after peak picking, replicate and blank filtering with given intensity thresholds. The left side (up to the dividing line) of part B and C shows, that a large proportion of features found by each software is unique and does not overlap with any other software feature list. The right most column of part B and C shows that only around 8 to 12 % are common features found with all software tools. Further, intersections of enviMass, Compound Discoverer and MZmine2 were higher compared with intersections including XCMS online.

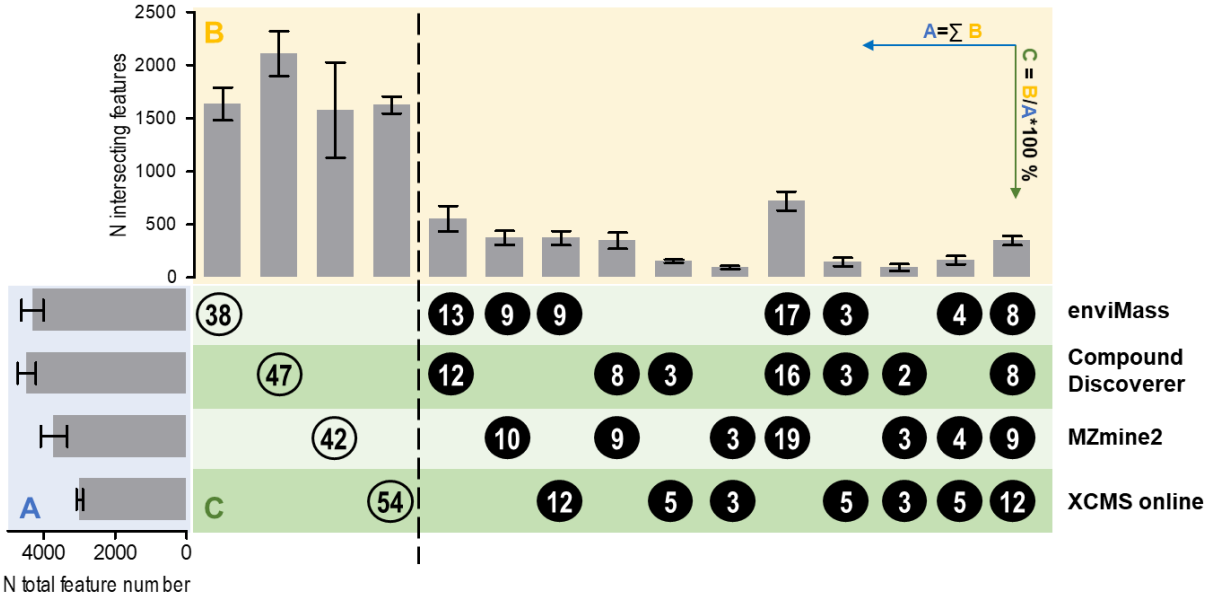


Figure 4-3: Overview of intersecting features of all programs of river water samples (N=6). A: the bar chart on the left shows the mean absolute feature numbers of input lists from different software. B: the upper bar chart shows the mean absolute feature number of non-intersecting areas (left part to dividing line) and intersection areas (right part). Errorbars in A and B show standard deviations C: the table below has two functions 1) it serves as legend, circles indicate overlaps between softwares to be compared, 2) the numbers in the circle show the ratio of overlapping features (B) to each input list (A). Again, the left part shows the percentage of non-overlapping features for each software (white circles) and the right part shows the percentages of overlapping features (black circles). Values in each row sum up to ≈100 % (deviations due to rounding). Reading example: The first bar right to the dividing line in Part B shows for example the commonly found feature by enviMass and Compound Discoverer which were 13 and 12 %.

The RT and m/z distribution of features that were common or exclusively found by different programs were exemplary examined but did not show obvious trends (figures S4-13/S4-14). Nevertheless, further detailed investigation on characteristics like retention time, intensity etc. of common or uncommon features is of interest for future work.

The low coherence of features extracted with different software tools is in accordance with the findings of Rafiei and Sleno, who found in their comparison an overlap of less than 10 % [127]. Myers et al. found an overlap of 60 % between XCMS and MZmine2 but unlike in this study, where different algorithms and further filtering steps after peak picking were included, Myers et al. compared the use of the same peak picking algorithm *centwave[123]*.

Several comparative studies in the metabolomics field suggested to use the overlap of different peak picking workflows for more rugged results [90,127]. However, this would be quite time consuming, especially if parameter optimization has to be performed for each software and might thus not to be a realizable option in practice.

It must be taken into account that even though commonly used parameter settings were selected and an individual optimization of data processing with each software was performed, other settings of data processing parameter for each software tool would probably influence

the presented findings. However, different settings of processing thresholds within one program also influence the efficiency of feature extraction which was shown for different workflows [103,142]. Also, the included software packages represent only a part of all available tools and these are regularly updated. Nevertheless, the main conclusion remains that different algorithms and settings do have a high influence on feature extraction and more attention should be dedicated to this in the future.

4.3.3 Overlapping features after intensity prioritization

Due to the vast number of features, a strict prioritization step is included in almost any data processing workflow to reduce feature numbers before further compound identification steps are performed. Often an intensity-based prioritization approach is included [143]. More intensive features are presumed to emerge from more abundant compounds in the sample, are more likely to produce high quality MS^2 spectra and are expected to contain less FP features or artefacts. Therefore, a closer look was taken on the 100 most intensive features found by each software and overlapping feature were determined. All nine samples showed the same trend (significant correlations, $p<0.05$) of relative intersections (part C). Results for all nine samples are summarized in figure 4-4.

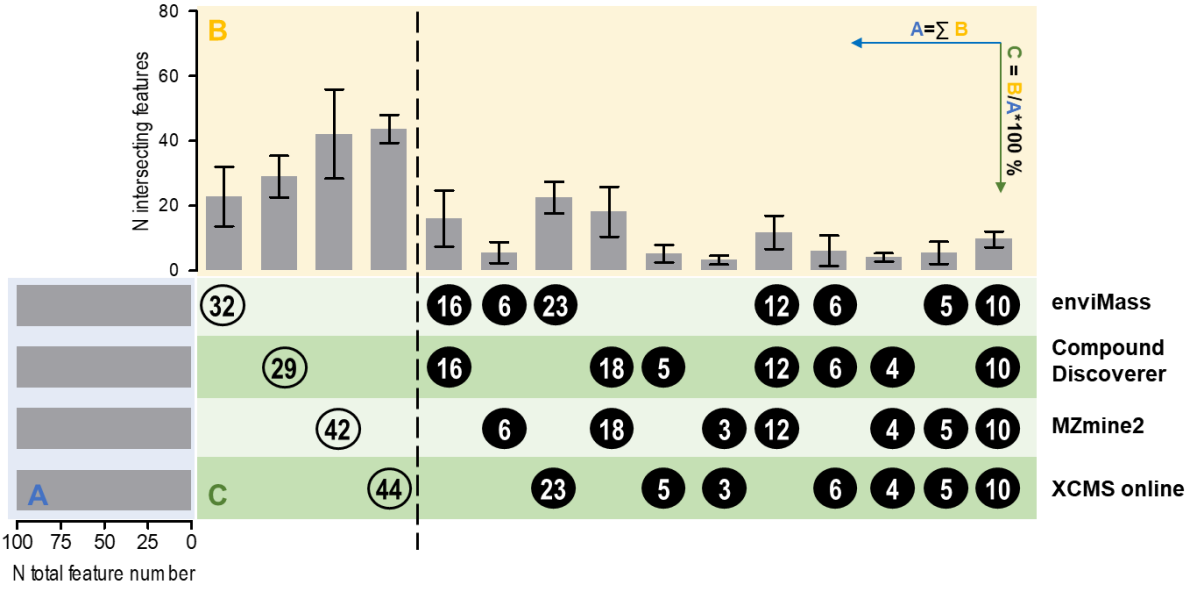


*Figure 4-4: Overview of intersections of 100 most intensive features for all programs of all samples (N=9). **A**: the bar chart on the left shows the mean absolute feature numbers of input lists from different software **B:** the upper bar chart shows the mean absolute feature number of non-intersection areas (left part) and intersection areas (right part). Errorbars show standard deviations **C**: the table below has two functions 1) it serves as legend, circles indicate overlaps of which software are displayed, 2) the numbers in the circle show the ratio of overlapping features (B) to each input list (A). Again, the left part shows the percentage of non-overlapping features for each software (white circles) and the right part shows the percentages of overlapping features (black circles). Values in each row sum up to ≈100 % (deviations due to rounding). Reading example: The first bar right to the dividing line in Part B shows for example the commonly found feature by enviMass and Compound Discoverer which were 16 %.*

The 100 most intensive feature did not show a higher ($p>0.05$) percentage of overlapping features for all four software (10 %) compared to the whole feature lists, according to the Mann–Whitney U-test. However, several different trends of detected overlaps between the whole feature lists and 100 most intensive features were observed and statistically evaluated. For example, non-overlapping features (white circles in figures 4-3 and 4-4) were skewed to lower values for all software except for MZmine2, indicating that for these programs the overlapping with other software increased. Further, the enhanced overlap between enviMass, Compound Discoverer and MZmine2 as observed for whole feature lists was not as distinctive for the Top 100 feature list. Two pairs of software, enviMass and XCMS online (23 %) as well as Compound Discoverer and MZmine2 (18 %) showed a higher overlap ($p<0.05$). A possible reason could be the same data input format of software pairs (centroidized mzXML for enviMass and XCMS online and profile raw data for Compound Discoverer and MZmine2). However, this trend was not observed for the whole feature lists in part 3.2. In addition, the overlapping features exhibit a wider distribution between different samples compared to whole feature lists. This becomes apparent from figure S4-12 that shows the box plot for all non-overlapping and all overlapping features of all samples for whole feature list (group 1) and 100 most intensive features (group 2). In general, the coherence of feature lists was not substantial higher for the more intensive features contrary to prior expectations. Some of the suspects were included in the lists of 100 most intensive features, however, not all of the top 100 features were identified and thus it was not confirmed if they are all relevant micropollutants

4.3.4 Exploration of observed discrepancies between software tools

As implementation of processing steps, such as blank and replicate filter differ between different programs, they are a potential source of the observed discrepancies. Thus, overlaps of feature lists before the implementation of these steps were compared exemplarily for one sample (Upstream 23.11). The intermediate feature lists before blank and replicate filter, after replicate but before blank filter, and after blank and replicate filter were separately compared with each other. With each filtering step the total number of features was reduced as shown in figure 4-5.

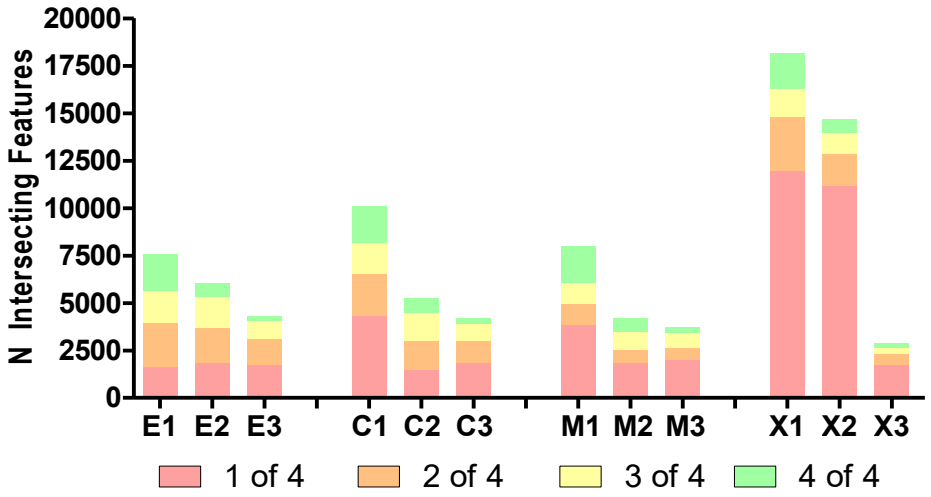


Figure 4-5: Comparison of feature numbers and intersections between different intermediate filtering steps for one sample file (Upstream 23.11). Abbreviations: E: enviMass, C: Compound Discoverer, M: MZmin2, X: XCMS online. 1: mean value for three individually processed replicates. 2: after merging of replicates, before blank subtraction. 3: after blank subtraction. Colours indicate if feature were found with one, two, three or all programs.

The reduction was especially strong for XCMS online, where by far the highest number of initial features was detected. A possible reason is that *centwave* algorithm detects a lot of FP features in samples with a lot of background signals, however these are also found in blank sample which leads to the extreme reduction of feature for this step (X2 to X3 in figure 4-5). Further, an extensive workflow optimization was not performed for XCMS online unlike for the other programs (see SI part 4.6.2). For MZmine2 and Compound Discoverer the replicate filter had a higher reducing effect than the blank filter. This might be due to the fact that this step was done manually after alignment of replicates unlike for enviMass and XCMS online where replicates were processed as a group from the beginning. EnviMass showed the lowest differences of feature numbers for different filtering stages. For enviMass and XCMS online blanks were filtered out with a ratio and thus not as strict as for other programs. Nevertheless, the blank filter had very high reducing effect on XCMS online, but feature reducing effect for enviMass was similar to the two other programs. For a better understanding of these filtering steps it would be of interest to further characterize parameters such as intensity or RT distribution of the features that are excluded by each step in future studies.

Having a closer look on the distribution of overlaps with other software, as indicated by bars in different colors in figure 4-5 for each step, several trends become apparent. For enviMass the amount of non-overlapping feature stays constant while features that overlap with all four programs are reduced with each step. This was also observed comparing steps 2 and 3 for MZmine2 and Compound Discoverer. If overlapping features relative to total feature numbers are evaluated (as in Part C of figure 4-3 and 4-4) for enviMass step 1 shows the best coherence

with other software because the percentage of non-overlapping features (red bars) is lowest and percentage of overlapping features with all programs (green bars) highest. For Compound Discoverer and MZmine2 the step after the replicate, and before the blank filter (2) showed the best consistency with other programs. XCMS online showed the overall highest percentage of non-overlapping features but this was drastically reduced by replicate and blank filter.

The comparison of feature overlaps at different processing steps after peak picking showed that these intermediate filtering steps have an influence on comparability of software tools. However, this influence was tested with one sample only which might influence the result. Further, relative overlaps of feature lists at different processing steps were compared with each other using statistical test of Wilcoxon signed-rank test and did not show a significant difference ($p>0.05$). Nevertheless, filtering steps should be employed carefully, and their implementation should be reported in all NTS studies.

The influence of the componentization step was not investigated in this study. At this moment a more comprehensive componentization including the merging of signals of in-source fragments, different charge states of an ion or other relevant groups of related features is of high interest and new publications of algorithms and tools are expected. A comparison of different approaches on this task could be of interest for future studies.

4.3.5 Comparison of identification with FOR-IDENT

After the comparison of feature lists, the next interesting question is if different data processing tools lead to a generally different interpretation of data and consequently drawn conclusions. There are many strategies on data evaluation as mentioned in the introduction. In this study a suspect screening type approach with the FOR-IDENT[141] database was carried out to test whether final results would have a better comparability than feature lists. Figure 4-6 shows that for each sample and software tool a different number of suspect hits were proposed from data base search. Around 20-40 % of suspect hits were commonly found with all four software tools. For enviMass a lower number of suspect hits was reported but these showed a higher correspondence with other software tools. For the WWTP effluent sample more than 50 % of the suspect hits were found with at least two software tools, and for the surface water sample above 30 %, respectively.

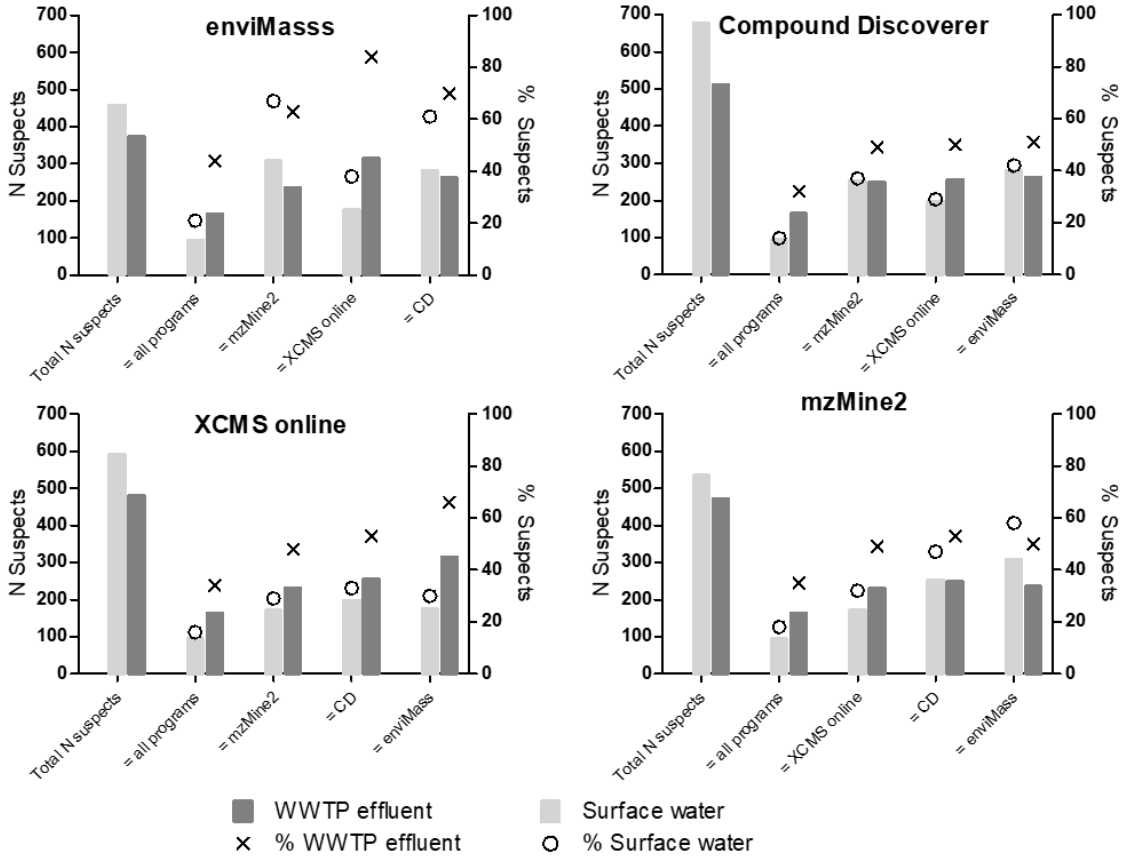


Figure 4-6: FOR-IDENT suspect search for all four programs for one WWTP sample and one surface water sample. Each graph from left to right, for both samples: total number of putative suspects, number of commonly found suspects with all four programs as absolute and relative values, overlaps with other three programs as absolute and relative values.

Of course, usually more effort should be taken to check plausibility of suspect search hits and final identification of compounds. Nevertheless, this attempt was made to show that final results and interpretation do not differ as extremely between software tools as overlap of plain feature lists. However, a comparison of further data analysis approaches such as a time trend analysis or other chemometric evaluation would be of interest

4.4 Conclusion

This study gives an overview on the different algorithm and filtering steps that are used by different data processing programs commonly used in environmental NTS and shows how these differences influence the comparability of extracted features. It was shown that the extracted features had a rather low coherence among different data processing tools. This incoherence was not significantly improved by focussing on the most intensive features only. Thus, the intensity of features as prioritization criteria should be used with care. In addition, the implementation of filtering steps such as replicate and blank filter can vary between different workflows. It was shown that this variation was a considerable reason for observed divergences of feature lists. The influence of the componentization step was not investigated in this study but is of interest for future studies as further developments and improvements in this area are expected. In addition, a more detailed investigation on how final data interpretation is influenced by different processing software would be of interest. Finally, the reported findings of this study can be influenced by the limited number and type of wastewater and surface water samples. Therefore, an extended study of different matrices and larger sample numbers is needed to test if the observed trends are indeed impacted by the sample selection. In general, it was noted that many parameters lack a description of their functionality and impact on data extraction which makes data processing a “black box” to the majority of users/operators. There is a need for further research on a better understanding and optimization of these algorithms.

Each of the investigated programs showed weaknesses and advantages for specific processing steps or included further promising tools that were not implemented in this study. Thus, a ranking of their performance is not meaningful at this point. Depending on the research question different tools can be favourable. Also, a standardization of data processing is not feasible, as different instruments with specific vendor software are used and each sample type needs an optimization of parameter settings. In addition, NTS can be understood as an explorative tool where standardization is not the goal and diversity of workflows can result in more discoveries of environmental critically compounds.

However, we want to encourage more awareness on data processing as a crucial step in the NTS workflow. A detailed documentation on the workflow including implementation of filtering steps, all parameter settings and criteria should be mandatory in all future studies. A commonly used framework for reporting of data [144] could be a way to increase transparency.

4.5 Acknowledgements

We thank Dr. André Liesener and Dr. Jan Lisec for valuable comments on the manuscript. This study was performed within the joint project FUTURE WATER founded by the Ministry of Culture and Science of the state of North-Rhine Westphalia (NRW), Germany.

4.6 Supporting information

4.6.1 Details on measurement

Table S 4-1: HESI source parameter

Sheath gas flow rate	37
Aux gas flow rate	15
Sweep gas flow rate	1
Spray Voltage	3.5
Capillary Temp.	320
S-Lens RF level	50
Aux gas heater Temp.	50

Table S 4-2: Settings of the Full MS/ddMS2(Top 5) method.

Method	Full MS – $ddMS^2$	
Polarity	positive	
Runtime	28 min	
	Full MS	**dd-MS2**
Resolution	70,000	17,500
AGC Target	1E+06	5E+04
Max. injection time	100 ms	50 ms
Scan Range	100 – 1000 m/z	
Loop Count		5*
Isolation window		1.4 m/z
NCE		30, 60
Intensity Threshold		2E+03

Table S 4-3: Overview of suspect compounds and spiked internal standards with m/z, retention time (RT) and molecular formula.

Compound	m/z	Molecular formula	RT
	Suspects		
Amisulpride	370.1795	C17H27N3O4S	9.9
Atrazine	216.1011	C8H14ClN5	13.8
Benzotriazole	120.0556	C6H5N3	10.2
Bisoprolol	326.2329	C18H31NO4	12.0
Boscalid	343.0399	C18H12Cl2N2O	14.6
Caffeine	195.0877	C8H10N4O2	10.2
Candesartan	441.167	C24H20N6O3	14.2
Carbamazepine	237.1022	C15H12N2O	13.4
Chlortoluron	213.0789	C10H13ClN2O	13.7
Clarithromycin	748.4842	C38H69NO13	14.2
Clozapine	327.1371	C18H19ClN4	12.2
Cyanazine	241.0963	C9H13ClN6	10.9
Diclofenac	296.0239	C14H11Cl2NO2	15.4
Diflufenican	395.0813	C19H11F5N2O2	15.9
Isoproturon	207.1492	C12H18N2O	13.9

Metformin	130.1087	C4H11N5	0.6
Metoprolol	268.1907	C15H25NO3	11.0
Morphine	268.1438	C17H19NO3	12.3
Sulfamethoxazole	254.0594	C10H11N3O3S	10.6
Terbutryn	242.1434	C10H19N5S	13.8
Trimethoprim	291.1452	C14H18N4O3	9.8
Valsartan	436.2343	C24H29N5O3	14.8
Internal standards			
Carbamazepine_D10	247.165	C15D10H2N2O	13.3
Claritromycin_D3	751.503	C38H66D3NO13	14.2
Diclopheanac_D4	300.0491	C14H7D4Cl2NO2	15.4
Diuron_D6	239.0619	C9H4D6Cl2N2O	14.0
Metoprolol_D7	275.2346	C15H18D7NO3	11.0

Table S 4-4: Overview of samples

Sample	Sampling place	Sampling date	N technical replicates
171123DiemO	River Diemel upstream	23.11.2017	3
171123DiemU	River Diemel downstream	23.11.2017	3
171123AbF	WWTP (Warburg) effluent	23.11.2017	3
171129DiemO	River Diemel upstream	29.11.2017	3
171129DiemU	River Diemel downstream	29.11.2017	3
171129AbF	WWTP (Warburg) effluent	29.11.2017	3
171206DiemO	River Diemel upstream	06.12.2017	3
171206DiemU	River Diemel downstream	06.12.2017	3
171206AbF	WWTP (Warburg) effluent	06.12.2017	3
MBlank	SPE Blank		3

4.6.2 Optimization of data processing

4.6.2.1 Optimization of data processing with MZmine2

The workflow described by Verkh et al. [132] was used as described in table S4-5. Parameters that were not changed are marked as “V”, parameters that were changed due to given conditions of the instrument and measurement parameters are indicated as given “G”, parameters that were changed to harmonize workflows between different software are marked as fixed “F”. Further parameters that were optimized are marked as “O” and will be explained in the following.

Table S 4-5: Overview of workflow steps and settings for data processing with MZmine2.

Step	Workflow Parameter		Category
	mass detection		
	retention time	0-28 min	F
	MS level	1	G
	Polarity	+	G
	Spectrum type	profile	G
	Mass detector	Exact mass	V
	Noise level	1.00E+05	F
	FTMS shoulder peak filter		
	Mass resolution	70,000	G
	Peak model function	Lorentzian extended	V
A	**Chromatogramm builder**		
	Retention time	0-28 min	F
	MS level	1	G
	Polarity	+	G
	Spectrum type	profile	G
	Min time span (min)	0.03	V
	Min height	1.00E+05	F
	m/z tolearance	0.001 m/z or 5 ppm	F
B	**Smoothing**		
	Filter width	7	V
C	**Chromatogramm deconvolution**		
	Algorithm	Noise amplitude	
	Min peak height	3.00E+05	O
	Peak duration range	0.05-1.0 min	O
	Amplitude of noise	1.00E+05	
D	**Isotopic peak filter**		
	m/z tolearance	0.001 m/z or 5 ppm	F
	Retention time tolerance	0.03 min	V
	Maximum charge	1	F
	Represntative isotope	lowest m/z	V
E	**Dublicate filter**		
	m/z tolerance	0.001 m/z or 5 ppm	V

Step	Workflow Parameter		Category
	RT tolerance	0.01 min	O
F	**Reduction False Positives**		
	Delete rows that do not meet the criteria:		
	Data Points	8 to 50	O
	isotope pattern	1	
G	*Search for adducts and delete rows:*		
	RT tolerance	0.03 min	V
	m/z tolerance	0.001 m/z or 5 ppm	V
	max relative adduct	50%	V

Alignment of feature lists:

For the implementation of the replicate filter and blank subtraction, feature lists were aligned with RANASAC algorithm as described by Verkh et al.

Algorithm	RANSAC	V
m/z	5 ppm	V
Rt tolerance	2 min	V
Rt tolerance after correction	1 min	V
iterations	0	V
minimum number of points	60%	V

H	**Replicate filter**	
	Align replicates with RANSAC aligner	
	Delete rows that do not meet criteria:	
	minimum peaks in a row	3
I	**Blank subtraction**	
	Align replicate list with individual processed blank lists	
	Delete rows that contain Blank peaks	

The optimization of several parameters was done by comparing total feature number and recall rates of 21 suspect compounds and 5 internal standards (see table S4-3) for a surface water sample (171123DiemO table S4-4). The results for steps (A to I) of the workflow are summarized in figure S4-1.

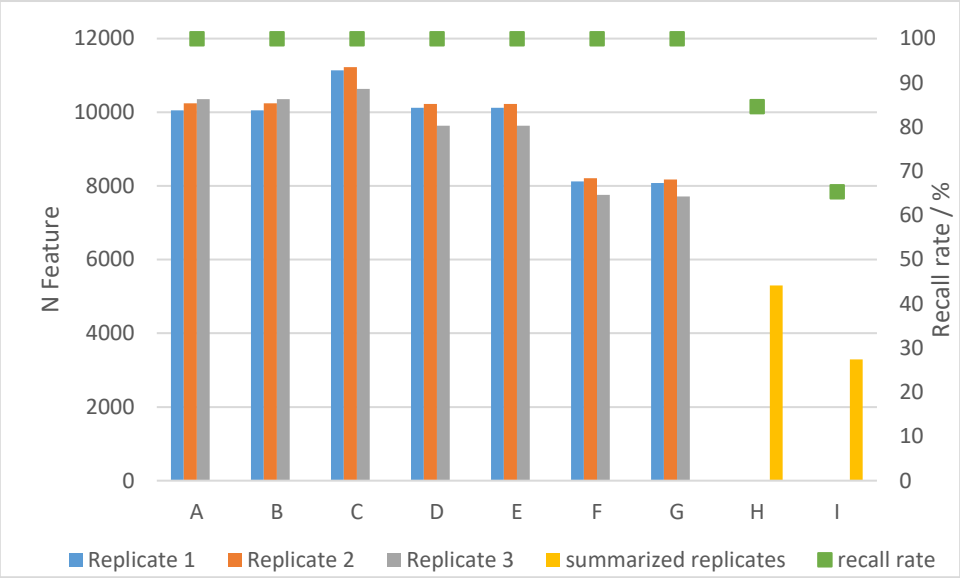


Figure S 4-1: Total feature number and recall rate for workflow steps A-I for MZmine2 optimization

Step C: Chromatogram deconvolution

The parameter given by Verkh et al. [132] for this step were:

- peak duration: 0.05 to 8 min
- min peak height: 3.0E+03
- Amplitude of noise: 1.0E+03

As the intensity threshold was set to 1.0E+05 for all software tools, the parameter “amplitude of noise” and “min peak height” was set to 1.0E+05 for the first try. Than “min peak height” was set to 3.0E+05 to have the same ratio as described by Verkh et al. As can be seen in figure S4-2 this led to a reduction of feature number and did not affect the recall rate. Suspects and IS had a height between 7.0E+05 to 4.0E+07, hence the parameter “min peak height” could have been set lower without effecting the recall rate. However, this was not done to avoid discrimination of low intensity peaks.

Figure S 4-2: Optimization of Step C for MZmine2 processing

As next step the peak duration rate was optimized. Suspects and IS had a peak duration between 0.08 to 0.3 min. By visual inspeciton it was observed, that peaks with a duration above 1 min were almost unexceptionally false positive peaks. For peaks with a duration of 1 min false positive as well as true positive peaks were observed. Figure S4-3 gives an example of a peak with a duration of 1 min and a peak with a duration of 8 min. Thus, the range 0.05 to 1.0 min was selected. As can be seen in figure S4-2 this slightly reduced the total feature number and did not affect the recall rate.

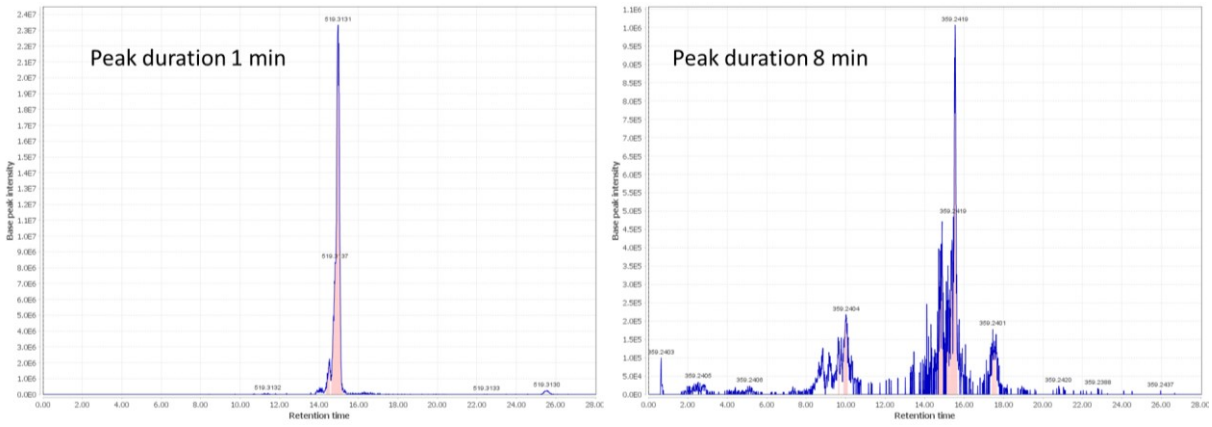


Figure S 4-3: Example of a peak with a duration of 1 min (left) and 8 min (right).

Step E – Duplicate Filter

For the duplicate filter several settings of m/z and RT tolerance were tested but did not change the number of features.

Step F – Reduction of False Positives

Verkh et al. used a criterion of 7 to 500 data points to filter out false positive peaks. Suspects and IS had between 8 and 25 data points. By visual inspection it was observed, that peaks

with more than 50 data were mainly false positive peaks. For peaks with around 50 data points a few true positive peaks were observed. Figure S4-4 shows an example of peak with 50 and 80 data points. Thus, the filter 7 to 50 data points was selected.

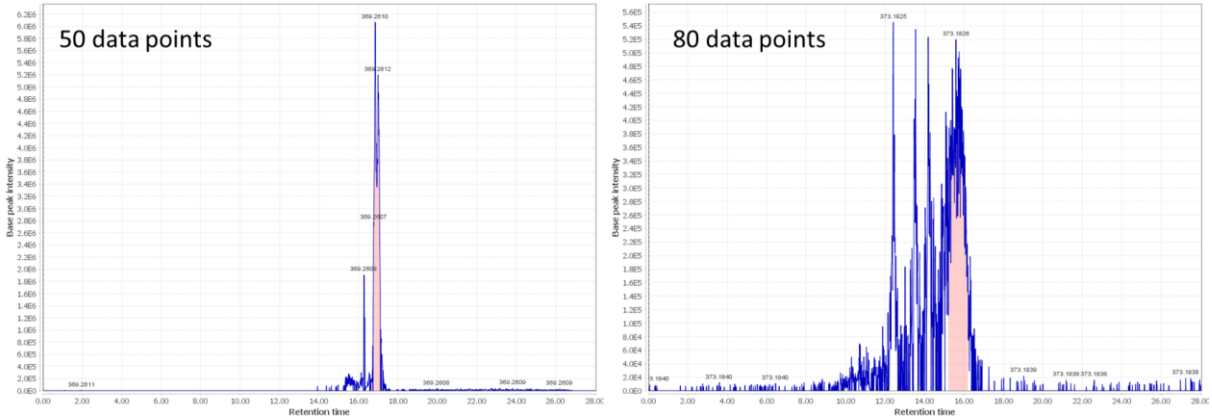


Figure S 4-4: Example of a peak with 50 data points (left) and 80 data points (right).

Figure S4-5 shows, that this filter led to a reduction of total feature number but did not affect the recall rate.

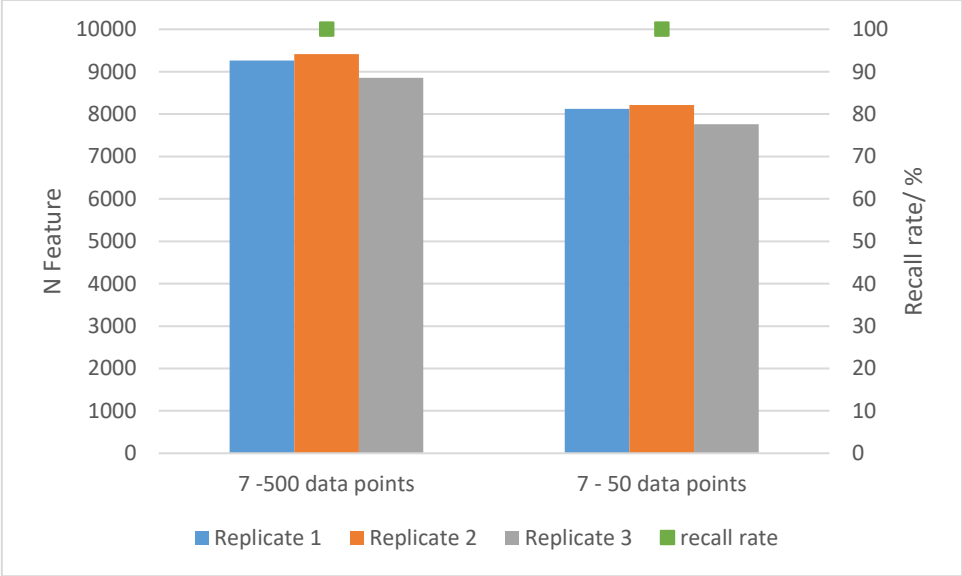


Figure S 4-5: Optimization of Step F for MZmine2 processing

Step H/I – Replicate filter and blank subtraction

As can be seen in figure S4-1, the replicate- and blank filter had a strong reducing effect on feature numbers. However, the recall rate also dropped as some suspects got lost during this filter steps.

For the replicate filter it was observed that after alignment with RANSAC algorithm, some suspects and IS were not aligned properly. In these cases, two rows for the same feature were created, both rows containing empty spaces for some replicates. Thus, these suspects were deleted as they were falsely seen as not present in all replicates. In a newer version (V.2.35)

published after the finalization of data processing for this study a filter was added with the option to form a consensus row for duplicate rows, which would avoid this problem.

For the blank filter the 5 internal standards that were also spiked into blank samples were correctly subtracted leading to a decrease of the recall rate.

4.6.2.2 Optimization of data processing with enviMass

Table S4-6 shows all settings of parameters of the workflow before optimization (draft settings) and the final settings after optimization for data processing with enviMass. In table S4-7 the included steps of the workflow options are shown. Optimization was performed on a surface water sample (171123DiemO table S4-4), 3 replicates were processed with a replicate tag and profiles were evaluated after result filtering as described in table S4-8.

In a first step the settings of the peak picking step were optimized, and all other settings were left as in the draft version. In the same way the settings for blind- and replicate filter, componentization and profile building were successively optimized.

*T*able S 4-6: Overview of draft and final settings of the data processing workflow in enviMass.

Parameter	**draft settings**	**final settings**
Peak picking		
RT range	0 to 28	0 to 28
mass range	100 to1000	100 to1000
Parameter estimation	yes	yes
Maximum RT gap in an EIC [s]	300	60
Maximum m/z deviation of centroid data points from its EIC mean [ppm]	3.5	5
Minimum number of centroid data points per peak…	4	10
within a given RT window [s]	8	10
Maximum RT gap length to be interpolated [s]	10	10
Maximum RT width of a single peak from its peak apex [s]	120	60
Minimum log 10 (intensity) threshold	4	5
Minimum Signal/Noise	5	3
Minium Signal/Base	2	2
Maximum possible number of peaks within a single EIC	3	3
Peak intensity: use peak area or peak intersoid?	intensoid	area (sum int.)
Peak mass definition	mean	mean
Upper log10(intensity) safety threshold	6.5	7
How often can a peak detection fail to end the recursion?	1	1
Weight for assigning centroid data points to a peak	1	1
Percentage of low-intense data points to discard	0	20
Blind		
Factor by which the sample peak intensity must exceed the blank/blind peak intensity to not be substracted	100	100
m/z tolerance (+/-) [ppm]	12	5

Parameter	draft settings	final settings
RT tolerance [s]	30	60
Replicates		
m/z tolerance (+/-) [ppm]	12	5
RT tolerance window of peaks caused by the same analyte across replicate samples [s]	30	30
Absolute log intensity tolerance X (log 10 scale, 10^X)	10	10
Componentization		
Profile componentization		
Minimum nr of files over which peaks of different profiles have to occur to check their intensity correlation	5	3
Minimum Pearson profile intensity correlation	0.9	0.9
RT tolerance window for co-occuring peaks of different profiles [s]	5	5
Restrict profiles componentization to isotopoloque and selected adduct relation only?	FALSE	FALSE
Restrict profile componentization to a set of latest files only?	FALSE	FALSE
File wise componentization		
Isotopoloque grouping		
m/z tolerance (+/-) [ppm]	8	5
RT tolerance of peaks within an isotopologue pattern [s]	2	2
Intensity tolerance %	50	50
Adduct grouping		
m/z tolerance (+/-) [ppm]	8	5
RT tolerance of peaks within an adduct pattern [s]	2	2
M+H, M+Na,M+K		
Profiles		
Peak mass deviation within profiles: m/z tolerance (+/-) [ppm]	12	5
Peak deviation within profiles: RT tolerance [s]	60	60
Omit files with table entry profiled = FALSE from profiling?	TRUE	TRUE

Table S 4-7: Enabled steps of the workflow options in enviMass.

Workflow settings	
only steps that are included area shown	
Blank/Blind peak detection	
Detect?	yes
Remove?	no
Replicate filter	
Include?	yes
Use in profiling	yes
LOD interpolation	
Include?	yes
Peak shpae correlation	
Include?	yes
File-wise componentization	
Include isotopologue grouping?	yes
Include adduct grouping?	yes
Profile Componentization	
Include?	yes
Profile extraction	
Include?	yes
Profile blind detection	
Include?	yes

Table S 4-8: Result filtering steps in enviMass.

Result Filtering	
To be able to export peaklists of merged replicates. Profiles were processed containing only replicates	
Filtering step	*setting*
including Blind substraction	yes
Filter profiles by mean sample vs. Blind intensity ratio	10
Rank (and filter) profile list by	mean intensity (decreasing)

The total number of peaks and of profiles as well as the recall rate of 21 suspects (the 5 IS were not included as they were present in the blank) were used to compare different settings and are summarized in figure S4-6. Both the total number of peaks and number of profiles was included as some steps such as componentization or profile building could affect both numbers differently.

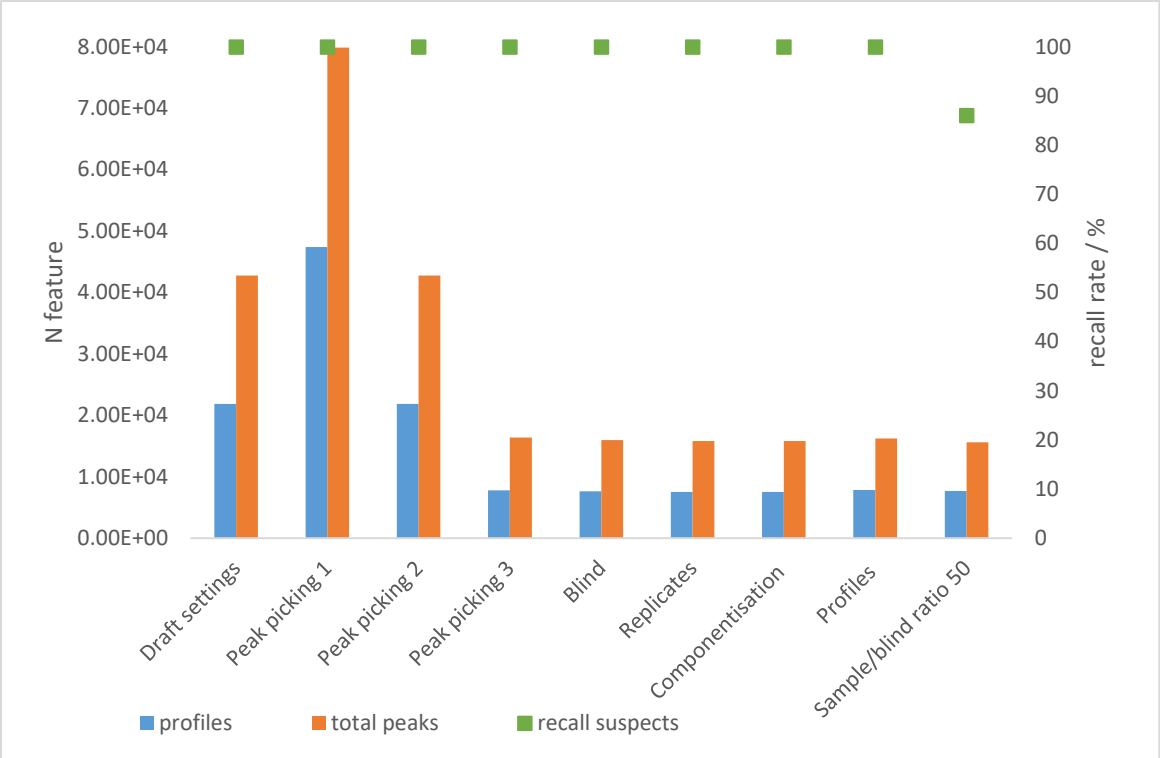


Figure S 4-6: Total feature and profile number and recall rate for different steps of enviMass processing optimization.

Peak picking

First the “parameter estimation” was set to “No” which lead to an increase in feature number and profile number as can be seen in figure S4-6 for peak picking 1. Thus, this parameter was reset to “Yes”. Further, the “minimum log 10 intensity threshold” was set to 5 and the “maximum m/z deviation of centroid data points from its EIC mean” was set to 5 ppm as commonly defined for all software. However, this did not affect the total feature/profile number (peak picking 2 in figure S4-6) compared to the draft settings. In a next step, several parameter were adjusted: “maximum RT gap in a EIC” was set to 60 s, “maximum number of centroid data points per peak” was set to 10 “within a given RT window” of 10 s, “Maximum RT width of a single peak from its peak apex” was set to 60 s, “minimum signal/noise” was set to 3, The “peak intensity” was set to “area (sum int.)”, the “upper log10(intensity) safety threshold” was set to 7, the “percentage of low-intense data pints to discard” was set to 20%. This led to a halving of total feature/profile number but no reduction of found suspects (figure S4-6, peak picking 3).

Blind filter

For the parameter “factor by which the sample peak intensity must exceed the blank/blind peak intensity to not be subtracted” was not adjusted, as blind subtraction was not enabled during the workflow but implemented during result filtering. Nevertheless, for parameters “m/z

tolerance (+/-)" and "RT tolerance", both affecting the blind detection, several combinations of settings were tested as shown in figure S4-7.

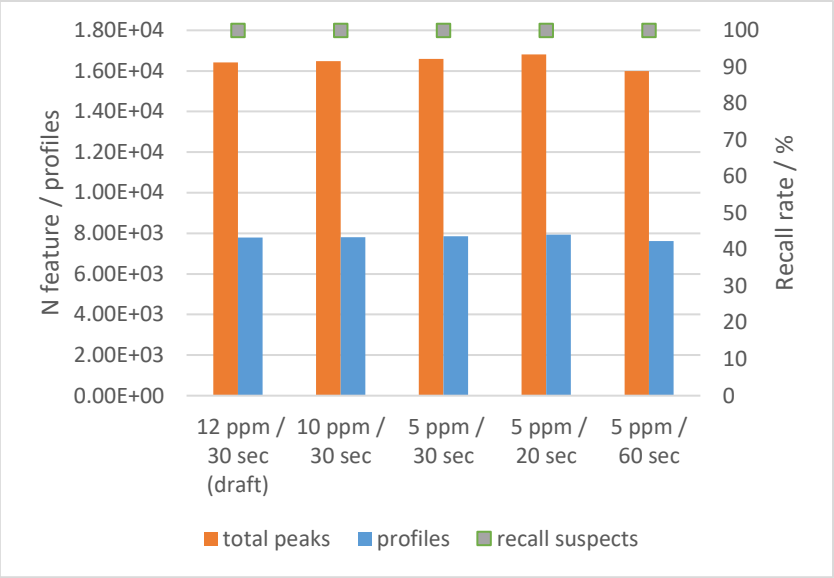


Figure S 4-7: Optimization of blind filter for enviMass processing.

The adjustment of the m/z tolerance did not affect the feature/profile number or recall of suspects. For the RT tolerance it was observed, that a narrower window increased the feature/profile number and a broader window reduced feature/profile number. Therefore, the combination of 5 ppm and 60 s was used.

Replicate filter

For the replicate filter the parameter "m/z tolerance (+/-)" was set to 5 ppm and for the parameter "Rt tolerance window of peaks caused by the same analyte across replicate samples" a window of 30 and 60 sec was tested. The narrower window of 30 showed a slightly lower feature/profile number as can be seen in figure S4-8 and was thus selected.

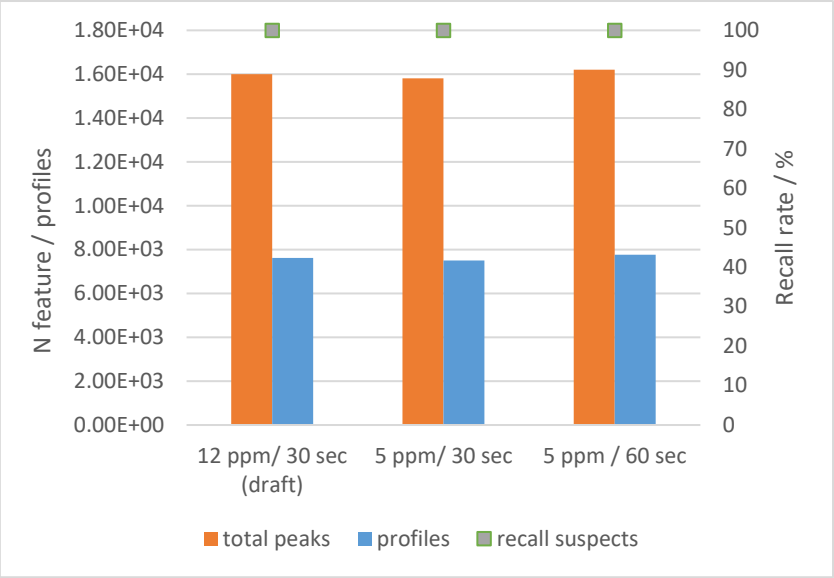


Figure S 4-8: Optimization of replicate filter for enviMass processing.

Componentization

The parameter „m/z tolerance (+/-)“ for the “isotopologue grouping” and “adduct grouping” were set to 5 ppm in two steps. However, this did not affect the total feature/profile number as well as recall of suspects.

Profile construction

For profile combination the parameter “Peak mass deviation within profiles: m/z tolerance (+/-)” was set to 5 ppm and “Peak deviation within profiles: RT tolerance” to 60 and 30 s. The broader RT window of 60 s slightly reduced the total feature/profile number, recall of suspects was not affected as can be seen in figure S4-9.

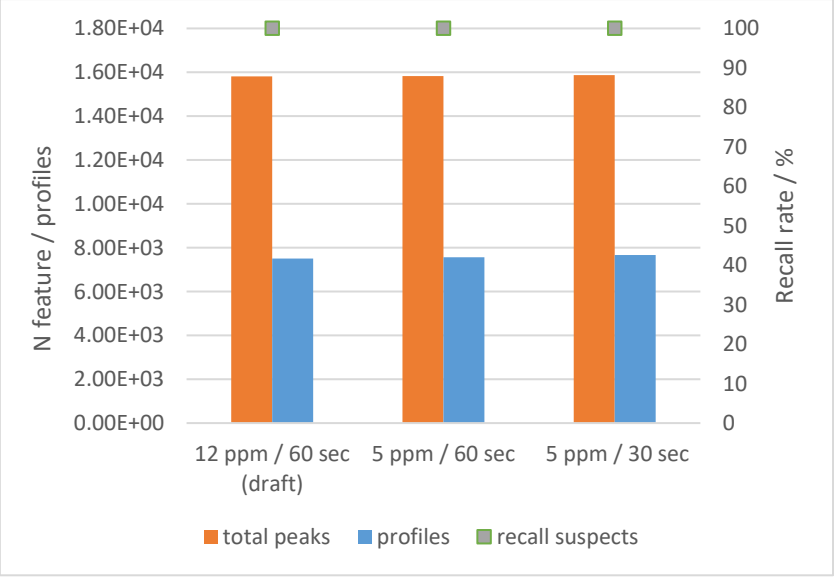


Figure S 4-9: Optimization of profile building for enviMass processing.

Sample/Blind ratio

For the result filtering the parameter “Filter profiles by mean sample vs. Blind intensity ratio” was set 50. As can be seen in figure S4-6 this led to a slight decrease of total peak/profile number but also affect the recall of some suspects. Therefore, the sample/blind ratio was kept at 10 for result filtering.

4.6.2.3 Data processing with Compound Discoverer

Data were processed with the workflow template “Environmental w Stats Unknown ID w Online and Local Database Searches” without database search as shown in figure S4-10. However, the results in the “compound” table are given as molecular weights not as m/z. As this would not be comparable with m/z results of other software the “Merged feature” result lists were used for the software comparison. Therefore, replicate- and blank filter had to be implemented manually by filtering the merged features list for features only contained in all replicates and not present in blank.

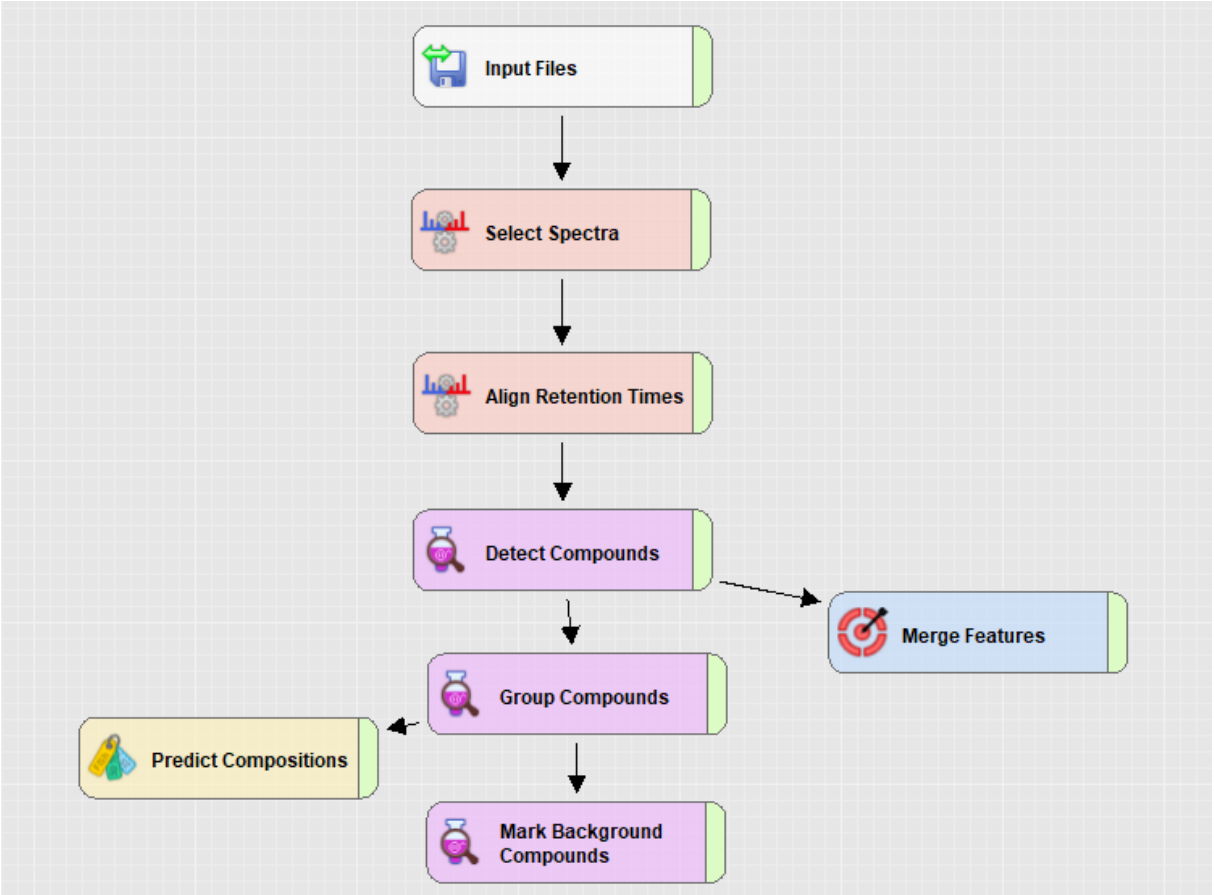


Figure S 4-10: Workflow nods of data processing with Compound Discoverer.

The parameter settings of different workflow nods are shown in table S4-9. Most parameter were given by instrumentation and measurement conditions or were fixed to harmonize workflows between different software. The recall rate of suspects was tested with these settings and was satisfying without further optimization.

Table S 4-9: Parameter settings of Compound Discoverer processing workflow.

Select Spectra	
1. General Settings	
Precursor Selection	Use MS(n-1) Precursor
Provide Profile Spectra	Automatic
2. Spectrum Properties Filter	
Lower RT Limit (min)	0
Upper RT Limit (min)	28
Min Precursor Mass (Da)	100
Max Precursor Mass (Da)	1000
Total Intensity Threshold	100000
3.Scan Event Filters	
Mass Analyzer	(Not specified)
MS Order	Is MS1; MS2
Activation Type	Is HCD
Min. Collision Energy	0
Max. Collision Energy	30
Scan Type	Is Full
Polarity Mode	Is +
4. Peak Filter	
S/N Threshold	3
5. Replacement for Unrecognized Properties	
Unrecognized Charge Replacements	1
Unrecognized mass analyzer Replacements	ITMS
Unrecognized MS Order Replacements	MS1
Unrecognized Activiation Type Replacement	HCD
Unrecognized Polarity Replacement	+
Unrecognized MS Resolution @ 200 Replacement	70000
Unrecognized MSn Resolution @ 200 Replacement	15000
Align Retention times	
Mass tolerance	5 ppm
Maximum shift	1 min
Alignment Model	Adaptive curve
Merge Features	
Mass Tolerance	5 ppm
RT Tolerance	0.1 min
Mark Background Compounds	
Hide Background	TRUE
Max. Blank/Sample	0
Max. Sample/Blank	5

4.6.2.4 Data processing with XCMS online

For data processing with XCMS online a predefined workflow for measurement with Orbitrap instruments at a resolution of 70,000 was used (ID 137). The parameters are shown in table S4-10, the commonly defined parameters were adjusted. A main advantage of this tool is, that only little previous knowledge is necessary, therefore it was assumed that most user would use predefined workflows, thus all other parameter were not changed. The triplicates of each sample were uploaded as sample group in mzXML format and were compared to the group of blank files.

Table S 4-10: Parameter settings of XCMS online data processing.

Feature Detection	
Method	centWave
ppm	5
Minimum peak width	10
Maximum peak width	60
S/N threshold	10
mzdiff	0.01
Integration method	1
Prefilter peaks	3
Prefilter intensity	100000
Noise filter	5000
Retention Time Correction	
Method	Obiwarp
profStep	1
Alignment	
mzwid	0.015
bw	5
minfrac	0.5
minsamp	1
max	100
Annotation	
Search for	Isotopes + adducts
m/z absolute error	0.015
ppm	5
Identification	
adducts	M+H, M+Na, M+K
ppm	5

4.6.3 Additional results

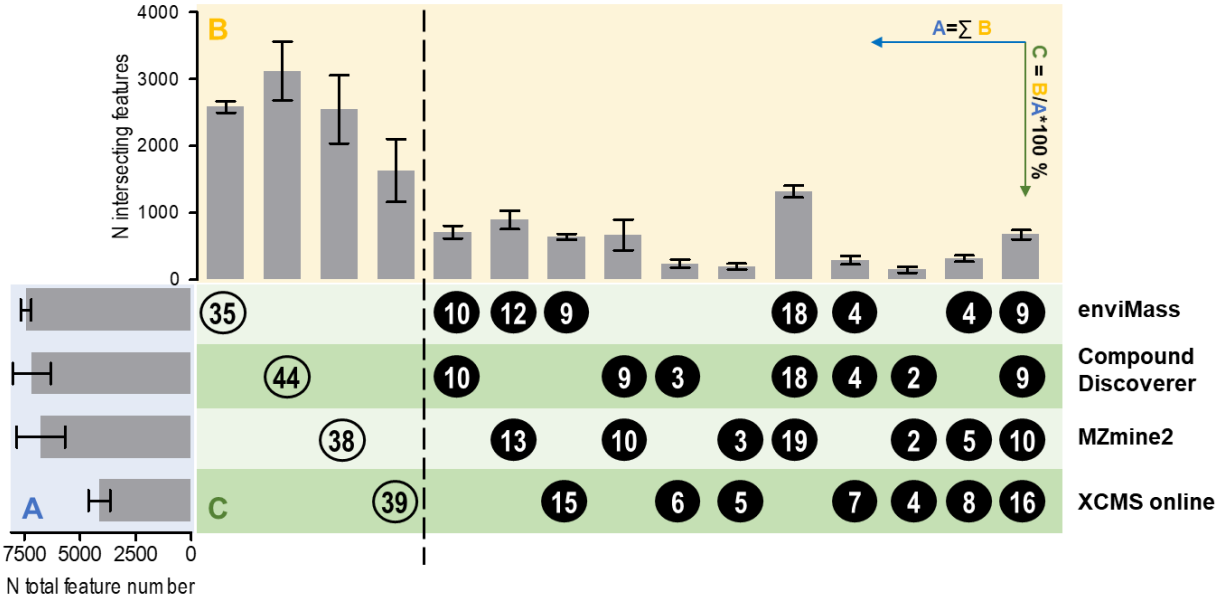


Figure S 4-11: Overview of intersecting features of all programs of WWTP effluent samples (N=3).

A: the bar chart on the left shows the mean absolute feature numbers of input lists from different software. B: the upper bar chart shows the mean absolute feature number of input lists from different software. **B:** the upper bar chart shows the mean absolute feature number of non-intersecting areas (left part to dividing line) and intersection areas (right part). **C**: the table below has two functions 1) it serves as legend, circles indicate overlaps between software to be compared, 2) the numbers in the circle show the ratio of overlapping features (B) to each input list (A). Again, the left part shows the percentage of non-overlapping features for each software (white circles) and the right part shows the percentages of overlapping features (black circles). Values in each row sum up to ≈100 % (deviations due to rounding). Reading example: The first bar right to the dividing line in Part B shows for example the commonly found feature by enviMass and Compound Discoverer which were 10 %.

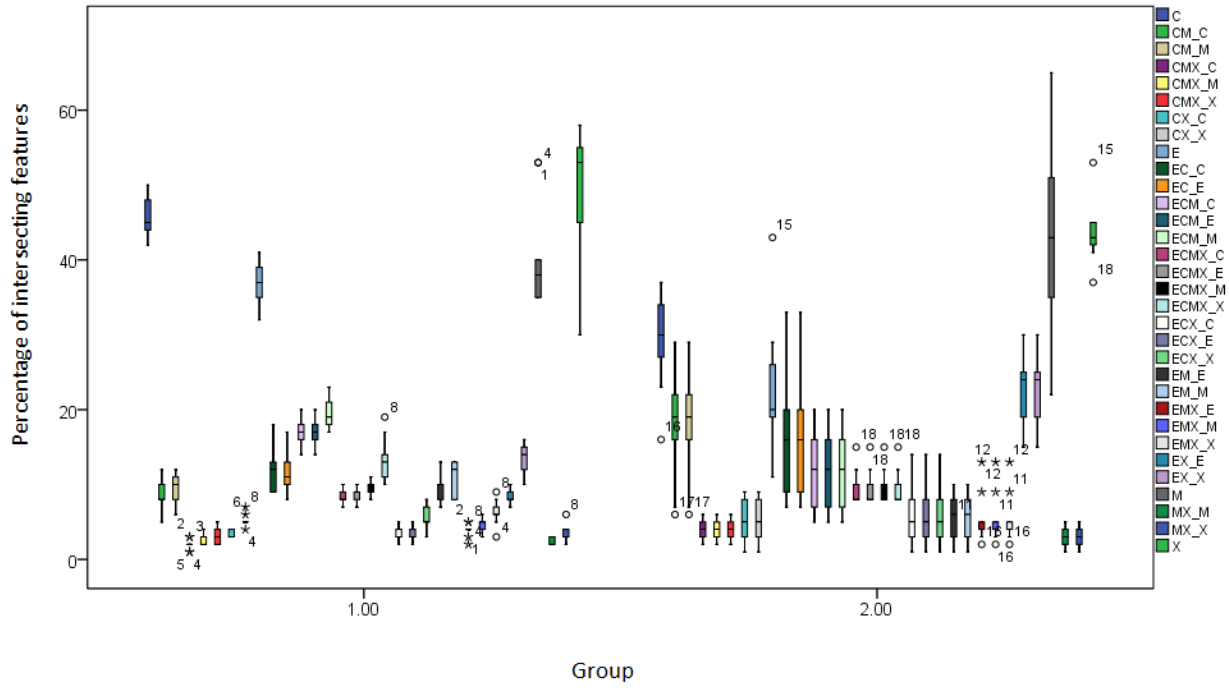


Figure S 4-12: Box plots for all non-overlapping and all overlapping features of all samples for whole feature list (group 1) and 100 most intensive features (group 2). C= Compound Discoverer, E= enviMass, M= MZmin2, X = XCMS online. Example: EMX_E overlap of enviMas

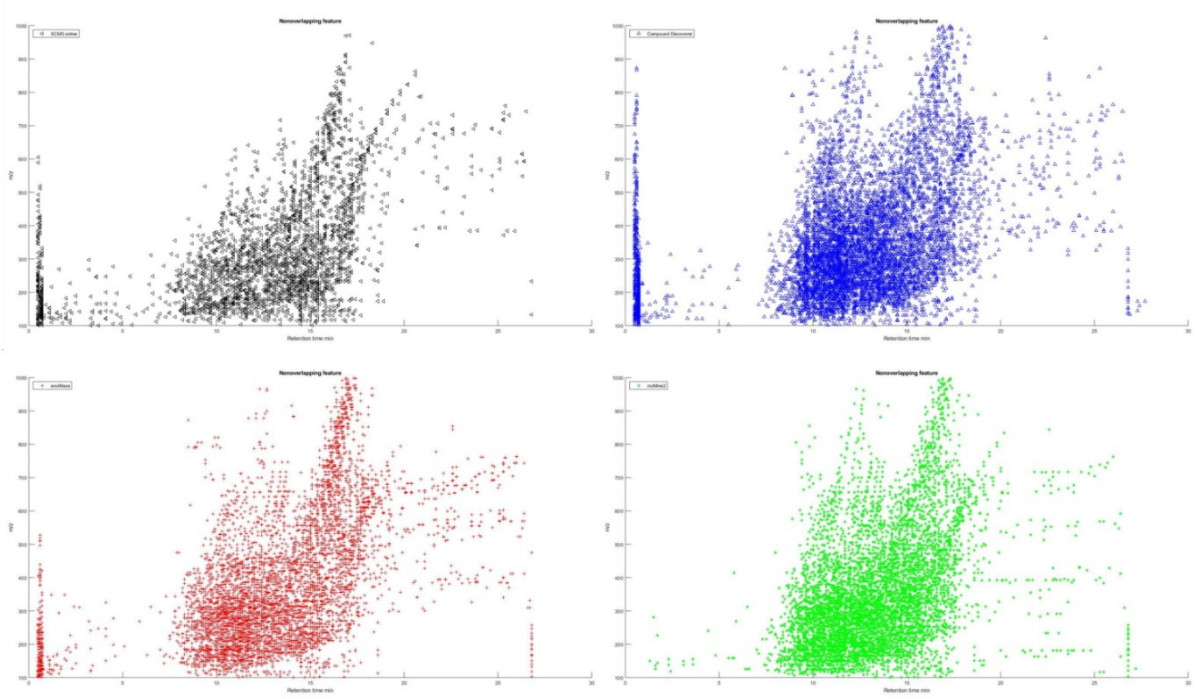


Figure S 4-13: Example of non-overlapping features as scatter (m/z vs. RT). Black: XCMS online, Blue: Compound Discoverer, Red: enviMass, Green: MZmine2.

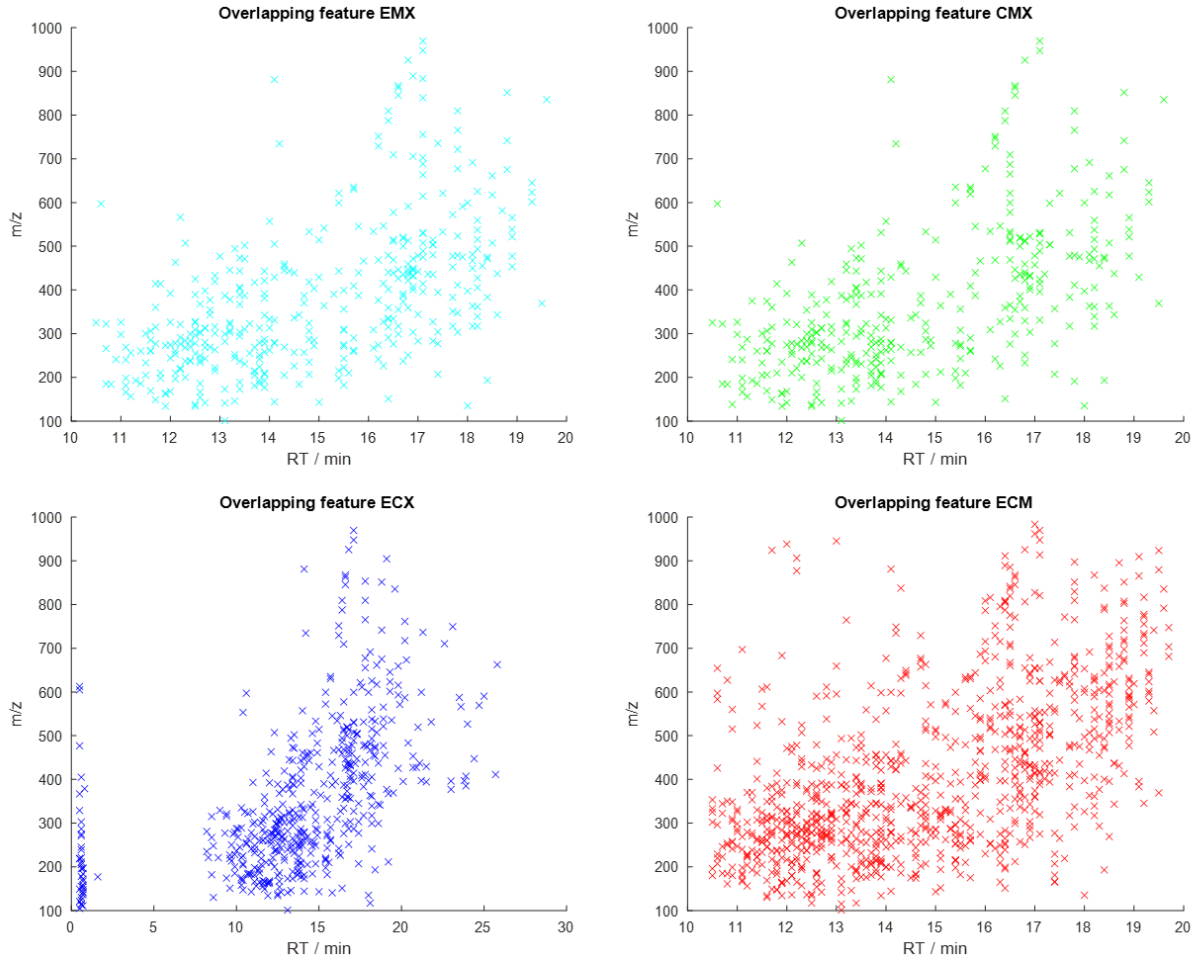


Figure S 4-14: Example of overlapping features as scatter (m/z vs. RT): E: enviMass, C: Compound Discoverer, X: XCMS online, M: MZmine2.

Figure S4-15 shows the effects on peak shape for suspects that show a very weak retention (Metformin, 0.6 min), middle retention (Aminosulprid, 9.9 min), (Carbamazepine, 13.4 min) and a strong retention (Diflufenican, 15.6min). The peak shapes of suspects were comparable, except for Metformin, which was however the only suspect eluting extremely early in the region of the injection peak.

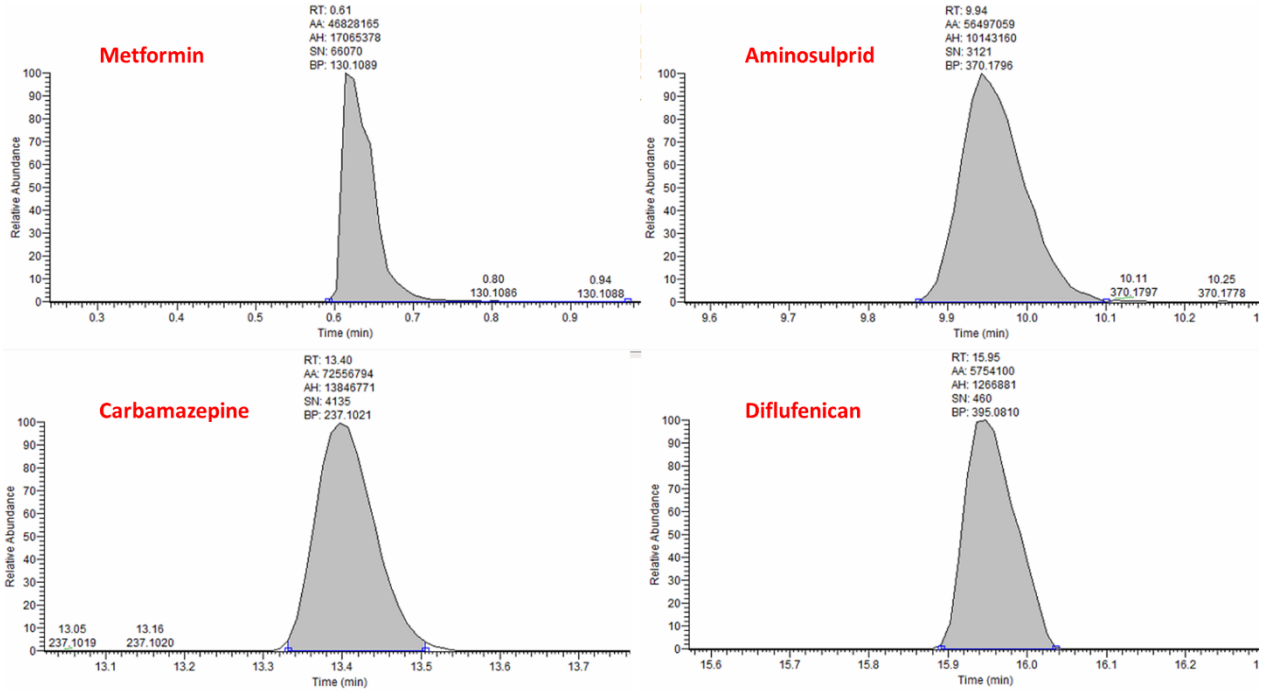


Figure S 4-15: Example of peak shapes of Metformin, Aminosulprid, Carbamazepine, Diflufenican extracted with Xcalibur Quan Browser.

Figure S 4-16 shows an example of the suspect carbamazepine and its variation of intensity/peak area over all samples with all four programs. As can be seen its presence in the environmental samples was not constant.

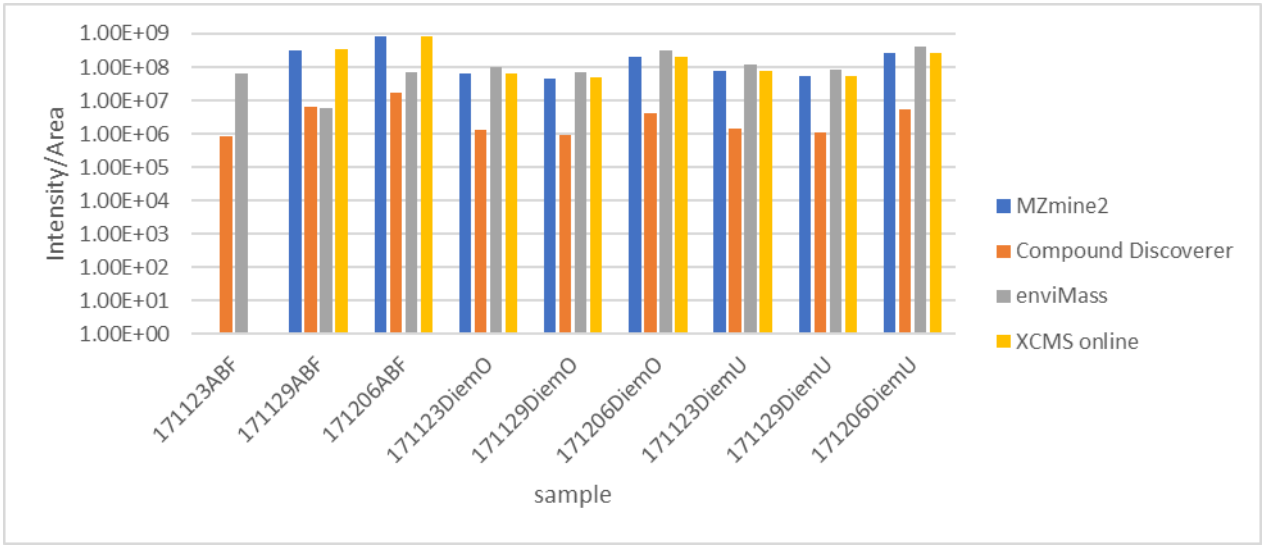


Figure S 4-16: Intensity/Area (logarithmic scale) of suspect carbamazepine over all 9 samples with four different programs.

Figure S 4-17 shows exemplarily the position of carbamazepine in the feature lists (sorted by intensity/area) with different software. As one can see, carbamazepine was in some cases included in the Top 100 in other cases not.

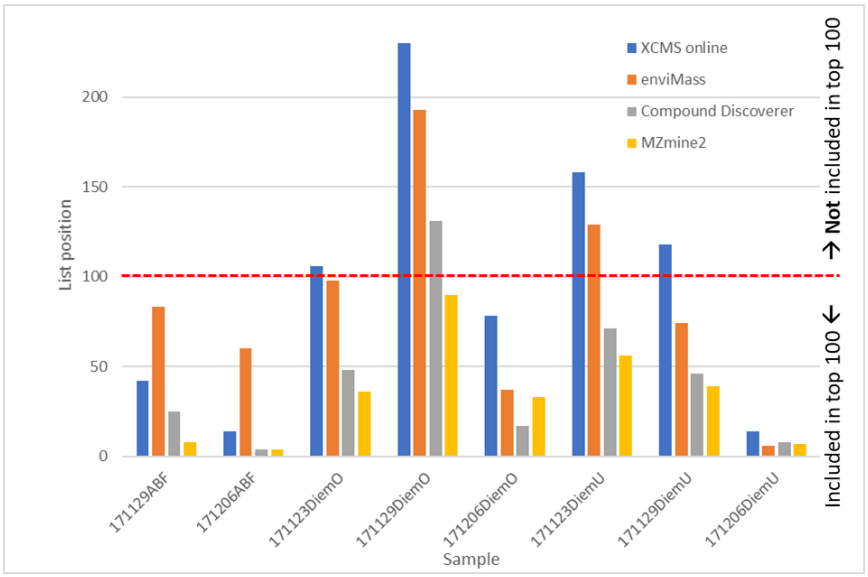


Figure S 4-17: List position of carbamazepine over 8 sample (171123ABF was excluded due to very high list position).

4.6.4 Statistical evaluation

4.6.4.1 Comparison of total and selected feature lists (part 3.2 and 3.3 of the paper)

Relative overlaps of feature lists between surface water samples and WWTP effluent samples are not significantly different from each other (Wilcoxon Signed Rank Test, p>0.05). Also, the correlation is significant through the bootstrapped Spearman´s rank correlation test. This is shown for whole feature lists and 100 most intensive feature lists in table S4-11.

Table S 4-11: Statistical evaluation of two group comparison including surface water and WWTP effluent samples.

Groups	Wilcoxon Signed Rank Test[c]		Correlations[d]	
	z	p-value	Correlation Coefficient	p-value
G1A and G1B[a]	-0.965	0.334	0.971	0.000
G2A and G2B[b]	-0.506	0.613	0.608	0.000

[a]G1A and G1B show two groups of surface water and WWTP samples, respectively, considering the whole feature list
[b]G2A and G2B show two groups of surface water and WWTP samples, respectively, considering 100 most intensive features
[c]The significance level is 0.05.
[d]Correlation is significant at the 0.01 level (2-tailed). Bootstrap results are based on 1000 bootstrap samples.

4.6.4.2 Comparison of total feature list with 100 most intensive features list

Table S4-12 shows the percentage overlapping and non-overlapping features of different software tools that were significantly different between total feature lists and 100 most intensive feature list according to the Mann–Whitney U-test.

Table S 4-12: Comparison of results of whole feature list and 100 most intensive feature list. C = Compound Discoverer, E = enviMass, M = MZmin2, X = XCMS online. Example: EMX_E overlap of enviMass, MZmine2 and XCMS online relative to total feature number of enviMas

Software[a]	C*	CM_C*	CM_M*	CMX_C*	CMX_M	CMX_X	CX_C	CX_X
Mann-Whitney U	0.000	14.500	16.000	5.000	11.500	14.000	16.500	27.000
p-value[b]	0.000	0.019	0.031	0.008	0.066	0.145	0.224	1.000
Software	E*	EC_C	EC_E	ECM_C	ECM_E	ECM_M*	ECMX_C	ECMX_E
Mann-Whitney U	6.000	24.000	25.000	12.500	11.500	4.500	12.500	16.000
p-value	0.012	0.776	0.864	0.088	0.066	0.005	0.088	0.224
Software	ECMX_M	ECMX_X	ECX_C	ECX_E	ECX_X	EM_E	EM_M*	EMX_E
Mann-Whitney U	25.000	11.000	19.000	19.000	26.500	11.000	7.500	18.500
p-value	0.864	0.066	0.388	0.388	0.955	0.066	0.018	0.328
Software	EMX_M	EMX_X	EX_E*	EX_X*	M	MX_M	MX_X	X*
Mann-Whitney U	27.000	18.500	0.000	0.500	26.000	18.000	26.000	2.000
p-value	1.000	0.328	0.000	0.000	0.955	0.328	0.955	0.002

[a] Relative overlaps of feature lists between different software
[b] The significance level is 0.05 (two-tailed).

4.6.4.3 Comparison of intermediate filtering steps (Part 3.4 of the paper)

Relative overlaps of feature lists at different processing steps were compared with each other using statistical test of Wilcoxon signed-rank test and did not show a significant difference ($p>0.05$) as given in table S4-13.

Table S 4-13: Wilcoxon signed ranks test for different intermediate steps.

Groups[a]	Wilcoxon Signed Rank Test[b]	
	z	p-value
STEP1 and STEP2	-0.128	0.898
STEP1 and STEP3	-0.083	0.934
STEP2 and STEP3	-0.785	0.432

[a]STEP1: individually processed replicates, STEP2: after replicate filter, STEP3 after blank and replicate filter.
[b] The significance level is 0.05 (two-tailed).

Chapter 5 Implementation of chemometric tools to improve data mining and prioritization in LC-high resolution mass spectrometry for nontarget screening of organic micropollutants in complex water matrices

This chapter was adapted from: L.L. Hohrenk, M. Vosough, T.C. Schmidt, Implementation of Chemometric Tools to Improve Data Mining and Prioritization in LC-HRMS for Nontarget Screening of Organic Micropollutants in Complex Water Matrixes, Anal. Chem. 91 (2019) 9213–9220. doi.org/10.1021/acs.analchem.9b01984

Abstract

One of the most critical steps in nontarget screening of organic micropollutants (OMP) in complex environmental samples is handling of massive data obtained from liquid chromatography coupled with high resolution mass spectrometry (LC-HRMS). Multivariate chemometric methods have brought about great progress in processing big data obtained from high dimensional chromatographic systems. This work aimed at a comprehensive evaluation of two LC-Q-Orbitrap mass spectrometry full-scan data sets for target and nontarget screening of OMPs in drinking and wastewater samples, respectively. For each dataset, following segmentation in the chromatographic dimension, at first multivariate curve resolution alternating least squares (MCR-ALS) was employed for simultaneous resolution of global matrices. The chromatographic peaks and the corresponding mass spectra of OMP were fully resolved in the presence of highly coeluting irrelevant and interfering peaks. Then, partial least squares-discriminant analysis (PLS-DA) was conducted to investigate the behaviour of MCR-ALS components in different water classes and selection of most relevant components. Further prioritization of features in wastewater before and after ozonation and their reduction to 24 micropollutants were then obtained by univariate statistics. Two-way information retrieved from MCR-ALS of LC-MS^1 data was also used to pick common precursor ions between predicted and measured data through data dependent acquisition. MS^1 and MS^2 spectral features were used for tentative identification of prioritized OMPs. This study indicates that the described strategy can be used as a promising tool to facilitate both feature selection through a reliable classification and interference-free identification of micropollutants in nontargeted and class-wise environmental studies

5.1 Introduction

Today, the use of liquid chromatography coupled with high resolution mass spectrometry (LC-HRMS) is the most powerful tool for screening of organic micropollutants (OMPs) and their transformation products (TPs) at environmentally relevant concentrations in complex environmental samples [145,146]. The main outline for LC-HRMS methods consists of three categories, the conventional target analysis which is used for quantification of target analytes with the help of reference standards and the qualitative suspect and nontarget screening, while the latter one has received much attention during the last years [147]. Nontarget screening approaches do not use any a priori information and thus provide a more comprehensive overview about the compounds present in a sample. Different data processing workflows are available in the literature for suspect and nontarget screening of unknown compounds, depending on the available software and the instrumentation [36,38,72,129,148–151]. The general procedure with different order in each workflow includes peak picking, removing the irrelevant peaks, componentization of isotopes and finally peak assignment steps. The main challenge with LC-HRMS instruments is the generation of massive amounts of data in full-scan acquisition mode of each chromatographic run. In fact, hundreds to thousands chromatographic peaks can be detected in each run and a huge number of chromatographic and mass-to-charge *features* are produced, accordingly. In addition, the ability of hybrid HRMS systems such as quadrupole-time-of-flight (Q-TOF) and Q-Orbitrap mass spectrometry (MS) for simultaneous full scan recording in MS^1 and MS^2 modes, increases the quantity of data in each run. Finally, the most complex situation arises when the simultaneous analysis of multiple chromatographic runs is considered. The huge amount of produced HRMS data necessitates the implementation of suitable post-acquisition data handling tools for proper background correction, peak detection, peak resolution, time-shift corrections and feature prioritization before an identification step.

As a very effective tool in this field, there are many different chemometrics methods that have been developed and grown to become a well-established high-dimensional data handling tool, with an extensive literature and increasing fields of applications [85,152]. In fact, different categorized chemometric methods such as multiway decomposition methods, supervised classification methods and so on, have a high potential to handle big data matrices in environmental analysis.

Multiway decomposition methods such as parallel factor analysis (PARAFAC) [153], PARAFAC2 [154], alternating trilinear decomposition (ATLD) [155] and multivariate curve resolution alternating least squares (MCR-ALS) are among the methods suitable for simultaneous resolution of “second-order” data such as hyphenated chromatographic data. MCR-ALS and PARAFAC2 have the extra advantage that one can retrieve the pure profiles of components even in the presence of retention time shifts and shape changes of components

between different chromatographic runs (non-trilinearity) [156].The former has been used successfully for modelling liquid chromatographic data coupled with diode array detection [60,157–161], fast-scanning fluorescence detection [157,161,162] and low resolution mass spectrometry (LRMS) [163–165]. Also, during the last few years MCR-ALS has successfully been utilized for analysis of LC-HRMS data in the context of targeted [166] or nontargeted metabolomics [42,61]and lipidomics [166] but not yet in environmental analysis.

Data arrays recorded from each LC-HRMS experiment can be processed to generate a data matrix with two modes of retention times and mass-to-charge (m/z) ratios. One of the LC-MS data processing and size reduction tools for efficient production of LC-MS data matrices has been developed by Tauler [167] and called region of interest (ROI) approach. Although this method has been successfully applied in metabolomics and lipidomics studies [168], it has not been implemented yet in environmental screening studies. The third dimension can be created by analysing the different LC-HRMS experiments obtained by replicate measurements of environmental samples acquired under different spatial, temporal and treatment conditions. These methods can be considered as an alternative to current peak detection approaches. In fact, while the most data processing strategies work based on analysing each m/z channel at a time, extracting ion chromatograms and grouping the features belonging to the same compound, multiway methods are able to simultaneously decompose the entire data matrices into the pure chromatograms, pure mass spectra and the relative quantities of all systematically oriented signals in different sample data matrices.

Also, the importance of utilizing these methods in case of resolving low intensity peaks, mathematically increasing the chromatographic selectivity without further LC method optimization was completely confirmed and their potential limitations were fully addressed previously [169,170].

Additionally, there exist a considerable number of multivariate chemometric methods which are well developed for exploratory data analysis such as principal component analysis (PCA) and hierarchical clustering as unsupervised methods or partial least squares-discriminant analysis (PLS-DA) [79] and support vector machines (SVM) [171] for supervised classification purposes. Here, the information provided from the multivariate methods can be used effectively for prioritization of features occurring in different classes of environmental samples.

During recent years, PCA has been used for inspection of relationships between different samples and to characterize samples' relevant unknown features [172,173] or creating a link between potential parent compounds and their TPs [71], in combination with the F-ratio method during the prioritization for nontarget analysis of environmental samples [174]. However, implementing supervised classification methods would lead to a very significant improvement in data classification and increasing specification. Finally, it has to be emphasized that the widespread chemometric methods need to be more comprehensively integrated into the

various processing steps for analysing the unknown samples in an environmental screening context and for tackling the relevant issues reported recently [100].

The main aim of this study is implementing multivariate chemometric tools for comprehensive evaluation of two LC-Q-Orbitrap MS data sets obtained through target and nontarget screening. The combination of the data compression method of regions of interest and multivariate curve resolution strategy (ROIMCR) described in this work allowed to overcome most of the encountered challenges, such as efficient data size reduction, alignment-free tracing of pollutants, tackling spectral interferences and highly overlapped OMPs elution profiles through one two-way decomposition step, feature prioritization and tentative identification of discriminatory pollutants between different classes of water samples.

5.2 Methods

5.2.1 Data collection

Details of the laboratory and data analysis practices are presented in the Supporting Information (SI). In brief, optimized solid-phase extraction-LC-Q-Orbitrap methods (tables S5-2 to S5-4) were performed on non-spiked and spiked (four concentration levels of 24 compounds, table S5-1) drinking water samples (target study) and also wastewater and river water samples (nontarget study). Sampling points for the later part of this work were obtained from influent-O_3 (after initial pre-treatment steps and before ozonation) and final effluent of the Warburg wastewater treatment plant as well as upstream of the receiving river Diemel. In each sampling point, three sampling dates, each with three replicates measurements were considered. In both parts of the study, three replicate measurements of a blank sample were performed.

5.2.2 Data evaluation

The resulting chromatographic data files were imported into the MATLAB environment (release 2013a, The Mathworks Inc., Natick, MA, U.S.A.). All matrices were individually and then simultaneously compressed by using the ROI approach. Then, the column-wise augmented matrices ($\mathbf{D}_{aug}$) containing 18 and 30 samples for the first and second part of this study were segmented and each segment was subjected to MCR-ALS analysis (figure 5-1).

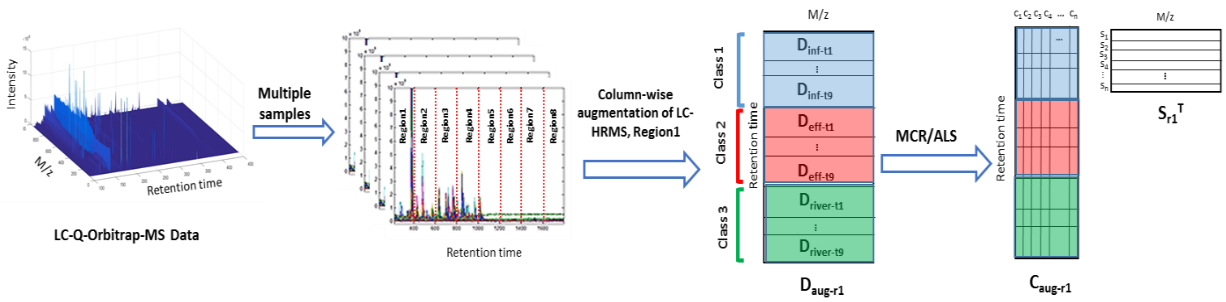


Figure 5-1: Representation of the multiple data collection, segmentation and column-wise augmentation of LC-Q-Orbitrap measurements of three different classes of influent-O3, effluent and river water samples. MCR-ALS resolution of the augmented matrix of first region (Daug-r1) into matrix of pure chromatographic profiles of Caug-r1 and pure mass spectral of Sr1T has been shown.

The areas under the MCR-ALS resolved chromatographic peaks in the $\mathbf{C}_{aug}$ matrix were transferred to a new data matrix, where its columns contain areas of every resolved component, and its rows contain the different water samples. The sizes of the matrices for target and nontarget parts of this study were 18×150 and 30×250, respectively. In order to process the matrices of peak areas, data cleaning was firstly performed in order to remove irrelevant chromatographic peaks from the rest of the data treatment pathway [132]. Among several data pre-treatment approaches, logarithmic transformation was selected. By this transformation, skewness was reduced giving normally distributed variables. In addition, data driven normalization based on row-wise scaling was applied and coefficients of variation (CV) of features among the replicates of a sample was decreased, accordingly [175].

The matrices were subjected to the further analysis of variance (ANOVA)-simultaneous component analysis (ASCA) [88] and PLS-DA [79]and univariate statistical studies. On the one hand, ASCA was applied to the nontarget screening part of the work to statistically study the effect of different locations and sampling times as possible sources of experimental variances. On the other hand, PLS-DA was implemented to discriminate among different sample groups. These groups in the target part of the study comprised spiked and non-spiked drinking water samples and in the nontarget part of the study influent-O3 samples, effluent samples and river water samples. The calculations involving MCR-ALS were performed using the MCR-ALS 2.0 toolbox (available at www.mcrals.info). ASCA and PLS-DA methods were applied using the PLS Toolbox 7.8 (Eigenvector Research Inc., Wenatchee, WA, U.S.A.) working under MATLAB.

5.3 Results and discussion

5.3.1 Resolution of LC−HRMS data

5.3.1.1 Target screening

With the aim of extraction of chromatographic and mass spectral profiles of all involved components in each sample data matrix, the extended MCR-ALS was implemented for

simultaneous analysis of all drinking water samples, without performing background correction or chromatographic peak alignment. As shown in figure S5-1 (A), the full exported LC-HRMS[1] data for each sample were partitioned into three regions in order to simplify the analysis (table S5-5). Chromatographic region I, II and III were resolved by an MCR-ALS model using 40, 50 and 60 components, respectively. The percentage of explained variance (R^2) for all models was higher than 99%, and lack of fit (LOF) lower than 4%, both highly acceptable considering the large number of the modelled components. Finally, a total number of 150 components were used to explain the variance of the whole data set. An example of the resolved elution profiles in the presence of irrelevant profiles is shown in figure S5-1 (B) for the MCR-ALS analysis of region II (containing 500 ng/L of target analytes). Fourteen target compounds were successfully resolved from interferences with mathematical separation. As can be seen from figure S5-1 (B), some micropollutants with low intensity signals such as 5-chloro-1H-benzotriazole that were embedded under other high intensity target compounds or interfering signals, were also completely resolved and detected using the MCR-ALS method. Each MCR-ALS component contains the successive pure elution profiles, pure mass spectra and peak area of that component that is retrieved from the original augmented matrix $\mathbf{D}_{\mathbf{aug\text{-}T}}$. So, all information regarding the mass features of each MCR-ALS component, such as isotopic peaks and adduct peaks can be collected and considered for identification purposes. Interfering signals in this part of the study imply detectable unknown drinking water compounds together with all irrelevant and noisy chromatographic peaks such as blank and instrumental contributions. Then, data cleaning of the peak area matrix was performed and the features with the following properties were removed from the rest of data analysis: square peaks which are artifact peaks with an area-to-height ratio >30, features with a CV >30% during triplicate analysis and features with very small standard deviation (SD) of the peak areas by pairwise comparison of spiked or non-spiked samples with a blank sample [132]. Finally, 49 from 150 resolved components were retained in the area table to be subjected for PCA and PLS-DA modelling.

5.3.1.2 Nontarget screening

Nontarget analysis of 30 data matrices containing influent-O3, effluent, river water and blank samples was attempted to demonstrate the utility of the ROIMCR strategy. Total ion chromatograms (TICs) of three water samples taken within one sampling time are indicated in figure S5-2 (A). Each chromatographic data set was partitioned into eight regions in order to simplify the resolution process (table S5-6). Then, MSROI subsets in 30 samples were column-wise augmented to establish the matrix $\mathbf{D}_{\mathbf{aug\text{-}NT}}$ for submission to MCR-ALS modeling. In figure S5-2 (B), $\mathbf{D}_{\mathbf{aug\text{-}NT}}$ for the second region of LC-MS data of different wastewater and river water samples is shown. The complexity of the chromatograms can be clearly observed and were, as expected, more serious in the influent-O_3. Then, individual models were built for each

augmented sub-matrix. An example of the resolution process using MCR-ALS analysis for region II is shown in figure 5-2. Here, the resolved LC profiles of 44 components (including OMPs, background and irrelevant peaks) in 27 matrices from three classes can be observed. The elution profiles resolved in the time region of the resolved mass spectra of tentatively compounds of acridine, 4- or 5-methyl-1H-benzotriazole and benzotriazole are shown in the matrices that contained the highest intensities. This figure clearly shows the resolution capability of the mentioned algorithm for a highly complex data matrix. So, MCR-ALS is a rather robust method since it allows working with highly complex scenarios such as large retention time shifts, highly intense and broad background contributions, and lower mass resolution, or combination of all issues. The mentioned decomposition process was successfully repeated for the other sub-matrices with various number of components. Finally, a total number of 250 MCR-ALS components were used to explain the variance of the whole data set in the samples (table S5-6).

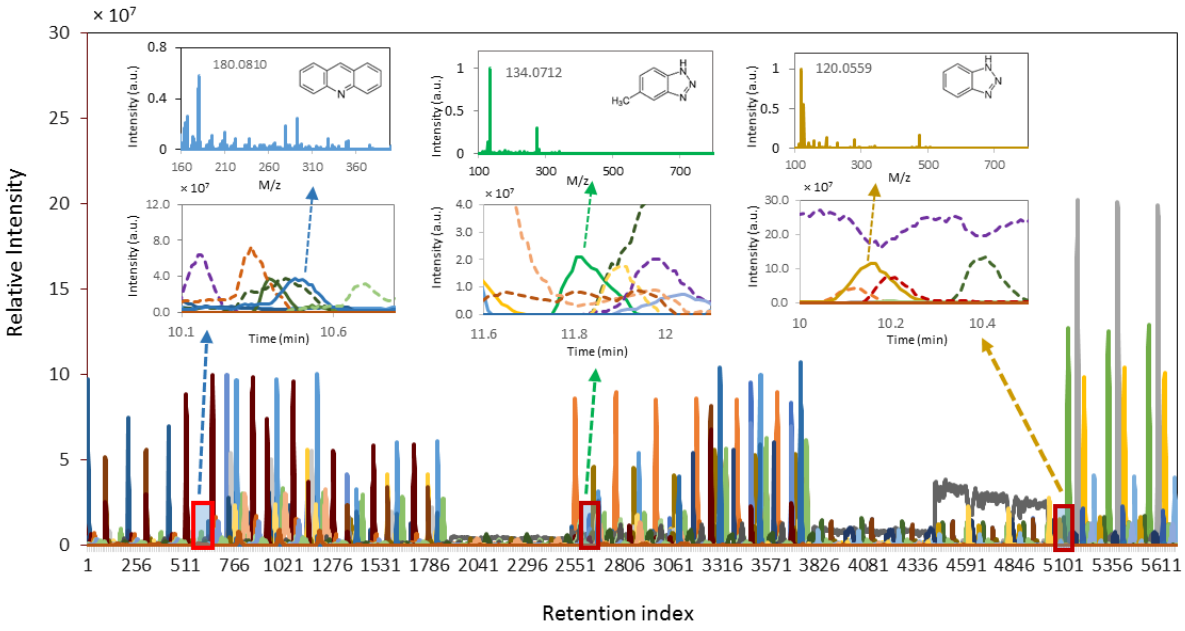


Figure 5-2: Successive 27 retrieved chromatographic profiles obtained through MCR-ALS analysis of the time region between 9.94 and 12.25 min, containing 44 components. The insets show the zoomed view of the elution profiles resolved in the time region of the three exhibited mass spectra.

R^2 values or all models were at least 98.6%, and the LOF values were lower than 6%, which is well acceptable considering the high number of components and the considered intensity threshold. Further, the resulting elution profiles were integrated and the areas of all resolved components were estimated for each time window. Area values were stored in one data table containing the information of the 8 chromatographic time windows, conformed by a number of rows equal to the number of studied samples and a number of columns corresponding to all the MCR-ALS resolved components. As many MCR-ALS resolved components could be assigned to solvent or background contributions, noise or artifacts, a cleaning of the peak area

matrix was performed, and the relevant components were selected. Finally, the cleaned matrix of area values which had the dimension of 27×101 was subjected to multivariate statistical analysis.

5.3.2 Statistical evaluations and prioritization of screening

5.3.2.1 Drinking water samples

In this section, the target compounds have a critical role in discriminating the sample classes. Prior to multivariate modelling, the peak areas were log-transformed to reduce skewness and give equal relevance to their possible change due to the different concentration levels in control and polluted samples. At first, PCA was applied to the peak area table to investigate the inherent structure of the data. Using PCA, separation of the non-spiked and spiked water samples (even with different concentration level of pollutions) could be achieved based on the score plot of the first two PCs (figure S5-3 (A)) and the peaks could be characterized based on water groups using their loading values. This is while even in case of retaining the solvent peaks during matrix cleaning step, the peak features attributed to solvent components could separate clearly the solvent class in the score plot. To make a more elaborate survey to reveal which features had the greatest influence on the discrimination among different samples, a PLS-DA analysis was performed to the peak area table considering two group classes of non-spiked and spiked water samples. The PLS-DA model was developed to investigate which micropollutant peak areas were more important in the discrimination between water samples. Two PLS-DA components were enough to explain most of the class variance (96.66 % variance) using the 61.10 % of X variance related to the changes on chromatographic peak areas of the resolved components (figure S5-3 (B)), with specificity and sensitivity values equal to 1 for each class. Also, through applying leave-one-out cross-validation, which is the most suitable strategy for a small number of samples, no outlier sample were observed. Then, the variable importance in projection (VIP) scores, which is a weighted sum of squares of PLS weights described for each predictor variable to the model, was calculated [176] . Among 25 spiked micropollutants, 3 targets were found with VIP<0.8, 4 targets with 0.8<VIP<1, 18 targets with VIP>1. Considering the fact that the main discriminating factor between the mentioned classes is the presence of micropollutants, a total detection rate of 72% (VIP>1) may be worth discussing. The three target compounds with VIP<0.8 are caffeine, terbutylazin-desethyl and ibuprofen. Caffeine was detected already in the drinking water sample, so it cannot play a discriminatory role and the other compounds had a highly significant within class variance (because of some measurement related issues), which caused the non-significant scores for discrimination. Considering the four micropollutants of chloridazon-desphenyl, metformin, toremifen and tamoxifene with VIP values between 0.8 and 1, the “greater than one” criterion could be re-adjusted to capture the features with less discriminatory power [177]. On the other hand, in case of retaining blank features in the spiked and non-spiked samples, all of them

were put into non-discriminatory features (VIP<0.5), so they can be discarded through this step.

5.3.2.2 Wastewater Samples

In this section, at first the statistical significance of the sampling time and sampling location factors was evaluated using a two-factor design using the ASCA method, which considers peak areas of all micropollutants in the same test. Results showed that both factors were statistically significant, with a p-value of 0.005 and 0.001 for effects of sampling time and sampling location, respectively. The interaction between factors was not statistically significant with a p-value of 0.17. Considering the noticeable variance of sampling location with respect to the other effects, this effect was further investigated on MCR-ALS resolved components patterns using the PLS-DA method. Here, PLS-DA was performed to find the components that showed significant differences among the three classes of influent-O3, effluent and upstream samples. The biplot depicted in figure 5-3A confirmed the significant difference between the three classes through 52.7% **X** data variance and 93.6% of the dependent variable **Y**. This model was constructed by performing venetian blind 10-fold cross validation and it demonstrated good modelling and prediction using two components. Through this modelling, no outlier sample was detected and the number of misclassified (NMC) samples was zero. Also, site-specific unknown compounds for all sampling locations could be found by selecting the discriminatory features for each class. The VIP score plot for two classes of influent-O3 and effluent water samples is shown in figure 5-3B. The more important features in this plot (54 features with VIP>1) revealed which of the detected micropollutants were more effective for distinction of wastewater samples. Further prioritization was performed by univariate analysis using a volcano plot (figure 5-3C) of selected classes which combines both fold-change and t-test criteria. This plot enabled us to select those relevant features that show both significant difference in magnitude and fold change (FC) through statistical tests (p-value<0.05). Here, 35 features shown in the volcano plot with dark blue colour, were found to be significant (with p-values<0.05 and FC>2 or FC<0.5). Eventually, 24 features which corresponded to 24 micropollutants met the overlapped area (with p-values<0.05 and FC>2 or FC<0.5, VIP>1). Also, it becomes evident from the figure that light blue points located in the grey area can be representative of the micropollutants with high potential of consistency during ozonation and biological treatment in the WWTP.

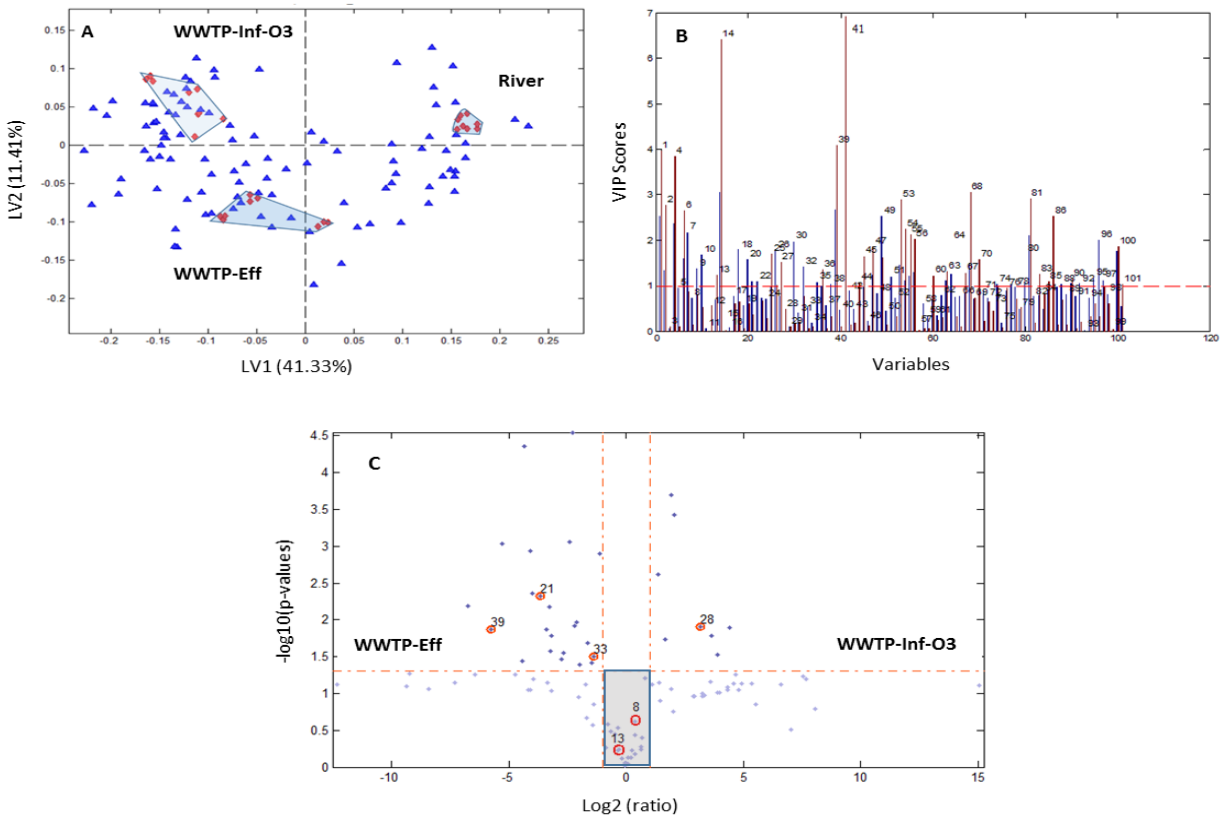


Figure 5-3: (A) Representation of score-loading biplot resulting from PLS-DA analysis of the preprocessed chromatographic peak areas obtained by MCR-ALS analysis of three classes of water samples. Red diamonds and blue triangles represent scores and loading, respectively. PLS-DA shows that there is an obvious difference between the pollution profiles of WWTP-Inf-O3, WWTP-Eff and upstream river samples with LV1 (41.33%) and LV2 (11.41%). The model was constructed by performing 10-fold cross validation with R2 of 93.33% and Q2 of 82.00%. (B) Projection (VIP scores) for each variable using the PLS-DA model. Horizontal red line shows the threshold value selecting the most important variables. (C) Volcano plot combining the statistical test (y-axis:- log (p-value)) and the magnitude of the change (log2 (FC)) of micropollutants on a scatter plot. Dark blue points represent the compounds with p-value < 0.05, and FC>2 or <0.5. Light blue points represent the compounds with p-value> 0.05. The grey area shows the micropollutants with the potential of consistency during treatment process in the WWTP. The numbers of 8, 13, 21, 28, 33 and 39 show the position of tentative identified compounds of trimethoprim, 4 or 5-methyl-1H-benzotriazole, 1H-benzotriazole, acridine, sulfamethazine and carbamazepine, respectively

5.3.3 Precursor ion selection using MCR-ALS resolved MS^1 spectra

To demonstrate the overall pattern of the precursor ions in DDA mode for producing MS^2 spectra, the scatter plot of all recorded precursor ions (more than 10000 MS^2 features) in each retention time of an influent-O_3 sample together with all resolved chromatograms is shown in figure S5-4. Since in DDA mode the five most intense ions in each scan are selected by the instrument, this does not necessarily result in capturing MS^2 of precursor ions of all micropollutants of interest. The reason is the low intensity value of some micropollutants in the samples and a highly intense and broad background in the retention time region of some compounds [178]. Also, because of the retention time shift of the compounds between different samples, their elution windows are not the same and the position and the width of each elution profile in different samples are different. The mentioned issues were resolved here by

combination of the information from the less complex samples with more complex ones through global analysis of all samples with MCR-ALS modelling, so the precursor ions of very low intensity peaks can also be captured using the mentioned strategy.

To do so, we used the two-way information retrieved from MCR-ALS of LC-MS^1 data, with the following advantages: utilizing the pure full spectrum of each micropollutant rather than relying on few mass fragments, independency on peak picking in the MS^2 domain that is problematic in highly complex samples, independency on chromatographic peak shapes because all pure chromatographic peaks have been resolved already. This can also be observed in figure S5-4, which shows that the selected precursor ions (red circles) have not always been located in the elution regions of all prioritized components for influent-O_3 sample. This way of MS^2 selection has a high level of confidence and low potential of false positive selection. The mentioned strategy of precursor ion selection will be further explored in the future in a more comprehensive way. It can also be extended to the more complex data-independent approach (DIA) due to their more considerable MS/MS coverage.

5.3.4 Identification and annotation of discriminating compounds

The resolved MS^1 and MS^2 spectral profiles associated with MCR-ALS resolved chromatographic profiles and peak areas allowed the tentative identification of the discriminatory or consistent compounds (table S5-7). In order to do this, at first the m/z of selected compounds were searched within a ± 5 ppm window in the database FOR-IDENT [141] where water relevant micropollutants, their transformation products and metabolites are listed. Further the m/z values were searched with the mzCloud database [179] using the peak search with a m/z accuracy of 0.001 Th and the MS^2 spectra searched with the spectrum search function. In addition, the MS^2 spectra were loaded into the Mass Bank database [180]. If the m/z and MS^2 search in all databases came to the same identification result, the compounds were classified as tentatively identified. As examples of the tentatively identified compounds considered to be consistent and associated with influent-O3 or effluent wastewater samples, the following micropollutants can be mentioned (see figure 5-3C). The accurate mass of compound 33 associated with effluent samples was 279.0918. This mass could be assigned to the precursor $[M+H]^+$ ion of the veterinary sulfonamide sulfamethazine (with 3 ppm relative error). Appearance of sulfamethazine in higher concentration in effluent samples can be the result of deacetylation of the metabolites (e.g., N4-acetylated- sulfamethazine) released in WWTPs and reforming the parent drug. A very recent study showed that acetylated sulfonamides which are much less stable than their nonacetylated forms may be re-transformed into their parent sulfonamides in effluent samples, which is in accordance with our findings [181].

Compound 21, which is associated with effluent samples, is tentatively identified as 1H-benzotriazole. This compound is a metal anticorrosive and applied in a wide range of

commercial and industrial applications. This compound can also be considered as a persistent pollutant if fold change interval is expanded in volcano plot. On the other hand, compound 13 with accurate mass of 134.0712 is one of the isomers of 4 and 5-methyl-1H-benzotriazole (with 2 ppm of relative error), lies in the middle part of the volcano plot and does not pass the statistical p-value and fold change criteria, so it can be considered as a persistent organic pollutant through the ozonation process and biological treatment. These results confirm the previous reports on these common forms of BZTs:1H-benzotriazole and tolyltriazole which are resistant to biodegradation and are only partially removed in wastewater treatment [182]. In fact, these pollutants were also detected with rather high peak intensities in upstream river water samples, which further confirms their persistence in water. Also, carbamazepine (compound 39) with an accurate mass of 237.1021 was classified as effluent indicator. This was also reported in literature which showed that higher signal intensities of carbamazepine in effluent samples can be attributed to the degradation of transformation products released in treatment plants or a higher amount of signal suppression in influent-O_3 samples [71]. Another important result is detection of acridine (with accurate mass of 180.081 for precursor $[M+H]^+$) which was considered as influent-O_3 classifier (compound 28). According to the literature, this mutagenic and carcinogenic compound is a biological transformation product of an acridine dye, so during the biological treatment has been reduced significantly [183]. Compound 8 with mass of 291.1469 represents the veterinary antibiotic trimethoprim in the middle part of the volcano plot and considered as a persistent micropollutant. A more comprehensive identification of all classified pollutants was not within the scope of the present study.

5.4 Conclusion

Multivariate chemometric tools can be efficiently employed for extracting a wealth of information in LC-HRMS(/MS) data. As a preliminary study in this field, it was found that the implementation of multivariate deconvolution and supervised classification of PLS-DA of LC-Q-Orbitrap data sets have a great potential to empower targeted and nontarget screening strategies for more reliable detection and discovery of micropollutants and their TPs without suffering from matrix interfering peaks. MCR-ALS modelling of all data matrices showed that all relevant and irrelevant components that have systematic signal variation can be retrieved in both ways of measurement. The main advantages include avoiding an uncontrolled loss of information by a separate baseline correction step, no need to perform chromatographic alignment, eliminating redundancies in data interpretation because of grouping of the multiple features of the same component and plausible tracing of all peaks across various matrices. Since most of the ambiguously assigned features are usually found in the retention time regions where there is a high probability of interfering profiles, the interference-free resolution of potential pollutants in highly coeluted and complex samples such as influent-O_3 and effluent

wastewater samples will mathematically recover the features without the need to further optimize the chromatographic method. ASCA of relative peak areas obtained by MCR-ALS showed a significant difference between the times and locations of sampling, so PLS-DA combined with an univariate statistical test was efficiently applied for selecting the most prioritized features to discriminate different site specific classes. Among 101 resolved compounds, 24 potential OMPs which were associated with influent-O_3 or effluent wastewater samples met both multivariate and univariate statistical criteria. Meanwhile, there was also the possibility of detecting another group of micropollutants that are resistant during treatment process in the WWTP. Considering the opportunities provided through implementation of different chemometric methods, this strategy can be applied to non-target screening of OMPs and their TPs in highly complex samples with different temporal or spatial patterns or when the water samples undergo a chemical or biological treatment process.

5.5 Acknowledgements

Financial support from the DFG (Germany) for initiation of this collaborative study is gratefully acknowledged (Grant No. SCHM 1372/12-1). We thank Gerrit Renner and Prof. Karl Molt for valuable comments on the manuscript.

5.6 Supporting Information

5.6.1 Chemicals and sampling

For the target screening part of this study, drinking water samples were spiked at concentrations of 0.1, 0.5, 1 and 10 µg/L with a standard mixture of 25 analytes. Targets were selected from various pollutant classes, such as pharmaceuticals, personal care products and pesticides (Table S5-1). Further, an unspiked drinking water sample and an ultra-pure water blank were measured. A triplicate analysis was performed on all samples. For the nontarget screening section, samples were taken at different treatment steps of a municipal waste water treatment plant (WWTP) in Warburg (Stadtwerke Warburg GmbH, Warburg, Germany), which is designed for 70,000 population equivalents and receives municipal as well as industrial waste water. The WWTP consists of an advanced treatment unit, which is designed as a two-line full-scale ozonation with a moving bed reactor as biological post-treatment. Grab samples were taken at specific ozone dosages of 0.3, 0.5 and 0.7 mgO_3/ mg DOC over a period of three weeks. Sampling points were obtained from influent-O_3 (after initial pretreatment steps and before ozonation) and effluent of Warburg wastewater treatment plant as well as upstream of the point source in the receiving river Diemel. In each sampling points, three sampling dates, each with three replicates measurements were considered.

5.6.2 Samples and sample preparation

A solid phase extraction (SPE) was carried out within 48 h after sampling. SPE cartridges (150 mg, 6 mL, Oasis HLB, Waters, Germany) were conditioned (2 x 5 mL methanol) and equilibrated (2 x 5 mL water, LC-MS grade). Cartridges were loaded with 1000 mL sample and subsequently dried under vacuum and stored at -18 °C until further sample preparation. The elution was done using 5 x 5 mL methanol (LC-MS grade), and the solvent was evaporated to complete dryness at 50 °C under a gentle nitrogen stream. Before analysis, the samples were re-dissolved in water (LC-MS grade). Further, blanks were processed the same way as described above with ultra-pure water instead of sample.

5.6.3 Instrumental conditions and analysis

Chromatographic separation for both target and nontarget screening was performed using a Dionex UltiMate 3000 HPLC system (Thermo Scientific, Bremen, Germany). The mobile phase consisted of eluent A: ultrapure water + 0.1% formic acid, and eluent B: methanol + 0.1% formic acid (both MS grade). Injection volume was 20 µL for both methods. All samples were measured in triplicates. The gradient elution programs, flow rates and chromatographic columns for target (method A) and nontarget (method B) screening are summarized in table S5-2. Mass spectrometric detection was performed on an Orbitrap mass spectrometer (QExactive Thermo Scientific, Bremen, Germany) using electrospray ionization in positive mode. The parameters of the ion source are summarized in table S5-3. The full scan HRMS

spectra acquisition (*m/z* 100-1000) with a resolution of 70,000 was followed by data dependent MS2 scans of the five most intense ions with a resolution of 17,500. The settings of the full scan/ddMS2 (Top 5) measurement are shown in table S5-4. Mass calibration was performed with the calibration solution (Pierce LTQ Velos Positive/Negative Ions Calibration Solution, Thermo Scientific, Bremen, Germany).

5.6.4 Data analysis strategy

5.6.4 1. Data arrangement and compression

All raw chromatographic data were acquired in profile mode using Xcalibur software (Thermo Fischer Vendor). These chromatograms were then converted to an open format, mzXML, employing the msConvertGUI software [140]. Then, these data files were imported into the MATLAB environment (release 2013a, The Mathworks Inc., Natick, MA, U.S.A.) by using *mzxmlread.m* and *mzxml2peaks.m* functions of the MATLAB Bioinformatics Toolbox (4.3.1 version). Finally, the MATLAB structure *mzXMLStruct* containing the peak information was created. For data compression, the ROI approach was employed in the present study. The ROIs reduce the MS data size and the computer storage requirements without any significant loss of spectral resolution and accuracy of *m/z* data [42,184,185]. Also, as MS instruments provide an irregular number of measured *m/z* and signal intensity for each scan number, employing ROIs has the ability to convert the initial arrays into a matrix of data which is suitable for multivariate data analysis. The choice of these regions depends on three parameters [167]: the mass intensity threshold (SNR_{thr}), the *m/z* error or the mass accuracy of the spectrometer and the minimum number of retention times to be considered in a ROI, which were set at 0.1% of the maximum MS signal intensity, 0.001 amu for the Orbitrap MS analyser and 15, respectively.

After individual compression of each sample data matrix, the matrices were arranged in a single column-wise augmented data matrix (figure 5-1). Then, the samples were simultaneously handled by the ROI function in order to find common and uncommon ROI m/z values between each pair of samples. This search was performed for all data matrices that were supposed to be handled simultaneously. The MSROI data matrices are augmented and they make a new augmented MSROI data matrix before MCR-ALS analysis.

5.6.4.2 MCR-ALS resolution of LC−HRMS water and wastewater data

In the present study, MCR-ALS was used for multivariate deconvolution of complex chromatographic data of water and wastewater samples into their pure LC profiles and their counterpart mass spectral profiles. An augmentation step is the first step of data analysis by MCR-ALS. This global data matrix of several chromatographic runs is created along the mode which is suspected to break the trilinearity of data structure. So, a column-wise augmented matrix ($\mathbf{D}_{aug}$) containing 18 and 30 samples for the first and second part of the study was

constructed, respectively. Simultaneous bilinear decomposition of each augmented data matrix $\mathbf{D}_{aug}$, containing *K* matrices was performed as shown in equation S5-1:

$$Equation\ S\ 5\text{-}1:\ \boldsymbol{D}_{aug} = \begin{bmatrix} D1 \\ D2 \\ D3 \\ . \\ . \\ . \\ Dk \end{bmatrix} = \begin{bmatrix} C1 \\ C2 \\ C3 \\ . \\ . \\ . \\ Ck \end{bmatrix} \boldsymbol{S}^{\boldsymbol{T}} + \begin{bmatrix} E1 \\ E2 \\ E3 \\ . \\ . \\ . \\ EK \end{bmatrix} = \boldsymbol{C}_{aug}\,\boldsymbol{S}^{\boldsymbol{T}} + \boldsymbol{E} \qquad K = 18,\ 30$$

where the rows in matrix $\mathbf{D}_{aug}$ (*I*(the number of elution times in each sample)×*(K*(number of samples)), *J* (number of m/z values)) contain the recorded MS_1 spectra as a function of time, the columns of $\mathbf{C}_{aug}$ (*I*×*K, N*) contain the elution time profiles of the compounds (*N*) involved in the process for all individual sub-matrices, the columns of $\mathbf{S}_T$ (*N,J*) represent their corresponding pure MS_1 spectra, and **E** (*I*×*K, J*) is a matrix of residuals not fitted by the model. The values for *I*, *J* and *N*, for each $\mathbf{D}_{aug}$ matrix in target and nontarget studies are indicated in tables S5-5 and S5-6, respectively. Before starting resolution, the number of components to each data matrix $\mathbf{D}_{aug}$ was estimated using singular value decomposition (SVD). The initial estimates were produced by SIMPLISMA (simple interactive self-modelling mixture analysis) which were the purest spectra of involved components [186]. Then, decomposition of $\mathbf{D}_{aug}$ and the estimation of $\mathbf{C}_{aug}$ and $\mathbf{S}_T$ matrices was performed by iterative least-squares minimization of ||**E**||, under constraints of non-negativity in the spectral and chromatographic profiles and spectral normalization (equal height). The mentioned constraints were selected according to the prior knowledge about the data sets and also through a previous MCR-ALS analysis of each augmented data matrix. In fact, according to the nontrilinear structure of the data sets and non-unimodal patterns of some elution profiles (such as background signals), the constraints of trilinearity and unimodality were put aside. About the correspondence criterion, although previous MCR/ALS analysis of data sets showed the presence/absence of some constituents in different experiments, the default setting which assumes the presence of all constituents in all samples were used in final MCR-ALS analysis. Also, by taking advantage of column-wise augmentation, the resolved mass spectra in $\mathbf{S}_T$ are enforced to be the same for the common constituents in the various samples, whereas elution profiles resolved in $\mathbf{C}_{aug}$ are flexible to be different (in retention time and peak shapes) in different samples.

Finally, by implementing the mentioned constraints and updating the $\mathbf{C}_{aug}$ and $\mathbf{S}^T$ profiles, the iterative optimization is continued, until the convergence criterion is fulfilled. This criterion is based on the comparison of the lack of fit (LOF) values obtained in two consecutive iterations. The lack of fit values are calculated according to the expression after equation S5-2:

$$Equation\ S\ 5\text{-}2:\ Lack\ of\ fit\ (\%) = \ 100\%\ x\ \sqrt{\frac{\sum ije_{ij}^2}{\sum ijd_{ij}^2}}$$

Another parameter used to indicate the fit quality of the MCR-ALS results is the percentage of explained variance (R2), calculated according to the equation S5-3:

$$Equation\ S\ 5\text{-}3:\ R^2 = 100\ x\ \frac{\sum ijd_{ij}^2 - \sum ije_{ij}^2}{\sum ijd_{ij}^2}$$

*(where each d_{ij} shows each experimental data matrix and each e_{ij} is the residual element of **E** matrix.)*

The number of components in MCR-ALS modelling was set as the number of singular values that larger than those already associated to noise (an estimation of this value can be obtained by inspecting the background parts of the chromatograms where low intensity signals were at the values below SNR_{Thr}) [187]. Then, repeated MCR-ALS analysis using different number of components were conducted and in case of decreasing in the LOF values (Eq. 2) and increasing in R^2 values (Eq. 3), the extra component(s) were included in the models. A complementary way to ensure the proper selection of the number of components is visual inspection of the chromatographic and mass spectra profiles, to avoid modelling LC-MS signal features related to noise in one way and to capture all features related to trace organic micropollutants, in another way.

Also, the quality of the MCR-ALS modelling was checked with the degree of rotational ambiguity in order to know if the solutions provided by the decomposition are practically unique. For this purpose, the MCR-BAND software presented by Jaumot and Tauler was implemented [188].

After MCR-ALS processing of $\mathbf{D_{aug}}$, the area under the resolved chromatographic profile in $\mathbf{C_{aug}}$, were used to obtain the relative quantitative information of different micropollutants in the different samples. Also, the resolved MS1 spectra were used for selecting the most prominent precursor ions necessary for MS2 spectral features. So, a matrix of **P** (I (retention times × 5 (number of precursor ions)) was constructed from each *mzXMLStruct* data sample. Then, the two-way information retrieved from MCR-ALS of LC-MS^1 data was utilized again. Here, the time windows of resolved elution profiles for each prioritized micropollutant in all samples was scanned in matrix **P** to find the common precursor ions (mass accuracy of 10 mDa) with the ones recovered as the five most intense mass fragments of resolved MS^1 through DDA mode.

5.6.4.3. Partial least squares – discriminant analysis (PLS-DA)

PLS-DA [189] is a discriminant method that establishes a multivariate regression model between a matrix of independent variables (**X**, predictor variables) and an array of dependent variables (**y**, predicted variables) that contains binary dummy variables indicating the class to which each sample belongs, where 1 indicates membership and 0 does not. In a multiclass problem, dummy vector of **y** (N×1) is converted into a dummy matrix of **Y** (N×G), containing

dummy codes 1 and 0 (G is the number of classes) [190,191]. In this method, PLS is first used to compute the scores matrix of **X**, i.e. **T**. Then, a LDA model is built by using **T** as input. The resulting model can be visualized in the latent variable space that captures the most variance of **X** that is relevant to **y** or **Y**. In practice, PLS1-DA and PLS2-DA algorithms respectively employ the dummy **y** and **Y** as output variables. In the particular cases studied in the present study, PLS1-DA and PLS2-DA algorithms were used respectively to discriminate

between unspiked and spiked water samples (target screening) and between three classes of influent-O_3, effluent, river water samples (nontarget screening). In this way, peak areas of every sample (**X**) were correlated with the vector describing the sample type class membership (**y** or **Y**). The models were cross-validated by venetian blinds, and the number of LVs was chosen based on the smallest cross-validation classification errors. Also, the statistical information obtained from this PLS-DA model can be used to determine which features (resolved organic micropollutants) of **X** are more important in determining class membership of **y** or **Y**. For this purpose, the variable importance on projection (VIP) scores [176], which are the weighted sum of squared PLS variable weights are computed. VIP scores measure the importance of each predictor variable into the final PLS model. The features with VIP > 1 were commonly considered to be important because the squared sum of all VIP values is equal to the number feature and thus, the average VIP would be equal to 1 [177].

Table S 5-1: Overview of spiked compounds with m/z, retention time and molecular formula.

Compound	Monoisotopic mass	Retention time	Molecular formula
17-a-Methyltestosteron	303.23186	14.8	$C_{20}H_{30}O_2$
1H-Benzotriazole	120.05562	7.9	$C_6H_5N_3$
5,6-Dimethyl-1H-Benzotriazol	148.08692	10.0	$C_8H_9N_3$
5-Chloro-1H-benzotriazole	154.01665	9.8	$C_6H_4ClN_3$
Acetanilide	136.07569	8.3	C_8H_9NO
Atenolol	267.17032	6.2	$C_{14}H_{22}N_2O_3$
Atrazine-desethyl	188.06975	8.9	$C_6H_{10}ClN_5$
Atrazine-desethyl-desisopropyl	146.02280	7.9	$C_3H_4ClN_5$
Atrazine-desisorpopyl	174.05410	10.8	$C_5H_8ClN_5$
Bisphenol S	251.03726	8.5	$C_{12}H_{10}O_4S$
Carbamazepine	237.10224	10.9	$C_{15}H_{12}N_2O$
Chloridazon-desphenyl	146.01157	5.9	$C_4H_4ClN_3O$
Diclofenac	296.02396	15.1	$C_{14}H_{11}Cl_2NO_2$
Caffeine	195.08765	7.5	$C_8H_{10}N_4O_2$
Melamine	127.07267	1.6	$C_3H_6N_6$
Metformin	130.10872	1.6	$C_4H_{11}N_5$
Metoprolol	268.19072	7.2	$C_{15}H_{25}NO_3$
Paracetamol	152.07061	6.8	$C_8H_9NO_2$
Propazine	230.11670	13.1	$C_9H_{16}ClN_5$
Simazine	202.08540	10.4	$C_7H_{12}ClN_5$
Sulfamethoxazole	254.05939	7.7	$C_{10}H_{11}N_3O_3S$
Tamoxifen	372.23219	13.2	$C_{26}H_{29}NO$
Terbutylazin-desethyl	202.08540	10.8	$C_7H_{12}ClN_5$
Toremifene	406.19322	12.7	$C_{26}H_{28}ClNO$

Table S 5-2: Details of the chromatographic separation for method A and B.

	Method A		**Method B**	
Gradient method	*time / min*	*% B*	*time / min*	*% B*
	0 - 2	0	0 - 5	5
	2 - 4	0 to 50	5 - 15	5 to 95
	4 - 17	50 to 98	15 - 25	95
	17 - 22	98	25 - 25.1	95 to 5
	22 -22.1	98 to 0	25.1 - 30	5
	22.1 - 30	0		
Flow rate	0.3 mL/min		0.35 mL/min	
Column	Atlantis T3 (2.1 mm x 150 mm, 3 µm particle size)		XSelect HSS T3 (2.1 mm x 75 mm, 3.5 µm particle size)	

Table S 5-3: HESI source parameter of method A and B.

	Method A	***Method B***
Sheath gas flow rate	37	40
Aux gas flow rate	15	5
Sweep gas flow rate	1	10
Spray Voltage	3.5 kV	4 kV
Capillary Temp.	320° C	350 °C
S-Lens RF level	50	60
Aux gas heater Temp.	50° C	100 °C

Table S 5-4: Settings of the Full MS/dd MS2(Top 5) method for method A and B.

	Method A		***Method B***	
	Full MS	*dd-MS2*	*Full MS*	*dd-MS²*
Resolution	70,000	17,500	70,000	17,500
AGC Target	1E+06	5E+04	1E+06	5E+04
Max. injection time	100 ms	50 ms	50 ms	50 ms
Scan Range	100 – 1000 m/z		100 – 1000 m/z	
Loop Count		5*		5*
Isolation window		1.4 m/z		2 m/z
NCE		30, 60		30
Intensity Threshold		2E+03		2E+03

* Loop Count specifies the number N of fragmentation scans

Table S 5-5: MCR/ALS results of target data set by segmentation of full scan chromatograms. LOF: lack of fit

Chromatographic region	Retention time (min)	No. of components	R^2	LOF (%)
I	1.30-7.02	40	99.34	3.25
II	6.89-11.73	50	99.21	2.43
III	11.61-16.0	60	99.67	3.63

Table S 5-6: MCR/ALS results of non-target data set by segmentation of full scan chromatograms.

Chromatographic region	Retention time (min)	No. of components	R^2	LOF (%)
I	7.82-10.05	21	99.34	5.36
II	9.94-12.25	44	98.81	4.65
III	12.14-14.41	94	98.60	3.12
IV	14.30-16.58	40	99.25	2.42
V	16.47-18.83	22	99.41	5.43
VI	18.72-21.12	16	98.92	2.87
VII	21.01-23.41	6	99.76	3.01
VIII	23.29-25.69	7	99.65	2.97

Table S 5-7: Tentative identification results for some prioritized compounds.

Exact mass	Retention time (min)	Compound name	Theoretical (m/z)	Molecular Formula	Relative mass error (ppm)	Associated class
134.0712	11.8	4-Methyl-1H-benzotriazole	134.07127	C7H7N3	-0.5	Influent-Effluent
178.1334	9.00	Bethanidine	178.13387	C10H15N3	-2.6	Influent
120.0559	10.2	Benzotriazole	120.05562	C6H5N3	2.3	Effluent
182.18989	11.2	Dicyclohexylamine	182.19033	C12H23N	-2.4	Influent
180.081	10.5	Acridine	180.0807	C13H9N	1.2	Influent
279.0918	14.3	Sulfamethazine	279.0910	C12H14N4O2S	3.0	Effluent
223.1439	13.8	Mexacarbate	223.1441	C12H18N2O2	-0.9	Effluent
237.1021	12.2	Carbamazepine	237.1022	C15H12N2O	-0.6	Effluent
291.1469	9.4	Trimethoprim	291.1451	C14H18N4O3	-6.0	Influent-Effluent
268.1908	11.6	Metoprolol	268.1907	C15H25NO3	-0.4	Influent-Effluent
291.2063	12.2	Verapamil metabolite D617	291.2067	C17H26N2O2	-1.4	Effluent
225.1959	14.6	N,N'-Dicyclohexylurea	225.1961	C13H24N2O	-1.1	Effluent

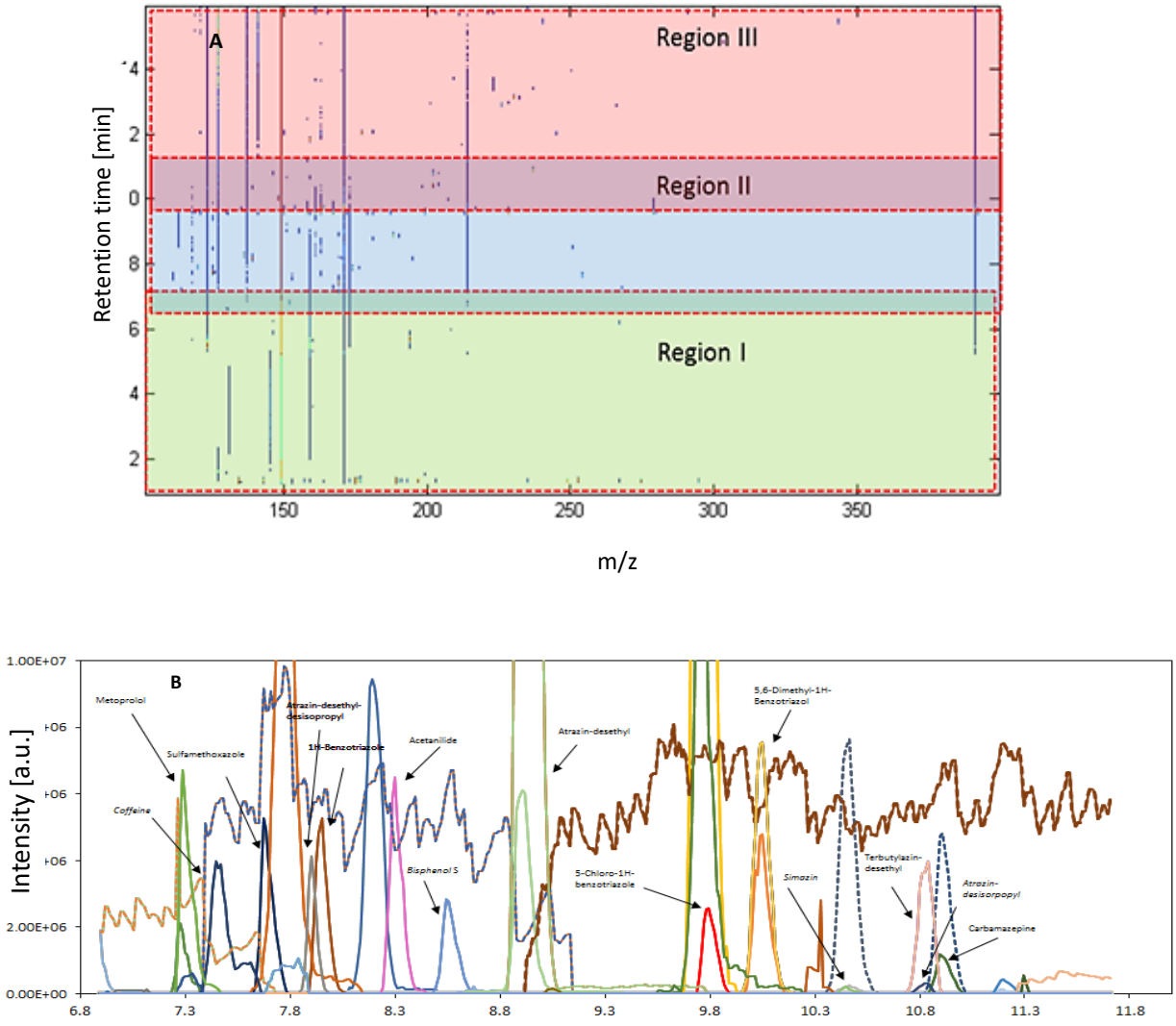


Figure S 5-1: (A) LC-HRMS contour plot obtained in the analysis of a drinking water sample spiked with 500 ng/L of targeted micropollutants, three chromatographic time regions used for MCR/ALS modelling are indicated, (B) Resolved chromatographic profiles for components of region II (6.89-11.73 min). Target compounds shown include metoprolol, caffeine, sulfamethoxazole, atrazin-desethyl-desisopropyl, 1H-benzotriazole, acetanilide, bisphenol S, atrazine-desethyl, 5-chloro-1H-benzotriazole, 5,6-dimethyl-1H-benzotriazole, simazine, terbutylazine-desethyl, atrazine-desisopropyl and carbamazepine.

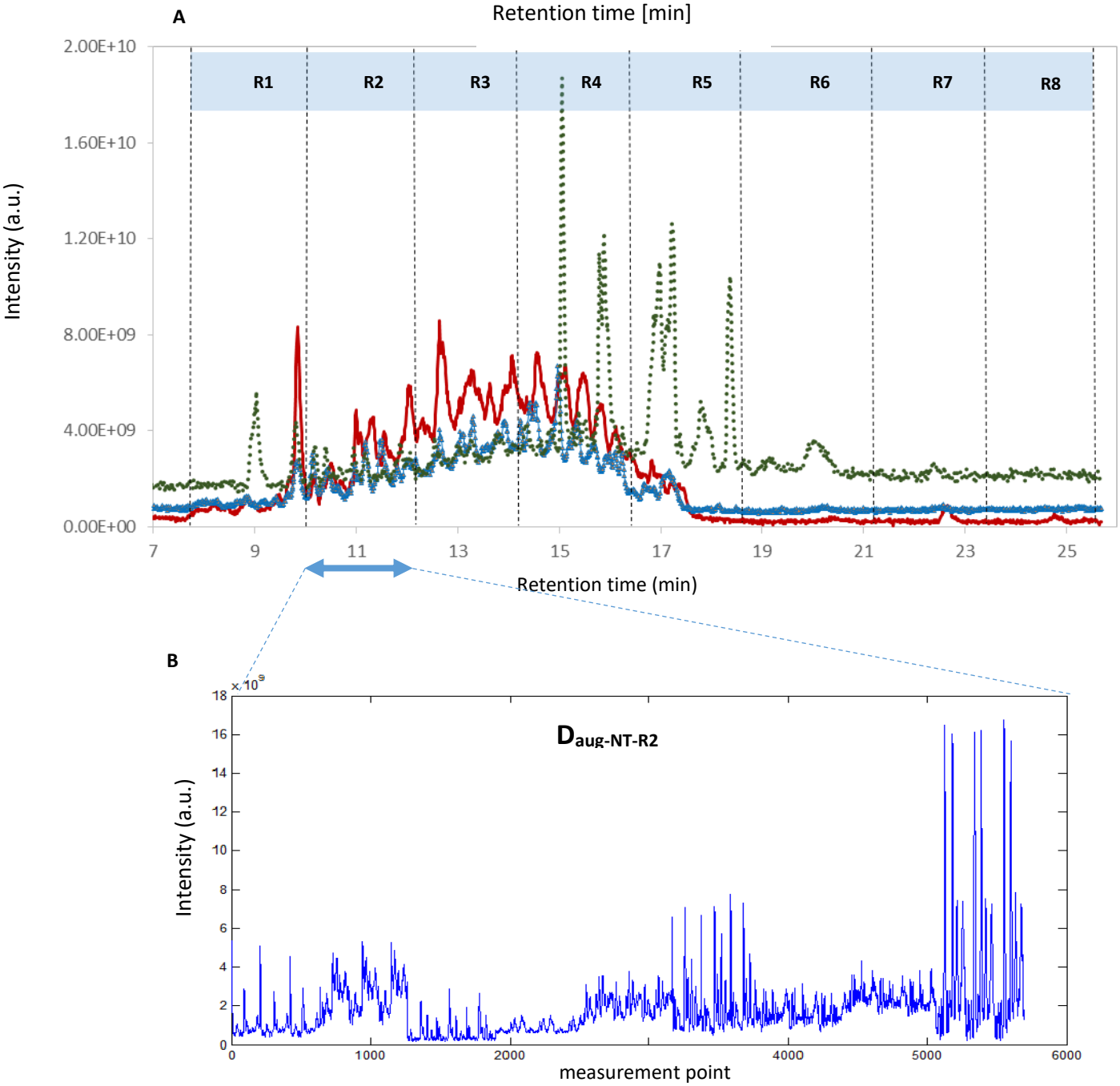


Figure S 5-2: (A) Total current ion chromatogram (TIC) of one Influent-O_3 (solid red line), effluent (dashed blue line) and river water sample (dotted green line), (B) shows 27 augmented chromatographic data for region II (Daug-NT-R2) to be processed with the MCR-ALS method.

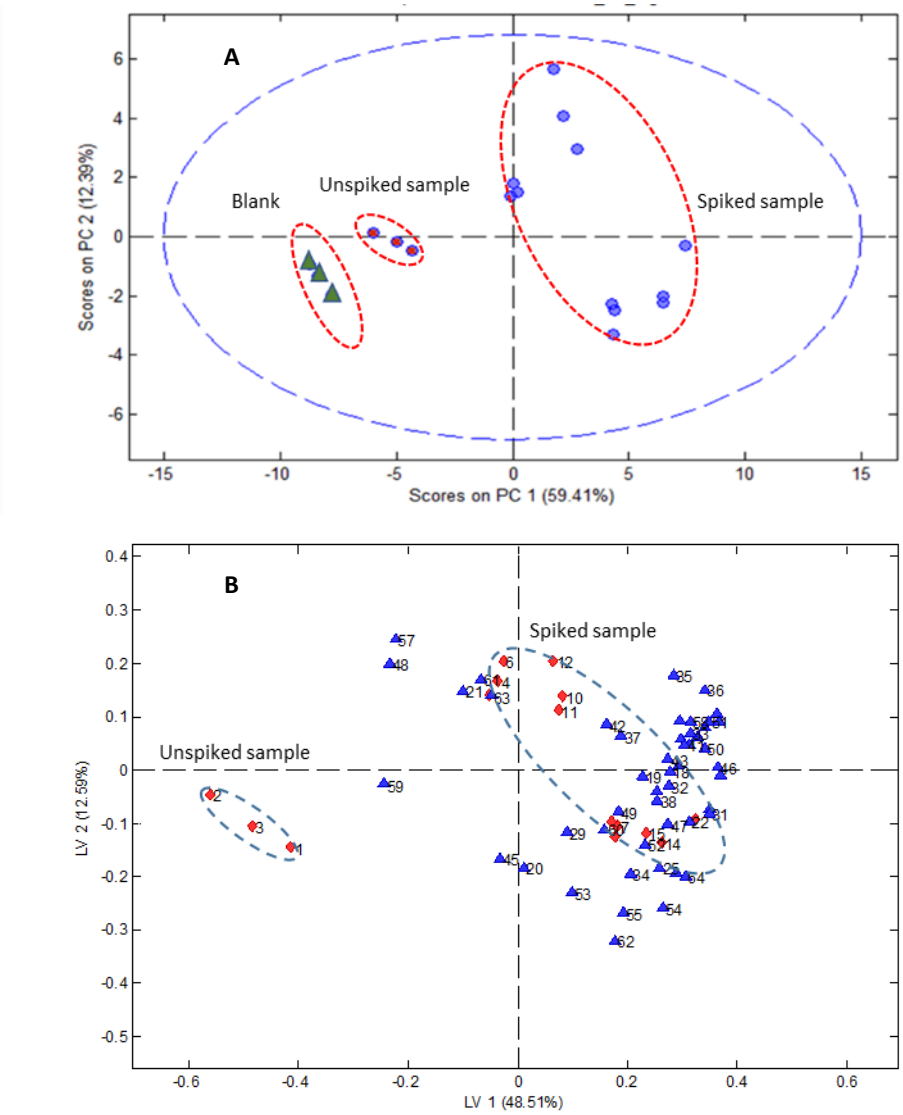


Figure S 5-3: (A) PC1 vs. PC2 scores plot obtained for grouping blanks (green triangles), unspiked (red circles) and spiked (blue circles) drinking water samples, measured in positive ionization mode (B) PLS-DA biplot obtained for target analysis of 15 unspiked (group1, diamonds No. 1-3) and spiked (group2, diamonds No. 4-15) drinking water samples. The superimposed loading values (blue triangles) show the discriminating features mostly in the right side of the plot.

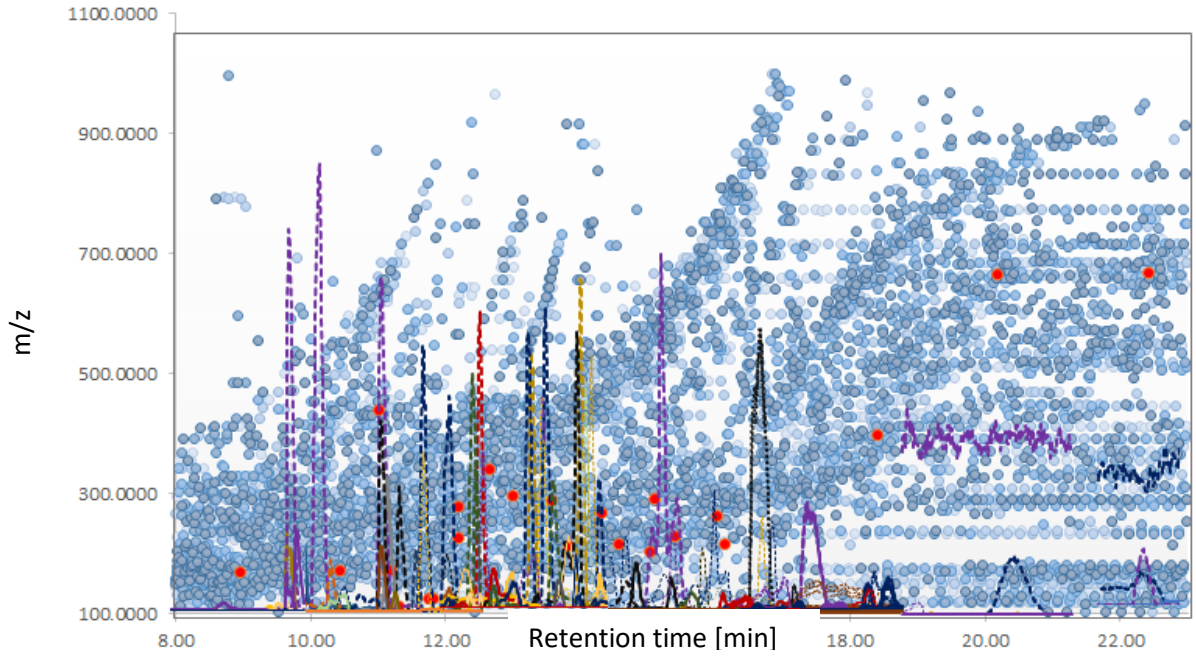


Figure S 5-4: Scatter plot of recorded (blue circles) precursor ions (m/z values) in each retention time for a LC-Q-Orbitrap run of the influent-O3 wastewater sample. The selected precursors are shown by red circles. All resolved chromatograms using MCR-ALS modelling of the current sample are superimposed on the plot (light colour dashed line) to show the accommodation of selected precursors with respect to the pure LC profiles.

Chapter 6 Non target analysis and chemometric evaluation of a passive sampler monitoring of small streams

This chapter was adapted from L.L. Hohrenk-Danzouma, M. Vosough, V.I. Merkus, F. Drees, T.C. Schmidt, Non-target Analysis and Chemometric Evaluation of a Passive Sampler Monitoring of Small Streams, Environmental Science and Technology (2022). https://doi.org/10.1021/acs.est.1c08014

Abstract

Complex multivariate datasets are generated in environmental non-target screening (NTS) studies covering different sampling locations and times. This study presents a comprehensive chemometric based data processing workflow to reveal hidden data patterns and to find a subset of discriminating features between samples. We used ANOVA simultaneous component analysis (ASCA) to disentangle the influence of spatial and seasonal effects as well as their interaction on a multi-class dataset. The dataset was obtained by a Chemcatcher® passive sampler (PS) monitoring campaign of three small streams and one major river over four sampling periods from spring to summer. Monitoring of small streams is important as they are impacted by non-point source introduction of organic micropollutants (OMPs). The use of PS provides a higher representativeness of sampling and NTS broadens the range of detectable OMPs. A comparison of ASCA results of target analysis and NTS showed for both datasets a dominant influence of different sampling locations and individual temporal pollution patterns for each river. With the limited set of target analytes, general seasonal pollution patterns were apparent, but NTS data provide a more holistic view on site-specific pollutant loads. The similarity of temporal pollution patterns of two geographically close small streams was revealed, which was not observed in undecomposed data analysis like PCA. With a complementary partial least squares-discriminant analysis (PLS-DA) and Volcano based prioritization strategy 223 site and 45 season specific features were selected and tentatively identified.

6.1 Introduction

Surface waters are impacted by a large variety of organic micropollutants (OMPs), including pharmaceuticals, pesticides, industrial compounds, and largely unknown transformation products (TPs). They can enter the aquatic environment from both point sources, such as wastewater effluent discharge, and non-point sources, such as agricultural runoff or street runoff [192]. Small streams can be especially vulnerable due to smaller dilution ratios and thus a high risk of peak exposures after heavy rainfalls. In a recent large-scale study agricultural non-point source pesticide pollution was identified as the major driver in reducing vulnerable insect populations in small streams [9]. Since concentrations of OMPs can vary over time, peak exposures might be missed when only grab sampling is applied. Passive samplers can be used to obtain a higher representativeness of sampling. They consist of a receiving phase suitable to accumulate OMPs during a certain time period in which they are exposed to the aquatic environment [193].

Target analysis is limited to preselected compounds and lacks an investigation of unknown or unexpected substances. Non-target analysis enables the screening of a broad range of OMPs at trace concentrations with high resolution mass spectrometry (HRMS). It provides the possibility for spatial or temporal trend analysis without previous restriction on defined analytes and thus has the potential to identify formerly unknown compounds including environmental TPs [100]. Several recent studies show how the combination of passive sampling and HRMS screening approaches can provide a more comprehensive picture of pollution loads in surface waters [20] compared to conventional monitoring with grab samples. Most of the studies are based on suspect screening or spectral library screening approaches [194,195]. Spatial and time trends analysis were employed in combination with compound risk assessment [196,197] and in relation to agricultural application and rainfall events [198]. Univariate statistical tools were applied to investigate stormwater infiltration to groundwater [199] and principal component analysis (PCA) score/loading scatter plot were used for visualization and to select site-specific markers [173]. However, exploiting the potential of advanced chemometric tools was not widely included. Due to the multivariate nature of data originating from HRMS, different state-of-the-art unsupervised and supervised pattern recognition methods can be employed. These approaches can provide deeper insights into pollution patterns or reveal overlooked potentially relevant pollutants as no prior restriction on the scope of included compounds e.g. by suspect lists or spectral library entries is made.

Moreover, the prioritization of relevant OMPs is especially crucial in non-target analysis of environmental samples, as usually feature lists with thousands of entries are obtained. As a complete identification process is extremely labour intensive it is essential to reduce and prioritize compounds of interest. In fact, chemometrics-based strategies have a high potential for an efficient selection of a subset of features that reveal similarities and/or differences

between samples [42], but are not yet widely developed in NTS studies. In some studies, hierarchical cluster analysis (HCA) and PCA were used for data exploration, inspecting sample relationships, or for the identification of TPs [71,200]. A few studies using advanced computationally based prioritization strategies with different prioritization aims have been recently published ([78] [201] [75] [77]) and show that the development of more sophisticated multivariate statistical tools is of utmost importance for a better understanding of the complexity of NTS data. Moreover, utilizing multiple modelling approaches, each derived from a different perspective, can be highly beneficial for efficient analysis of complex HRMS data sets where the number of variables often by far exceeds the number of samples.

In environmental monitoring studies often several factors like different sampling times or sampling sites play a role. As each factor can be individually influenced, e.g. by different point sources, pollution events, local agricultural pesticide application, rainfall events etc., disentangling the contribution of different factors of an experimental setup can thus be helpful to understand seasonal changes or location specific differences. The factors can be included in an ANOVA-simultaneous component analysis (ASCA), which is a more effective approach than PCA as it is able to analyse all factors within one model and separate the sources of variation in the total dataset associated to the different factors and their interactions [88].

In this study, we present a workflow of complementary chemometric approaches which can be applied to both target and non-target datasets obtained from designed studies including several interacting factors and discuss benefits and limitations of the approaches.

Starting by exploratory analysis with PCA and HCA of original matrices, we show how both, spatial and temporal effects as well as their interaction, can be evaluated by ASCA modelling for a multi-class environmental dataset obtained by passive sampling monitoring of three small streams and one major river over four sampling periods. In a subsequent step both univariate statistics and multivariate-based ranking of features through PLS-DA models are used to prioritize discriminating compounds, which from a classification point of view, are the features that have the greatest influence on the class differentiation of the samples. Separate models are built for both different sampling locations (factor 1) and different sampling periods (factor 2) to select and later identify features discriminating (by being present with higher relative intensities in certain samples) for each factor. In summary, we illustrate how information gaps that may arise from sampling, measurement and data evaluation can be closed by chemometric evaluation of HRMS passive sampling datasets.

6.2 Experimental section

Figure 6-1 shows the study design with the two factors sampling site (1) and sampling period (2) in part A and all workflow steps from sampling to chemometric evaluation in part B.

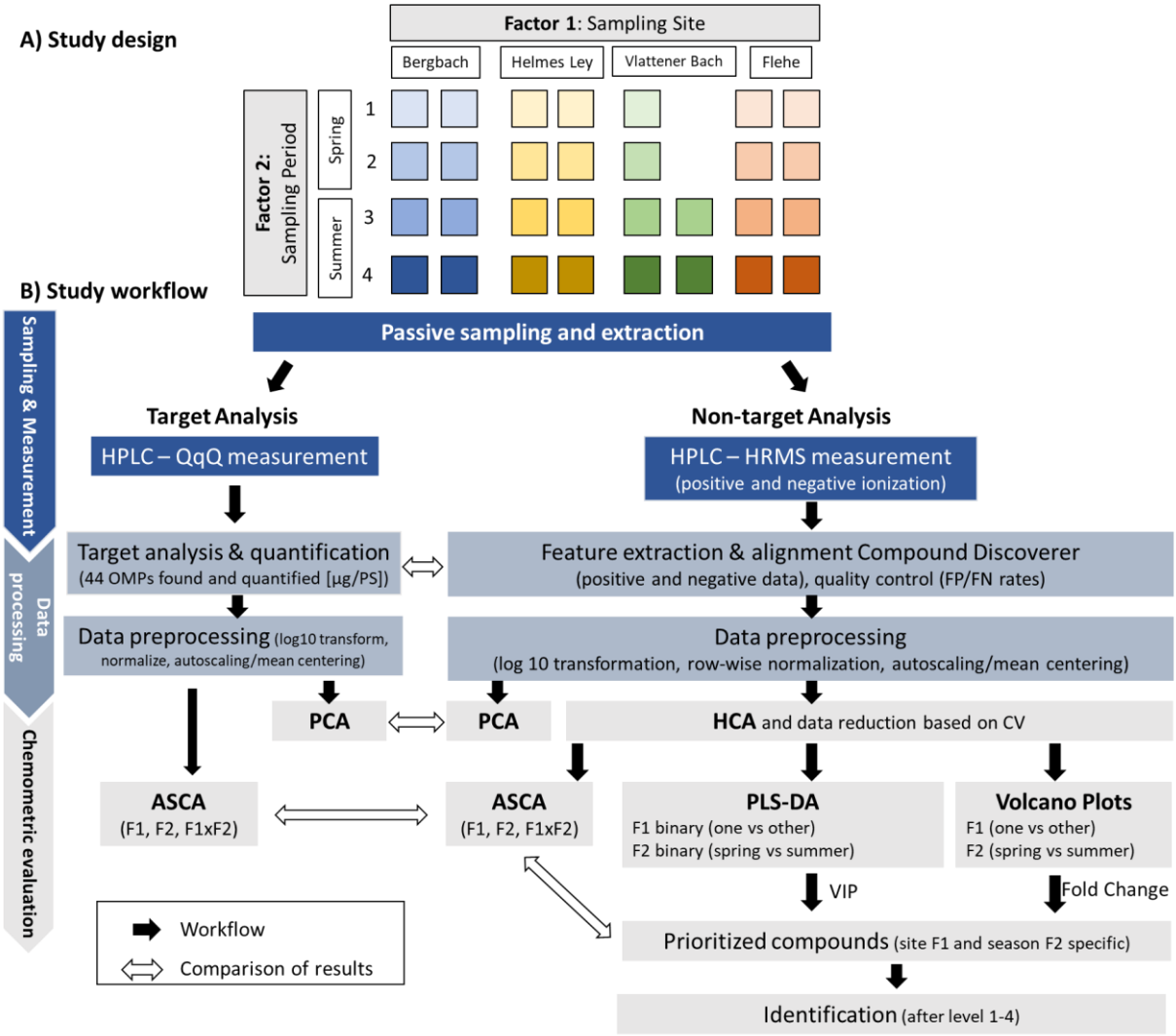


*Figure 6-1: **Part A**: Overview of sampling sites and periods, **Part B:** Study workflow from sampling, measurement to data evaluation. **Abbreviations**: ASCA: ANOVA simultaneous compound analysis, CV: coefficient of variance, F1 factor 1(sampling site), F2: factor 2 (sampling time), HCA: hierarchical cluster analysis, HPLC: high performance liquid chromatography, HRMS: high resolution mass spectrometry, OMP: organic micropollutant, PCA: Principal component analysis, PLS-DA: partial least square discriminant analysis, µg/PS: microgram/passive sampler; QqQ: triple quadrupole mass spectrometer, VIP: variable importance in projection (score)*

6.2.1 Sample collection and extraction

Chemcatcher® (T.E. Laboratories Ltd., Ireland) with a 47 mm diameter housing were used with Empore™ SDB-RPS Disks (3M) and a Supor® 200 PES-membrane filter 0.2 µm as protection. Details on sampler conditioning and extraction are summarized in SI Part 6.8.1.1. Passive samplers were employed in three different small streams Helmes Ley (*Helm*), Bergbach (*Berg*) and Vlattener Bach (*Vlat*) four times for a sampling period of four weeks from May to August 2018 (see table S6-1). All sampling sites had to fulfil the following criteria: (i)

catchment area < 30 km^2, (ii) proportion of intensive agriculture > 40% and (iii) no close point sources [8]. Further, one sampling location at river Rhine (*Flehe*) was used as control, representing a major river in the same time periods, however, with different sampling dates (exposure period 3 weeks). For each sampling point two passive sampling devices were deployed. In addition, a field control was taken for each sampling period and site, which was transported and treated as other samples, exposed to air at sampling sites but not deployed in the water.

6.2.2 Measurement condition for target and non-target analysis

For target analysis 99 analytes (mainly pesticides and related TPs) were quantified as µg/PS according to DIN 38497-36 [202] with HPLC-MS/MS by North Rhine-Westphalia Office of Nature, Environment and Consumer Protection. Before measurement internal standards were added at a concentration of 0.25 µg/L. Non-target analysis was performed with a HRMS full scan measurement on a Orbitrap mass spectrometer (Q Exactive Thermo Scientific, Bremen, Germany). Ionization in positive and negative electrospray ionization mode was used in separate runs and samples were measured in technical triplicates. Detailed description of measurement methods for target and non-target analysis are given SI Part 6.8.1.2 and 6.8.1.3.1.

6.2.3 Chemometrics data treatment workflow

6.2.3.1 Initial data arrangement and preprocessing

The general strategy proposed for chemometric analysis of target and non-target data is shown in figure 6-1B. Initial raw data were processed with Compound Discoverer version 3.1.1.12 for positive and negative data in separate workflows for untargeted environmental studies (details SI Part 6.8.1.3.2). Cleaned up data matrices consisting of 30 samples (rows) and 4661 and 4296 features (columns) for positive and negative ionization modes, respectively, were prepared with mean peak areas of triplicate measurements as matrix entries. For chemometric processing of target data, the mean peak areas for all target analytes that were detected at any sampling point/time (44 out of 99 targets detected during previous target quantification) were collected from Compound Discoverer. As a result, a 30×44 matrix was generated. Following, the generated data matrices were imported into MATLAB environment for further data processing. For data preprocessing, the matrices were first elementary log-scaled to correct the heteroscedasticity in the data, reducing data skewness and leading to normally distributed variables. Then, the log-transformed matrices were row-wise scaled (by total area normalization) to reduce the systematic variations in the experimental process, as a data-driven-based method. In the end, a column-wise mean-centering was applied on each data matrix to focus the fluctuating feature intensities around zero. For comparison purposes, in the

exploratory data analysis section (PCA), an autoscaling step was also carried out in combination with the other mentioned pre-treatment steps (for more information see SI 1.4.1). The same data processing (including feature extraction and preprocessing) was done for target and non-target data for better comparability of results.

6.2.3.2 Data reduction and HCA clustering

In the next step as a da*ta pre-filtering,* the features with a steady peak area profile over all samples, which were not of interest in the scope of this study, were omitted to reduce the matrices to a handier size. Data were reduced based on coefficient of variance (CV%) of peak areas of all samples without distorting the hidden and inherent data structure. Thresholds for CV were incrementally increased and features below it were omitted. Reduced data matrices were analysed by hierarchical cluster analysis (HCA) with Euclidean distance and changes of data structure and sub-cluster distances were assessed. This approach was continued until a remarkable change in the resulting dendrograms was observed. Reduced data matrices consisted of 1426 variables with CV above 161% for positive and 1448 variables with CV above 147% for negative ionization mode, respectively (details are given in SI Part 6.8.2.2.1).

6.2.3.3 Chemometric analysis with PCA, ASCA and PLS-DA

The data matrices of target and non-target analysis were subjected to several multivariate methods. PCA was employed for an initial data evaluation and ASCA for more specific factor-based data evaluations.

In ASCA, the power of ANOVA to separate variance sources is merged with the advantages of simultaneous component analysis (SCA) [88] (details SI Part 6.8.1.4.3). ASCA modelling was performed on preprocessed target data and preprocessed and reduced data matrix for non-target data of positive and negative ionization. To validate ASCA models, permutation tests were used to determine the statistical significance of the observed effects of factors and their interactions. The permutation test involved randomly permuting the original data matrix (1000 permutations) and recalculating the sum of squares due to the factors [203]). Then, the factor-related score matrices obtained from target and non-target datasets were subjected to procrustes analysis [204] to see how closely score data fit each other (SI Part 6.8.1.4.6).

Partial Least Squares Discriminant Analysis (PLS-DA) is a well-known supervised classification method for the prioritization of features discriminating different classes of environmental samples. In the current study, PLS-DA models were built for classification of non-target data in separated model for location (F1) and time (F2) classifying factors. For the location factor four binary models for each sampling site versus all other rivers grouped together were built. For PLS-DA models of time factor, the periods one and two were grouped as “S*pring*” and periods three and four were grouped as “*Summer*” and analysed in a binary model. All PLS-DA models were built using seven-fold venetian blind cross-validation, and the

statistical significance of the PLS-DA models was validated using 100 random permutations to avoid over-fitting. No outliers and misclassified samples were observed in any model (PCA, ASCA or PLS-DA) based on evaluation of Hotelling T^2 values versus the Q residuals.

6.2.3.4 Feature prioritization strategy

In order to prioritize compounds of interest to be subjected to the identification workflow, discriminating variables were selected based on their Variable Importance in Projection (VIP) scores obtained through PLS-DA models. Univariate statistical analysis using Volcano plots was also utilized in a further prioritizing step to compare relative intensity changes between different classes expressed as fold changes (FC) which can be graphically represented together with p-values. Volcano plots were constructed for each sampling site versus group of all other sites for the location factor and for *Spring* versus *Summer* samples for F2. Variables were prioritized for each sampling site and season based on their VIP score, fold change and p-value as follows: Threshold for F1 VIP > 2, Fold change > 5 and p-value < 0.05 and for F2 VIP > 1, Fold change > 2 and p-value < 0.05.

6.2.3.5 Identification of prioritized feature

For the identification of prioritized features several steps in Compound Discoverer including prediction and ranking of molecular formulae, database search ChemSpider and mzCloud were evaluated. In addition, m/z values were searched in FOR IDENT [141]. To confirm identification results, reference standards were measured if available and retention times, m/z and MS^2 spectra were compared. Identification confidence was classified in adaption to Schymanski et al. [205] (Table S6-9).

6.2.3.6 Software

Volcano plots, HCA, procrustes and the heatmap clustergram function is included in the Bioinformatic Toolbox™ and Statistics and Machine Learning Toolbox™ of MATLAB (The Mathworks Inc, version 9.9, 2020b). PCA, ASCA (plus version for unbalanced design) and PLS-DA modelling was performed in PLS Toolbox version 8.9 (Eigenvector Research Inc., Wenatchee, WA, U.S.A.) implemented in MATLAB environment.

6.3 Results and discussion

6.3.1 Target screening: PCA and ASCA modeling

As an initial data exploration, PCA was applied on the target data matrix. The score plot in figure 6-2 shows a distinction of all four sampling sites with 3 PCs, however a high within dispersion pattern caused by seasonal trends of different sampling times is evident. Then, in order to achieve a better understanding of the data variations and their corresponding features, the study design was incorporated in an ASCA modelling Here, each SCA sub-model represents only a fraction of the variance between the total observations that would be otherwise masked by other sources of variation in PCA exploration of original undecomposed data. The results of the ASCA model for the matrix of target data considering sum-of-squares correction for unbalanced data [89] is shown in table 6-1. From the ANOVA point of view, these results show that the overall effect of all factors (F1 sampling location, F2 sampling time and their interaction, F1 x F2) are significant (p-values<0.001) confirmed by permutation test. In ASCA for each SCA sub-model, score and loading plots are obtained in a similar way to PCA corresponding to the different factors and their interaction.

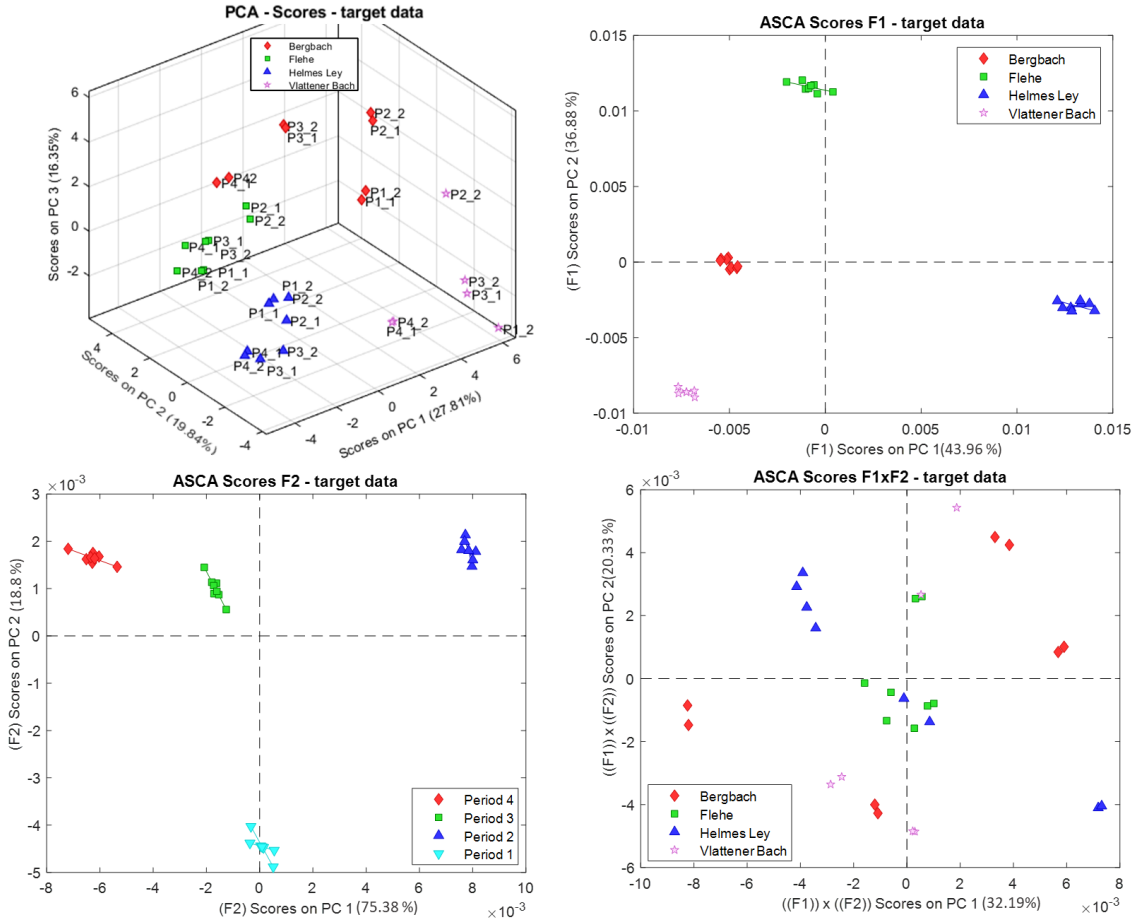


Figure 6-2: PCA score plot including 3 PCs (upper left) coloured by sampling location labelled with P1-P4 indicating sampling periods, ASCA score plots of factor 1 location (upper right) and factor 2 time (bottom left) and ASCA score plot of interaction of factor 1 and 2 (bottom right) including 2 PCs and coloured by sampling location of target data (30 samples, 44 variables, mean peak areas).

ASCA score plots (figure 6-2) show a clear separation of all four sampling sites for factor 1 using two PCs and thus more clearly than with PCA. In addition, a separation of all four sampling periods was possible for factor 2. Along the first PC a separation between sampling site *Helm* and the other locations for factor 1 and a clear separation for the second sampling period with respect to the others for factor 2 is apparent. Considering the interaction effects (figure 6-2, bottom right), a considerable differential temporal trend for the scores in each location is present, which is especially distinctive for small rivers *Helm, Berg* and *Vlat* in comparison to *Flehe* (Rhine) which accumulates in the center of the plot. Previous research has shown that small streams are mainly influenced by discontinuous diffusive sources [8] in contrast to big streams mainly impacted by almost continuous OMP introduction by WWTPs [7]. These observations can be confirmed by our chemometric data evaluation showing highly differing temporal trends. Furthermore, the targets that are more responsible for differentiating factor levels can be determined using the loadings plots in the established ASCA model (figure S6-7). Examples on loadings interpretation are given in the SI part 6.8.2.1.

6.3.2 Non-target screening

6.3.2.1 Exploratory data analysis (PCA and HCA)

In a first step, a regular PCA and HCA as initial unsupervised exploration tools were applied to preprocessed non reduced data matrices in both ionization modes. Figure 6-3 shows excellent separation between all four sampling sites obtained with PCA.

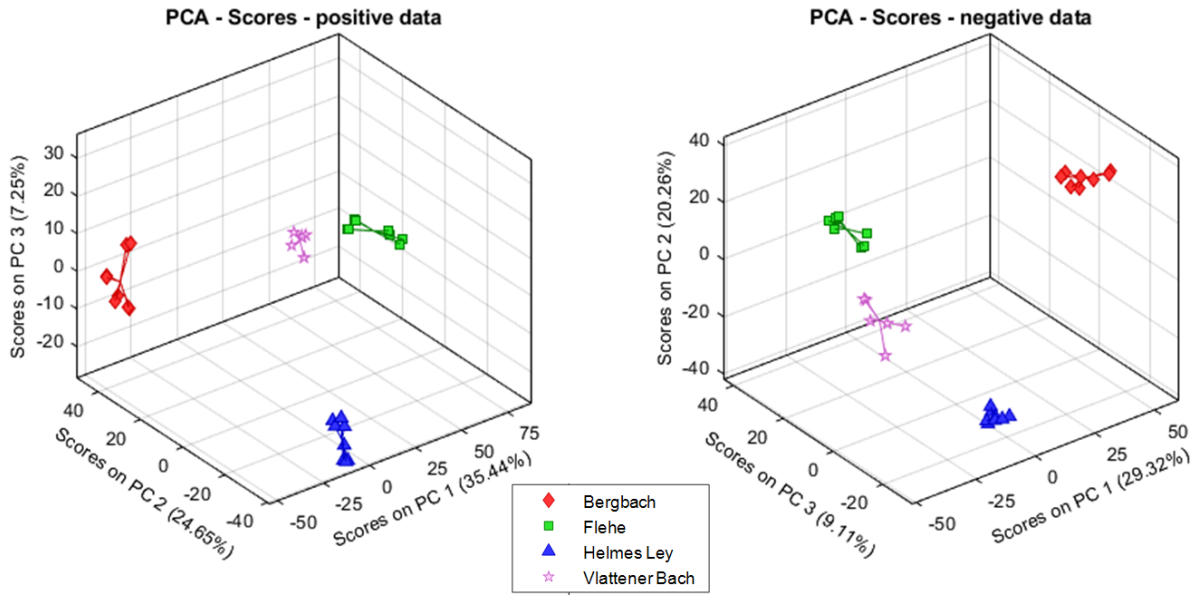


Figure 6-3: PCA score plots including 3 PCs for positive ionization (left) and negative ionization (right) non-target data, colorized by sampling location.

Three PCs are required to separate all four sampling site groups and within group variations (due to time effects) are much less pronounced, compared with target data (figure 6-2, top left). PCA score plots, as well as HCA analysis displayed in figures S6-9 and S6-10 show that

Flehe samples are the most dissimilar class to the other sampling sites, followed by the *Berg* samples. The two sample groups *Helm* and *Vlat* are the most similar classes to each other. The mentioned PCA and HCA trends hold true for the data obtained in both ionization modes of measurements. The lower group dispersion for non-target data comparing PCA scores plots to target data (figure 6-2) can be explained by the fact that target data include only a selective category of pollutants (mainly pesticides/herbicides and TPs). The non-target data, on the other hand, provide a more holistic view of the data dispersion pattern. In a next step, data matrices were reduced based on CV values as described in SI part 6.8.2.2.1. HCA clustergrams in figures S6-9 and S6-10 show that data inherent structure was not disturbed by data reduction and the following analyses were made with reduced matrices.

6.3.2.2 ASCA results (main and interactive factor effects)

In order to disentangle the contribution of temporal and spatial OMP profiles and to investigate the possible interaction of both factors, ASCA modelling was applied to the reduced and preprocessed data matrices. Here, for the first time, the potential of ASCA for high-dimensional environmental LC-HRMS data (with >1400 feature profiles for each ionization mode) was investigated in depth. Table 6-1 shows the results of ASCA models, considering sum-of-squares corrections for unbalanced data. The effect of both factors and their interaction were significant (p-values <0.01) for both positive and negative data matrices utilizing permutation tests.

ASCA score plots of factor location, time and their interaction are given in figure 6-4, figure S6-12 and in figure S6-13 their corresponding loading plots are displayed. The effect of each factor on the variation in the dataset is quantitatively given in Table 6-1.

Table 6-1: ASCA results for target and non-target (positive and negative ionization) dataset: Significance and partitioning of the total variance into the individual terms corresponding to factors and interaction.

	Factor	Cum EV[a]	Percentage of variation[b]	Significance (p-value)
Target analysis	Location[c]	14.34	64.21	<0.001
	Time[d]	3.40	14.87	<0.001
	Location×Time[e]	4.31	19.47	<0.001
	Residuals	-	1.46	-
Non-target analysis Positive	Location[c]	50.75	68.71	<0.001
	Time[d]	7.70	10.03	<0.001
	Location×Time[e]	13.18	17.80	<0.001
	Residuals	-	3.47	-
Non-target analysis Negative	Location	49.18	67.15	<0.001
	Time	6.05	8.07	<0.001
	Location×Time	14.01	19.28	<0.001
	Residuals	-	5.50	-

[a] Cumulative Eigenvalue (×1e7).
[b] Percentage of variation expressed as sum of squared deviations from the overall mean.
[c] Sampling location factor with four levels: sampling places of Berg, Fleh, Helm and Vlat rivers.
[d] Sampling time factor with four levels: sampling periods of 1, 2 ,3 and 4.
[e] Sampling location and time interaction.

As can be seen, modelling of data obtained by positive and negative ionization showed a very consistent picture, the between-location variations with roughly 70% are the dominant part of the variations, larger than the between-time or interaction effects. Figure S6-14 showing a superimposed graph of all scores plots of different factors also visualizes the domination of factor 1 as it surrounds factor 2 and factor 1x2. This leads to the conclusion that composition of OMPs for the different small streams is highly individual and exceeds temporal variations. The lower contribution of seasonal variations (factor 2) of non-target data compared to target data, that was already apparent in dispersion patterns of PCA score plots, is also reflected by percentages of variation of factor 2 in ASCA (~15 % compared to 10 and 8 %). Nevertheless, apart from the different extent to which seasonal changes contribute to the total variance, very similar temporal patterns were obtained with target and non-target analysis as shown by procrustes analysis of F2 score plots of both approaches described in SI Part 6.8.2.2.4. This demonstrates that even with the limited number of target compounds a general seasonal trend can be illustrated, which endorses the use of chemometric methods on target data. In contrast, the spatial patterns differ from target to non-target datasets which indicates that target data do not comprise all location relevant substances and non-target analysis is necessary to draw a comprehensive picture of pollution.

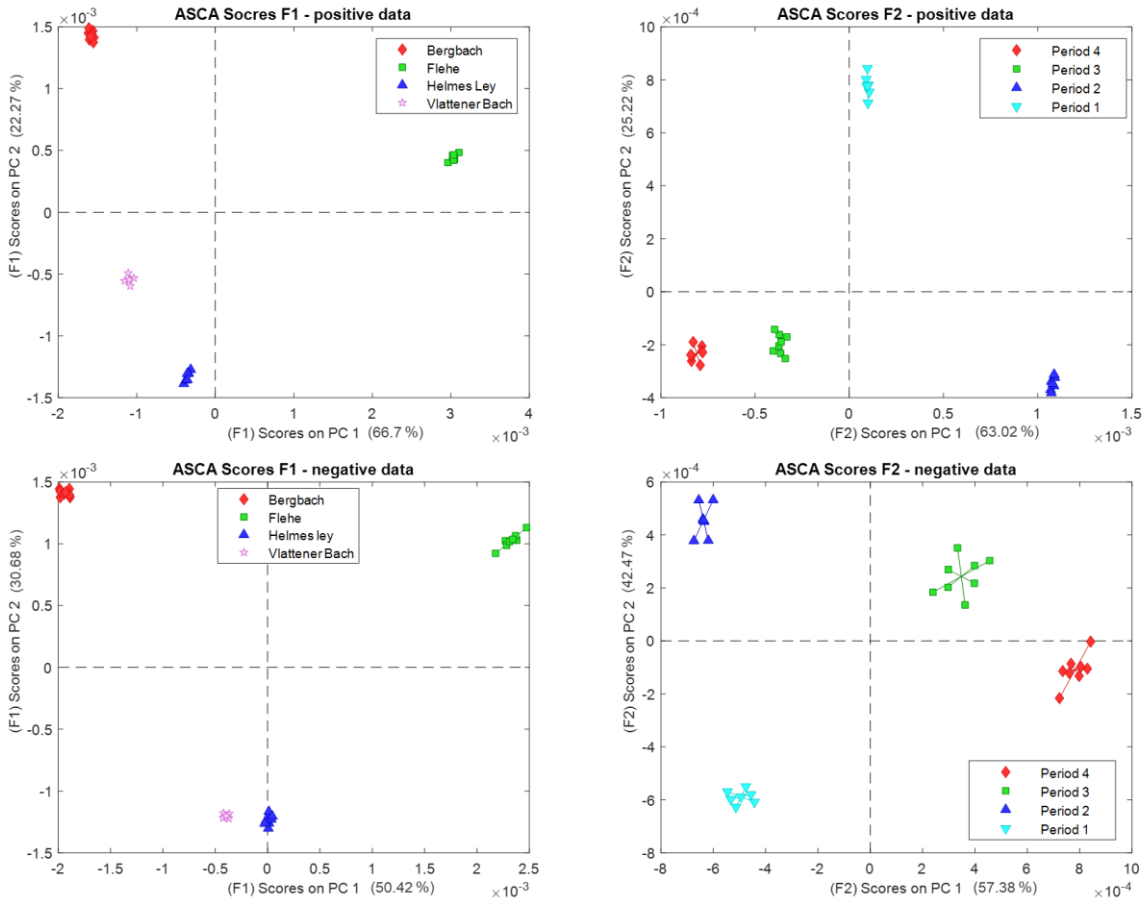


Figure 6-4: ASCA Score plots including 2 PCs of positive (upper plots) and negative (bottom plots) non-target data for factor 1 location (left plots) and factor 2 time (right plots).

The effect of sampling location can now be visualized using two PCs and thus again more clearly than with PCA as was the case for target data. In the score plots for the factor 'location', sampling site *Flehe* shows the greatest difference from the other three sampling sites. This trend is in accordance with findings of PCA, HCA and it also meets our expectations as *Flehe* was the only sampling site at a large river (Rhine). Sampling site *Berg* is separated from the other two sites in the direction of PC2 and the two small rivers *Helm* and *Vlat* show the highest similarity among all sampling sites, especially for data in negative mode. Reasons for higher deviations of *Berg* samples compared to other sampling sites located at small streams where not completely revealed but could be explained by discharges of intensive livestock farming close by. Looking at the time factor, the score plots show that between 'sampling period' variations are more distinct for sampling period 1 and sampling period 2, relative to the two other sampling periods.

In addition, table 6-1 shows that a significant fraction of variation is related to the interaction of the factors location and time, so both effects were not independent from each other. The radar plots of the first two PCs for the interaction effects (figure S6-15) show that among all four sampling sites, *Berg* and *Vlat* show a more similar temporal trend than the other sampling

locations. The similarity of *Berg* and *Vlat* regarding the time factor, not to be confused with similarity of *Helm* and *Vlat* regarding location factor alone, can be explained by the fact that both streams are geographically close by (around 1 km) and thus are impacted by weather conditions like rainfall events or agricultural pesticide application in a similar way. However, figure S6-15 shows that besides *Berg* and *Vlat* the different sampling locations reveal predominantly dissimilar temporal trends. Therefore, different OMPs are relevant to characterize the individual temporal variation of each sampling site compared with isolated evaluation of time factor. This can be visualized with the help of the loading plots in figure S6-13 which reflect the different orientations of the variable groups associated with each experimental factor. Especially for the loading diagrams for the interaction factor a remarkable star shape can be noted. However, the analysis of loadings can become complicated when dealing with datasets with a large number of variables and potentially noisy measurements, such as the data often obtained in non-target environmental screening. Nevertheless, features with high absolute loading or high leverage in a specific PC are those that follow the behaviour described by this component and are mostly affected by each factor. In a two-dimensional PC space, these are the most peripheral variables which surround the bulk of variables.

As discussed above the sampling sites *Berg* and *Vlat* showed a similar temporal pattern examining the interaction of time and local trends. Thus, it could be of interest to identify the OMPs responsible for this pattern. For this purpose, several outstanding loadings for F1xF2 PC1 were exemplary selected to illustrate benefits of ASCA interaction factor loadings interpretation. Figure S6-16 shows several selected variables with high loading values on F1xF2 PC1 and thus discriminative properties for *Berg* and *Vlat* period 2. In figure S6-16 C mean peak areas over all samples are shown that confirm high abundances of the selected exemplary features in the mentioned samples. One of them (f9) was tentatively identified as pethoxamid, a herbicide. Another coeluting m/z (f42) is probably an in-source fragment of pethoxamid and a third one (f125) shows a similar intensity profile but was not identified. Moving toward the center of the loading plot, the contribution of other corresponding subsets of features will increase. For example, f46 also shows a similar temporal pattern for *Berg* and *Vlat* but was also detected with lower intensities in other samples. The same deduction can be made for other variables located at elongated edges in each direction of the plot which show higher concentration for certain sampling locations and times in similar positions of the ASCA score plots. For example, f30 and f36 were specific for *Helm* sampling period 3 and 4 or f17 and f9 were found in high concentrations in *Flehe* samples for period 4 for negative ionization data (see figure S6-17). This shows, how the ASCA model can be used to better understand the interacted influence of two factors which would not be possible by an isolated investigation of both factors.

6.3.2.3 PLS-DA, statistical evaluation and feature prioritization

Considering the statistical significance of factors 'sampling site' and 'sampling period', PLS-DA was used to further narrow down which variables (features) were responsible for the observed differences. For factor 1, perfect modelling and classification were obtained for all binary class models including two PLS latent variables with the range of 56.6 to 63.1% for positive and 42.3% to 55.5% for negative data. PLS-DA results show that the pollution profiling of each sampling site is significantly different from the second group which is the collection of all other sites, which confirms ASCA results. Unlike for the factor 1, none of the binary and multi-class models constructed by sampling time were significant using the permutation test (at the 95% of confidence level). So, another binary model was constructed by joining samples of periods 1 and 2 as new class 1 (*Spring*) and sampling periods 3 and 4 as new class 2 (*Summer*). All quality measures of different models are summarized in table S6-8. Following PLS-DA model construction, the VIP values for 1426 and 1448 variables were calculated for positive and negative data for each PLS-DA model.

After the construction of PLS-DA models and calculation of VIP scores values a complementary univariate statistical test using a Volcano plot was performed to discern the individual effect of each feature for distinction between each pair of the mentioned classes, considering both fold-change and t-test criteria. Figure S6-21 and S6-22 show Venn diagrams of overlaps of prioritized features meeting the Volcano criteria (FC, p-Value) and PLS-DA-VIP criteria for all different sampling locations and seasons for positive and negative data. As can be seen, for some sampling sites VIP criteria were limiting the number of prioritized features, for others the Volcano criteria reduced this number. Figure 6-5 summarizes the final numbers of prioritized features of each sampling site, sampling season *Spring* and S*ummer* for both ionization modes.

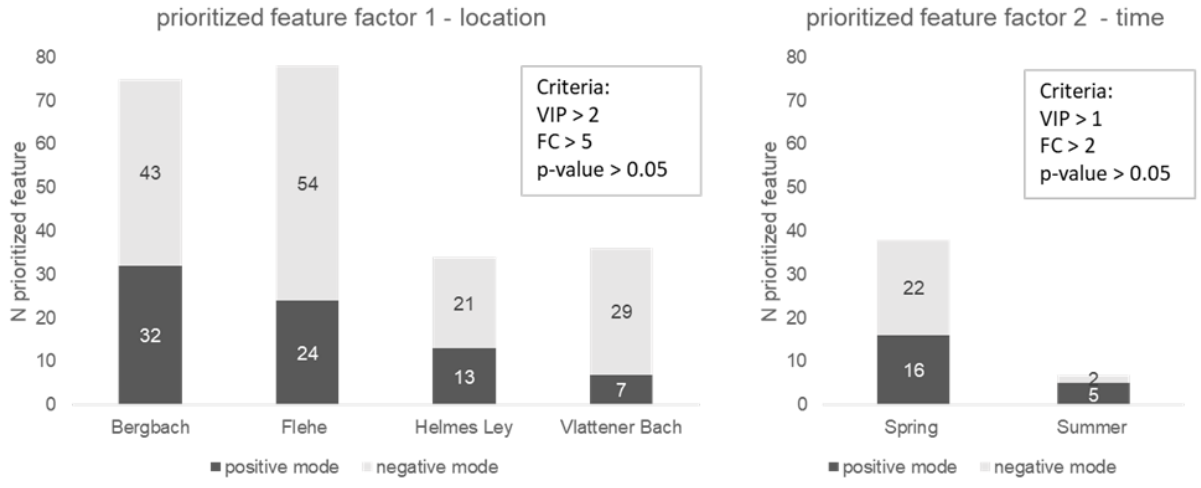


Figure 6-5: Overview of the number of prioritized features with stated criteria for four sampling sites (left) and two sampling seasons (right)

As can be seen the two sample classes *Berg* and *Flehe* with highest distinction to other classes in different chemometric models turned up with the highest number of prioritized features. Furthermore, *Spring* samples showed a higher number of prioritized features compared to *Summer* which might be explained by higher pesticide applications in this period. Unexpectedly, more features were prioritized for data in negative ionization mode compared to positive mode, which is surprising as many NTS monitoring studies include only measurements in positive mode [20,76,195,197]. This emphasizes that measurements in negative ionization mode should not be neglected in NTS studies.

Correct assignment of prioritized features to different sample classes was additionally verified by simultaneous hierarchical clustering of samples and prioritized features using heat map clustering plots as described in SI part 6.8.2.3.1. The selected features show excellent discrimination properties for four different sampling locations and two different sampling seasons as shown in figure S6-24 A-D which suggests that the correct subset of feature was selected and confirms results of supervised classification methods with an unsupervised approach.

6.3.2.4 Comparison of ASCA, PLS-DA and Volcano plots

Comparing the prioritized features for each sampling location and time by PLS-DA-VIP and Volcano plots with results of ASCA modelling can be used as validation of both types of approaches and additional confirmation of prioritized features. Figure 6-6 presents the position and orientation of the ASCA loadings with prioritized feature of combined PLS-DA-VIP and Volcano plot approach indicated by different colours for all models in both ionization modes.

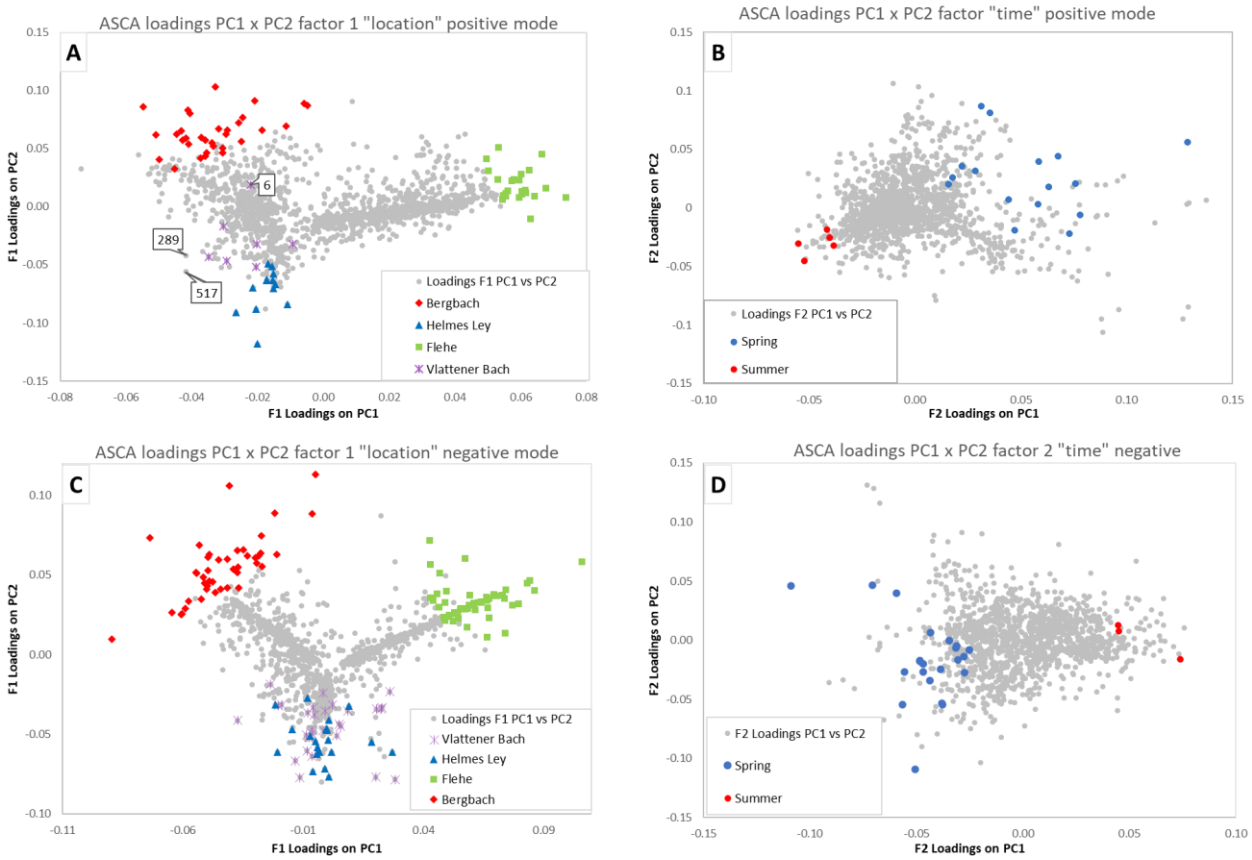


Figure 6-6: ASCA loading plots with prioritized features indicated in colour for different sampling locations and seasons for positive (A&B) and negative (C&D) data.

From figure 6-6 some important points can be derived: The features selected by PLS-DA-VIP-Volcano are located mostly in the extreme parts of the ASCA loading plots and their position in 2D space is related to the different sampling site and sampling season derived from ASCA scores plots (figure 6-4), confirming its results. For factor 1 an assignment of features to sampling sites *Berg* and *Flehe* was highly congruent with all different approaches. However, for the two sampling sites *Helm* and *Vlat,* showing a weaker class separation in ASCA and consequently no clear distinction in loading plots, a more precise assignment of location specific OMPs was possible by the complementary approach. For example, f6 (which was also identified as target isoproturon) located at the center of the ASCA loading plot was identified as *Vlat* related compound by the PLS-DA-VIP-Volcano approach. In addition, not all extreme loadings in ASCA are specific to just one location, e.g., f517 and f289 in the down left corner of figure 6-6 A close to the *Vlat*-specific area were not selected by PLS-DA-VIP. Figure S6-23 shows that they were found with elevated peak areas in *Vlat* samples but also occurred in *Berg* and *Helm* samples. In conclusion, the PLS-DA-VIP-Volcano approach shows a higher confident in selecting one-class specific compounds and ASCA has the advantage of revealing compounds commonly found in different sample classes (with differential intensities) or influenced by several factors. Therefore, it can be confirmed that the methods ASCA and PLS-

DA and Volcano plots in a non-target multi-class environmental dataset complement each other in a perfect manner.

6.3.3 Identification of prioritized pollutants and environmental impact

Figure 6-7A shows the level of identification confidence for all prioritized features. As shown, almost one third was identified with level 1 or 2A/B (27%) but another third (28%) remains without any hint of their identity at level 4 (level description in SI table 6-9).

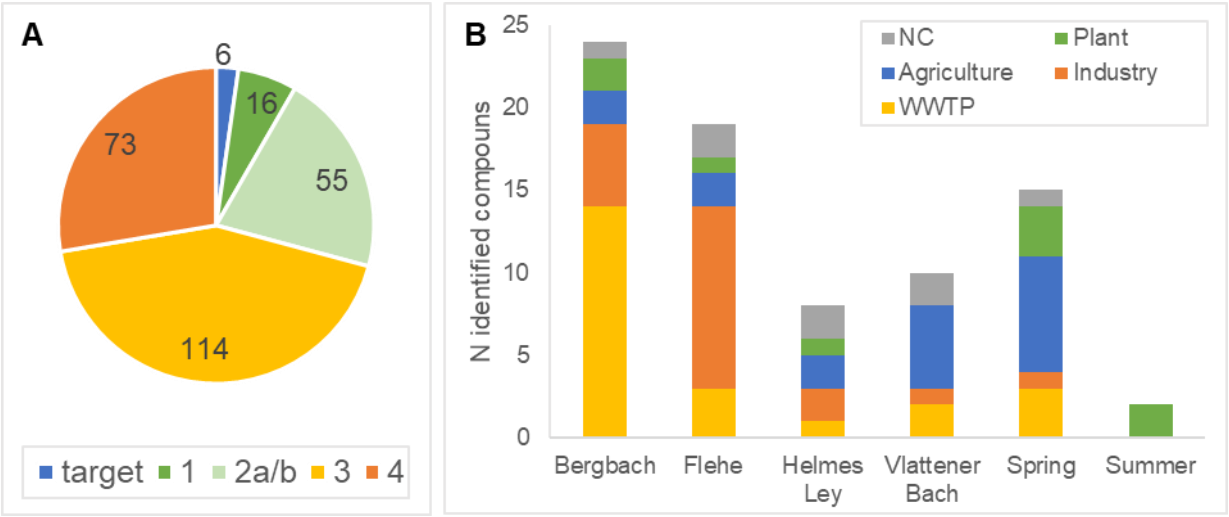


Figure 6-7: A: Level of identification confidence of prioritized features for all classes and both ionization modes. B: Summary of pollutant categories of prioritized features identified at level 1 and 2. Categories include: Agriculture: herbicides, pesticides, fungicides and related TPs, Industry: educts for synthesis, REACH chemicals, NC: no category, Plant: naturally occurring substances, phytotoxins, flavonoids, WWTP: Pharmaceuticals, related TPs, surfactants, PFAS, personal care products.

For the compounds identified with level 1 and 2 different categories of probable origin (agriculture, WWTP effluent, industry, natural occurring) were assigned and are summarized in figure 6-7B. Details on categorization are given in SI Part 6.8.2.4.2 but it has to be taken into account that not all included features were unambiguously identified and emission pathways were not verified. For sampling site *Berg* several known pharmaceuticals and related TPs were identified (e.g. carbamazepine, oxcarbamazepine, amisulpride, tiapride-N-oxide, valsartan and ibersartan), which was in contrast to other sampling sites located at small streams. This is an interesting observation especially in comparison to *Vlat,* as both sampling sites are less than one kilometre apart, and the influence of a municipal WWTP was excluded. Sampling site *Berg* is even closer to the origin of the stream than *Vlat* (~9 km vs ~22 km) but is possibly influenced due to the discharges of an intensive livestock farming close by. Nevertheless, similar temporal influences, e.g., by pesticide application, stormwater runoff etc. was evident and discussed in Part 4.2.2. For *Helm* and *Vlat* mainly agriculture related compounds and their TPs were identified. Besides well-known herbicides like isoproturon or diuron, which were already included in target analysis, several TPs like monodesmethylisoproturon and 3-(2,4-

dichlorophenyl)-6-fluoro-quinazoline-2,4(1H,3H)-dione (TP fluquinconazole) [63] were found. In contrast, for the sampling site located at river Rhine (*Flehe*) a range of different pollutants from various classes but primary industry associated compounds were found. For the two sampling periods in *Spring* mainly agriculture related compounds such as quinmerac, quizalofop or bixafen and TPs like desamino metamitron were identified. The trend of higher pesticide concentration in *Spring* matches with findings from target analysis. In addition, two *Spring* and one *Summer* specific compound were tentatively identified as imperatorin, hederagenin and erucifoline N-oxide, naturally occurring phytotoxins previously reported by [206]. As identification is still limited by information available in databases, many prioritized features could not be sufficiently identified.

6.4 Implications and limitations for non-target environmental screening

We present a non-target chemometric based prioritization workflow for comprehensive investigation of spatial and temporal trends to underline the benefits of implementing multiple advanced chemometric methods for NTS data evaluation. In this study, the use of a variety of graphical tools to interpret different multivariate statistical results of NTS data in a clear way is emphasized.

Unsupervised data exploration tools like PCA and HCA are helpful to obtain a general overview on data structure in NTS studies. However, by comparing its results to ASCA analysis, it becomes evident how additional information on hidden data patterns can be revealed by disentangling different sources of variation, when influences of interacting factors need to be investigated. The similarity of temporal pollution patterns of two geographically close small streams, which was not observed in undecomposed data analysis like PCA, was shown by comparison of ASCA interaction factor F1xF2 scores. In addition, we propose comparing target and non-target data sets using the effect matrices obtained after ASCA multi-factorial matrix decomposition. To this end, each pair of effect matrices obtained from each data set can be subjected to a Procrustes analysis to reveal their factor-based common trends and correlations. A similar temporal trend was observed for the limited number of targets compared to non-target data. However, the comparison further showed that non-target data provide a more holistic picture on site-specific pollutant loads, which underlines the relevance of broad untargeted OMP screening approaches. However, while with ASCA a significant amount of information can be retrieved with a small number of samples due to the connection with study design, the high-dimensionality of data sets such as those we deal with in NTS (including thousands of features) makes feature loading interpretations more challenging. Thus, we show how the data processing workflow can be complemented with feature prioritization based on PLS-DA and Volcano plots for a comprehensive examination of a non-target multi-class environmental dataset. The subsequent tentative identification of prioritized features revealed a so far unexpected occurrence of pharmaceuticals at one sampling site.

Nevertheless, the NTS data can provide qualitative information only and the scope of the applied sampling approach with passive samplers is limited and can display only a part the of pollution loads in the environment. Furthermore, identification of OMPs is highly dependent on available information in databases and thus the detection of unknowns such as TPs or unexpected compounds is unfortunately strongly limited to this day. We showed that with an impartial chemometric based prioritization workflow without any a priori limitation by library or database entries a high number of completely unknown features appear as relevant. This emphasizes the urge of including non-target screening approaches in environmental monitoring and putting more efforts to extend databases and develop new techniques for

unknown identification. With this study we demonstrate how the information obtained from different complementary supervised and unsupervised chemometric approaches can help the environmental scientist to gain a deeper insight into collected data and obtain more reliable and consistent results than using one method alone. Each method contributes unique aspects and differing perspectives on the same environmental question, by that coping with increasingly complex NTS datasets. The presented workflow can be transferred to many other environmental datasets and different research questions including combined spatial and temporal investigations. We want to encourage a more intensive use of chemometric based prioritization workflows in future monitoring studies to complete the field of OMP screening methods.

6.7 Acknowledgement

We thank North Rhine-Westphalia Office of Nature, Environment and Consumer Protection for providing samples, data of target analysis and valuable discussion of results. This study was performed within the joint project FUTURE WATER funded by the Ministry of Culture and Science of the state of North-Rhine Westphalia (NRW), Germany

6.8. Supporting Information

6.8.1 Experimental section

6.8.1.1Sampling

Chemcatcher® (T.E. Laboratories Ltd., Ireland) with a 47 mm diameter housing were used with Empore™ SDB-RPS Disks (3M) and a Supor® 200 PES-membrane filter 0.2 µm as protection. SDB-RPS -Disks were conditioned for 30 min with methanol and for 30 min with ultra-pure water on a shaker at 100 rpm. The PES-membrane was also conditioned with methanol and ultra-pure water for 30 min without shaking After sampling, discs were stored in 7 mL acetone at -18°C until measurement.

After acclimatization, discs in storing solution (7 mL acetone) were extracted on the shaker for 30 min. Extraction of discs was repeated with 7 mL methanol. Storing tubes were rinsed twice with 5 mL methanol. All aliquots were merged in a volumetric flask which was filled up to 25 mL. For analysis 200 µL of extract were evaporated and reconstituted in 1 mL of ultrapure water.

Small streams in agricultural areas are categorized as especially vulnerable to high loads of pesticide pollution. Criteria for the selection of sampling sites are discussed in detail in a study of the German environmental agency [207] and were also applied in other studies [8] [9]. Details on sampling periods and sampling sites are given in table S6-1 and figure 6-S1. The influence of a municipal wastewater treatment plant (WWTP) to the sampling sites was checked and excluded in *Elwas Web* (https://www.elwasweb.nrw.de/elwas-web/index.jsf). In addition, a sampling site at river Rhine (Flehe) was used as “control site” representing the opposite of a small agricultural stream as it is a large river influenced by many different point sources. Sampling site selection, sampling, sample preparation and target analysis was performed by North Rhine-Westphalia Office of Nature, Environment and Consumer Protection.

Table S 6-1: Details on sampling campaign

Sampling period	Date	Helmes Ley	Bergbach	Vlattener Bach	Date Flehe	Flehe
1	03.05.-24.05.2018	H01/021	B01/02	V01	11.04.-09.05.2018	F25/26
2	24.05.-13.06.2018	H03/04	B03/04	V03*	09.05.-21.06.2018	F29/F30
3	13.06.-03.07.2018	H05/06	B05/06	V05/06	21.06.-20.07.2018	F31/F32
4	3.07.-24.07.2018	H07/08	B07/08	V07/08	20.07.-27.08.2018	F33/F34

Sampling Site	Helmes Ley	Bergbach	Vlattener Bach	Flehe
Latitude	51.575335	50.654203	50.666680	51.18608152778109
Longitude	6.355987	6.634771	6.553860	6.785835442019569
Distance to origin (km)	6.0	7.9	21.8	584.7
Site characteristics	Small agricultural stream. No influence of municipal WWTP. Influence of small sewage plants, storm water overflow and agricultural runoff possible			River Rhine (right side) sampling station in Düsseldorf. Influenced by several WWTPs, industrial discharge, commercial shipping etc.

* V02 and V04 were covered in sediment and had to be discarded

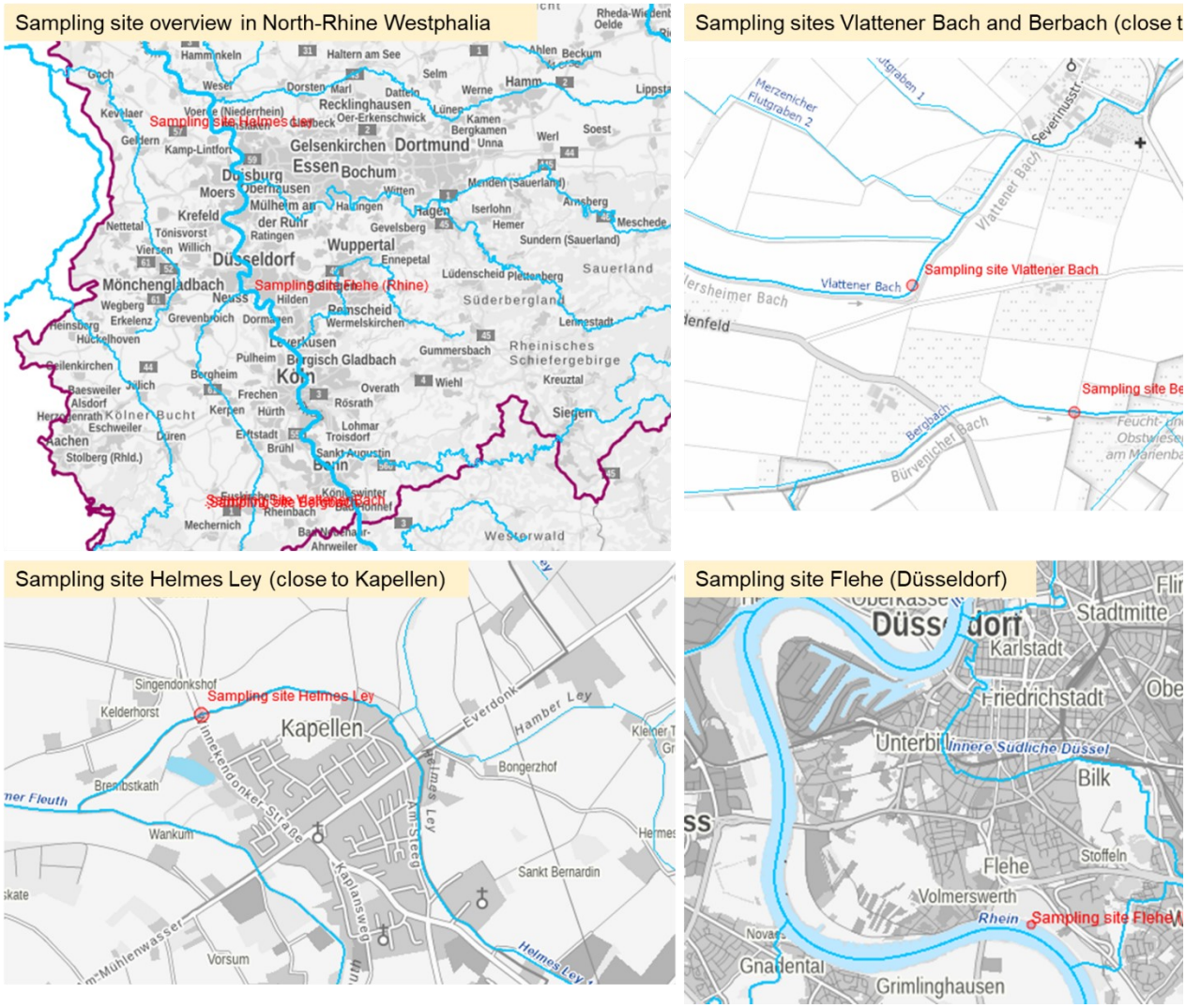


Figure S 6-1: Sections from maps obtained from Elwas web showing the location of different sampling sites in North-Rhine Westphalia, Germany

6.8.1.2 Target analysis

Chromatographic separation was performed on a HPLC-1200 SL Agilent Technologies Deutschland, Böblingen, with a Synergi 2.5 u Hydro-RP 100 A (100 mm * 3 mm) column. Eluent A consisted of water and eluent B of methanol, both with 0.0025 mol/L ammonium formiate and 0.005 % formic acid. A gradient elution method as given in table S6-2 was employed at 0.3 mL/min. The injection volume was 100 µL and column temperature 40°C. Analytes were detected with an TSQ-Quantum Ultra Thermo Scientific, Waltham USA. For each analyte two precursor and product ion mass transitions were measured. For quantification only one mass transition was evaluated. Quantification was performed according to DIN 38407-36:2014-09 [202]. For each sequence a calibration with at least 3 calibration points in a range of 0.005 µg/L – 0.5 µg/L, 0.01 µg/L – 1.0 µg/L and for some TPs and ethofumesat 0.05 µg/L - 5.0 µg/L was included. Concentrations were calculated as accumulated mass per passive sampler disk (µg/PS) considering the volume of the extract and the dilution factor. For the compounds with corresponding internal standard, a further internal

calibration was performed. An overview of all targets and internal standards is given in table S6-3.

Table S 6-2: Gradient elution program target analysis

time	% B
0	5
20	100
25	100
25.5	5
35	5

Table S 6-3: Overview of all targets and internal standards measured in target analysis.

Name	Formula	CAS-Nr.	Precursor Ion	Product Ion 1	Product Ion 2
2,6-Dichlorbenzamid	$C_7H_5Cl_2NO$	2008-58-4	190.0	173.0	175.0
Acetamiprid	$C_{10}H_{11}ClN_4$	135410-20-7	223.0	126.0	90.0
Acetochlor	$C_{14}H_{20}ClNO_2$	34256-82-1	223.1	132.1	
Alachlor	$C_{14}H_{20}ClNO_2$	15972-60-8	270.0	162.2	132.1
Ametryn	$C_9H_{17}N_5S$	834-12-8	228.0	186.1	68.0
Amidosulfuron	$C_9H_{15}N_5O_7S_2$	120923-37-7	370.0	261.0	218.0
Anthranilic acid isopropylamide	$C_{10}H_{14}N_2O$	30391-89-0	179.0	120.0	92.0
Atrazine	$C_8H_{14}ClN_5$	1912-24-9	216.0	174.1	67.9
Atrazine-desethyl	$C_6H_{10}ClN_5$	6190-65-4	188.0	146.0	104.0
Atrazine-desisopropyl	$C_5H_8ClN_5$	1007-28-9	174.0	67.9	104.0
Azoxystrobin	$C_{22}H_{17}N_3O_5$	131860-33-8	404.0	372.1	344.1
Boscalid	$C_{18}H_{12}Cl_2N_2O$	88425-85-6	343.0	307.1	140.0
Bromacil	$C_9H_{13}BrN_2O_2$	314-40-9	260.9	205.0	187.9
Carbamazepine	$C_{15}H_{12}N_2O$	298-46-4	237.0	194.2	193.1
Carbendazim	$C_9H_9N_3O_2$	10605-21-7	192.0	160.1	132.0
Carbetamide	$C_{12}H_{16}N_2O_3$	16118-49-3	236.9	192.1	72.0
Chlorfenvinfos	$C_{12}H_{14}Cl_3O_4P$	470-90-6	358.8	155.3	99.0
Chloridazon	$C_{10}H_8ClN_3O$	1698-60-8	222.0	77.0	65.0
Chloridazon-desphenyl	$C_4H_4ClN_3O$	6339-19-1	146.0	117.0	101.0
Chloridazon-methyldesphenyl	$C_5H_6ClN_3O$	17254-80-7	160.0	88.0	116.9
Chlorpyrifos	$C_9H_{11}Cl_3NO_3PS$	2921-88-2	349.8	96.9	198.0
Chlortoluron	$C_{10}H_{13}ClN_2O$	15545-48-9	213.0	72.0	71.9
Climbazol	$C_{15}H_{17}ClN_2O_2$	38083-17-9	293.0	197.1	225.0
Clomazone	$C_{12}H_{14}ClNO_2$	81777-89-1	240.0	125.0	127.0
Clothianidin	$C_6H_8ClN_5O_2S$	210880-92-5	250.0	169.1	132.0
Cyproconazol	$C_{15}H_{18}ClN_3O$	94361-06-5	292.1	70.0	125.0
DEET	$C_{12}H_{17}NO$	134-62-3	192.1	119.0	91.0
Desmetryn	$C_8H_{15}N_5S$	1014-69-3	214.0	172.1	82.0
Diflubenzuron	$C_{14}H_9ClFN_2O_2$	35367-38-5	310.9	158.0	141.0
Diflufenican	$C_{19}H_{11}F_5N_2O_2$	83164-33-4	395.0	266.1	246.0
Dimefuron	$C_{15}H_{19}ClN_4O_3$	34205-21-5	338.9	71.9	166.9

Name	Formula	CAS-Nr.	Precursor Ion	Product Ion 1	Product Ion 2
Dimethachlor-oxalamic acid	$C_{13}H_{17}NO_4$	1086384-49-7	252.1	220.1	132.1
Dimethachlor-ethane sulfonic acid	$C_{13}H_{18}NO_5S$	1231710-75-0	302.0	270.0	132.0
Dimethenamid	$C_{12}H_{18}ClNO_2S$	87674-68-8	276.0	244.0	168.1
Dimethoat	$C_5H_{12}NO_3PS_2$	60-51-5	230.0	198.9	125.0
Dimoxystrobin	$C_{19}H_{22}N_2O_3$	149961-52-4	327.0	205.0	116.0
Diphenylsulfon	$C_{12}H_{10}O_2S$	127-63-9	218.9	76.9	140.9
Diuron	$C_9H_{10}Cl_2N_2O$	330-54-1	232.9	72.0	72.0
Dimethylsulfamid	$C_2H_8N_2O_2S$	3984-14-3	201.0	92.0	65.0
Epoxiconazol	$C_{17}H_{13}ClFN_3O$	106325-08-1	330.0	121.1	101.0
Ethidimuron	$C_7H_{12}N_4O_3S_2$	30043-49-3	265.0	207.9	114.0
Ethofumesat	$C_{13}H_{18}O_5S$	26225-79-6	287.0	121.1	259.1
Fenamidone	$C_{17}H_{17}N_3OS$	161326-34-7	312.1	92.0	236.1
Fenpropimorph	$C_{20}H_{33}NO$	67564-91-4	304.1	147.1	117.0
Florasulam	$C_{12}H_8F_3N_5O_3S$	145701-23-1	359.9	129.0	
Flufenacet	$C_{14}H_{13}F_4N_3O_2S$	142549-58-3	364.1	152.0	194.1
Flufenacet-ethane sulfonic acid	$C_{11}H_{14}FNO_4$	947601-87-8	276.0	234.0	112.0
Flurtamone	$C_{18}H_{14}F_3NO_2$	96525-23-4	334.0	247.1	178.1
Hexazinon	$C_{12}H_{20}N_4O_2$	51235-04-2	253.1	171.1	71.1
Imidacloprid	$C_9H_{10}ClN_5O_2$	13826-41-3	256.0	209.1	175.1
Irgarol M1	$C_8H_{15}N_5S$	30125-65-6	214.0	158.1	68.0
Irgarol	$C_{11}H_{19}N_5S$	28159-98-0	254.1	198.1	83.0
iso-Chloridazon	$C_{10}H_8ClN_3O$	1698-61-9	222.0	104.0	77.0
Isoproturon	$C_{12}H_{18}N_2O$	34123-59-6	207.1	72.0	165.1
Lenacil	$C_{13}H_{18}N_2O_2$	21640-08-1	235.0	153.1	136.1
Linuron	$C_9H_{10}Cl_2N_2O_2$	330-55-2	248.9	160.0	182.1
Metalaxyl	$C_{15}H_{21}NO_4$	57837-19-1	280.1	220.2	192.1
Metalaxyl-CA	$C_{14}H_{19}NO_4$	75596-99-5	266.1	192.1	160.1
Metamitron	$C_{10}H_{10}N_4O$	41394-05-2	203.0	175.1	104.0
Metazachlor	$C_{14}H_{16}ClN_3O$	67129-08-2	278.0	134.1	210.1
Metazachlor-ESA	$C_{14}H_{17}N_3O_4S$	67129-08-2	324.1	134.2	256.1
Metconazol	$C_{17}H_{22}ClN_3O$	125116-23-6	320.1	70.0	125.0
Methabenzthiazuron	$C_{10}H_{11}N_3OS$	18691-97-9	222.0	165.1	150.1
Methiocarb	$C_{11}H_{15}NO_2S$	2032-65-7	226.1	169.1	121.1
Metobromuron	$C_9H_{11}BrN_2O_2$	3060-89-7	258.9	170.0	148.1
Metolachlor	$C_{15}H_{22}ClNO_2$	51218-45-2	284.1	252.2	176.2
Metolachlor OA	$C_{15}H_{21}NO_4$	152019-73-3	280.1	248.2	
Metolachlor-ESA	$C_{15}H_{22}NO_5S$	171118-09-5	330.0	298.1	202.1
Metoxuron	$C_{10}H_{13}ClN_2O_2$	19937-59-8	229.0	72.0	156.0
Metribuzin	$C_8H_{14}N_4OS$	21087-64-9	215.0	84.0	187.1
Monolinuron	$C_9H_{11}ClN_2O_2$	1746-81-2	215.0	126.0	99.0
Napropramide	$C_{17}H_{21}NO_2$	15299-99-7	272.0	171.1	129.1
Nitenpyram	$C_{11}H_{15}ClN_4O_2$	150824-47-8	271.0	225.2	126.0
Norflurazon	$C_{12}H_9ClF_3N_3O$	27314-13-2	304.0	284.0	286.0
Omethoate	$C_5H_{12}NO_4PS$	1113-02-6	214.0	124.9	182.8

Name	Formula	CAS-Nr.	Precursor Ion	Product Ion 1	Product Ion 2
Penconazol	$C_{13}H_{15}Cl_2N_3$	66246-88-6	284.0	70.0	159.0
Pencycuron	$C_{19}H_{21}ClN_2O$	66063-05-6	329.0	125.0	127.0
Pendimethalin	$C_{13}H_{19}N_3O_4$	40487-42-1	282.1	212.1	194.0
Picoxystrobin	$C_{18}H_{16}F_3NO_4$	117428-22-5	367.8	145.1	205.1
Pirimicarb	$C_{11}H_{18}N_4O_2$	23103-98-2	239.1	72.0	182.1
Prochloraz	$C_{15}H_{16}Cl_3N_3O_2$	67747-09-5	376.1	308.0	310.0
Prometryn	$C_{10}H_{19}N_5S$	7287-19-6	242.1	158.0	200.1
Propazin	$C_9H_{16}ClN_5$	139-40-2	230.0	146.0	104.0
Propiconazol	$C_{15}H_{17}Cl_2N_3O_2$	60207-90-1	342.0	159.0	161.0
Propyzamid	$C_{12}H_{11}Cl_2NO$	23950-58-5	256.0	190.0	172.9
Prosulfocarb	$C_{14}H_{21}NOS$	52888-80-9	252.1	91.0	65.0
Prothioconazol-desthio	$C_{14}H_{15}Cl_2N_3O$	120983-64-4	312.3	70.1	125.0
Pyraclostrobin	$C_{19}H_{18}ClN_3O_4$	175013-18-0	388.1	194.1	163.1
Quinoxyfen	$C_{15}H_8Cl_2FNO$	124495-18-7	307.9	197.0	162.0
Simazin	$C_7H_{12}ClN_5$	122-34-9	202.0	132.1	104.0
Tebuconazol	$C_{16}H_{22}ClN_3O$	107534-96-3	308.1	70.0	70.0
Tebutam	$C_{15}H_{23}NO$	35256-85-0	234.1	91.0	65.0
Terbumeton	$C_{10}H_{19}N_5O$	33693-04-8			
Terbuthylazin	$C_9H_{16}ClN_5$	5915-41-3	230.0	174.1	176.1
Terbuthylazin-desethyl	$C_7H_{12}ClN_5$	30125-63-4	202.0	146.0	148.1
Terbutryn	$C_{10}H_{19}N_5S$	886-50-0	242.1	186.1	68.0
Thiabendazole	$C_{10}H_7N_3S$	148-79-8	202.0	175.0	131.1
Thiacloprid	$C_{10}H_9ClN_4S$	111988-49-9	253.0	126.0	90.0
Thiacloprid-SA	$C_{10}H_{13}ClN_4O_5S$	-	337.2	126.0	320.2
Thiamethoxam	$C_8H_{10}ClN_5O_3S$	153719-23-4	292.1	211.2	181.1
	internal standards				
Atrazine 13C3	$C_8H_{14}ClN_5$	-	219.0	177.1	
Atrazine-desethyl d6	$C_6H_{10}ClN_5$	-	194.0	147.0	
Atrazine-desisopropyl d5	$C_5H_8ClN_5$	-	179.0	69.0	
Chloridazon-desphenyl 15N2	$C_4H_4ClN_3O$	-	148.0	117.0	
Chlorpyrifos d10	$C_9HCl_3D_{10}NO_3PS$	-	362.0	98.8	
Chlortoluron d6	$C_{10}D_6H_7ClN_2O$		219.0	78.1	
Diuron d6	$C_9D_6H_4Cl_2N_2O$	-	239.0	78.1	
Isoproturon d6	$C_{12}D_6H_{12}N_2O$	-	213.1	78.2	
Metolachlor d6	$C_{15}D_6H_{16}ClNO_2$	-	290.1	258.1	
Pendimethalin d5	$C_{10}D_5H_8ClN_2O$	-	287.1	213.1	
Simazin d5	$C_7D_5H_7ClN_5$	-	207.0	129.1	
Tebuconazol d6	$C_{16}D_6H_{16}ClN_3O$	-	214.1	72.0	
Terbuthylazin d5	$C_9D_5H_{11}ClN_5$	-	235.1	179.1	

6.8.1.3 Non-target analysis

6.8.1.3.1 Sample measurement

Chromatographic separation was performed using a Dionex UltiMate 3000 HPLC system (Thermo Scientific, Bremen, Germany). Gradient elution was carried out on a XSelect HSS T3 (2.1 mm x 100 mm, 2.5 µm particle size) column and XSelect HSS T3 XP VanGuard (2.1 mm x 5 mm, 2.5 µm particle size) precolumn from Waters (Milford, MA, USA). The mobile phase consisted of eluent A: ultrapure water + 0.1 % formic acid, and eluent B: methanol + 0.1 % formic acid. A gradient elution method as given in table S6-4 was employed at a flow rate 0.3 mL/min. 100 µL of sample were injected and all samples were measured in technical triplicates for positive and negative ionization mode. Detection was performed on an Orbitrap mass spectrometer (QExactive). The full scan HRMS spectra (*m/z* 80-1000) with a resolution of 70,000 was followed by five data dependent MS^2 scans of the most intense ions with a resolution of 17,500 for each scan. Details of electrospray conditions and measuring parameters are given in table S6-5 and S6-6. For identification samples were remeasured with m/z of prioritized compounds in an inclusion list to generate MS^2 spectra.

Table S 6-4: Elution gradient method for non-target analysis.

time	**% B**
0	0
2	0
4	50
17	98
22	98
22.1	0
30	0

*T*able S 6-5: Ionization parameter of heated electrospray ionization (HESI) for both ionization modes.

Parameter	**Positive**	**Negative**
Sheath gas flow rate	37	40
Aux gas flow rate	15	15
Sweep gas flow rate	1	0
Spray Voltage (k.V)	3.5	3.0
Capillary Temp.(°C)	300	360
S-Lens RF level	50	50
Aux gas heater Temp. (°C)	50	360

Table S 6-6: Parameter of acquisition method Full MS/dd-MS2 for both ionization modes.

Parameter	Positive	Negative
General		
Polarity	Positive	Negative
Runtime	28 min	28 min
Chromatographic peak width	6 s	6 s
Full MS		
Resolution	70,000	70,000
AGC Target	1E+06	1E+06
Maximum injection time	100 ms	50 ms
Scan Range	80 – 1000 m/z	80 – 1000 m/z
dd-MS^2		
Resolution	17,500	17,500
AGC Target	5E+04	5E+04
Maximum IT	50 ms	50 ms
Loop Count	5*	5*
Isolation window	1.4 m/z	1.4 m/z
NCE	30,60,80	30,60,80
Intensity Threshold	1.6E+04	1.6E+04
Minimum AGC target	8.0E+02	8.0E+02
Dynamic exclusion	3.0 s	3.0 s

6.8.1.3.2 Data processing Compound Discoverer

Non-target data extraction was performed with Compound Discoverer version 3.1.1.12 for positive and negative data in separate workflows for untargeted environmental studies as shown in figure S6-2. In short, features were detected and aligned over several files. If several adducts were formed, they were grouped into so called “Compounds”, background signals were marked with sample/blank ratio of 10, gaps were filled with noise, elemental composition of compounds was predicted and MS^1 were automatically searched in Chemspider and MS^2 in mzCloud, respectively. Field controls were used as blank samples. All settings of processing parameter are summarized in table S6-7.

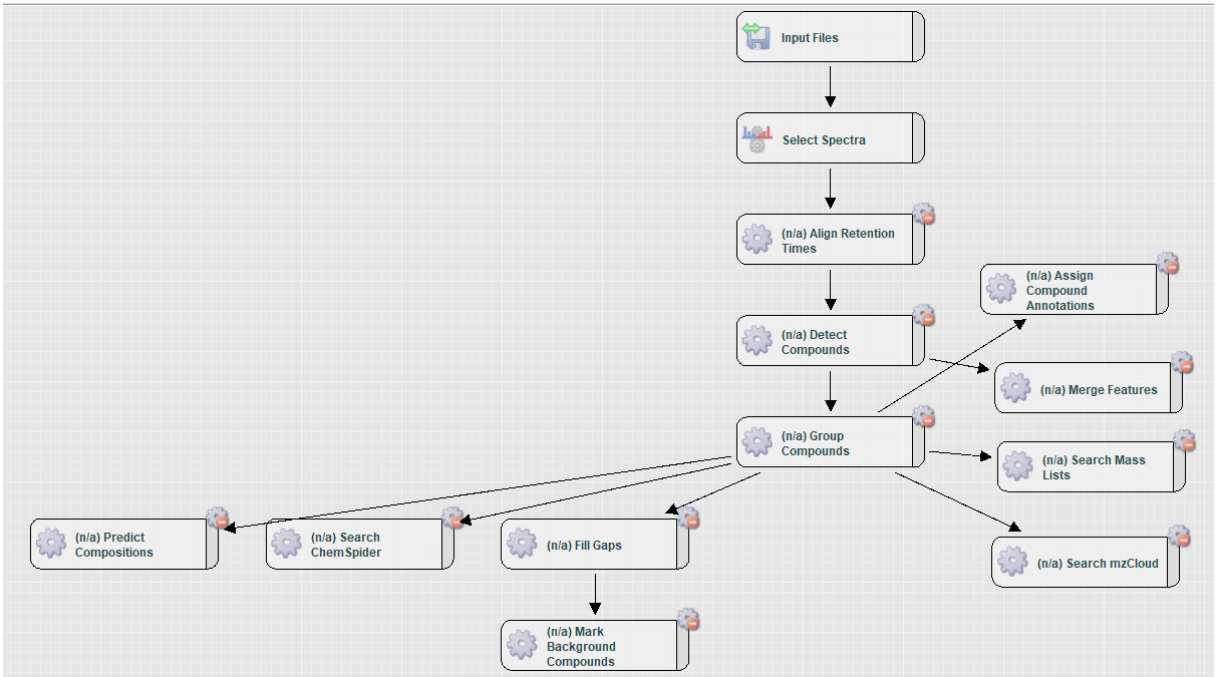


Figure S 6-2: Overview of all processing steps included in Compound Discoverer NTS workflow.

Table S 6-7: Details on processing parameter of NTS data processing workflow in Compound Discoverer for both ionization modes.

Select Spectra		
	Positive Ionization	**Negative Ionization**
	1. General Settings	
Precursor Selection	Use MS(n-1) Precursor	Use MS(n-1) Precursor
Provide Profile Spectra	Automatic	Automatic
	2. Spectrum Properties Filter	
Lower RT Limit (min)	1	1
Upper RT Limit (min)	24	24
Min Precursor Mass (Da)	80	80
Max Precursor Mass (Da)	1000	1000
Total Intensity Threshold	50000	10000
Minimum Peak Count	5	5
	3.Scan Event Filters	

Mass Analyzer	Any	Any
MS Order	Any	Any
Activation Type	Is HCD	Is HCD
Min. Collision Energy	0	0
Max. Collision Energy	100	100
Scan Type	Is Full	Is Full
Polarity Mode	Is +	Is -
4. Peak Filter		
S/N Threshold	3	3
5. Replacement for Unrecognized Properties		
Unrecognized Charge Replacements	1	1
Unrecognized mass analyzer Replacements	ITMS	ITMS
Unrecognized MS Order Replacements	MS1	MS1
Unrecognized Activiation Type Replacement	HCD	HCD
Unrecognized Polarity Replacement	+	-
Unrecognized MS Resolution @ 200 Replacement	70000	70000
Unrecognized MSn Resolution @ 200 Replacement	15000	15000
Align Retention times		
Mass tolerance	5 ppm	5 ppm
Maximum shift	0.3 min	0.3 min
Alignment Model	Adaptive curve	Adaptive curve
Detect compounds		
Intensity tolerance	30%	30%
Mass tolerance	5 ppm	5 ppm
Max element counts	C90 H190 Br3 Cl4 F6 K2 N10 Na2 O18 P3 S5	C90 H190 Br3 Cl4 F6 K2 N10 Na2 O18 P3 S5
Min element counts:	CH	CH
ions	M+H, M+K, M+Na	M-H, M-2H, 2H+K, M-H-H2O
Min. peak intensity	100000	30000
S/N Threshold	3	3
Group Compounds		
Mass tolerance	5 ppm	5 ppm
RT Tolerance	0.1 min	0.1 min
Preferred Ions	M+H	M-H
Merge Features		
Mass Tolerance	5 ppm	5 ppm
RT Tolerance	0.1 min	0.1 min
Predict Composition		
1. Prediction Settings		

Mass Tolerance	5 ppm	5 ppm
Min element counts:	CH	CH
Max. Element Counts	C90 H190 Br3 Cl8 F18 N10 O18 P3 S5	C90 H190 Br3 Cl8 F18 N10 O18 P3 S5
Min. RDBE	0	0
Max. RDBE	40	40
Min H/C	0.1	0.1
Max. H/C	3.5	3.5
Max. # Candidates	10	10
2. Pattern Matching		
Intensity Threshold	0.10%	0.10%
Intensity Tolerance	30%	30%
S/N Threshold	3	3
Use Dynamic Recalibration	True	True
3. Fragments Matching		
Mass Tolerance	5 ppm	5 ppm
S/N Threshold	3	3
Use Framents Matching	TRUE	TRUE
Mark Background Compounds		
Hide Background	TRUE	TRUE
Max. Blank/Sample	0	0
Max. Sample/Blank	10	10
Fill Gaps		
Mass Tolerance	5 ppm	5 ppm
S/N Threshold	1.5	1.5
Use Real Peak Detection	True	True
Search ChemSpider		
Database	DrugBank; EAWAG Biocatalysis/Biodegradation Database; enviPath; EPA DSSTox; EPA Toxcast	DrugBank; EAWAG Biocatalysis/Biodegradation Database; enviPath; EPA DSSTox; EPA Toxcast
Mass Tolerance	5 ppm	5 ppm
Search Mode	By Formula or Mass	By Formula or Mass
Search mzCloud		
Compound Classes	All; Counterfeit Drug (Therapeutic); Drugs of Abuse/Illegal Drugs; Endogenous Metabolites; Excipients/Additives/Colorants; Extractables/Leachables; Industrial Chemicals; Natural Toxins; Personal Care Products/Cosmetics; Pesticides/Herbicides; Small Molecule Chemicals	All; Counterfeit Drug (Therapeutic); Drugs of Abuse/Illegal Drugs; Endogenous Metabolites; Excipients/Additives/Colorants; Extractables/Leachables; Industrial Chemicals; Natural Toxins; Personal Care Products/Cosmetics; Pesticides/Herbicides; Small Molecule Chemicals
FT Fragment Mass Tolerance	10 ppm	10 ppm
IT Fragment Mass Tolerance	0.4 Da	0.4 Da
Post Processing	Recalibrated	Recalibrated

Precursor Mass Tolerance	10 ppm	10 ppm
	DDA Search	
Activation Energy Tolerance	20	20
Apply Intensity Threshold	True	True
Match Activation Energy	Any	Any
Match Factor Threshold	60	60
Identity Search	HighChemHighRes	HighChemHighRes
Similarity Search	Similarity Forward	Similarity Forward

Data filtering

Several filtering steps were employed to reduce the total number of features and rate of false positive detected peaks (FP rate) by ensuring that targets (true positives, TP) were not lost. False positives were determined as follows: 100 random numbers between 1 and N feature were selected and features with these index numbers were visually inspected. The first filtering step included a background filter with a sample/blank ratio of 10. For each sampling point (time and location) an individual field blank was measured and processed. However, at this point a high number of false positive features (FP) was observed mainly due to erroneous integration of noise. In a next step, features which did not show a replicate group peak area above 1.0 E+04 in any sample and features which were not present in at least three files (replicate filter) were filtered out. In a third step features were reduced based on their coefficient of variance (CV%) and HCA clustering pattern of data as described in SI Part 6.8.2.2.1.

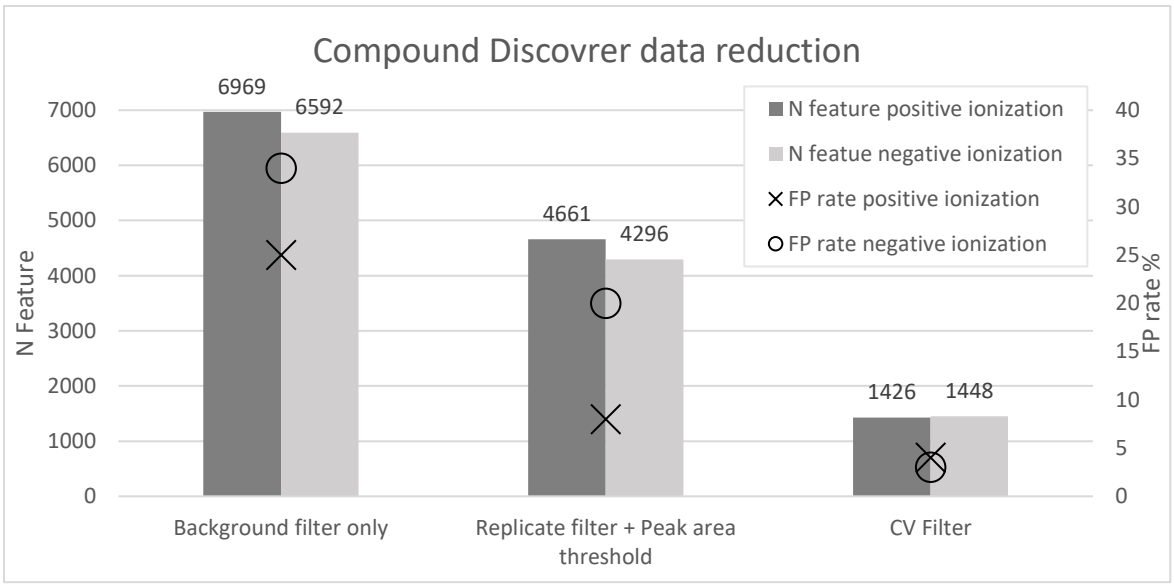


Figure S 6-3: Overview of feature reduction (left y-axis) and false positive (FP) rates (right y-axis) for different filtering steps for positive and negative data.

In figure S6-3 data reduction for positive and negative datasets are displayed for different filtering steps. As can be seen total feature numbers for both ionization modes were reduced in a similar way for both positive and negative data. It was noted that FP rates for negative

data were initially higher compared to positive data but were reduced to below 5% for both ionization modes after application of the CV filter.

Quality assurance

To ensure stability of passive sampler extracts and quality of both HPLC-HRMS measurement and data processing, concentration profiles over all 30 samples of targets quantified during target analysis were compared to profiles of peak areas of targets detected in non-target measurement. The correlation analysis showed a high degree of correlation ($p<0.05$) between both data matrices. From a total of 44 target compounds 40 showed a Pearson Correlation coefficient >0.9 as shown in figure S6-4. For Carbendazim and Imidacloprid concentration profiles were similar but showed differences in a few samples. For chloridazon-Desphenyl and climbazol concentration profiles were different. This is probably due to instability of these analytes in the stored extracts in acetone/methanol. Nevertheless, overall stability of samples, measurement and recall of targets in data processing was satisfying.

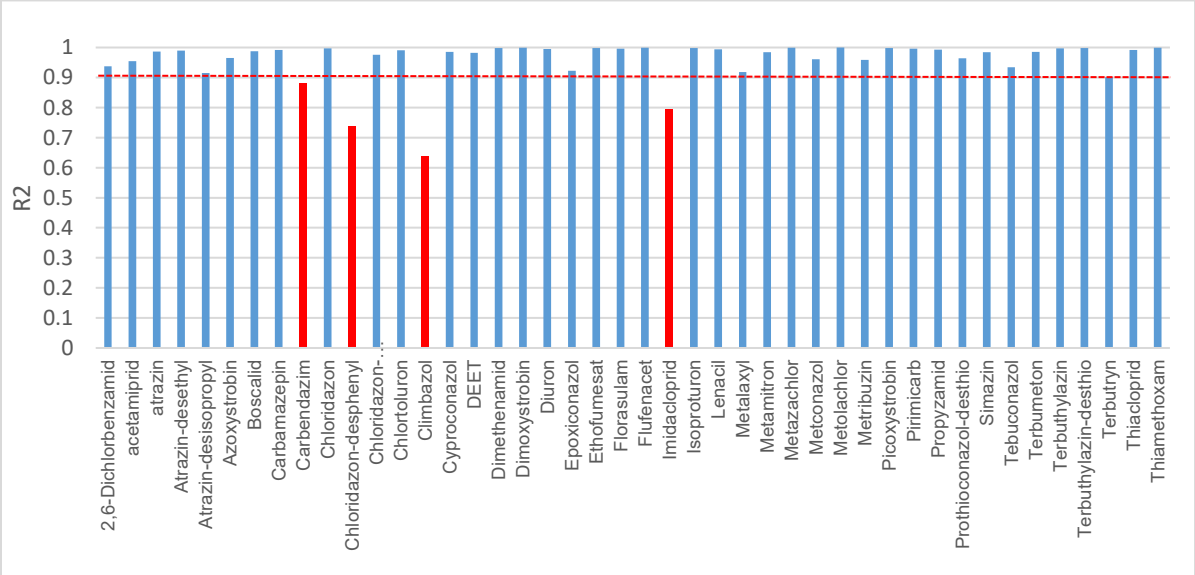


Figure S 6-4: Correlation given as R^2 of peak areas obtained by non-target analysis and concentration (µg/PS) obtained by target analysis for all target compounds over all 30 samples.

6.8.2 Summary of employed chemometrics tools in this workflow

In the following sections, a brief overview of outlined chemometric methods is given in order to highlight their strengths and limitations in the area of big analytical data, as well as to provide some rationale for utilizing multiple data processing approaches as performed in this study. We only discuss their implementation in the context and scope of the presented workflow; the methods can be used in other studies with different aims or may be developed further in the future. This is far from a systematic review of the available methods, and more in-depth information can be found elsewhere [90] [81]. The goal is to support environmental scientists

who are less familiar with chemometric data evaluation and to provide an impression of possible fields of application.

6.8.2.1 Data preprocessing

Before performing different chemometric approaches data are generally preprocessed to remove confounding variations of ion intensities while the relevant variation containing information on environmental differences among samples should be preserved.

Figure S6-5 shows the high differences of signal magnitudes in the data, which is usually the case for LC-HRMS environmental data. As shown, a few intense peaks overlay other signals, which can adversely influence the result. Thus, the matrices were first elementary log-scaled reducing data skewness and leading to normally distributed variables.

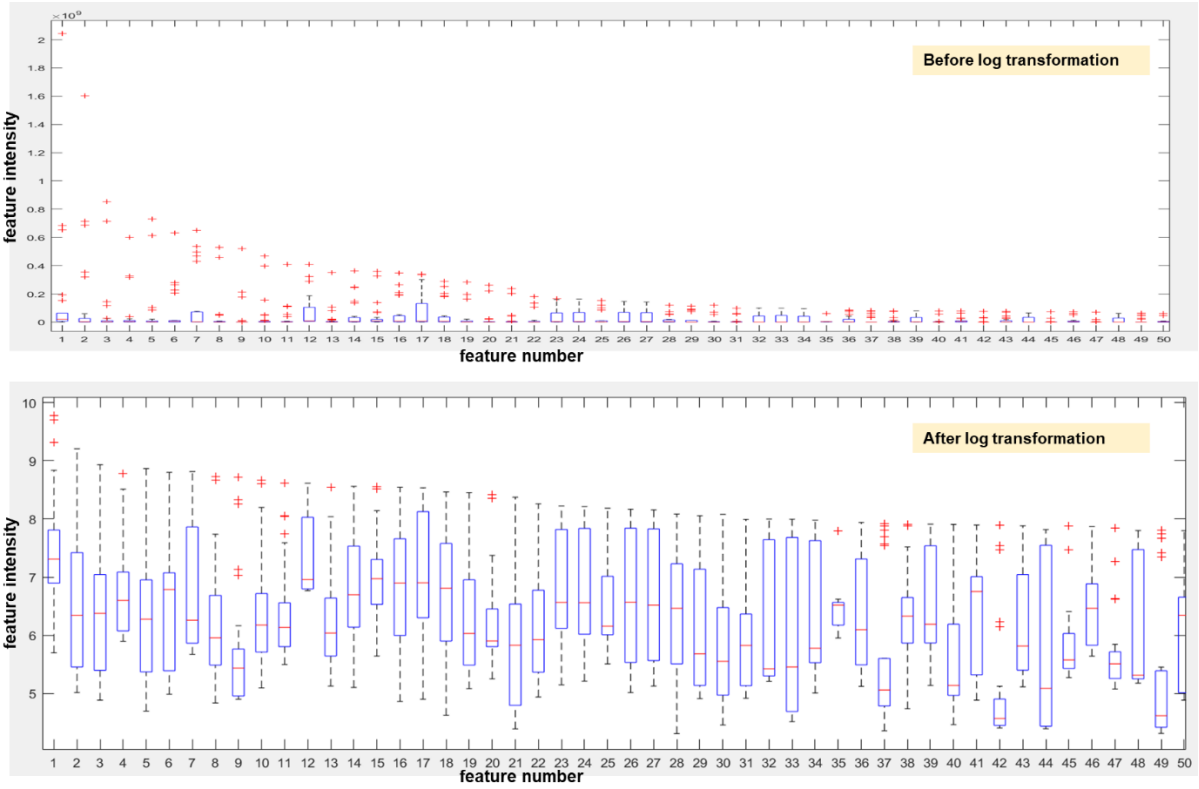


Figure S 6-5: Example of effect of log transformation for a selection of features (50 most intensive) in positive ionization mode.

In the following, the impact of the choice of different preprocessing methods on non-target data variation in sample space is exemplary shown. Only a small selection of available preprocessing methods is compared. To this end, both original and reduced data matrices in positive mode of ionization are subjected to PCA with the following preprocessing methods: (A) raw data without preprocessing, (B) sample-wise normalization/autoscaling, (C) log-scaling/sample-wise normalization/autoscaling, and (D) log-scaling/sample-wise normalization/mean centering. Figure S6-6 shows that in general, PCA score plots for all the preprocessed data matrices show a class-wise pattern (by sampling site) of the data sets. This is more obvious for preprocessing D for all classes, and more highlighted for original data

relative to reduced data. The latter is due to the lower percentage of influential variables in original data than for the reduced data matrix. Furthermore, autoscaling has a significant overall effect on increasing the contribution of temporal effects to total variance (this is specially the case for Flehe river, as indicated by higher within class separations in figure S6-6 D). Simultaneously, this affects the between-site separations. However, in this study ASCA has been utilized further to visualize multivariate patterns in the space of each effect (time and location) of statistical models linked to the experimental design.

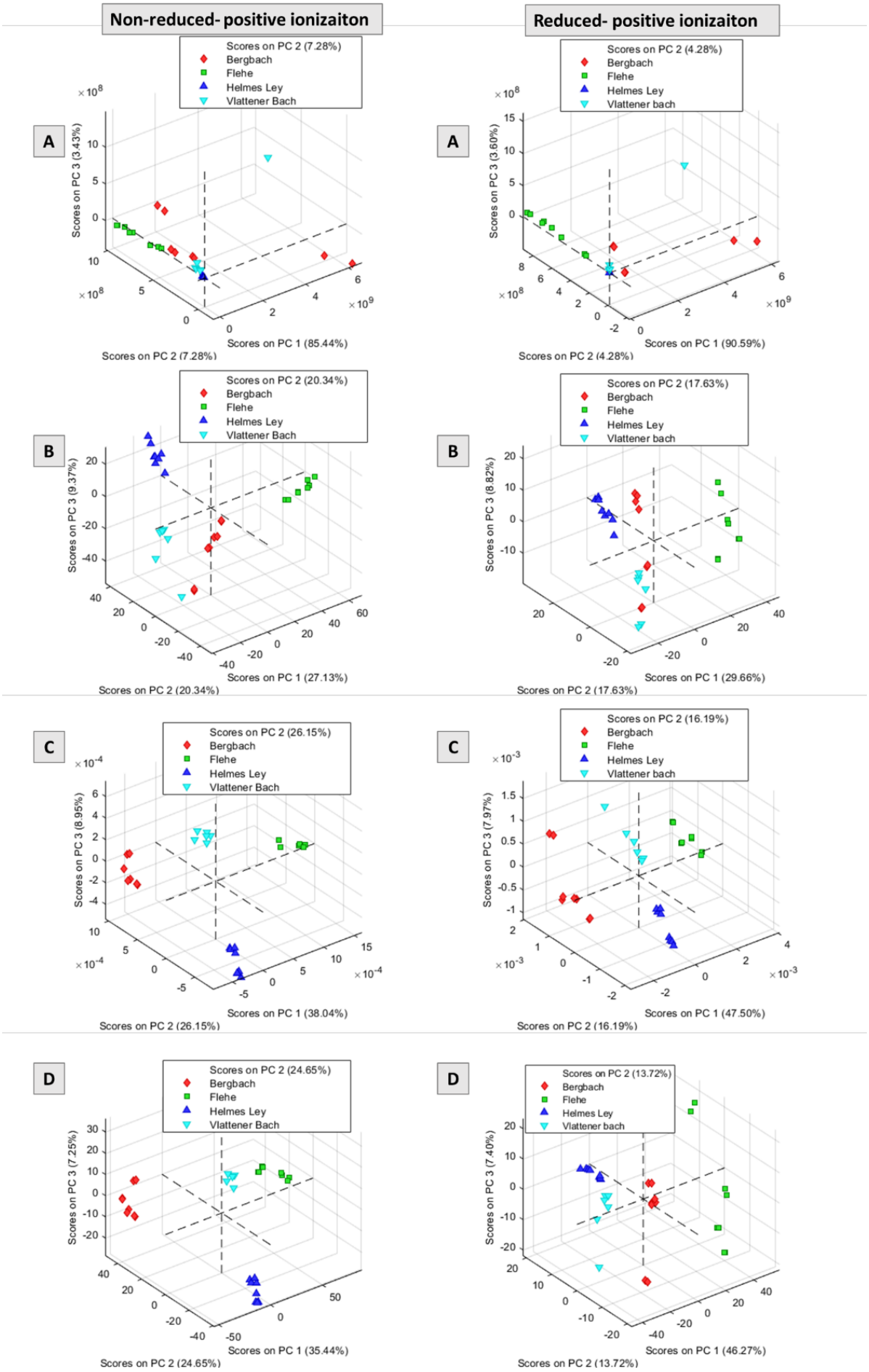

Non-reduced- positive ionizaiton
Reduced- positive ionizaiton
A
Scores on PC 2 (7.28%)
Bergbach
Flehe
Helmes Ley
Vlattener Bach
Scores on PC 3 (3.43%)
Scores on PC 1 (85.44%)
Scores on PC 2 (7.28%)
A
Scores on PC 2 (4.28%)
Bergbach
Flehe
Helmes Ley
Vlattener bach
Scores on PC 3 (3.60%)
Scores on PC 1 (90.59%)
Scores on PC 2 (4.28%)
B
Scores on PC 2 (20.34%)
Bergbach
Flehe
Helmes Ley
Vlattener Bach
Scores on PC 3 (9.37%)
Scores on PC 1 (27.13%)
Scores on PC 2 (20.34%)
B
Scores on PC 2 (17.63%)
Bergbach
Flehe
Helmes Ley
Vlattener bach
Scores on PC 3 (8.82%)
Scores on PC 1 (29.66%)
Scores on PC 2 (17.63%)
C
Scores on PC 2 (26.15%)
Bergbach
Flehe
Helmes Ley
Vlattener Bach
Scores on PC 3 (8.95%)
Scores on PC 1 (38.04%)
Scores on PC 2 (26.15%)
C
Scores on PC 2 (16.19%)
Bergbach
Flehe
Helmes Ley
Vlattener bach
Scores on PC 3 (7.97%)
Scores on PC 1 (47.50%)
Scores on PC 2 (16.19%)
D
Scores on PC 2 (24.65%)
Bergbach
Flehe
Helmes Ley
Vlattener Bach
Scores on PC 3 (7.25%)
Scores on PC 1 (35.44%)
Scores on PC 2 (24.65%)
D
Scores on PC 2 (13.72%)
Bergbach
Flehe
Helmes Ley
Vlattener bach
Scores on PC 3 (7.40%)
Scores on PC 1 (46.27%)
Scores on PC 2 (13.72%)

Figure S 6-6: Examples of influence of preprocessing on PCA results of reduced and non-reduced non-target data obtained in positive ionization mode. A: no preprocessing; B: row wise normalization and autoscaling; C: log transformation, row wise normalization and mean centering; D: log transformation, row wise normalization and autoscaling.

6.8.2.2 Data exploration

Principal component analysis (PCA) is a good starting point of a chemometric analysis as it provides an overview of general data patterns, specific characteristics, and possible outliers. As it is an unsupervised method, no predefined sample classes need to be known. Here, data visualization is based on scores and loading plots and can get rather complex for high-dimensional datasets. The benefits of detailed analysis of a dataset in a factorial-based decomposition (as for ASCA) is highlighted in the manuscript.

Hierarchical cluster analysis (HCA) can be used, like PCA, to provide an overview of data inherent structure with the help of a dendrogram. Different measures of distance between pairs of observations can be used and subclusters can be evaluated. In this study, HCA was used for data reduction purposes as described in section 6.2.2.1.

In addition, a heat map can also be helpful to visualize the tracking of variables in different samples by inspection of cluster formation. But since graphical interpretations are highly cumbersome for original non-target data sets, it would be more informative for data sets including subsets of variables. Thus, in this study, following the previous analysis by PLS-DA and volcano tests a simultaneous hierarchical clustering was performed on the statistically significant features, resulting in the demonstration of clear differences among the four sampling locations and two seasonal cases.

6.8.2.3 ANOVA simultaneous component analysis (ASCA)

In ASCA, the power of ANOVA to separate variance sources is merged with the advantages of simultaneous component analysis (SCA). A useful tutorial and review can be found in [87]. ASCA is generally expressed as:

$$X = 1m^T + X_{F1} + X_{F2} + X_{(F1F2)} + E$$

where $1m^T$ is the matrix of means, X_{F1} and X_{F2} are the effect matrices of the different factors F1 and F2 and $X_{(F1F2)}$, describes the interaction between them. E is the residual matrix and contains the variation that cannot be described by the model. The effect matrix X_{F1} is decomposed as:

$$X_{F1} = T_{F1} P^T_{F1} + E_{F1}$$

where T_{F1} and P^T_{F1} denote the scores and loading matrices for the effect F1, respectively, and E_{F1} contains the residuals. By that, scores and loadings plots of the different factors and interactions of them can be obtained to visualize different patterns and relations of samples and variables. The importance of the factors is assessed based on the evaluation of the sum of squares of the effect matrices and comparing its value to the corresponding null hypothesis-based distribution, estimated by means of a randomization test [88]. However, because in

many a balanced experimental design cannot be covered, employing the conventional ASCA method would result in a biased estimate of the factor effects. Due to this limitation, a modified variant of ASCA (ASCA+), which is based on the theory of general linear models (GLMs) is used. For more details, refer to [87].

ASCA has the great benefit that several study factors can be analyzed within one model. This can help to disentangle the contribution to variance of each factor and the interaction of them on the data set. Besides sampling time and location other environmentally relevant factors could be included (e.g., sample types, temperatures, model organisms, sampling devices etc.) However, the dataset or study has to be of a "designed" structure, which means samples for all combinations of factors (e.g. time/location) each with several classes (e.g. river 1, river 2, ...) need to be taken. In addition, as for many chemometric methods, interpretation of effect loadings for datasets with an extremely high number of variables becomes more difficult. While in the present study, the loadings were selected based on ad-hoc decision thresholds followed by visual inspection of the most intense values, other tools maybe utilized, as well, such as constructing bootstrapped confidence intervals for loading values or performing sparse PCA decomposition [90].

6.8.2.4 Partial least squares – discriminant analysis (PLS-DA)

PLS-DA is a supervised linear discriminant analysis that establishes a multivariate regression model between a matrix of independent variables (**X**, predictor variables) and an array of dependent variables (**y**, predicted variables) that contains binary dummy variables (0 and 1) indicating the class to which each sample belongs [79]. In the current study, PLS-DA was used to select variables of interest. For this purpose, separated models for each study factor (location and time) have to be build. For each model, peak areas of samples were correlated with the vector describing the sample type class membership (**y**). The models were cross-validated by repeated venetian blinds, and the number of latent variables (LVs) was chosen based on the smallest cross-validation classification errors and permutation test. Also, the statistical information obtained from each model can be utilized for feature ranking and to determine which features (micropollutants) of **X** are more associated with the class membership of **y.** For this purpose, the variable importance on projection (VIP) scores, which are the weighted sum of squared PLS variable weights are computed [91]. While the features with VIP>1 are commonly considered to be important (since they are larger than the average of squared VIP values), other cutoff thresholds (2, 3 or the average of VIP values) may worth trying in different studies according to attributes such as percentage of the number of relevant variables and their correlations, magnitude of mean difference of relevant variables between groups, and sample size) [93].

6.8.2.5 Class comparison using complementary univariate statistics

In order to visualize the significant of changes in feature intensities, univariate statistical analysis can also be utilized. The commonly used univariate analysis methods for differential feature analysis include coefficient of variation analysis, t-test, ANOVA and volcano scatter plots. With volcano plots features are selected based on fold changes and p-values (significance level) in an independent way. This is especially useful when only two groups (e.g. with effect/without effect) need to be compared. However, because univariate statistics necessitate significance testing of hundreds of features for non-targeted data sets, multiple test corrections should be considered as an integral part of this method to protect against the increasing risk of false positives [95]. As for PLS-DA, the thresholds can be set depending on the study purpose and a combination with a multivariate-based prioritization approach can be used to discern environmentally relevant information in a more efficient way.

6.8.2.6 Procrustes Analysis

Procrustes Analysis is an effective way of unravelling correlations between measured data. To this end, rather than rotating the data itself, the corresponding principal components (obtained through ASCA in this case) are being rotated. If T1 and T2 are the scoring matrices for two original data sets (X1 and X2, respectively), one of them, T1, is fixed and the other, T2, is transformed to match T1. In the current study, T1 and T2 are score matrices obtained by ASCA modelling of target (fixed as reference data) and non-target data sets (comparison data), respectively. The rotation is performed geometrically by translating, rotating, and then stretching/shrinking to minimize the total of squared distances (D^2) between the elements of T1 and the corresponding elements of T2. The smaller the value of D^2, the closer the two configurations are. A $D^2=0$ is the result of a perfect match. The significance of the fit between the two methods can be assessed empirically, where a common approach is to employ Monte Carlo methods using many times (e.g. 10000 times) randomly cuting the samples in one of the blocks and calculating distance measures. More details regarding Procrustes Analysis can be found elsewhere [204].

6.8.3 Results

6.8.3.1 Target analysis: additional figures of ASCA modelling

The loading plots for PC 1 and PC 2 for each factor (figure S6-7) show which targets are characterized by a significant concentration change between samples and are thus responsible for a distinction of different levels for each factor in the ASCA model. For example, 2,6-dichlorobenzamide (t1), chloridazon-desphenyl (t11) and chloridazon-methyldesphenyl (t12), all transformation products of herbicides, with the highest positive loading values for PC 1 factor 1, show the highest mean concentration values in *Helm*, while they are also detected at other sampling sites (figure S6-8). In a similar way, the most important targets for PC 2 of

factor 1 or the most discriminating targets for factor 2, showing different temporal trends, can be evaluated based on loading values. As an example, metolachlor (t31) is highly associated with sampling period 2. Targets with high or low loadings of the interaction factor (F1xF2) indicate a differential temporal concentration pattern among different sampling locations. This effect was most pronounced for propyzamide (t35) that exhibits completely different temporal occurrence patterns among the four sampling locations (figure S6-8).

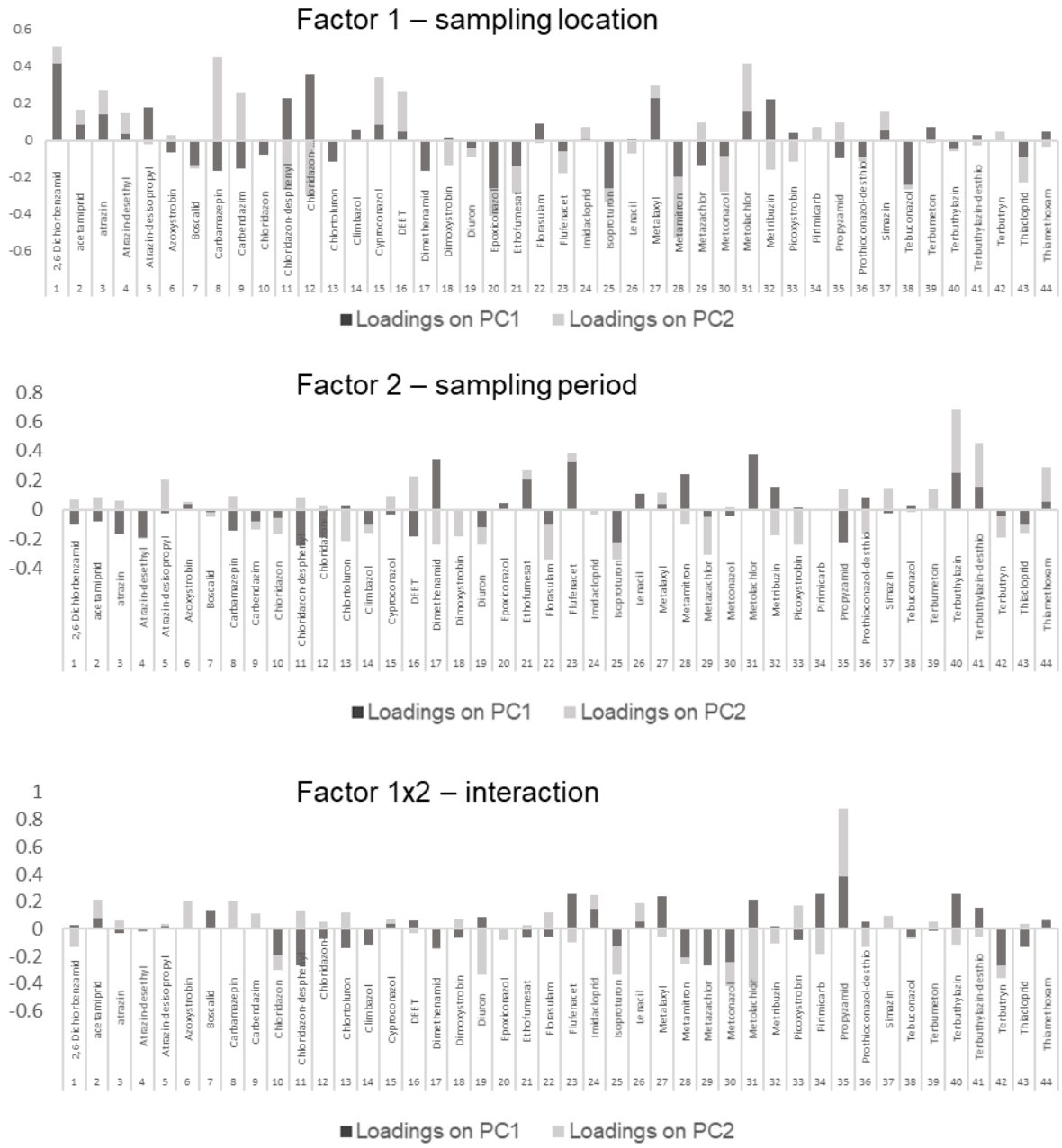


Figure S 6-7: ASCA loading plots of first two PCs of target data for factor 1, 2 and interaction of both.

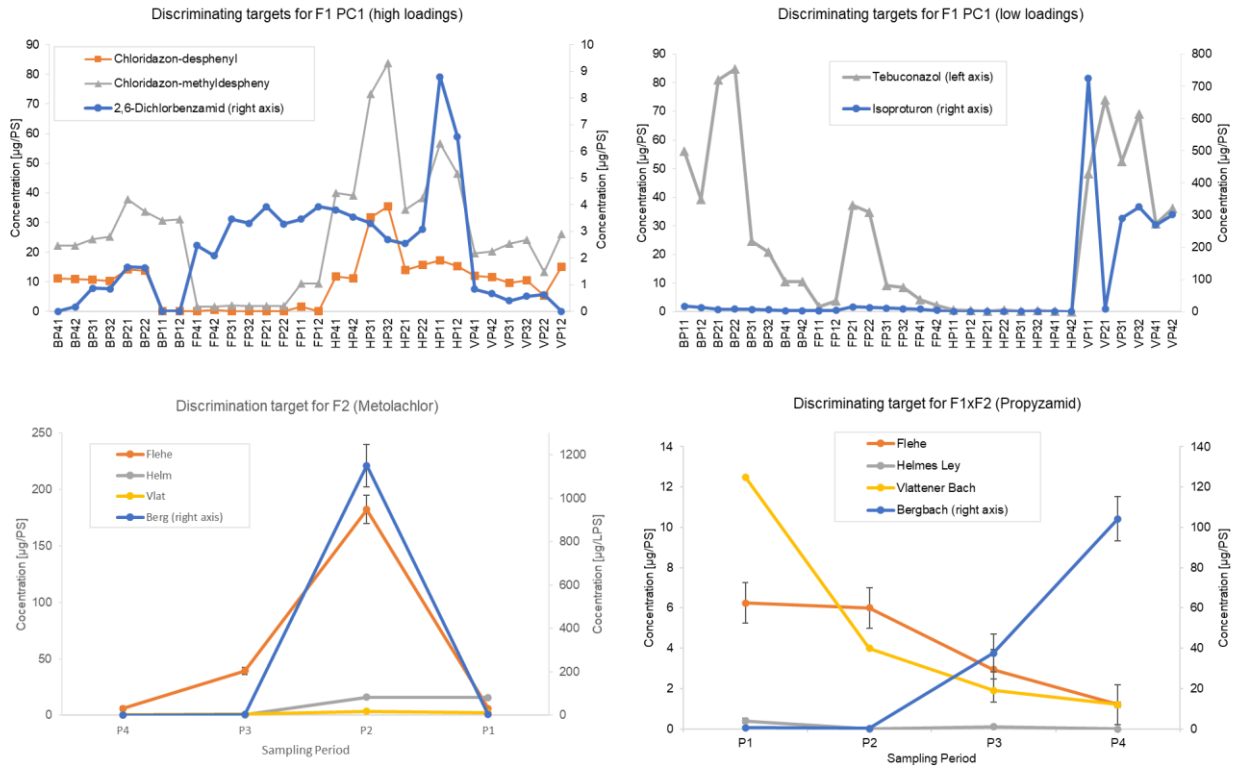


Figure S 6-8: Exemplary concentration profiles of targets with discriminating properties. B: Berg; F: Flehe; H: Helm; V: Vlat; P1: sampling period 1; P2: sampling period 2; P3: sampling period 3; P4: sampling period 4.

6.8.3.2 Non target analysis

6.8.3.2.1 Data reduction based on CV filter

Before employing advanced chemometric tools a reduction of variables/feature was necessary to reduce the matrices to a manageable size. For this purpose, features with a steady intensity profile over all samples, which were not of interest in the scope of this study, were discarded. Features with low intensity variations over all samples might be background signals or ubiquitous pollutants without discriminating properties and thus non-informative for the research question. This data reduction approach was chosen to filter out features based on their intensity variation over samples and not based on intensities to avoid discrimination of low intensity features. Data were reduced based on coefficient of variance (CV%) of peak areas of all samples without distorting the hidden and inherent data structure. Thresholds for CV were incrementally increased and features below it were omitted. Reduced data matrices were analysed by hierarchical cluster analysis (HCA) with Euclidean distance and changes of data structure and sub-cluster distances were assessed. This approach was continued until a remarkable change in the resulting dendrograms were observed. Reduced data matrices consisted of 1426 variables with a CV above 161% for positive ionization mode and 1448 variables with CV above 147% for negative ionization mode, respectively. In figure S6-9 and S6-10 HCA dendrograms for original and final reduced data matrices are shown for positive and negative data. As can be seen, Flehe samples were the most dissimilar samples from

others and Helm and Vlat were the most similar ones for all different data matrices. Besides some trivial rotation data inherent structure was not disturbed.

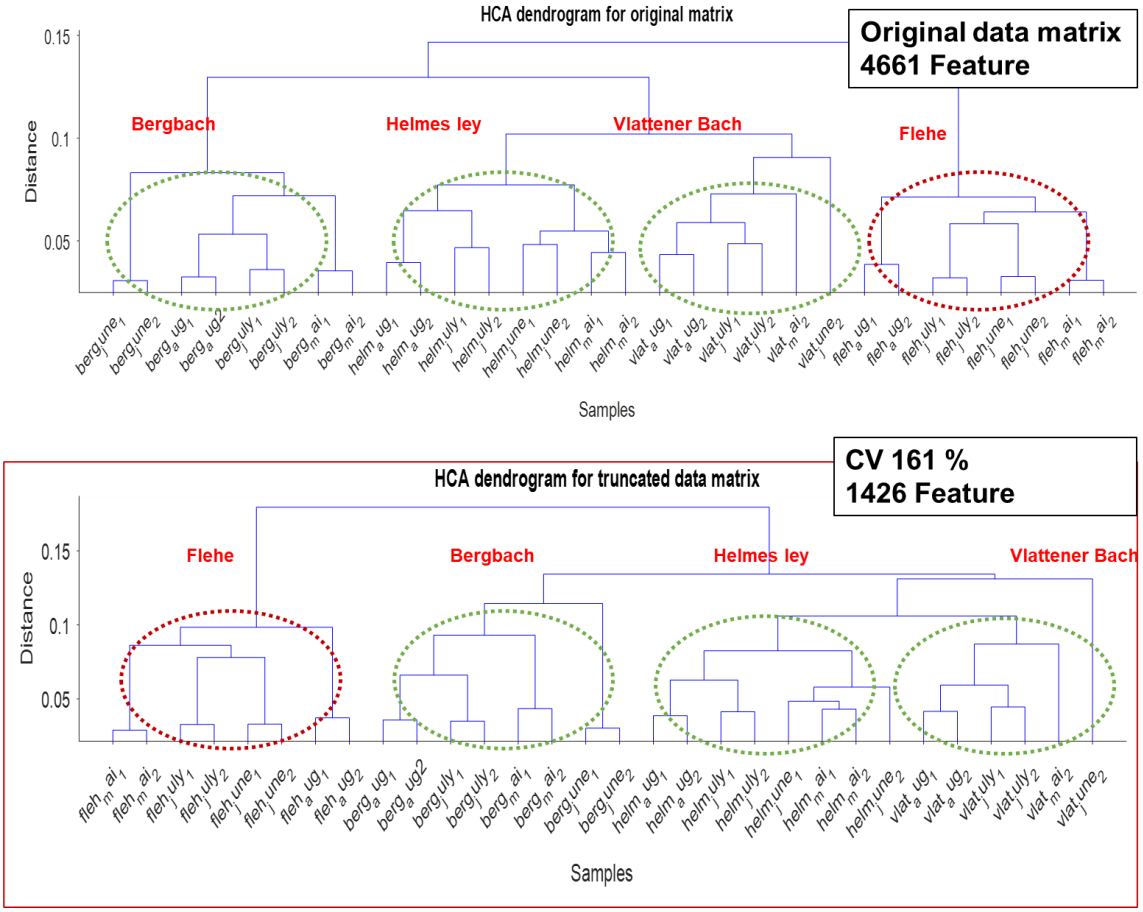


Figure S 6-9: HCA dendrogram analysis of original data matrix (top) and reduced data matrix based on CV 161% (bottom) for positive data.

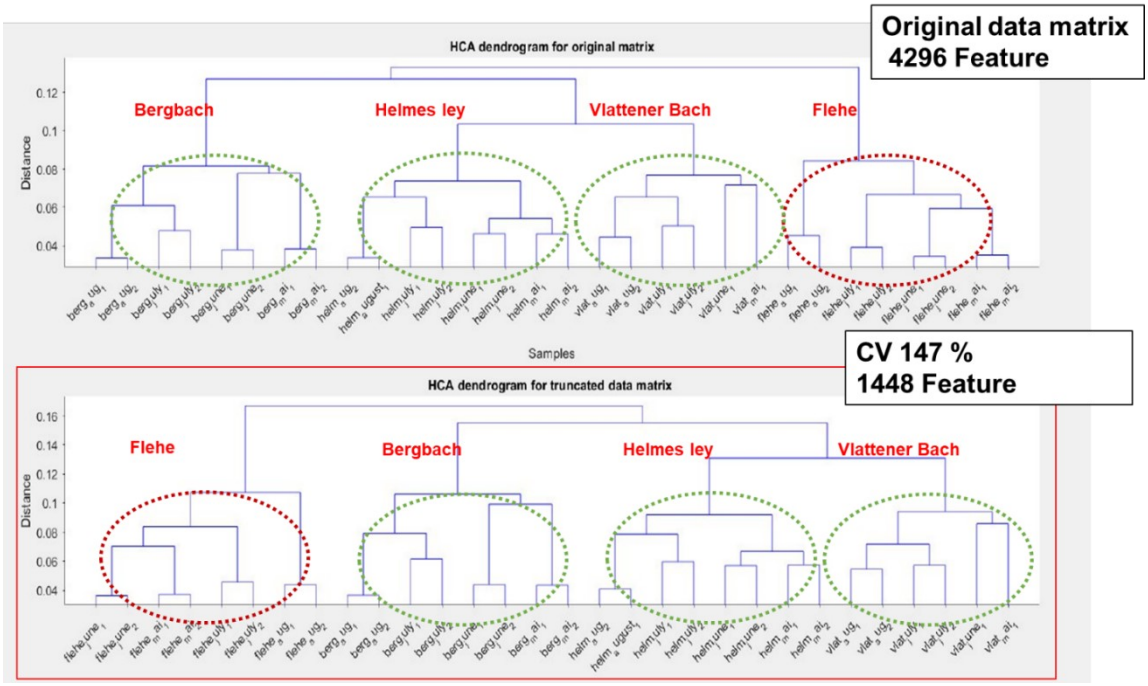


Figure S 6-10: HCA dendrogram analysis of original data matrix (top) and reduced data matrix based on CV 148% (bottom) for negative data.

6.8.3.2.2 Additional figures of PCA modelling

PCA modelling was implemented as explorative tool firstly on non-reduced data as presented in the manuscript figure 6-3 and showed a consistent picture with HCA clustering. After data reduction, PCA was performed again and results are shown in figure S6-11. Especially for positive data, within-class dispersions are higher for reduced data (due to the fact that the score values are now obtained by a linear combination of a limited number, but more informative subset to original variables). However, no cross-wise dislocation between class elements existed.

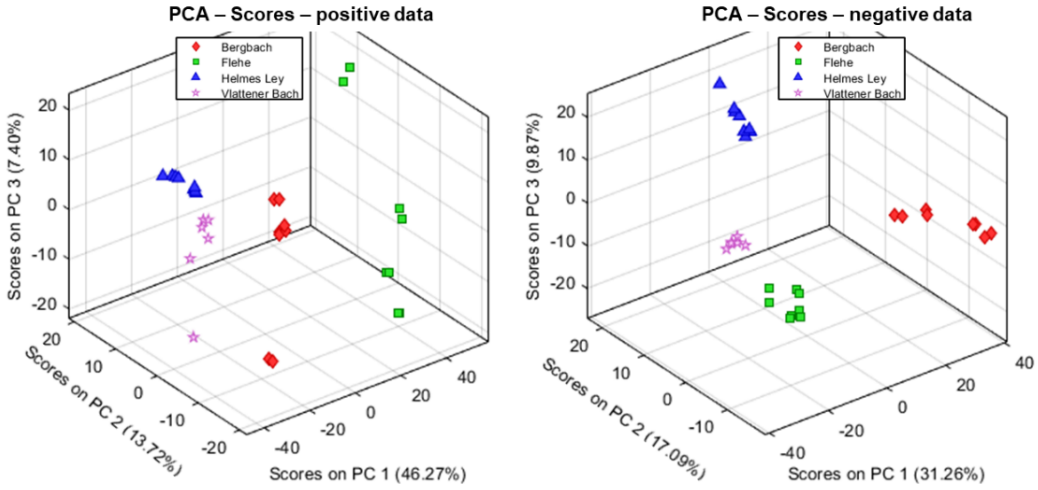


Figure S 6-11: PCA score plots including 3 PCs for positive ionization (left) and negative ionization (right) reduced non-target data, colorized by sampling location.

6.8.3.2.3 Additional figures of ASCA modelling

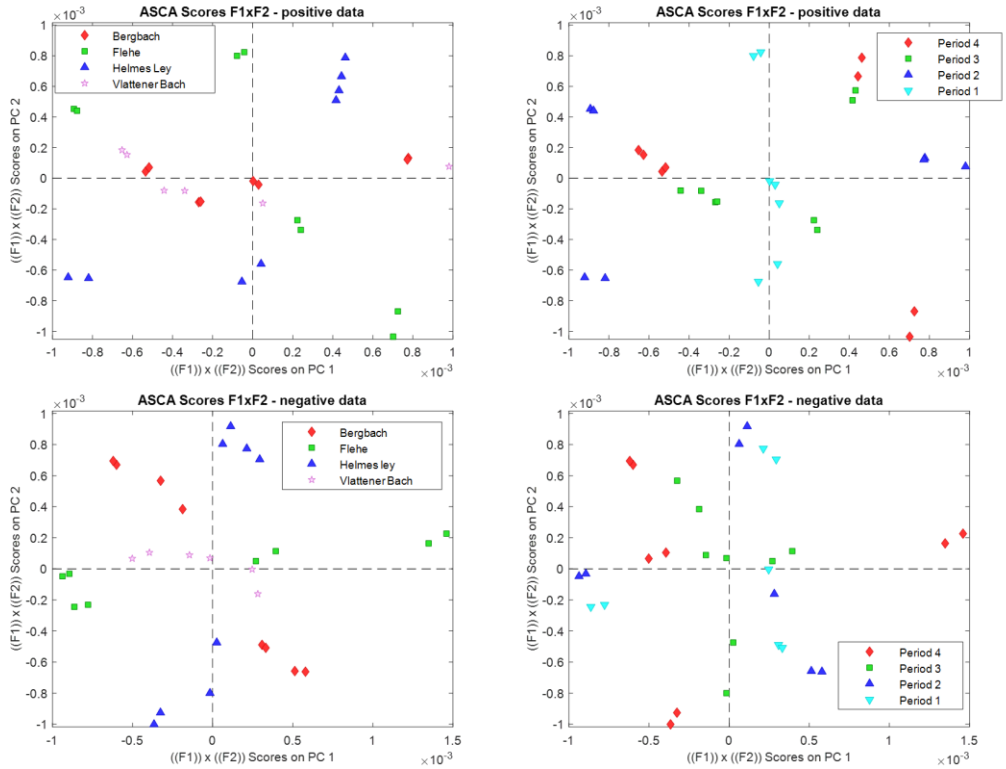


Figure S 6-12: Scores plots of interaction factor for positive (upper plots) and negative (bottom plots) non-target data coloured by location (left plots) and sampling period (right plots).

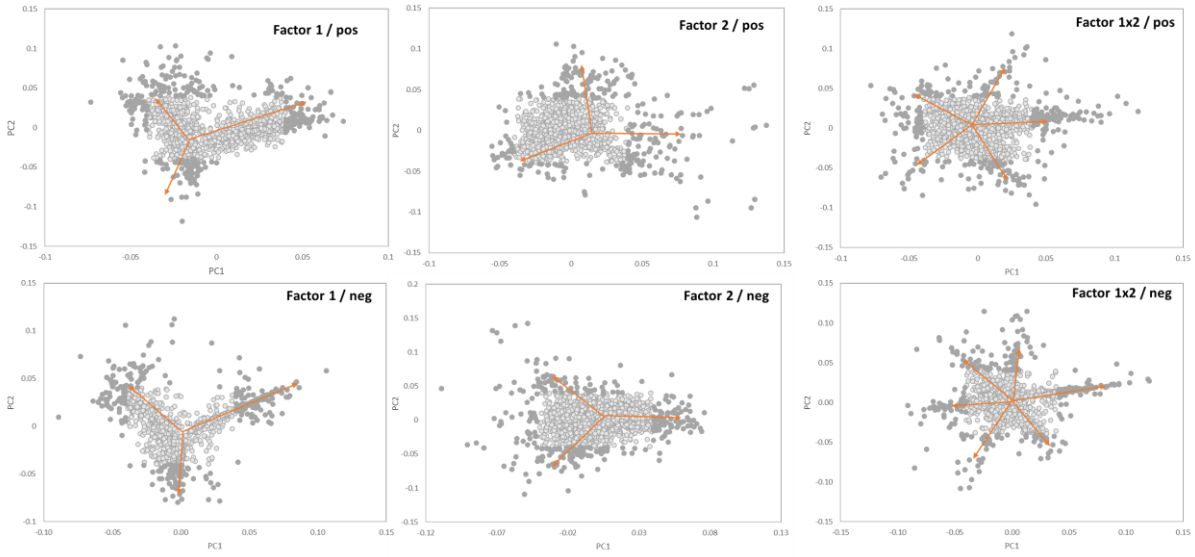


Figure S 6-13: ASCA loading plots of all factors for positive (upper plots) and negative (bottom plots) data. Features above ad hoc threshold value of +/-0.04 are highlighted in darker grey. For F1, F2 and F1x2, 23.6%, 19.4% and 23% of the features for positive mode.

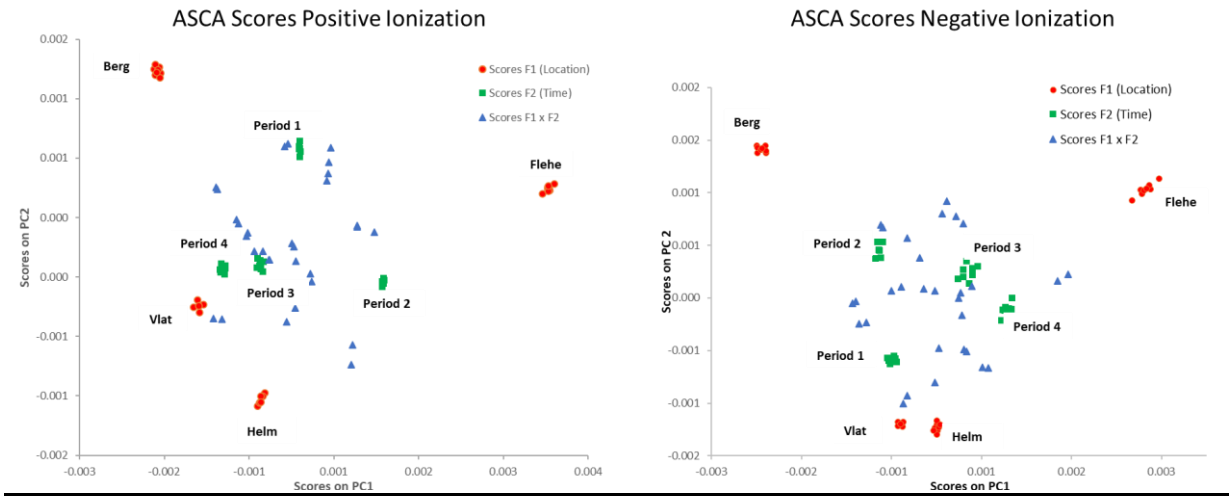


Figure S 6-14: Scores plots of interaction factor for positive (upper plots) and negative (bottom plots) non-target data coloured by location (left plots) and sampling period (right plots).

The radar plots of the first two PCs for the interaction effects (figure S6-15) show that among all four sampling sites, *Berg* and *Vlat* show a more similar temporal trend than the others, increasing the score values in period 2 for positive data (PC1 and PC 2), but in period 3 (PC1) and periods 1 and 2 (PC2) for negative data. This pattern is in strong contrast to what is seen for *Helm* and *Flehe*.

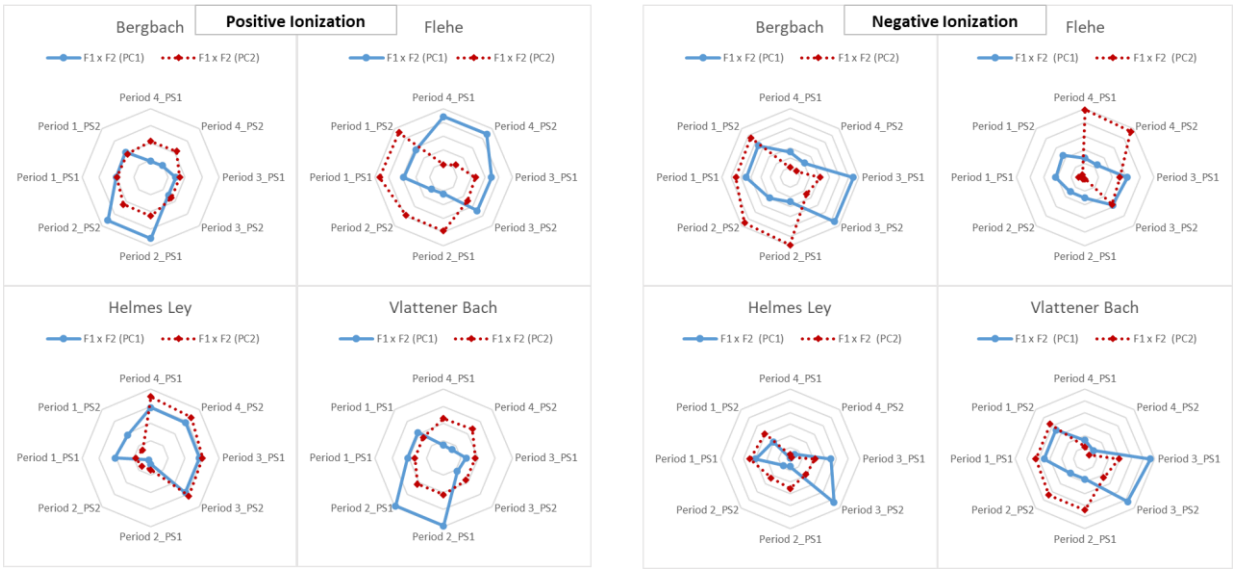


Figure S 6-15: Radar plots of scores of first two PCs of interaction factor for each sampling location. The scale of the plots is in arbitrary units and was removed for better readability. PS1 and PS2 represent replicates of passive sampling devices for each sampling

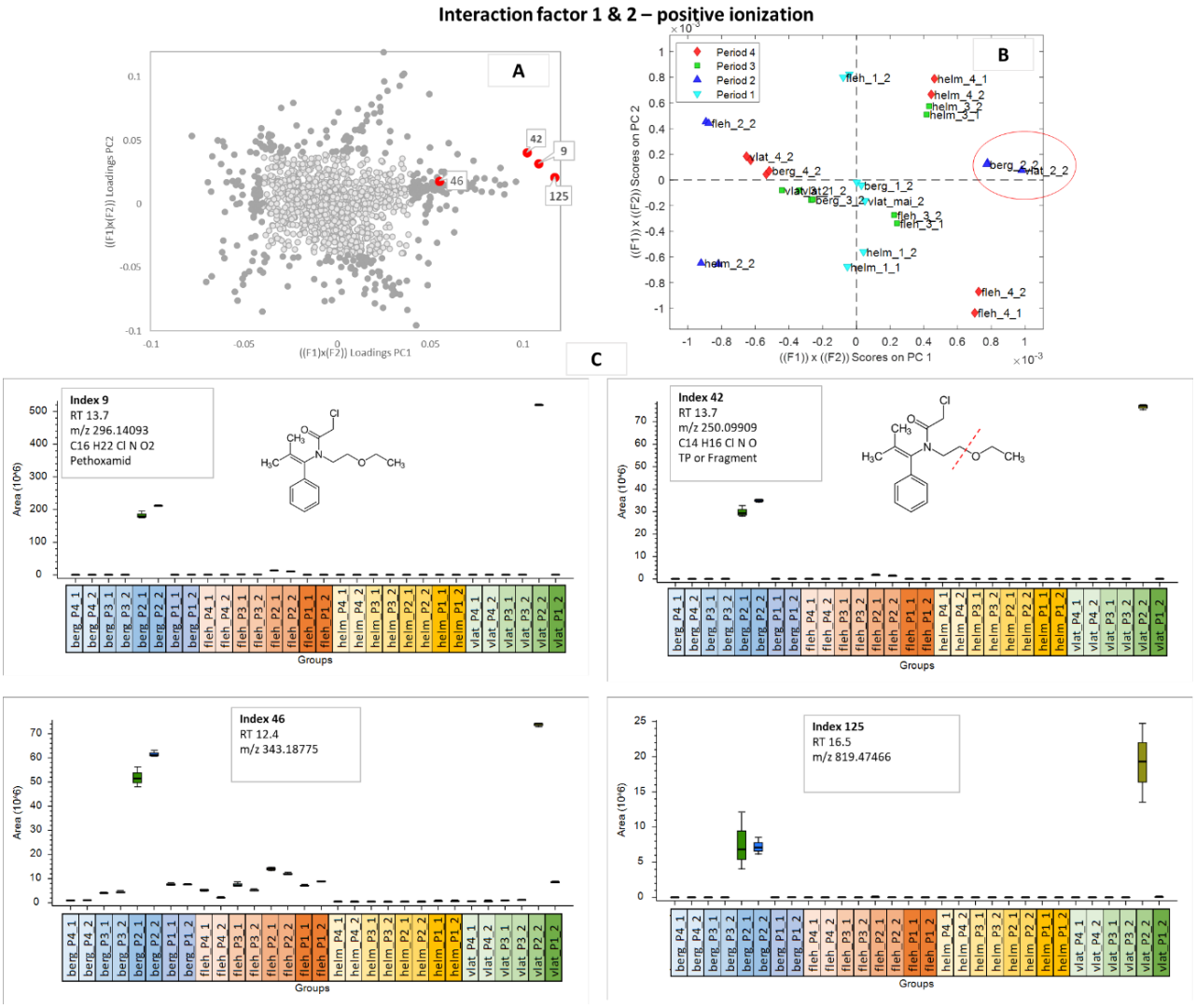


Figure S 6-16: Examples of features with discriminating properties for Vlat and Berg samples of sampling period 2 for positive non-target data. A: ASCA loadings plot for F1xF2, B: ASCA scores plot for F1xF2, C: Peak area profiles for exemplary features f9, f42, f46 and f135 over all samples.

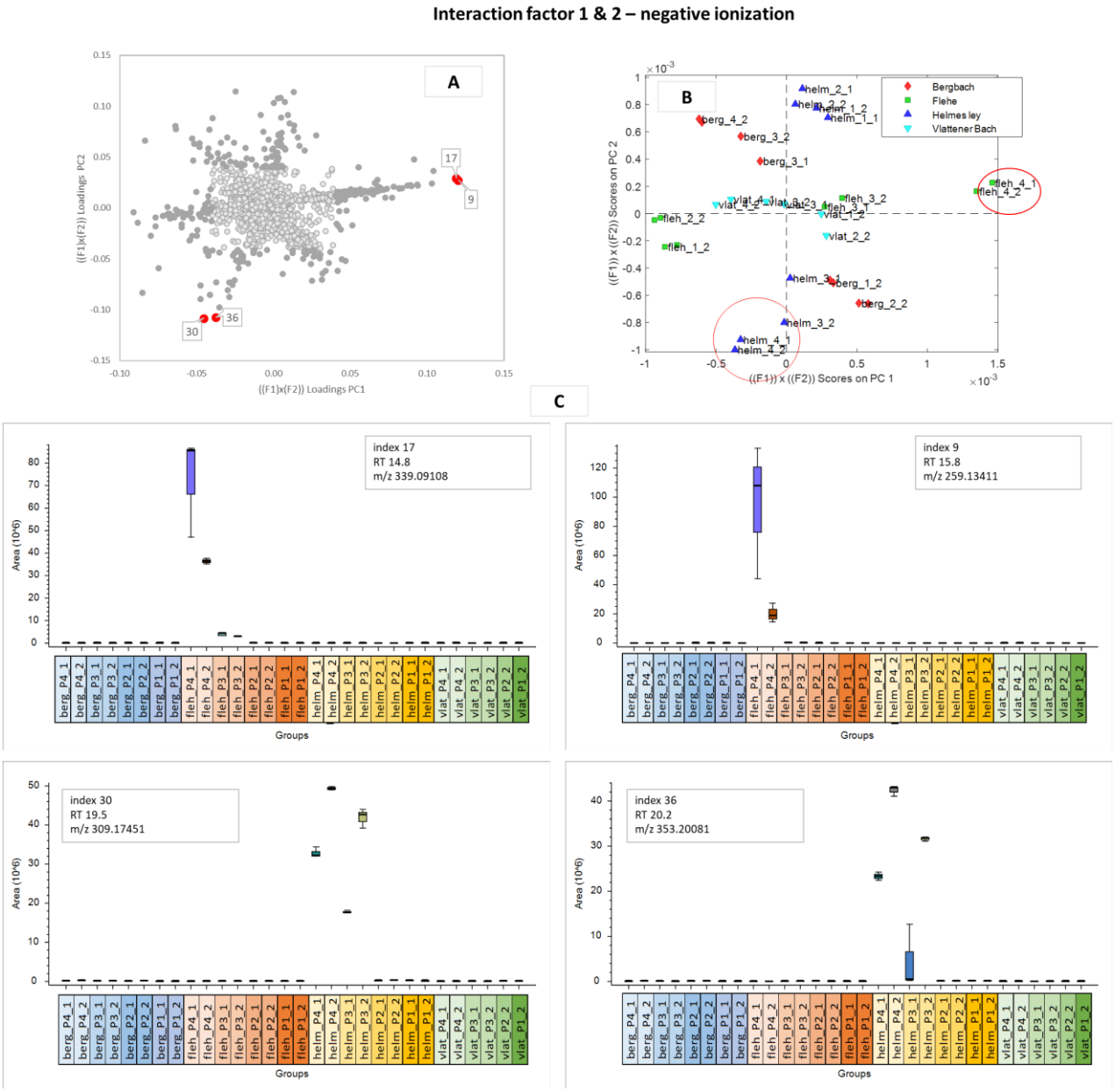


Figure S 6-17:: Examples of features with discriminating properties for interaction of F1 and F2 for negative non-target data. A: ASCA loadings plot for F1xF2, B: ASCA scores plot for F1xF2, C: Peak area profiles for exemplary features f17, f9, f30 and f36 over all samples.

6.8.3.2.4 Comparison of ASCA results of target and non-target data

In order to compare the similarity of data patterns obtained with ASCA by target and non-target data a procrustes analysis was performed on scores plots in matlab. After scaling and rotation of non-target data, the dissimilarity measure was calculated. This measure gives a quantitative value on how well different shapes (in this case scores in PC1xPC2 space) overlay. The results are shown in figure S6-18, the dissimilarity of scores plots of factor 1 showing the locational pattern (0.1575) is much higher compared to dissimilarity of temporal pattern represented with factor 2 (0.0442). The good overlay of temporal patterns can also be visually recognized in the figure S6-18 (right), sampling period 2 is separated on the first PC, period 1 on the second PC and periods 3 and 4 are close together for both types of data modelling. For the location factor (left graph) Flehe samples are separated from the three small rivers within PC1 for non-target data which is not observed for target data where Helm samples are the most dissimilar group.

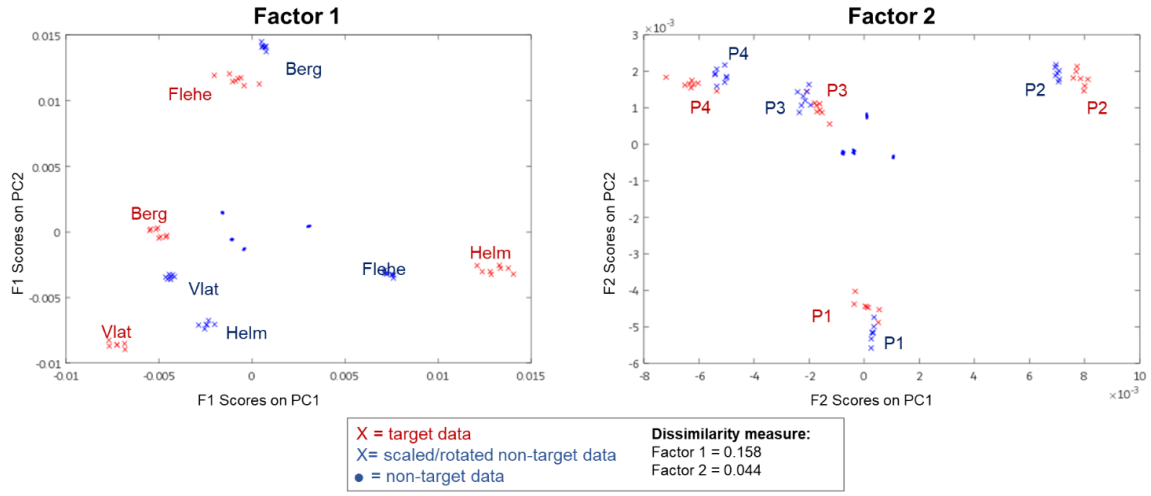


Figure S 6-18: Procrustes analysis of ASCA Scores plots of non-target data (positive mode) and target data for factor location (left) and time (right).

6.8.3.2.5 PLS-DA additional figures and table

All quality measures of different models are summarized in table S6-8. A passed permutation test shows that the probability of models to occur randomly was less than 0.01. No outliers or misclassified samples were observed in any model.

Table S 6-8: Quality measures of all different PLS-DA models.

			cross-validation			
	Nr. LV	explained variation [%]	sensitivity Y1;Y2*	specificity Y1;Y2*	class error rate Y1;Y2*	permutation test passed
	Factor 1 positive ionization					
Flehe	2	63.1	1;1	1;1	0;0	x
Berg	2	63.0	1;1	1;1	0;0	x
Helm	2	61.3	1;1	1;1	0;0	x
Vlat	3	70.4	1;1	1;1	0;0	x
	Factor 1 negative ionization					
Flehe	2	55.5	1;1	1;1	0;0	x
Berg	2	55.5	1;1	1;1	0;0	x
Helm	2	33.4	1;1	1;1	0;0	x
Vlat	2	42.3	1;1	1;1	0;0	x
	Factor 2 positive ionization					
with Flehe	3	67.5	1;0.93	0.93;1	0.04;0.04	x
without Flehe	3	61.0	0.9;1	1;0.9	0.05;0.05	x
	Factor 2 negative ionization					
with Flehe	3	59.3	1;0.93	0.93;1	0.04;0.04	x
without Flehe	3	64.3	1;1	1;1	0;0	x

* results for different classes of binary model (Y1&Y2) are given separated by ; *Nr. LV:number of latent variables

6.8.3.2.6 Seasonal modelling (factor 2) with and without Flehe samples

Because of the larger disparity between *Flehe* samples and the rest of the sampling locations, which was consistent with the data's inherent structure and different sampling dates for *Flehe* samples, additional seasonal PLS-DA modelling was done by excluding the *Flehe* samples from the data matrix. For both types of seasonal modelling VIP > 1 were selected and compared. Based on the results of this cross modelling different conclusions can be drawn: On the one hand, a group of pollutants with significant seasonal changes are only prioritized by the model including *Flehe* samples (f46, f164, f348, f625 and f1011) are given in figure S6-19 (A). Several of them show highly correlated concentration profiles (R^2>0.95), however, had different m/z and retention times which indicates this could be transformation products of the same compound. On the other hand, several OMPs were prioritized only in the model on the data matrix not including *Flehe* samples. For these features significant seasonal changes were observed which were covered up by high concentration fold change (FC) and opposing trends in *Flehe* samples figure S6-19 B).

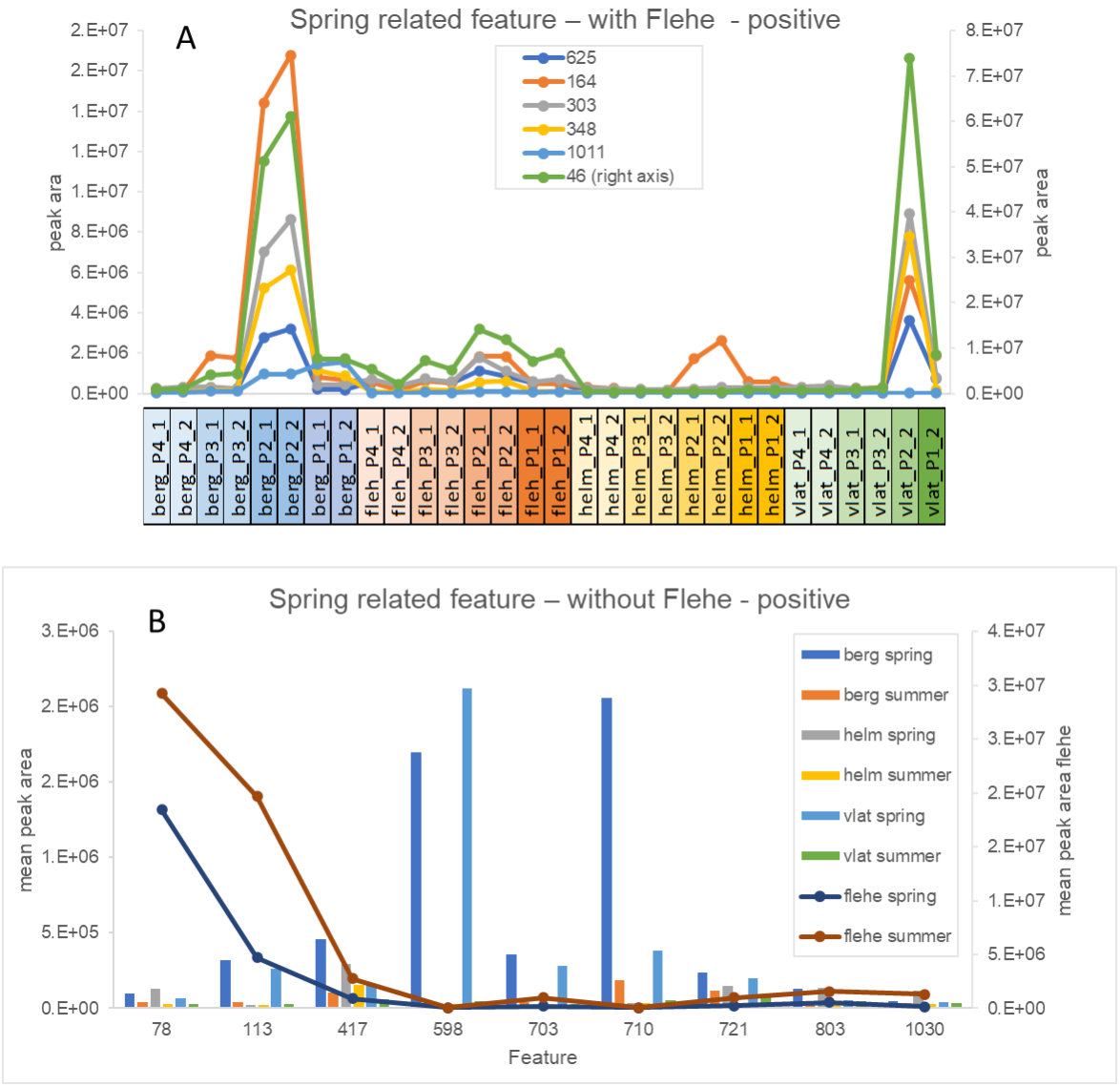


Figure S 6-19: Peak area profiles of features prioritized in models including Flehe samples (A) and without Flehe samples (B).

6.8.3.2.7 Feature prioritization

During the prioritization workflow and later identification, several groups of correlated features and doublings were discovered. Feature groups had highly correlated concentration profiles over different samples and had either same RT but different m/z (adducts or in-source fragments) or same m/z and same RT (falsely aligned features) and were thus caused by imperfection of the peak picking workflow in Compound Discoverer. Further, some groups with different m/z and slightly different RT (possible homologues) were discovered. To avoid double counting, these groups were summarized and only one “representative” feature was used for evaluation of prioritized features in figures and discussion about prioritized features and their identification.

In figure S6-20 two Volcano plots are exemplarily depicted for binary class modelling of *Berg* versus the rest of samples and *Spring* vs *Summer* samples for positive ionization mode. Accordingly for figure S6-20 (A) the upregulated features (upper right part of the plot) with colours yellow or orange are those selected as specific for *Berg*. As can be seen several features are located in the upper right corner and thus selected by Volcano criteria (green colour) but not meeting VIP criteria. In figure S6-20 (B) *Spring* specific (upregulated, upper right corner) and *Summer* specific (downregulated, upper left corner) are shown. Here in contrast to the Volcano plot of *Berg* modelling several features with VIP > 2 (yellow to red) were not selected by Volcano criteria, especially due to high p-values (below -1.3 on log 10 scale).

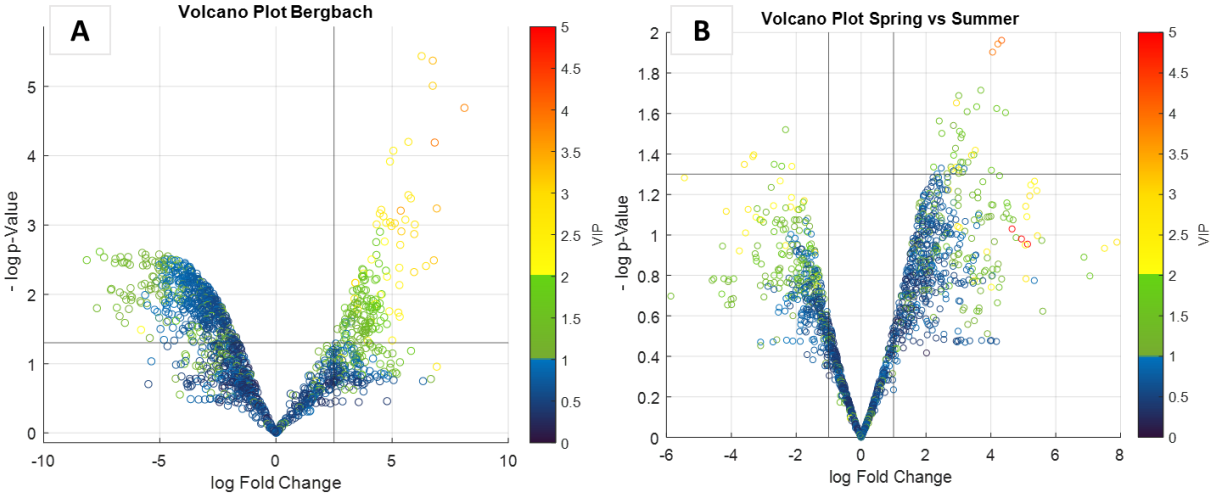
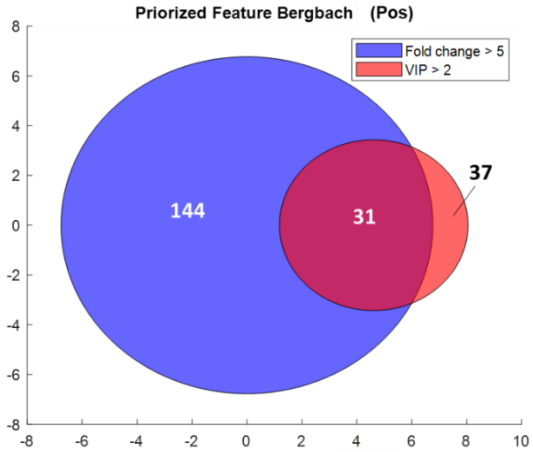
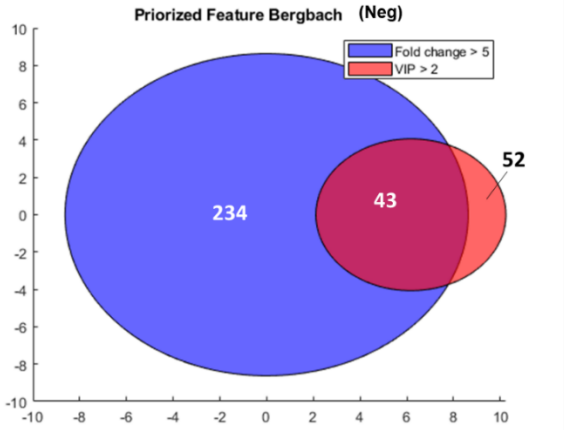
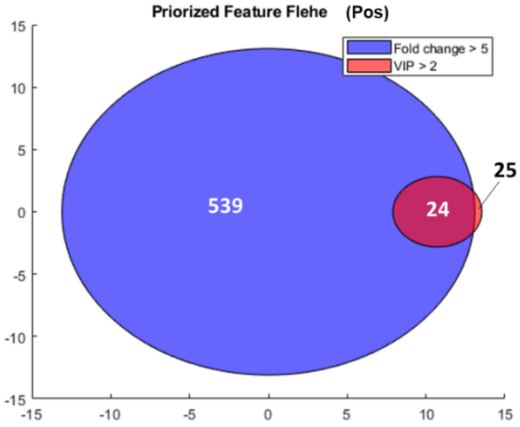
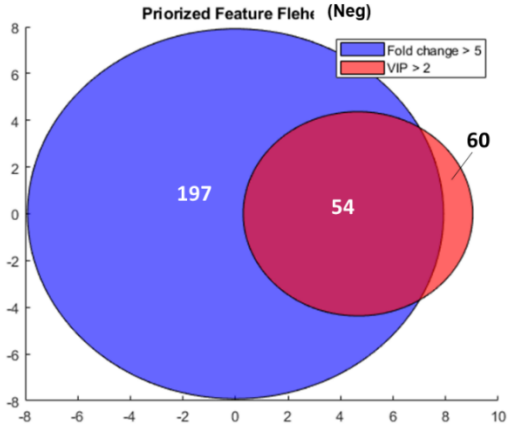


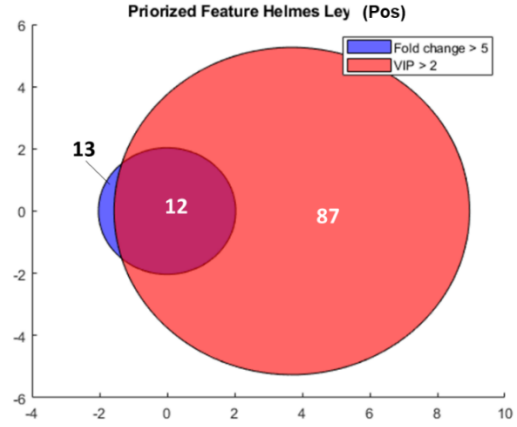
Figure S 6-20: Exemplary Volcano plots for Berg (A) and Spring vs Summer (B) for positive data. Color bars represent VIP-scores. Indicated lines represent thresholds for p-value 0.05 and Fold change 5 (A) and 2 (B).

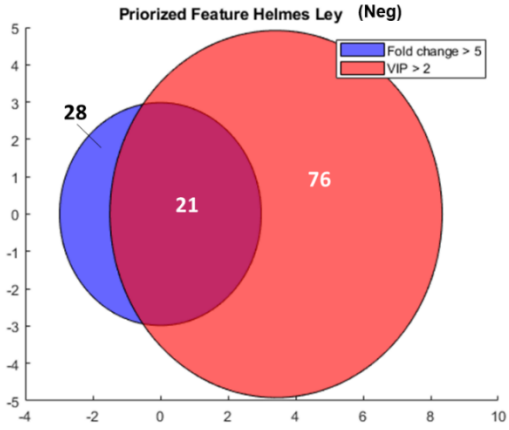

Priorized Feature Bergbach (Pos)
Fold change > 5
VIP > 2
144
31
37

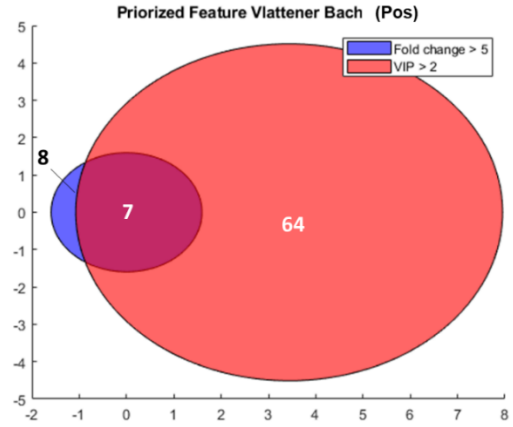

Priorized Feature Bergbach (Neg)
Fold change > 5
VIP > 2
234
43
52


Priorized Feature Flehe (Pos)
Fold change > 5
VIP > 2
539
24
25


Priorized Feature Flehe (Neg)
Fold change > 5
VIP > 2
197
54
60


Priorized Feature Helmes Ley (Pos)
Fold change > 5
VIP > 2
13
12
87


Priorized Feature Helmes Ley (Neg)
Fold change > 5
VIP > 2
28
21
76


Priorized Feature Vlattener Bach (Pos)
Fold change > 5
VIP > 2
8
7
64

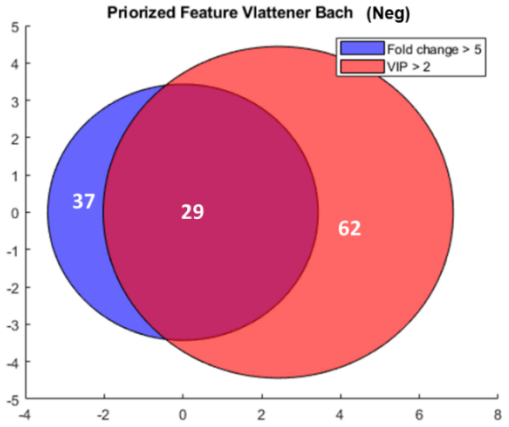

Priorized Feature Vlattener Bach (Neg)
Fold change > 5
VIP > 2
37
29
62

Figure S 6-21: Venn diagrams showing overlap of features meeting VIP criteria > 2 and Volcano criteria (Fold change > 5, p-value < 0.05) for different sampling sites for positive (left) and negative (right) data.

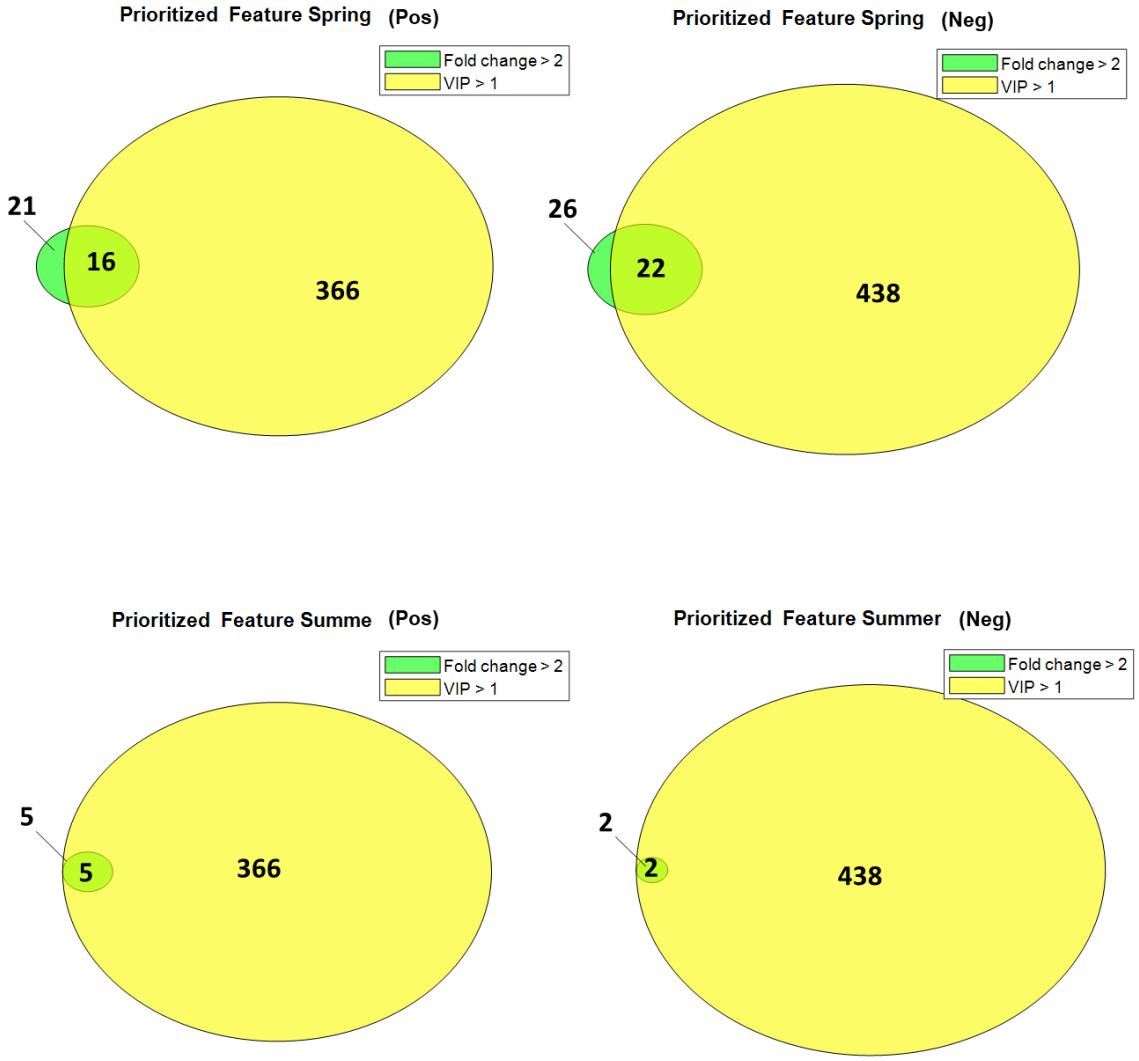


Figure S 6-22: Venn diagrams showing overlap of features meeting VIP criteria > 1 and Volcano criteria (Fold change > 2, p-value < 0.05) for different sampling seasons for positive (left) and negative (right) data.

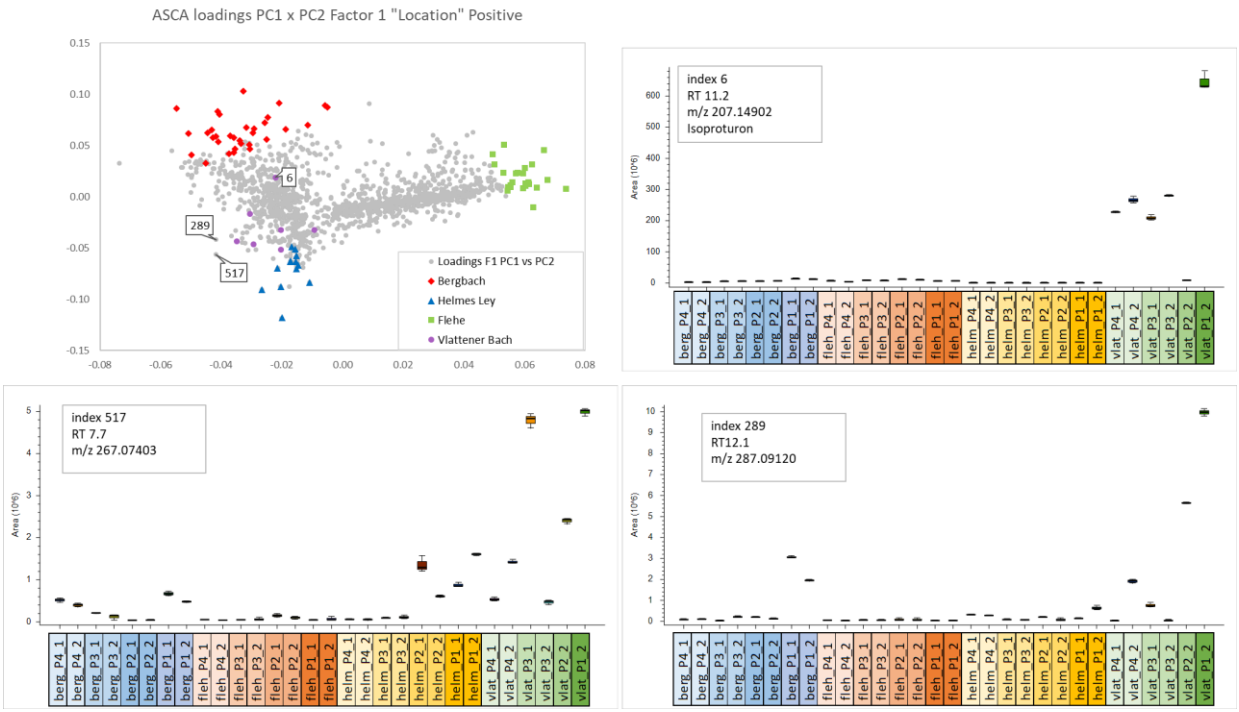


Figure S 6-23: Examples of peak area profiles over all samples of features with index 6, 289 and 517.

6.8.3.2.8 Heat map of prioritized feature and samples

The peak areas of the prioritized features for both factors and in both modes of ionization were analysed by simultaneous hierarchical clustering presented in a heat map (clustergram) type of plot. In the rows of the clustergram plot (y-axis) the samples are displayed and in the columns (x-axis) the prioritized features are listed. Results are displayed in figure S6-24 for factor 1 positive data (A) and negative data (B) and factor 2 positive data (C) and negative data (D): The heat map analysis of prioritized pollutants by PLS-DA and Volcano test show great discrimination among different sampling locations and sampling seasons. In figure S6-24 A and B, location-related pollutants have been categorized into four classes which is in accordance with four sampling sites. In addition, two blocks of large and two blocks of small rectangles can be detected for both modes of ionization. The large subsets of features associated with significantly higher intensities are allocated to Berg and Flehe with 31 and 24 pollutants for positive and 43 and 54 pollutants for negative data. The smaller subsets belong to the small rivers Helm and Vlat with 12 and 7 features for positive and 21 and 29 features for negative mode. In figure S6-24 C and D, the pollutants with drastic concentrations changes between different sampling seasons and predefined criteria are highlighted.

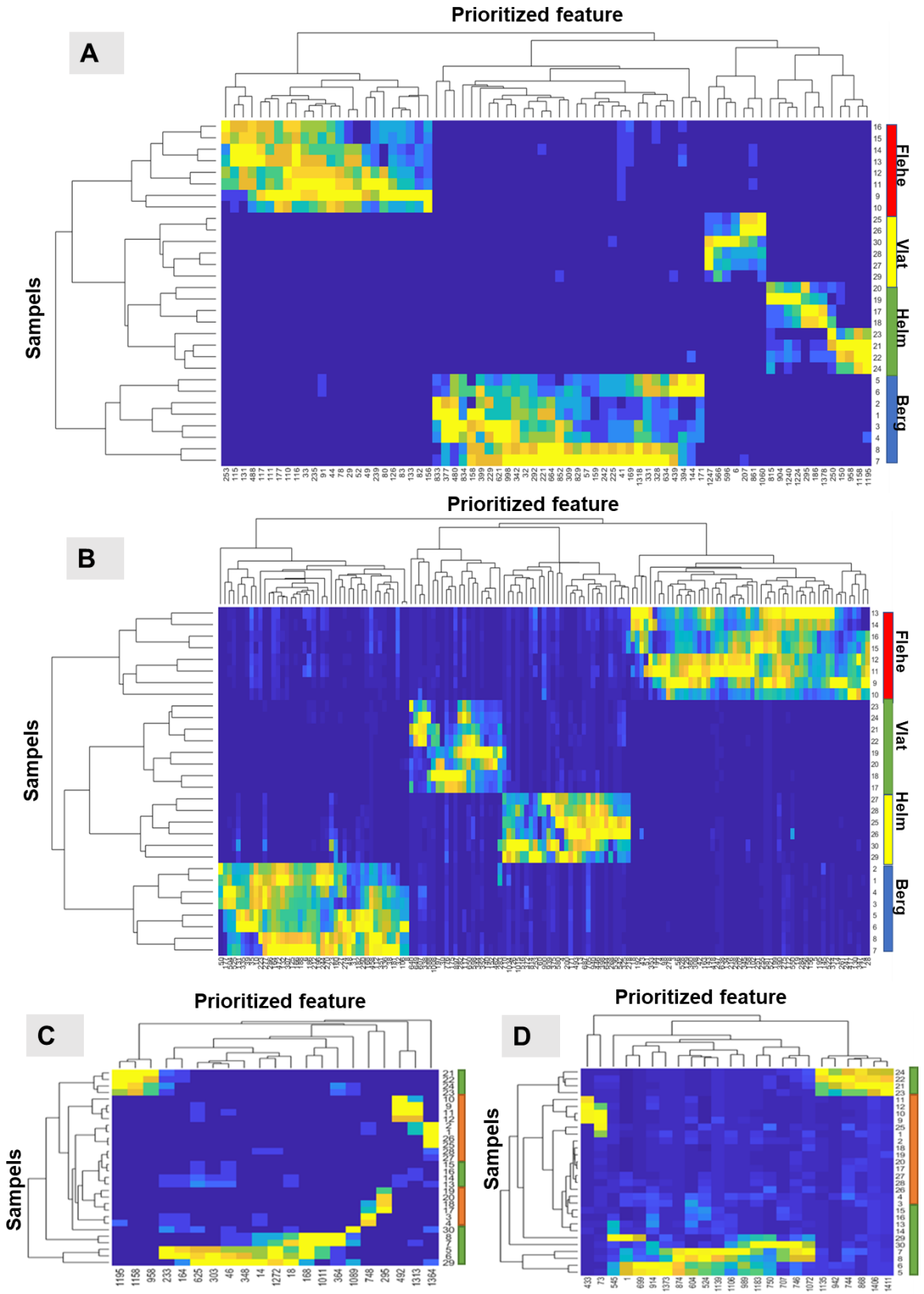


Figure S 6-24: Heatmaps of relative feature intensity changes for factor "sampling location", (A) and (B) and for factor "sampling season" (C) and (D). Dendrograms according to simultaneous h-clustering of both pollutants and samples are shown on the top and left of the heatmaps for positive (A,C) and negative (B,D) modes. The row and column indices show the samples (y-axis) and prioritized pollutants (x-axis), respectively. Colour bars on the right show the different sampling locations (blue: Berg, red: Flehe, gree: Helm and and yellow: Vlat) and different seasons (green for spring and orange for summer).

6.8.3.3 Identification

6.8.3.3.1 Identification level

To show the confidence of identification of prioritized features five identification levels were defined as shown in table S6-9. Molecular formulas were predicted with Compound Discoverer and only formulas with Δ ppm error below 2 ppm, without missed isotopes, and high percentage of coverage of theoretical and measured isotopic pattern (>80%) were included. Exact masses were searched in ChemSpider and ForIdent databases with a 5 ppm window. MS^2 spectra were searched in mzCloud with defined precursor masses.

Table S 6-9: Description of confidence level of identification.

1	RT and MS^2 match with reference standard
2A	Molecular formula, probable structure, MS^2 data base match
2B	Molecular formula, probable structure, match in several data bases, no MS^2 recorded
3	Molecular formula, several possible structures
4	m/z and RT

6.8.3.3.2 Micropollutant categories

For all prioritized features with identification level 1 and 2 information about their application and possible origin were collected. The m/z of prioritized features were additionally searched in suspect lists published by [206] including naturally occurring phytotoxins and by [63] including new relevant pesticide transformation products. Subsequently, level 1 and 2 identified features were assigned to different categories of pollutant origin to evaluate the overall influence of different point and non-point introduction pathways. The category “agriculture” includes herbicides, pesticides, fungicides, and related TPs, “WWTP” includes pharmaceuticals, related TPs, surfactants, PFAS and personal care products, “industry” includes known educts and intermediates for organic synthesis, and REACH chemicals, “plant” includes naturally occurring substances like phytotoxins or flavonoids. Pollutants in the categories “agriculture” and “plant” are most probably originating from non-point sources and those in the categories “industry” or “WWTP” are introduced by point sources. It has to be taken into account, that not all included prioritized features were unambiguously identified and emission pathways were not verified. Thus, only a rough estimation on dominance of different routes can be made.

Chapter 7 General conclusions and outlook

Non-target screening (NTS) approaches have a high potential to enable more holistic environmental monitoring. This thesis showed that the implementation and validation of sophisticated data processing strategies during the whole workflow are essential to exploit the full potential of complex NTS data. In different chapters, several aspects of NTS data processing have been studied.

For the step of feature extraction, a variety of commercial and open-source software tools are available. Although all of them pursue the same aim to extract analyte signals from raw data and many of them use similar strategies, the implementation of different algorithms and in addition, setting of parameters and thresholds can highly influence the outcome as shown in chapter 4. The observed inconsistencies between different software tools can impair data quality and the transferability of results. However, a general harmonization or standardization of data processing approaches is not feasible, as different instruments and sample types show different requirements regarding data processing. In addition, NTS is applied to many different research questions and used in different sectors such as academia, by regulative authorities, and in the (water) industry sector, all having different requirements on and resources for NTS. In addition, as has been pointed out previously, NTS can be applied as an explorative approach where a diversity of methodologies can lead to more discoveries of critical compounds [144]. Yet, a higher awareness of the impact of the feature extraction step and a transparent and detailed reporting of the data processing workflow, e.g., in a standardized reporting frame including implementation of filtering steps, all parameter settings, and criteria are necessary.

Furthermore, there is a lack of harmonized criteria that can be used to optimize and validate the feature extraction step. The reporting of false positive and false negative (FP/FN) detected features has been applied more broadly lately [49,208]. However, the differentiation between FP and true positive (TP) signals is so far mostly done on a subjective basis. An agreement on data quality estimation and the development of additional data quality criteria would thus be beneficial. The introduction of an NTS Guideline (2019) [54] has highly improved quality assurance during sampling, measurement, and reporting of identification confidence in the NTS community. The “Non-target Screening” expert committee of the German Water Chemistry Society provides a good platform to discuss possibilities to harmonize data quality reporting as well.

Another conclusion that can be drawn from the comparison of different feature extraction tools performed in this thesis is, that too little knowledge on the influence of different parameter settings for each approach exists or is available to the majority of end-users. Even for open-

access tools, the functionality of algorithms can often only be understood with high programming knowledge. Thus, a higher robustness of data processing tools and better documentation or guidance for end-users would be desirable. Tools for automated optimization of processing parameters [209], or minimization of false-positive peaks [210] have been proposed in the *metabolomic* field and can be a helpful support in this regard, however, they are unfortunately not available for all software tools.

Another aspect that can be derived from chapter 4 is that the "componentization" step, meaning the grouping of signals related to the same analyte, shows large differences between distinct programs. Indeed, this processing step has great potential to reduce the number of redundant signals in a data set [208]. However, the grouping of related signals is a challenging task, as during electrospray ionization besides the presence of isotopologues and formation of adducts, several unpredictable masses, e.g., formed by in-source fragmentation, can occur [52]. New advancements, for reliable and robust componentization, e.g. using elution profile or other correlation measures, would be of high interest to improve feature extraction.

An alternative feature extraction approach, which circumvents the step of componentization, was presented in chapter 5. Here, multivariate deconvolution of liquid chromatography- high-resolution mass spectrometry (LC-HRMS) raw data based on the combination of methods region of interest (ROI) and multivariate curve resolution alternating least squares (MCR-ALS) showed several advantages for the application in NTS. At the end of the ALS optimization process, the pure mass spectra of all relevant MCR-ALS components (similar to a feature) are obtained, which eliminates the requirement for grouping several features of the same compound. In addition, there is no need for chromatographic alignment or retention time shift corrections, as data matrices of different samples are column-wise augmented before resolution process. Complex set of samples containing coelution of several relevant signals and high interference of matrix constitutes can be simultaneously resolved mathematically by that approach [60].

MCR-ALS components are determined by iterative optimization cycles based on a set of initial estimates of components and under the action of certain constraints. An adequate selection of initial estimates and constraints is crucial to obtain reliable solutions [44] but can be challenging for scientists less experienced in chemometrics. To complete knowledge and give general recommendations on useful methods to estimate an initial component set and the selection of adequate constraints for this type of data, further studies are necessary. Those studies need to implement the ROI/MCR-ALS approach on NTS data obtained from different sample matrices and different HRMS systems. For a broad application of this promising feature extraction approach in NTS, its implementation in a user-friendly software program or platform and/or stronger collaboration with the chemometric community is necessary. In addition, in

chapter 5 several multivariate chemometric methods for feature prioritization were implemented, providing a comprehensive data processing workflow.

Apart from robust and reliable feature extraction, the prioritization of relevant features is crucial in NTS to extract the information of interest. Possibilities of prioritization strategies are just as diverse as the scope of application of NTS to different research questions. Besides a selection of relevant features based on suspect lists or based on feature properties or occurrence frequencies etc. as discussed in detail in the introduction part 2.2.1 due to the multivariate nature and complexity of NTS data chemometrics-based strategies have a high potential for an efficient selection of a subset of features that reveal similarities and/or differences between samples.

In chapter 6 the benefits of feature prioritization based on multiple complementary multivariate chemometric methods for NTS data were emphasized. Temporal and spatial trends on a data set were analysed with different complementary unsupervised (PCA: principal component analysis and HCA: hierarchical cluster analysis) and supervised (ASCA: ANOVA simultaneous component analysis and PLS-DA: partial least squares discriminant analysis) chemometric approaches. The contribution of each approach to an overall deeper understanding of samples and hidden pollution patterns was illustrated. Chapter 6 also highlighted the importance of a well-designed experiment according to a multi-factorial scenario. In fact, chemometric data analysis and modelling requires an adequate set of samples to be able to make reliable interpretations on the real impact of involved factors especially for high-dimensional data.

Recently, further, highly interesting multivariate prioritization approaches have been published using hierarchical cluster analysis [68,211], a feature-based molecular networking [76], neural network and multiple factor analysis [74], multivariate empirical Bayes approach [78], a group-wise PCA [77], group-wise ASCA [75]. This emerging trend in NTS shows that a diversity of methods can be used for data mining and the development and/or application of new approaches is expected which offers the opportunity for a more comprehensive evaluation of environmental samples.

Before implementation of chemometric tools, data have to be pre-treated to make them comparable by steps like transformation, scaling, normalization, etc. to reduce the skewness of data and/or systematic/random variation. This aspect has received very little attention so far for LC-HRMS data in water analysis. However, as these data pre-treatment steps are crucial before the implementation of any chemometric method [81], a detailed investigation and comparison of existing methods and the development of new methods for NTS data would be highly necessary. In addition, NTS datasets often show difficulties like a large number of missing values and in most cases a by far higher number of variables (features) compared to

samples. For the imputation of missing values, several approaches can be applied [90], however, a detailed comparison of them for NTS data has not been done so far.

An open access dataset obtained from different HRMS systems and including different sample matrices, with an as a high as possible extent of known properties and other meta-data, would be highly useful to evaluate new data processing tools focussing on any remaining challenge of feature extraction, prioritization, data pre-treatment, identification etc. discussed in this chapter

In addition, for both feature extraction as well as feature prioritization in other scientific communities using LC-HRMS data (*metabolomics, proteomics, lipidomics, foodomics*, ...) different advanced chemometric strategies are much more established. An interdisciplinary scientific exchange can thus greatly benefit the advancement of LC-HRMS screening and data processing [38].

Finally, in chapter 6 it was shown that with an impartial chemometric based prioritization workflow without any a priori limitation by library or database entries a high number of completely unknown features appear as relevant which can ultimately not be identified. This shows that unknown identification remains a bottleneck in NTS data processing, despite the fact that for identification of unknown water relevant OMPs great developments have been achieved in the last years. Spectral databases like *MassBank* or *mzCloud* have been largely extended for LC-HRMS spectra and databases like *ForIdent*, with a focus on water-relevant OMPs, have been introduced. In addition, *In silico* fragmentation prediction tools as, e.g., included in *MetFrag,* can assist in confirmation of possible candidates where no library spectra are available [97]. However, all databases are limited to known structures, and so-called “unknown unknowns” or “true unknowns”, e.g., most transformation products, are not included and remain unidentified. Orthogonal analytical techniques like nuclear magnetic resonance spectroscopy (NMR) can be powerful tools for the improvement of true unknown identification, but usually require high substance amounts and purity grades which can be limiting factors for the application of NMR on OMPs in water [39]. Further, MS^2 spectra detection e.g. with data-independent methods (DIA) and the development of advanced DIA-MS^2 spectra deconvolution algorithms have a high potential to improve feature identification [35]. Promising approaches using DIA spectra without creation of instrument specific spectrum libraries have been proposed for *proteomics* data [212].

Not to forget, there exist further fields of development in NTS, besides data processing strategies. In chapter 3 of this thesis, instrumental aspects of data acquisition have been studied. Even with the best NTS methods, some chemicals remain outside the analytical window, such as those that elute too early or late from the column, are poorly ionized by existing ionization methods, or are not captured during sampling. Instrumental, sampling and

sample treatment developments with the purpose of closing those gaps, e.g. focussing on more polar compounds [39], have been reported.

Finally, coupling NTS data with further scientific disciplines like bioanalytics will be a field of development in the future and a completion to advanced prioritization strategies. For example, the coupling with effect-directed analysis (EDA), e.g. using cell-based in vitro bioassays to capture mixture effects, has been presented as a promising prioritization strategy [110,213]. In addition, merging NTS data with further "omics" data such as *metagenomics, transcriptomics*, etc. could be of high interest in the future. However, this type of "data fusion" will bring up further challenges to data processing [90].

In conclusion, exciting developments in the field of NTS, focussing on different remaining challenges of data processing at all workflow stages, are to be expected in the future. With these advancements, the full potential of NTS as a tool to complete our knowledge and understanding of environmental pollution patterns, risks, remediation strategies, etc. will be achieved. This thesis contributes to these advancements by highlighting the importance of the validation of data processing and the potential of the implementation of advanced multivariate chemometric tools for both feature extraction as well as prioritization.

References

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

[2] R. Loos, B.M. Gawlik, G. Locoro, E. Rimaviciute, S. Contini, G. Bidoglio, EU-wide survey of polar organic persistent pollutants in European river waters, Environ. Pollut. 157 (2009) 561–568. https://doi.org/10.1016/j.envpol.2008.09.020.

[3] M. Ruff, M.S. Mueller, M. Loos, H.P. Singer, Quantitative target and systematic non-target analysis of polar organic micro-pollutants along the river Rhine using high-resolution mass-spectrometry--Identification of unknown sources and compounds, Water Res. 87 (2015) 145–154. https://doi.org/10.1016/j.watres.2015.09.017.

[4] R. Loos, G. Locoro, S. Comero, S. Contini, D. Schwesig, F. Werres, P. Balsaa, O. Gans, S. Weiss, L. Blaha, M. Bolchi, B.M. Gawlik, Pan-European survey on the occurrence of selected polar organic persistent pollutants in ground water, Water Res. 44 (2010) 4115–4126. https://doi.org/10.1016/j.watres.2010.05.032.

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

[6] S. Schulze, D. Zahn, R. Montes, R. Rodil, J.B. Quintana, T.P. Knepper, T. Reemtsma, U. Berger, Occurrence of emerging persistent and mobile organic contaminants in European water samples, Water Res. 153 (2019) 80–90. https://doi.org/10.1016/j.watres.2019.01.008.

[7] Y. Luo, W. Guo, H.H. Ngo, L.D. Nghiem, F.I. Hai, J. Zhang, S. Liang, X.C. Wang, A review on the occurrence of micropollutants in the aquatic environment and their fate and removal during wastewater treatment, Sci. Total Environ. 473-474 (2014) 619–641. https://doi.org/10.1016/j.scitotenv.2013.12.065.

[8] K. Halbach, M. Möder, S. Schrader, L. Liebmann, R.B. Schäfer, A. Schneeweiss, V.C. Schreiner, P. Vormeier, O. Weisner, M. Liess, T. Reemtsma, Small streams-large concentrations? Pesticide monitoring in small agricultural streams in Germany during dry weather and rainfall, Water Res. 203 (2021) 117535. https://doi.org/10.1016/j.watres.2021.117535.

[9] M. Liess, L. Liebmann, P. Vormeier, O. Weisner, R. Altenburger, D. Borchardt, W. Brack, A. Chatzinotas, B. Escher, K. Foit, R. Gunold, S. Henz, K.L. Hitzfeld, M. Schmitt-Jansen, N. Kamjunke, O. Kaske, S. Knillmann, M. Krauss, E. Küster, M. Link, M. Lück, M. Möder, A. Müller, A. Paschke, R.B. Schäfer, A. Schneeweiss, V.C. Schreiner, T. Schulze, G. Schüürmann, W. von Tümpling, M. Weitere, J. Wogram, T. Reemtsma, Pesticides are the dominant stressors for vulnerable insects in lowland streams, Water Res. 201 (2021) 117262. https://doi.org/10.1016/j.watres.2021.117262.

[10] Y. Zhang, C.F. Marrs, C. Simon, C. Xi, Wastewater treatment contributes to selective increase of antibiotic resistance among Acinetobacter spp, Sci. Total Environ. 407 (2009) 3702–3706. https://doi.org/10.1016/j.scitotenv.2009.02.013.

[11] J. Kuckelkorn, R. Redelstein, T. Heide, J. Kunze, S. Maletz, P. Waldmann, T. Grummt, T.-B. Seiler, H. Hollert, A hierarchical testing strategy for micropollutants in drinking water regarding their potential endocrine-disrupting effects-towards health-related indicator values, Environ. Sci. Pollut. Res. Int. 25 (2018) 4051–4065. https://doi.org/10.1007/s11356-017-0155-3.

[12] P.A. Neale, G. Braun, W. Brack, E. Carmona, R. Gunold, M. König, M. Krauss, L. Liebmann, M. Liess, M. Link, R.B. Schäfer, R. Schlichting, V.C. Schreiner, T. Schulze, P. Vormeier, O. Weisner, B.I. Escher, Assessing the Mixture Effects in In Vitro Bioassays of Chemicals Occurring in Small Agricultural Streams during Rain Events, Environ. Sci. Technol. 54 (2020) 8280–8290. https://doi.org/10.1021/acs.est.0c02235.

[13] I. Delpla, A.-V. Jung, E. Baures, M. Clement, O. Thomas, Impacts of climate change on surface water quality in relation to drinking water production, Environ. Int. 35 (2009) 1225–1233. https://doi.org/10.1016/j.envint.2009.07.001.

[14] Z. Wang, G.W. Walker, D.C.G. Muir, K. Nagatani-Yoshida, Toward a Global Understanding of Chemical Pollution: A First Comprehensive Analysis of National and Regional Chemical Inventories, Environ. Sci. Technol. 54 (2020) 2575–2584. https://doi.org/10.1021/acs.est.9b06379.

[15] W. Brack, D. Barcelo Culleres, A.B.A. Boxall, H. Budzinski, S. Castiglioni, A. Covaci, V. Dulio, B.I. Escher, P. Fantke, F. Kandie, D. Fatta-Kassinos, F.J. Hernández, K. Hilscherová, J. Hollender, H. Hollert, A. Jahnke, B. Kasprzyk-Hordern, S.J. Khan, A. Kortenkamp, K. Kümmerer, B. Lalonde, M.H. Lamoree, Y. Levi, P.A. Lara Martín, C.C. Montagner, C. Mougin, T. Msagati, J. Oehlmann, L. Posthuma, M. Reid, M. Reinhard, S.D. Richardson, P. Rostkowski, E. Schymanski, F. Schneider, J. Slobodnik, Y. Shibata, S.A. Snyder, F. Fabriz Sodré, I. Teodorovic, K.V. Thomas, G.A. Umbuzeiro, P.H. Viet, K.G. Yew-Hoong, X. Zhang, E. Zuccato, One planet: One health. A call to support the initiative on a global science-policy body on chemicals and waste, Environ Sci Eur 34 (2022) 21. https://doi.org/10.1186/s12302-022-00602-6.

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

[17] U. Hübner, U. von Gunten, M. Jekel, Evaluation of the persistence of transformation products from ozonation of trace organic compounds - a critical review, Water Res. 68 (2015) 150–170. https://doi.org/10.1016/j.watres.2014.09.051.

[18] 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.

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

[20] B. Schulze, D. van Herwerden, I. Allan, L. Bijlsma, N. Etxebarria, M. Hansen, S. Merel, B. Vrana, R. Aalizadeh, B. Bajema, F. Dubocq, G. Coppola, A. Fildier, P. Fialová, E. Frøkjær, R. Grabic, P. Gago-Ferrero, T. Gravert, J. Hollender, N. Huynh, G. Jacobs, T. Jonkers, S. Kaserzon, M. Lamoree, J. Le Roux, T. Mairinger, C. Margoum, G. Mascolo, E. Mebold, F. Menger, C. Miège, J. Meijer, R. Moilleron, S. Murgolo, M. Peruzzo, M. Pijnappels, M. Reid, C. Roscioli, C. Soulier, S. Valsecchi, N. Thomaidis, E. Vulliet, R. Young, S. Samanipour, Inter-laboratory mass spectrometry dataset based on passive sampling of drinking water for non-target analysis, Sci. Data 8 (2021) 223. https://doi.org/10.1038/s41597-021-01002-w.

[21] M.C. McMaster, LC/MS: A practical user's guide, John Wiley, Hoboken, N.J, 2005.

[22] S. Kromidas, Optimization in HPLC, Wiley, 2021.

[23] R. Romero-Gonzáles, A.G. Frenich (Eds.), Applications in High Resolution Mass Spectrometry: Food Safety and Pesticide Residue Analysis, Elsevier, Amsterdam, Oxford, Cambridge, 2017.

[24] 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.
[25] A. Beschnitt, M. Schwikowski, T. Hoffmann, Towards comprehensive non-target screening using heart-cut two-dimensional liquid chromatography for the analysis of organic atmospheric tracers in ice cores, J. Chromatogr. A 1661 (2022) 462706. https://doi.org/10.1016/j.chroma.2021.462706.
[26] S. Bieber, G. Greco, S. Grosse, T. Letzel, RPLC-HILIC and SFC with Mass Spectrometry: Polarity-Extended Organic Molecule Screening in Environmental (Water) Samples, Anal. Chem. 89 (2017) 7907–7914. https://doi.org/10.1021/acs.analchem.7b00859.
[27] O. Höcker, D. Flottmann, T.C. Schmidt, C. Neusüß, Non-targeted LC-MS and CE-MS for biomarker discovery in bioreactors: Influence of separation, mass spectrometry and data processing tools, Sci. Total Environ. 798 (2021) 149012. https://doi.org/10.1016/j.scitotenv.2021.149012.
[28] T. de Vijlder, D. Valkenborg, F. Lemière, E.P. Romijn, K. Laukens, F. Cuyckens, A tutorial in small molecule identification via electrospray ionization-mass spectrometry: The practical art of structural elucidation, Mass Spectrom. Rev. 37 (2018) 607–629. https://doi.org/10.1002/mas.21551.
[29] B.O. Keller, J. Sui, A.B. Young, R.M. Whittal, Interferences and contaminants encountered in modern mass spectrometry, Anal. Chim. Acta 627 (2008) 71–81. https://doi.org/10.1016/j.aca.2008.04.043.
[30] K.K. Murray, R.K. Boyd, M.N. Eberlin, G.J. Langley, L. Li, Y. Naito, Definitions of terms relating to mass spectrometry (IUPAC Recommendations 2013), Pure and Applied Chemistry 85 (2013) 1515–1609. https://doi.org/10.1351/PAC-REC-06-04-06.
[31] A.G. Brenton, A.R. Godfrey, Accurate mass measurement: Terminology and treatment of data, J. Am. Soc. Mass Spectrom. 21 (2010) 1821–1835. https://doi.org/10.1016/j.jasms.2010.06.006.
[32] S. Eliuk, A. Makarov, Evolution of Orbitrap Mass Spectrometry Instrumentation, Annu. Rev. Anal. Chem. (Palo Alto Calif) 8 (2015) 61–80. https://doi.org/10.1146/annurev-anchem-071114-040325.
[33] A. Agüera, A. Martı´nez-Piernas, M. Campos-Mananas, Analytical Strategies Used in HRMS, in: R. Romero-Gonzáles, A.G. Frenich (Eds.), Applications in High Resolution Mass Spectrometry: Food Safety and Pesticide Residue Analysis, Elsevier, Amsterdam, Oxford, Cambridge, 2017.
[34] D.C. Liebler, Shotgun mass spec goes independent, Nat. Methods 1 (2004) 16–17. https://doi.org/10.1038/nmeth1004-16.
[35] J. Guo, T. Huan, Comparison of Full-Scan, Data-Dependent, and Data-Independent Acquisition Modes in Liquid Chromatography-Mass Spectrometry Based Untargeted Metabolomics, Anal. Chem. 92 (2020) 8072–8080. https://doi.org/10.1021/acs.analchem.9b05135.
[36] C. Hug, N. Ulrich, T. Schulze, W. Brack, M. Krauss, Identification of novel micropollutants in wastewater by a combination of suspect and nontarget screening, Environ. Pollut. 184 (2014) 25–32. https://doi.org/10.1016/j.envpol.2013.07.048.
[37] E.L. Schymanski, H.P. Singer, J. Slobodnik, I.M. Ipolyi, P. Oswald, M. Krauss, T. Schulze, P. Haglund, T. Letzel, S. Grosse, N.S. Thomaidis, A. Bletsou, C. Zwiener, M. Ibáñez, T. Portolés, R. de Boer, M.J. Reid, M. Onghena, U. Kunkel, W. Schulz, A. Guillon, N. Noyon, G. Leroy, P. Bados, S. Bogialli, D. Stipaničev, P. Rostkowski, J. Hollender, Non-target screening with high-resolution mass spectrometry: Critical review using a collaborative trial on water analysis, Anal. Bioanal. Chem. 407 (2015) 6237–6255. https://doi.org/10.1007/s00216-015-8681-7.

[38] J. Hollender, E.L. Schymanski, H.P. Singer, P.L. Ferguson, Nontarget Screening with High Resolution Mass Spectrometry in the Environment: Ready to Go?, Environ. Sci. Technol. 51 (2017) 11505–11512. https://doi.org/10.1021/acs.est.7b02184.
[39] F. Menger, P. Gago-Ferrero, K. Wiberg, L. Ahrens, 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 (2020) e00102. https://doi.org/10.1016/j.teac.2020.e00102.
[40] V. Hinnenkamp, P. Balsaa, T.C. Schmidt, Quantitative screening and prioritization based on UPLC-IM-Q-TOF-MS as an alternative water sample monitoring strategy, Anal. Bioanal. Chem. 411 (2019) 6101–6110. https://doi.org/10.1007/s00216-019-01994-w.
[41] B. Schulze, Y. Jeon, S. Kaserzon, A.L. Heffernan, P. Dewapriya, J. O'Brien, M.J. Gomez Ramos, S. Ghorbani Gorji, J.F. Mueller, K.V. Thomas, S. Samanipour, An assessment of quality assurance/quality control efforts in high resolution mass spectrometry non-target workflows for analysis of environmental samples, TrAC Trends in Analytical Chemistry 133 (2020) 116063. https://doi.org/10.1016/j.trac.2020.116063.
[42] E. Gorrochategui, J. Jaumot, S. Lacorte, R. Tauler, Data analysis strategies for targeted and untargeted LC-MS metabolomic studies: Overview and workflow, TrAC Trends in Analytical Chemistry 82 (2016) 425–442. https://doi.org/10.1016/j.trac.2016.07.004.
[43] R. Stolt, R.J.O. Torgrip, J. Lindberg, L. Csenki, J. Kolmert, I. Schuppe-Koistinen, S.P. Jacobsson, Second-order peak detection for multicomponent high-resolution LC/MS data, Anal. Chem. 78 (2006) 975–983. https://doi.org/10.1021/ac050980b.
[44] E. Gorrochategui, J. Jaumot, R. Tauler, ROIMCR: A powerful analysis strategy for LC-MS metabolomic datasets, BMC Bioinformatics 20 (2019) 256. https://doi.org/10.1186/s12859-019-2848-8.
[45] T. Pluskal, S. Castillo, A. Villar-Briones, M. Oresic, MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data, BMC Bioinformatics 11 (2010) 395. https://doi.org/10.1186/1471-2105-11-395.
[46] Martin Loos, Envimass Version 3.1, Zenodo, 2016.
[47] C.A. Smith, E.J. Want, G. O'Maille, R. Abagyan, G. Siuzdak, XCMS: Processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification, Anal. Chem. 78 (2006) 779–787. https://doi.org/10.1021/ac051437y.
[48] R. Tautenhahn, G.J. Patti, D. Rinehart, G. Siuzdak, XCMS Online: A web-based platform to process untargeted metabolomic data, Anal. Chem. 84 (2012) 5035–5039. https://doi.org/10.1021/ac300698c.
[49] C. Dietrich, A. Wick, T.A. Ternes, Open-source feature detection for non-target LC-MS analytics, Rapid Commun. Mass Spectrom. 36 (2022) e9206. https://doi.org/10.1002/rcm.9206.
[50] R. Helmus, T.L. Ter Laak, A.P. van Wezel, P. de Voogt, E.L. Schymanski, patRoon: Open source software platform for environmental mass spectrometry based non-target screening, J. Cheminform. 13 (2021) 1. https://doi.org/10.1186/s13321-020-00477-w.
[51] M. Katajamaa, M. Oresic, Data processing for mass spectrometry-based metabolomics, J. Chromatogr. A 1158 (2007) 318–328. https://doi.org/10.1016/j.chroma.2007.04.021.
[52] C. Kuhl, R. Tautenhahn, C. Böttcher, T.R. Larson, S. Neumann, CAMERA: An integrated strategy for compound spectra extraction and annotation of liquid chromatography/mass spectrometry data sets, Anal. Chem. 84 (2012) 283–289. https://doi.org/10.1021/ac202450g.
[53] M. Oehme, U. Berger, S. Brombacher, F. Kuhn, S. Kölliker, Trace analysis by HPLC-MS: Contamination problems and systematic errors, TrAC Trends in Analytical Chemistry 21 (2002) 322–331. https://doi.org/10.1016/S0165-9936(02)00503-4.

[54] W. Schulze, T. Lucke, Non-Target Screening in Water Analysis: Guideline for the application of LC-ESI-HRMS for screening, first ed., Mülheim an der Ruhr, 2019.
[55] T. Bader, W. Schulz, K. Kümmerer, R. Winzenbacher, General strategies to increase the repeatability in non-target screening by liquid chromatography-high resolution mass spectrometry, Anal. Chim. Acta 935 (2016) 173–186. https://doi.org/10.1016/j.aca.2016.06.030.
[56] R. Tauler, A. Smilde, B. Kowalski, Selectivity, local rank, three-way data analysis and ambiguity in multivariate curve resolution, J. Chemometrics 9 (1995) 31–58. https://doi.org/10.1002/cem.1180090105.
[57] C. Perez-Lopez, A. Ginebreda, M. Carrascal, D. Barcelò, J. Abian, R. Tauler, Non-target protein analysis of samples from wastewater treatment plants using the regions of interest-multivariate curve resolution (ROIMCR) chemometrics method, Journal of Environmental Chemical Engineering 9 (2021) 105752. https://doi.org/10.1016/j.jece.2021.105752.
[58] I. Santana, M. Breitkreitz, L. Pinto, Multivariate Curve Resolution Alternating Least Squares applied to Chromatographic Data: From the Basics to the Recent Advances, Braz. J. Anal. Chem. 8 (2021). https://doi.org/10.30744/brjac.2179-3425.RV-30-2021.
[59] A. de Juan, J. Jaumot, R. Tauler, Multivariate Curve Resolution (MCR). Solving the mixture analysis problem, Anal. Methods 6 (2014) 4964–4976. https://doi.org/10.1039/c4ay00571f.
[60] M. Vosough, M. Rashvand, H.M. Esfahani, K. Kargosha, A. Salemi, Direct analysis of six antibiotics in wastewater samples using rapid high-performance liquid chromatography coupled with diode array detector: a chemometric study towards green analytical chemistry, Talanta 135 (2015) 7–17. https://doi.org/10.1016/j.talanta.2014.12.036.
[61] M. Navarro-Reig, J. Jaumot, A. García-Reiriz, R. Tauler, Evaluation of changes induced in rice metabolome by Cd and Cu exposure using LC-MS with XCMS and MCR-ALS data analysis strategies, Anal. Bioanal. Chem. 407 (2015) 8835–8847. https://doi.org/10.1007/s00216-015-9042-2.
[62] B. Khakimov, J.M. Amigo, S. Bak, S.B. Engelsen, Plant metabolomics: Resolution and quantification of elusive peaks in liquid chromatography-mass spectrometry profiles of complex plant extracts using multi-way decomposition methods, J. Chromatogr. A 1266 (2012) 84–94. https://doi.org/10.1016/j.chroma.2012.10.023.
[63] K. Kiefer, A. Müller, H. Singer, J. Hollender, New relevant pesticide transformation products in groundwater detected using target and suspect screening for agricultural and urban micropollutants with LC-HRMS, Water Res. 165 (2019) 114972. https://doi.org/10.1016/j.watres.2019.114972.
[64] B.F. Günthardt, J. Hollender, M. Scheringer, K. Hungerbühler, M.Y. Nanusha, W. Brack, T.D. Bucheli, Aquatic occurrence of phytotoxins in small streams triggered by biogeography, vegetation growth stage, and precipitation, Sci. Total Environ. 798 (2021) 149128. https://doi.org/10.1016/j.scitotenv.2021.149128.
[65] N. Park, Y. Choi, D. Kim, K. Kim, J. Jeon, Prioritization of highly exposable pharmaceuticals via a suspect/non-target screening approach: A case study for Yeongsan River, Korea, Sci. Total Environ. 639 (2018) 570–579. https://doi.org/10.1016/j.scitotenv.2018.05.081.
[66] A.C. Chiaia-Hernandez, E.L. Schymanski, P. Kumar, H.P. Singer, J. Hollender, Suspect and nontarget screening approaches to identify organic contaminant records in lake sediments, Anal. Bioanal. Chem. 406 (2014) 7323–7335. https://doi.org/10.1007/s00216-014-8166-0.
[67] M.P. Schlüsener, U. Kunkel, T.A. Ternes, Quaternary Triphenylphosphonium Compounds: A New Class of Environmental Pollutants, Environ. Sci. Technol. 49 (2015) 14282–14291. https://doi.org/10.1021/acs.est.5b03926.
[68] L.-M. Beckers, W. Brack, J.P. Dann, M. Krauss, E. Müller, T. Schulze, Unraveling longitudinal pollution patterns of organic micropollutants in a river by non-target screening

and cluster analysis, Sci. Total Environ. 727 (2020) 138388. https://doi.org/10.1016/j.scitotenv.2020.138388.

[69] J.M. Weiss, E. Simon, G.J. Stroomberg, R. de Boer, J. de Boer, S.C. van der Linden, P.E.G. Leonards, M.H. Lamoree, Identification strategy for unknown pollutants using high-resolution mass spectrometry: Androgen-disrupting compounds identified through effect-directed analysis, Anal. Bioanal. Chem. 400 (2011) 3141–3149. https://doi.org/10.1007/s00216-011-4939-x.

[70] J.E. Schollee, E.L. Schymanski, J. Hollender, Statistical Approaches for LC-HRMS Data To Characterize, Prioritize, and Identify Transformation Products from Water Treatment Processes, in: J.E. Drewes, T. Letzel (Eds.), Assessing transformation products of chemicals by non-target and suspect screening: Strategies and workflows, American Chemical Society; Distributed in print by Oxford University Press, Washington, DC, Oxford, 2016]-, pp. 45–65.

[71] J.E. Schollée, E.L. Schymanski, S.E. Avak, M. Loos, J. Hollender, Prioritizing Unknown Transformation Products from Biologically-Treated Wastewater Using High-Resolution Mass Spectrometry, Multivariate Statistics, and Metabolic Logic, Anal. Chem. 87 (2015) 12121–12129. https://doi.org/10.1021/acs.analchem.5b02905.

[72] T. Bader, W. Schulz, K. Kümmerer, R. Winzenbacher, LC-HRMS Data Processing Strategy for Reliable Sample Comparison Exemplified by the Assessment of Water Treatment Processes, Anal. Chem. 89 (2017) 13219–13226. https://doi.org/10.1021/acs.analchem.7b03037.

[73] F. Itzel, N. Baetz, L.L. Hohrenk, L. Gehrmann, D. Antakyali, T.C. Schmidt, J. Tuerk, Evaluation of a biological post-treatment after full-scale ozonation at a municipal wastewater treatment plant, Water Res. 170 (2020) 115316. https://doi.org/10.1016/j.watres.2019.115316.

[74] V. Nikolopoulou, R. Aalizadeh, M.-C. Nika, N.S. Thomaidis, TrendProbe: Time profile analysis of emerging contaminants by LC-HRMS non-target screening and deep learning convolutional neural network, J. Hazard. Mater. 428 (2022) 128194. https://doi.org/10.1016/j.jhazmat.2021.128194.

[75] R. Lotfi 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, Microchemical Journal 160 (2021) 105641. https://doi.org/10.1016/j.microc.2020.105641.

[76] D. Oberleitner, R. Schmid, W. Schulz, A. Bergmann, C. Achten, Feature-based molecular networking for identification of organic micropollutants including metabolites by non-target analysis applied to riverbank filtration, Anal. Bioanal. Chem. 413 (2021) 5291–5300. https://doi.org/10.1007/s00216-021-03500-7.

[77] K. Purschke, M. Vosough, J. Leonhardt, M. Weber, T.C. Schmidt, Evaluation of Nontarget Long-Term LC-HRMS Time Series Data Using Multivariate Statistical Approaches, Anal. Chem. 92 (2020) 12273–12281. https://doi.org/10.1021/acs.analchem.0c01897.

[78] N.A. Alygizakis, P. Gago-Ferrero, J. Hollender, N.S. Thomaidis, Untargeted time-pattern analysis of LC-HRMS data to detect spills and compounds with high fluctuation in influent wastewater, J. Hazard. Mater. 361 (2019) 19–29. https://doi.org/10.1016/j.jhazmat.2018.08.073.

[79] D. Ballabio, V. Consonni, Classification tools in chemistry. Part 1: Linear models. PLS-DA, Anal. Methods 5 (2013) 3790. https://doi.org/10.1039/c3ay40582f.

[80] J.M. Roger, J.-C. Boulet, M. Zaiter, D.N. Rutledge, 3.01 Pre-processing Methods, in: S. Brown, R. Tauler, B. Walczak (Eds.), Comprehensive chemometrics: Chemical and biochemical data analysis, second ed., Elsevier, 2020, pp. 1–75.

[81] R. Di Guida, J. Engel, J.W. Allwood, R.J.M. Weber, M.R. Jones, U. Sommer, M.R. Viant, W.B. Dunn, Non-targeted UHPLC-MS metabolomic data processing methods: A comparative investigation of normalisation, missing value imputation, transformation and scaling, Metabolomics 12 (2016) 93. https://doi.org/10.1007/s11306-016-1030-9.
[82] H. Jonsson, J. Gabrielsson, 3.02 Evaluation of Preprocessing Methods, in: S. Brown, R. Tauler, B. Walczak (Eds.), Comprehensive chemometrics: Chemical and biochemical data analysis, second ed., Elsevier, 2020, pp. 77–82.
[83] J.N. Miller, J.C. Miller, Statistics and chemometrics for analytical chemistry, sixthth edition, Prentice Hall/Pearson, Harlow, England, New York, 2010.
[84] K. Esbensen, P. Geladi, 2.02 Principal Component Analysis: Concept, Geometrical Interpretation, Mathematical Background, Algorithms, History, Practice, in: S. Brown, R. Tauler, B. Walczak (Eds.), Comprehensive chemometrics: Chemical and biochemical data analysis, second ed., Elsevier, 2020, pp. 3–15.
[85] R.G. Brereton, J. Jansen, J. Lopes, F. Marini, A. Pomerantsev, O. Rodionova, J.M. Roger, B. Walczak, R. Tauler, Chemometrics in analytical chemistry-part I: History, experimental design and data analysis tools, Anal. Bioanal. Chem. 409 (2017) 5891–5899. https://doi.org/10.1007/s00216-017-0517-1.
[86] I. Lee, J. Yang, 2.25 Common Clustering Algorithm, in: S. Brown, R. Tauler, B. Walczak (Eds.), Comprehensive chemometrics: Chemical and biochemical data analysis, second ed., Elsevier, 2020, pp. 531–564.
[87] C. Bertinetto, J. Engel, J. Jansen, ANOVA simultaneous component analysis: A tutorial review, Anal. Chim. Acta X 6 (2020) 100061. https://doi.org/10.1016/j.acax.2020.100061.
[88] A.K. Smilde, J.J. Jansen, H.C.J. Hoefsloot, R.-J.A.N. Lamers, J. van der Greef, M.E. Timmerman, ANOVA-simultaneous component analysis (ASCA): A new tool for analyzing designed metabolomics data, Bioinformatics 21 (2005) 3043–3048. https://doi.org/10.1093/bioinformatics/bti476.
[89] M. Thiel, B. Féraud, B. Govaerts, ASCA+ and APCA+: Extensions of ASCA and APCA in the analysis of unbalanced multifactorial designs, J. Chemometrics 31 (2017) e2895. https://doi.org/10.1002/cem.2895.
[90] L. Yi, N. Dong, Y. Yun, B. Deng, D. Ren, S. Liu, Y. Liang, Chemometric methods in data processing of mass spectrometry-based metabolomics: A review, Anal. Chim. Acta 914 (2016) 17–34. https://doi.org/10.1016/j.aca.2016.02.001.
[91] S. Wold, E. Johansson, M. Cocchi, PLS: Partial Least Squares Projections to Latent Structures, DEU; LEIDEN; KLUWER ESCOM Science Publisher, 1993.
[92] M. Farrés, S. Platikanov, S. Tsakovski, R. Tauler, Comparison of the variable importance in projection (VIP) and of the selectivity ratio (SR) methods for variable selection and interpretation, J. Chemometrics 29 (2015) 528–536. https://doi.org/10.1002/cem.2736.
[93] N. Akarachantachote, S. Chadcham, K. Saithanu, CUTOFF THRESHOLD OF VARIABLE IMPORTANCE IN PROJECTION FOR VARIABLE SELECTION, Int. J. of Pure and Appl. Math. 94 (2014). https://doi.org/10.12732/ijpam.v94i3.2.
[94] G.H. Tinnevelt, U.F.H. Engelke, R.A. Wevers, S. Veenhuis, M.A. Willemsen, K.L.M. Coene, P. Kulkarni, J.J. Jansen, Variable Selection in Untargeted Metabolomics and the Danger of Sparsity, Metabolites 10 (2020). https://doi.org/10.3390/metabo10110470.
[95] D.I. Broadhurst, D.B. Kell, Statistical strategies for avoiding false discoveries in metabolomics and related experiments, Metabolomics 2 (2007) 171–196. https://doi.org/10.1007/s11306-006-0037-z.
[96] S. Wolf, S. Schmidt, M. Müller-Hannemann, S. Neumann, In silico fragmentation for computer assisted identification of metabolite mass spectra, BMC Bioinformatics 11 (2010) 148. https://doi.org/10.1186/1471-2105-11-148.

[97] C. Ruttkies, E.L. Schymanski, S. Wolf, J. Hollender, S. Neumann, MetFrag relaunched: Incorporating strategies beyond in silico fragmentation, J. Cheminform. 8 (2016) 3. https://doi.org/10.1186/s13321-016-0115-9.
[98] A. Lai, R.R. Singh, L. Kovalova, O. Jaeggi, T. Kondić, E.L. Schymanski, Retrospective non-target analysis to support regulatory water monitoring: From masses of interest to recommendations via in silico workflows, Environ Sci Eur 33 (2021) 4110. https://doi.org/10.1186/s12302-021-00475-1.
[99] 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.
[100] T.C. Schmidt, Recent trends in water analysis triggering future monitoring of organic micropollutants, Anal. Bioanal. Chem. 410 (2018) 3933–3941. https://doi.org/10.1007/s00216-018-1015-9.
[101] F. Menger, P. Gago-Ferrero, K. Wiberg, L. Ahrens, 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 (2020) e00102. https://doi.org/10.1016/j.teac.2020.e00102.
[102] W. Schulz, T. Lucke, Non-Target Screening in Water Analysis: Guideline for the application of LC-ESI-HRMS for screening (2019), first.0, Mülheim an der Ruhr, 2019.
[103] G. Nürenberg, M. Schulz, U. Kunkel, T.A. Ternes, Development and validation of a generic nontarget method based on liquid chromatography - high resolution mass spectrometry analysis for the evaluation of different wastewater treatment options, J. Chromatogr. A 1426 (2015) 77–90. https://doi.org/10.1016/j.chroma.2015.11.014.
[104] H.-J. Kuss, S. Kromidas (Eds.), Quantification in LC and GC: A practical guide to good chromatographic data, Wiley-VCH-Verl., Weinheim, 2009.
[105] M.M. Martin, Optimization Strategies in LC – MS Method Development, Wiley, 2021.
[106] F. Wode, P. van Baar, U. Dünnbier, F. Hecht, T. Taute, M. Jekel, T. Reemtsma, Search for over 2000 current and legacy micropollutants on a wastewater infiltration site with a UPLC-high resolution MS target screening method, Water Res. 69 (2015) 274–283. https://doi.org/10.1016/j.watres.2014.11.034.
[107] A. Kalli, G.T. Smith, M.J. Sweredoski, S. Hess, Evaluation and optimization of mass spectrometric settings during data-dependent acquisition mode: Focus on LTQ-Orbitrap mass analyzers, J. Proteome Res. 12 (2013) 3071–3086. https://doi.org/10.1021/pr3011588.
[108] Thermo Scientific, Normalilzed Collision Energy Technology, 2012.
[109] Á. Révész, H. Hevér, A. Steckel, G. Schlosser, D. Szabó, K. Vékey, L. Drahos, Collision energies: Optimization strategies for bottom-up proteomics, Mass Spectrom. Rev. (2021) e21763. https://doi.org/10.1002/mas.21763.
[110] B.I. Escher, H.M. Stapleton, E.L. Schymanski, Tracking complex mixtures of chemicals in our changing environment, Science 367 (2020) 388–392. https://doi.org/10.1126/science.aay6636.
[111] EPA United States Environmental Protection Agency, Definition and Procedure for the Determination of the Method Detection Limit, 2nd ed., 2016.
[112] B. Magnusson, U. Örnemark, Eurachem Guide: The Fitness for Purpose of Analytical Methods – A Laboratory Guide to Method Validation and Related Topics, second ed., 2014.
[113] DIN Deutsches Institut für Normung e.V, Chemische Analytik - Nachweis-, Erfassungs- und Bestimmungsgrenze unter Wiederholbedingungen - Begriffe, Verfahren, Auswertung, 2008.
[114] M. Otto, Chemometrics: Statistics and computer application in analytical chemistry, Third edition, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, 2017.

[115] R. Romero-Gonzáles, A.G. Frenich (Eds.), Applications in High Resolution Mass Spectrometry: Food Safety and Pesticide Residue Analysis, Elsevier, Amsterdam, Oxford, Cambridge, 2017.

[116] C. Moschet, A. Piazzoli, H. Singer, J. Hollender, Alleviating the reference standard dilemma using a systematic exact mass suspect screening approach with liquid chromatography-high resolution mass spectrometry, Anal. Chem. 85 (2013) 10312–10320. https://doi.org/10.1021/ac4021598.

[117] C. Moschet, B.M. Lew, S. Hasenbein, T. Anumol, T.M. Young, LC- and GC-QTOF-MS as Complementary Tools for a Comprehensive Micropollutant Analysis in Aquatic Systems, Environ. Sci. Technol. 51 (2017) 1553–1561. https://doi.org/10.1021/acs.est.6b05352.

[118] J. Leonhardt, T. Teutenberg, J. Tuerk, M.P. Schlüsener, T.A. Ternes, T.C. Schmidt, A comparison of one-dimensional and microscale two-dimensional liquid chromatographic approaches coupled to high resolution mass spectrometry for the analysis of complex samples, Anal. Methods 7 (2015) 7697–7706. https://doi.org/10.1039/c5ay01143d.

[119] V. Hinnenkamp, J. Klein, S.W. Meckelmann, P. Balsaa, T.C. Schmidt, O.J. Schmitz, Comparison of CCS Values Determined by Traveling Wave Ion Mobility Mass Spectrometry and Drift Tube Ion Mobility Mass Spectrometry, Anal. Chem. 90 (2018) 12042–12050. https://doi.org/10.1021/acs.analchem.8b02711.

[120] S.G. Villas-Bôas, Metabolome analysis: An introduction, Wiley-Interscience, Hoboken, N.J, 2007.

[121] C.A. Hastings, S.M. Norton, S. Roy, New algorithms for processing and peak detection in liquid chromatography/mass spectrometry data, Rapid Commun. Mass Spectrom. 16 (2002) 462–467. https://doi.org/10.1002/rcm.600.

[122] M. Hermansson, A. Uphoff, R. Käkelä, P. Somerharju, Automated quantitative analysis of complex lipidomes by liquid chromatography/mass spectrometry, Anal. Chem. 77 (2005) 2166–2175. https://doi.org/10.1021/ac048489s.

[123] O.D. Myers, S.J. Sumner, S. Li, S. Barnes, X. Du, Detailed Investigation and Comparison of the XCMS and MZmine 2 Chromatogram Construction and Chromatographic Peak Detection Methods for Preprocessing Mass Spectrometry Metabolomics Data, Anal. Chem. 89 (2017) 8689–8695. https://doi.org/10.1021/acs.analchem.7b01069.

[124] Z. Li, Y. Lu, Y. Guo, H. Cao, Q. Wang, W. Shui, Comprehensive evaluation of untargeted metabolomics data processing software in feature detection, quantification and discriminating marker selection, Anal. Chim. Acta 1029 (2018) 50–57. https://doi.org/10.1016/j.aca.2018.05.001.

[125] J.B. Coble, C.G. Fraga, Comparative evaluation of preprocessing freeware on chromatography/mass spectrometry data for signature discovery, J. Chromatogr. A 1358 (2014) 155–164. https://doi.org/10.1016/j.chroma.2014.06.100.

[126] G. Gürdeniz, M. Kristensen, T. Skov, L.O. Dragsted, The Effect of LC-MS Data Preprocessing Methods on the Selection of Plasma Biomarkers in Fed vs. Fasted Rats, Metabolites 2 (2012) 77–99. https://doi.org/10.3390/metabo2010077.

[127] A. Rafiei, L. Sleno, Comparison of peak-picking workflows for untargeted liquid chromatography/high-resolution mass spectrometry metabolomics data analysis, Rapid communications in mass spectrometry: RCM 29 (2015) 119–127. https://doi.org/10.1002/rcm.7094.

[128] P. Gago-Ferrero, E.L. Schymanski, A.A. Bletsou, R. Aalizadeh, J. Hollender, N.S. Thomaidis, Extended Suspect and Non-Target Strategies to Characterize Emerging Polar Organic Contaminants in Raw Wastewater with LC-HRMS/MS, Environ. Sci. Technol. 49 (2015) 12333–12341. https://doi.org/10.1021/acs.est.5b03454.

[129] A.A. Deeb, S. Stephan, O.J. Schmitz, T.C. Schmidt, Suspect screening of micropollutants and their transformation products in advanced wastewater treatment, Sci. Total Environ. 601-602 (2017) 1247–1253. https://doi.org/10.1016/j.scitotenv.2017.05.271.

[130] M. Solliec, A. Roy-Lachapelle, S. Sauvé, Development of a suspect and non-target screening approach to detect veterinary antibiotic residues in a complex biological matrix using liquid chromatography/high-resolution mass spectrometry, Rapid communications in mass spectrometry: RCM 29 (2015) 2361–2373. https://doi.org/10.1002/rcm.7405.

[131] C. Moschet, E.L.M. Vermeirssen, H. Singer, C. Stamm, J. Hollender, Evaluation of in-situ calibration of Chemcatcher passive samplers for 322 micropollutants in agricultural and urban affected rivers, Water Res. 71 (2015) 306–317. https://doi.org/10.1016/j.watres.2014.12.043.

[132] Y. Verkh, M. Rozman, M. Petrovic, Extraction and cleansing of data for a non-targeted analysis of high-resolution mass spectrometry data of wastewater, MethodsX 5 (2018) 395–402. https://doi.org/10.1016/j.mex.2018.04.008.

[133] M. Hu, M. Krauss, W. Brack, T. Schulze, Optimization of LC-Orbitrap-HRMS acquisition and MZmine 2 data processing for nontarget screening of environmental samples using design of experiments, Anal. Bioanal. Chem. 408 (2016) 7905–7915. https://doi.org/10.1007/s00216-016-9919-8.

[134] M. Loos, enviMass LC-HRMS trend detection workflow - R package, Zenodo, 2018.

[135] F. Itzel, L. Gehrmann, H. Bielak, P. Ebersbach, A. Boergers, H. Herbst, C. Maus, A. Simon, E. Dopp, M. Hammers-Wirtz, T.C. Schmidt, J. Tuerk, Investigation of full-scale ozonation at a municipal wastewater treatment plant using a toxicity-based evaluation concept, J. Toxicol. Environ. Health Part A 80 (2017) 1242–1258. https://doi.org/10.1080/15287394.2017.1369663.

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

[137] M.A. Fischler, R.C. Bolles, Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography, in: Readings in Computer Vision, Elsevier, 1987, pp. 726–740.

[138] M.J. Loos, Mining of High-Resolution Mass Spectrometry Data to Monitor Organic Pollutant Dynamics in Aquatic Systems, ETH Zurich, 2015.

[139] R. Tautenhahn, C. Böttcher, S. Neumann, Highly sensitive feature detection for high resolution LC/MS, BMC Bioinformatics 9 (2008) 504. https://doi.org/10.1186/1471-2105-9-504.

[140] R. Adusumilli, P. Mallick, Data Conversion with ProteoWizard msConvert, Methods Mol. Biol. 1550 (2017) 339–368. https://doi.org/10.1007/978-1-4939-6747-6_23.

[141] S. Grosse, T. Letzel, User Manual for FOR-IDENTDatabase, third ed., 2017.

[142] L. Tian, J. Verreault, M. Houde, S. Bayen, Suspect screening of plastic-related chemicals in northern pike (Esox lucius) from the St. Lawrence River, Canada, Environ. Pollut. 255 (2019) 113223. https://doi.org/10.1016/j.envpol.2019.113223.

[143] P. Gago-Ferrero, E.L. Schymanski, J. Hollender, N.S. Thomaidis, Nontarget Analysis of Environmental Samples Based on Liquid Chromatography Coupled to High Resolution Mass Spectrometry (LC-HRMS), in: Applications of Time-of-Flight and Orbitrap Mass Spectrometry in Environmental, Food, Doping, and Forensic Analysis, Elsevier, 2016, pp. 381–403.

[144] S. Samanipour, J.W. Martin, M.H. Lamoree, M.J. Reid, K.V. Thomas, Letter to the Editor: Optimism for Nontarget Analysis in Environmental Chemistry, Environ. Sci. Technol. 53 (2019) 5529–5530. https://doi.org/10.1021/acs.est.9b01476.

[145] M. Zedda, C. Zwiener, Is nontarget screening of emerging contaminants by LC-HRMS successful? A plea for compound libraries and computer tools, Anal. Bioanal. Chem. 403 (2012) 2493–2502. https://doi.org/10.1007/s00216-012-5893-y.

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

[147] E.L. Schymanski, H.P. Singer, J. Slobodnik, I.M. Ipolyi, P. Oswald, M. Krauss, T. Schulze, P. Haglund, T. Letzel, S. Grosse, N.S. Thomaidis, A. Bletsou, C. Zwiener, M. Ibáñez, T. Portolés, R. de Boer, M.J. Reid, M. Onghena, U. Kunkel, W. Schulz, A. Guillon, N. Noyon, G. Leroy, P. Bados, S. Bogialli, D. Stipaničev, P. Rostkowski, J. Hollender, Non-target screening with high-resolution mass spectrometry: critical review using a collaborative trial on water analysis, Anal. Bioanal. Chem. 407 (2015) 6237–6255. https://doi.org/10.1007/s00216-015-8681-7.

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

[149] R. Bade, A. Causanilles, E. Emke, L. Bijlsma, J.V. Sancho, F. Hernandez, P. de Voogt, Facilitating high resolution mass spectrometry data processing for screening of environmental water samples: An evaluation of two deconvolution tools, Science of the Total Environment 569-570 (2016) 434–441. https://doi.org/10.1016/j.scitotenv.2016.06.162.

[150] Y. Verkh, M. Rozman, M. Petrovic, A non-targeted high-resolution mass spectrometry data analysis of dissolved organic matter in wastewater treatment, Chemosphere 200 (2018) 397–404. https://doi.org/10.1016/j.chemosphere.2018.02.095.

[151] S. Samanipour, M.J. Reid, K. Bæk, K.V. Thomas, Combining a Deconvolution and a Universal Library Search Algorithm for the Nontarget Analysis of Data-Independent Acquisition Mode Liquid Chromatography-High-Resolution Mass Spectrometry Results, Environmental Science and Technology 52, 2018.

[152] R.G. Brereton, J. Jansen, J. Lopes, F. Marini, A. Pomerantsev, O. Rodionova, J.M. Roger, B. Walczak, R. Tauler, Chemometrics in analytical chemistry—part II: modeling, validation, and applications, Anal. Bioanal. Chem. 410 (2018) 6691–6704. https://doi.org/10.1007/s00216-018-1283-4.

[153] J. Zhao, H.-L. Wu, J.-F. Niu, Y.-J. Yu, L.-L. Yu, C. Kang, Q. Li, X.-H. Zhang, R.-Q. Yu, Chemometric resolution of coeluting peaks of eleven antihypertensives from multiple classes in high performance liquid chromatography: a comprehensive research in human serum, health product and Chinese patent medicine samples, J. Chromatogr. B Analyt. Technol. Biomed. Life Sci. 902 (2012) 96–107. https://doi.org/10.1016/j.jchromb.2012.06.032.

[154] H.A.L. Kiers, J.M.F. ten Berge, R. Bro, PARAFAC2—Part I. A direct fitting algorithm for the PARAFAC2 model, J. Chemometrics 13 (1999) 275–294. https://doi.org/10.1002/(SICI)1099-128X(199905/08)13:3/4<275:AID-CEM543>3.0.CO;2-B.

[155] H.-L. Wu, M. Shibukawa, K. Oguma, An alternating trilinear decomposition algorithm with application to calibration of HPLC-DAD for simultaneous determination of overlapped chlorinated aromatic hydrocarbons, J. Chemometrics 12 (1998) 1–26. https://doi.org/10.1002/(SICI)1099-128X(199801/02)12:1<1:AID-CEM492>3.0.CO;2-4.

[156] R. Tauler, A. de Juan, Multivariate Curve Resolution for Quantitative Analysis, in: Fundamentals and Analytical Applications of Multiway Calibration, Elsevier, 2015, pp. 247–292.

[157] C.M. Teglia, P.M. Peltzer, S.N. Seib, R.C. Lajmanovich, M.J. Culzoni, H.C. Goicoechea, Simultaneous multi-residue determination of twenty one veterinary drugs in poultry litter by

modeling three-way liquid chromatography with fluorescence and absorption detection data, Talanta 167 (2017) 442–452. https://doi.org/10.1016/j.talanta.2017.02.030.

[158] F. Rezaei, M. Sheikholeslami, M. Vosough, M. Maeder, Handling of highly coeluted chromatographic peaks by multivariate curve resolution for a complex bioanalytical problem: Quantitation of selected corticosteroids and mycophenolic acid in human plasma, Talanta 187 (2018) 1–12. https://doi.org/10.1016/j.talanta.2018.04.089.

[159] L. Pinto, C.H. Díaz Nieto, M.A. Zón, H. Fernández, M.C.U. de Araujo, Handling time misalignment and rank deficiency in liquid chromatography by multivariate curve resolution: Quantitation of five biogenic amines in fish, Anal. Chim. Acta 902 (2016) 59–69. https://doi.org/10.1016/j.aca.2015.10.043.

[160] O. Monago-Maraña, R.L. Pérez, G.M. Escandar, A. La Muñoz de Peña, T. Galeano-Díaz, Combination of Liquid Chromatography with Multivariate Curve Resolution-Alternating Least-Squares (MCR-ALS) in the Quantitation of Polycyclic Aromatic Hydrocarbons Present in Paprika Samples, J. Agric. Food Chem. 64 (2016) 8254–8262. https://doi.org/10.1021/acs.jafc.6b03852.

[161] S.A. Bortolato, A.C. Olivieri, Chemometric processing of second-order liquid chromatographic data with UV-vis and fluorescence detection. A comparison of multivariate curve resolution and parallel factor analysis 2, Anal. Chim. Acta 842 (2014) 11–19. https://doi.org/10.1016/j.aca.2014.07.007.

[162] A. Mancha de Llanos, M.M. de Zan, M.J. Culzoni, A. Espinosa-Mansilla, F. Cañada-Cañada, A. La Muñoz de Peña, H.C. Goicoechea, Determination of marker pteridines in urine by HPLC with fluorimetric detection and second-order multivariate calibration using MCR-ALS, Anal. Bioanal. Chem. 399 (2011) 2123–2135. https://doi.org/10.1007/s00216-010-4071-3.

[163] E. Peré-Trepat, R. Tauler, Analysis of environmental samples by application of multivariate curve resolution on fused high-performance liquid chromatography-diode array detection mass spectrometry data, Journal of Chromatography A 1131 (2006) 85–96. https://doi.org/10.1016/j.chroma.2006.07.047.

[164] E. Peré-Trepat, S. Lacorte, R. Tauler, Alternative calibration approaches for LC-MS quantitative determination of coeluted compounds in complex environmental mixtures using multivariate curve resolution, Anal. Chim. Acta 595 (2007) 228–237. https://doi.org/10.1016/j.aca.2007.04.011.

[165] C. Dantas, R. Tauler, M.M.C. Ferreira, Exploring in vivo violacein biosynthesis by application of multivariate curve resolution on fused UV-VIS absorption, fluorescence, and liquid chromatography-mass spectrometry data, Anal. Bioanal. Chem. 405 (2013) 1293–1302. https://doi.org/10.1007/s00216-012-6507-4.

[166] M.M. Sinanian, D.W. Cook, S.C. Rutan, D.S. Wijesinghe, Multivariate Curve Resolution-Alternating Least Squares Analysis of High-Resolution Liquid Chromatography − Mass Spectrometry Data (2016). https://doi.org/10.1021/acs.analchem.6b03116.

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

[168] N. Dalmau, C. Bedia, R. Tauler, Validation of the Regions of Interest Multivariate Curve Resolution (ROIMCR) procedure for untargeted LC-MS lipidomic analysis, Anal. Chim. Acta 1025 (2018) 80–91. https://doi.org/10.1016/j.aca.2018.04.003.

[169] J.M. Amigo, T. Skov, R. Bro, ChroMATHography: solving chromatographic issues with mathematical models and intuitive graphics, Chem. Rev. 110 (2010) 4582–4605. https://doi.org/10.1021/cr900394n.

[170] M. Vosough, Current challenges in second-order calibration of hyphenated chromatographic data for analysis of highly complex samples, J. Chemometrics 32 (2018) e2976. https://doi.org/10.1002/cem.2976.
[171] T. Hastie, R. Tibshirani, J. Friedman, Springer Series in Statistics The Elements of, 2017.
[172] A. Masiá, J. Campo, C. Blasco, Y. Picó, Ultra-high performance liquid chromatography–quadrupole time-of-flight mass spectrometry to identify contaminants in water: An insight on environmental forensics, Journal of Chromatography A 1345 (2014) 86–97. https://doi.org/10.1016/j.chroma.2014.04.017.
[173] C. Soulier, C. Coureau, A. Togola, Environmental forensics in groundwater coupling passive sampling and high resolution mass spectrometry for screening, Sci. Total Environ. 563-564 (2016) 845–854. https://doi.org/10.1016/j.scitotenv.2016.01.056.
[174] S. Samanipour, M.J. Reid, K.V. Thomas, Statistical Variable Selection: An Alternative Prioritization Strategy during the Nontarget Analysis of LC-HR-MS Data, Anal. Chem. 89 (2017) 5585–5591. https://doi.org/10.1021/acs.analchem.7b00743.
[175] B. Li, J. Tang, Q. Yang, X. Cui, S. Li, S. Chen, Q. Cao, W. Xue, N. Chen, F. Zhu, Performance Evaluation and Online Realization of Data-driven Normalization Methods Used in LC/MS based Untargeted Metabolomics Analysis, Sci. Rep. 6 (2016) 38881. https://doi.org/10.1038/srep38881.
[176] S. Wold, E. Johansson, M. Cocchi, Theory, methods and applications, ESCOM, Leiden, 1993.
[177] I.G. Chong, C.H. Jun, Performance of some variable selection methods when multicollinearity is present, Chemometrics and Intelligent Laboratory Systems 78 (2005) 103–112. https://doi.org/10.1016/j.chemolab.2004.12.011.
[178] C.D. Broeckling, E. Hoyes, K. Richardson, J.M. Brown, J.E. Prenni, Comprehensive Tandem-Mass-Spectrometry Coverage of Complex Samples Enabled by Data-Set-Dependent Acquisition, Anal. Chem. 90 (2018) 8020–8027. https://doi.org/10.1021/acs.analchem.8b00929.
[179] mzCloud, HighChem.
[180] H. Horai, M. Arita, S. Kanaya, Y. Nihei, T. Ikeda, K. Suwa, Y. Ojima, K. Tanaka, S. Tanaka, K. Aoshima, Y. Oda, Y. Kakazu, M. Kusano, T. Tohge, F. Matsuda, Y. Sawada, M.Y. Hirai, H. Nakanishi, K. Ikeda, N. Akimoto, T. Maoka, H. Takahashi, T. Ara, N. Sakurai, H. Suzuki, D. Shibata, S. Neumann, T. Iida, K. Tanaka, K. Funatsu, F. Matsuura, T. Soga, R. Taguchi, K. Saito, T. Nishioka, MassBank: a public repository for sharing mass spectral data for life sciences, Journal of Mass Spectrometry 45 (2010) 703–714. https://doi.org/10.1002/jms.1777.
[181] S.-F. Yuan, Z.-H. Liu, H. Yin, Z. Dang, P.-X. Wu, N.-W. Zhu, Z. Lin, Trace determination of sulfonamide antibiotics and their acetylated metabolites via SPE-LC-MS/MS in wastewater and insights from their occurrence in a municipal wastewater treatment plant, Science of the Total Environment 653 (2019) 815–821. https://doi.org/10.1016/j.scitotenv.2018.10.417.
[182] S.D. Richardson, T.A. Ternes, Water analysis: emerging contaminants and current issues, Anal. Chem. 83 (2011) 4614–4648. https://doi.org/10.1021/ac200915r.
[183] T. Kosjek, H.R. Andersen, B. Kompare, A. Ledin, E. Heath, Fate of carbamazepine during water treatment, Environ. Sci. Technol. 43 (2009) 6256–6261. https://doi.org/10.1021/es900070h.
[184] M. Navarro-Reig, J. Jaumot, A. Baglai, G. Vivó-Truyols, P.J. Schoenmakers, R. Tauler, Untargeted Comprehensive Two-Dimensional Liquid Chromatography Coupled with High-Resolution Mass Spectrometry Analysis of Rice Metabolome Using Multivariate Curve

Resolution, Anal. Chem. 89 (2017) 7675–7683. https://doi.org/10.1021/acs.analchem.7b01648.

[185] E. Ortiz-Villanueva, F. Benavente, B. Piña, V. Sanz-Nebot, R. Tauler, J. Jaumot, Knowledge integration strategies for untargeted metabolomics based on MCR-ALS analysis of CE-MS and LC-MS data, Anal. Chim. Acta 978 (2017) 10–23. https://doi.org/10.1016/j.aca.2017.04.049.

[186] W. Windig, J. Guilment, Interactive self-modeling mixture analysis, Anal. Chem. 63 (1991) 1425–1432. https://doi.org/10.1021/AC00014A016.

[187] M. Navarro-Reig, J. Jaumot, R. Tauler, An untargeted lipidomic strategy combining comprehensive two-dimensional liquid chromatography and chemometric analysis, J. Chromatogr. A 1568 (2018) 80–90. https://doi.org/10.1016/J.CHROMA.2018.07.017.

[188] J. Jaumot, R. Tauler, MCR-BANDS: A user friendly MATLAB program for the evaluation of rotation ambiguities in Multivariate Curve Resolution, Chemometrics and Intelligent Laboratory Systems 103 (2010) 96–107. https://doi.org/10.1016/j.chemolab.2010.05.020.

[189] M. Barker, W. Rayens, Partial least squares for discrimination, J. Chemometrics 17 (2003) 166–173. https://doi.org/10.1002/cem.785.

[190] A.L. Pomerantsev, O.Y. Rodionova, Multiclass partial least squares discriminant analysis: Taking the right way-A critical tutorial, J. Chemometrics 32 (2018) e3030. https://doi.org/10.1002/cem.3030.

[191] L.C. Lee, C.-Y. Liong, A.A. Jemain, Partial least squares-discriminant analysis (PLS-DA) for classification of high-dimensional (HD) data: a review of contemporary practice strategies and knowledge gaps, Analyst 143 (2018) 3526–3539. https://doi.org/10.1039/c8an00599k.

[192] R.I.L. Eggen, J. Hollender, A. Joss, M. Schärer, C. Stamm, Reducing the discharge of micropollutants in the aquatic environment: The benefits of upgrading wastewater treatment plants, Environ. Sci. Technol. 48 (2014) 7683–7689. https://doi.org/10.1021/es500907n.

[193] K. Godlewska, P. Stepnowski, M. Paszkiewicz, Pollutant analysis using passive samplers: Principles, sorbents, calibration and applications. A review, Environ Chem Lett 19 (2021) 465–520. https://doi.org/10.1007/s10311-020-01079-6.

[194] R. Guibal, S. Lissalde, A. Charriau, G. Poulier, N. Mazzella, G. Guibaud, Coupling passive sampling and time of flight mass spectrometry for a better estimation of polar pesticide freshwater contamination: Simultaneous target quantification and screening analysis, J. Chromatogr. A 1387 (2015) 75–85. https://doi.org/10.1016/j.chroma.2015.02.014.

[195] C. Rimayi, L. Chimuka, A. Gravell, G.R. Fones, G.A. Mills, Use of the Chemcatcher® passive sampler and time-of-flight mass spectrometry to screen for emerging pollutants in rivers in Gauteng Province of South Africa, Environ. Monit. Assess. 191 (2019) 388. https://doi.org/10.1007/s10661-019-7515-z.

[196] A.C. Taylor, G.R. Fones, A. Gravell, G.A. Mills, Use of Chemcatcher® passive sampler with high-resolution mass spectrometry and multi-variate analysis for targeted screening of emerging pesticides in water, Anal. Methods 12 (2020) 4015–4027. https://doi.org/10.1039/d0ay01193b.

[197] A.C. Taylor, G.A. Mills, A. Gravell, M. Kerwick, G.R. Fones, Passive sampling with suspect screening of polar pesticides and multivariate analysis in river catchments: Informing environmental risk assessments and designing future monitoring programmes, Sci. Total Environ. 787 (2021) 147519. https://doi.org/10.1016/j.scitotenv.2021.147519.

[198] L. Curchod, C. Oltramare, M. Junghans, C. Stamm, M.A. Dalvie, M. Röösli, S. Fuhrimann, Temporal variation of pesticide mixtures in rivers of three agricultural watersheds during a major drought in the Western Cape, South Africa, Water Res. X 6 (2020) 100039. https://doi.org/10.1016/j.wroa.2019.100039.

[199] L. Pinasseau, L. Wiest, A. Fildier, L. Volatier, G.R. Fones, G.A. Mills, F. Mermillod-Blondin, E. Vulliet, Use of passive sampling and high resolution mass spectrometry using a suspect screening approach to characterise emerging pollutants in contaminated groundwater and runoff, Sci. Total Environ. 672 (2019) 253–263. https://doi.org/10.1016/j.scitotenv.2019.03.489.

[200] A.C. Chiaia-Hernández, B.F. Günthardt, M.P. Frey, J. Hollender, 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 (2017) 12547–12556. https://doi.org/10.1021/acs.est.7b03357.

[201] L.L. Hohrenk, M. Vosough, T.C. Schmidt, Implementation of Chemometric Tools To Improve Data Mining and Prioritization in LC-HRMS for Nontarget Screening of Organic Micropollutants in Complex Water Matrixes, Anal. Chem. 91 (2019) 9213–9220. https://doi.org/10.1021/acs.analchem.9b01984.

[202] DIN Deutsches Institut für Normung e. V., Deutsche Einheitsverfahren zur Wasser-, Abwasser- und Schlammuntersuchung - Gemeinsam erfassbare Stoffgruppen (Gruppe F) - Teil 36: Bestimmung ausgewählter Pflanzenschutzmittelwirkstoffe und anderer organischer Stoffe in Wasser - Verfahren mittels Hochleistungs-Flüssigkeitschromatographie und massenspektrometrischer Detektion (HPLC-MS/MS bzw. -HRMS) nach Direktinjektion (F 36), Berlin, 2014.

[203] D.J. Vis, J.A. Westerhuis, A.K. Smilde, J. van der Greef, Statistical validation of megavariate effects in ASCA, BMC Bioinformatics 8 (2007) 322. https://doi.org/10.1186/1471-2105-8-322.

[204] J.M. Andrade, M.P. Gómez-Carracedo, W. Krzanowski, M. Kubista, Procrustes rotation in analytical chemistry, a tutorial, Chemometrics and Intelligent Laboratory Systems 72 (2004) 123–132. https://doi.org/10.1016/j.chemolab.2004.01.007.

[205] E.L. Schymanski, H.P. Singer, P. Longrée, M. Loos, M. Ruff, M.A. Stravs, C. Ripollés Vidal, J. Hollender, Strategies to characterize polar organic contamination in wastewater: Exploring the capability of high resolution mass spectrometry, Environ. Sci. Technol. 48 (2014) 1811–1818. https://doi.org/10.1021/es4044374.

[206] B.F. Günthardt, J. Hollender, K. Hungerbühler, M. Scheringer, T.D. Bucheli, Comprehensive Toxic Plants-Phytotoxins Database and Its Application in Assessing Aquatic Micropollution Potential, J. Agric. Food Chem. 66 (2018) 7577–7588. https://doi.org/10.1021/acs.jafc.8b01639.

[207] A. Wick, B. Bänsch-Baltruschat, M. Keller, A. Scharmüller, R. Schäfer, K. Foit, S. Maaßen, Umsetzung des Nationalen Aktionsplans zur nachhaltigen Anwendung von Pestiziden Teil 2: Konzeption eines repräsentativen Monitorings zur Belastung von Kleingewässern in der Agrarlandschaft, Dessau-Roßlau, 2019.

[208] T. Köppe, K.S. Jewell, C. Dietrich, A. Wick, T.A. Ternes, Application of a non-target workflow for the identification of specific contaminants using the example of the Nidda river basin, Water Res. 178 (2020) 115703. https://doi.org/10.1016/j.watres.2020.115703.

[209] G. Libiseller, M. Dvorzak, U. Kleb, E. Gander, T. Eisenberg, F. Madeo, S. Neumann, G. Trausinger, F. Sinner, T. Pieber, C. Magnes, IPO: a tool for automated optimization of XCMS parameters, BMC Bioinformatics 16 (2015) 118. https://doi.org/10.1186/s12859-015-0562-8.

[210] B.C. DeFelice, S.S. Mehta, S. Samra, T. Čajka, B. Wancewicz, J.F. Fahrmann, O. Fiehn, Mass Spectral Feature List Optimizer (MS-FLO): A Tool To Minimize False Positive Peak Reports in Untargeted Liquid Chromatography-Mass Spectroscopy (LC-MS) Data Processing, Anal. Chem. 89 (2017) 3250–3255. https://doi.org/10.1021/acs.analchem.6b04372.

[211] B. Du, Z. Tian, K.T. Peter, E.P. Kolodziej, C.S. Wong, Developing Unique Nontarget High-Resolution Mass Spectrometry Signatures to Track Contaminant Sources in Urban Waters, Environ. Sci. Technol. Lett. 7 (2020). https://doi.org/10.1021/acs.estlett.0c00749.

[212] L.K. Pino, S.C. Just, M.J. MacCoss, B.C. Searle, Acquiring and Analyzing Data Independent Acquisition Proteomics Experiments without Spectrum Libraries, Mol. Cell. Proteomics 19 (2020) 1088–1103. https://doi.org/10.1074/mcp.P119.001913.

[213] T.J.H. Jonkers, J. Meijer, J.J. Vlaanderen, R.C.H. Vermeulen, C.J. Houtman, T. Hamers, M.H. Lamoree, High-Performance Data Processing Workflow Incorporating Effect-Directed Analysis for Feature Prioritization in Suspect and Nontarget Screening, Environ. Sci. Technol. 56 (2022) 1639–1651. https://doi.org/10.1021/acs.est.1c04168.

Appendix

I. List of Abbreviations

AbF	WWTP effluent
AGC	Automated gain control
ALS	Alternating least squares
ANOVA	Analysis of variance
APCI	Atmospheric pressure chemical ionization
APPI	Atmospheric pressure photoionization
ASCA	ANOVA simultaneous component analysis
Berg	Bergbach
CE	Capillary electrophoresis
CV	Coefficient of variance
CWT	Continuous Wavelet Transform
DDA	Data-dependent acquisition
DIA	Data-independent acquisition
DiemO	River Diemel upstream
DiemU	River Diemel downstream
EDA	Effect-directed analysis
EFA	Evolving Factor Analysis
ESI	Electrospray ionization
F1	Factor 1
F2	Factor 2
FC	Fold change
Flehe	Rhine (sampling point Flehe)
FN	False negative
FP	False positive
FWHM	Full width at half maximum
HCA	Hierarchical cluster analysis
HCD	Higher energy collisional dissociation
Helm	Helmes Ley
HESI	Heated electrospray ionization
HILIC	Hydrophilic interaction liquid chromatography
HPLC	High-performance liquid chromatography
HRMS	High-resolution mass spectrometry
k	Retention time factor
LC	Liquid chromatography
LOD	Limit of detection
LOF	Lack of fit
log D	Distribution coefficient
LOQ	Limit of quantification
LRMS	low resolution mass spectrometry
m/z	Mass-to-charge ratio
MCR	Multivariate curve resolution
MDL	Method detection limit

MS	Mass spectrometry
MS^2	Tandem mass spectrometry
n.i	Not indicated
n.s.	Not stated
NCE	Normalized collision energy
NMR	Nuclear magnetic resonance spectroscopy
NPOC	Non-purgeable organic carbon
NTS	Non-target screening
OMPs	Organic micropollutants
PARAFAC	Parallel factor analysis
PC	Principal component
PCA	Principal component analysis
PLS-DA	Partial least squares-discriminant analysis
Q-TOF	Quadrupole-time-of-flight
R	Resolution
RANSAC	Random sample consensus
ROI	Regions of interest
RP	Reversed phase
RSD	Relative standard deviations
RT	Retention time
S/N	Signal to noise ratio
SCA	Simultaneous component analysis
SD	Standard deviation
SFC	Supercritical fluid chromatography
SIMPLISMA	Simple-to-use Interactive Self-modeling Mixture Analysis
SNR_{thr}	Mass intensity threshold
SPE	Solid phase extraction
SR	Selectivity ratio
SVM	Support vector machines
T	Tailing factor
TIC	Total ion chromatograms
TLD	Trilinear decomposition
TP	True positive
TPs	Transformation products
UPW	Ultra-pure water
VIP	Variable importance on projection
Vlat	Vlattener Bach
WWTP	Wastewater treatment plant

II. List of Figures

III. List of Figure S

IV. List of Tables

V. List of Tables S

VI. List of Publications

Articles in Peer-Reviewed Journals as first author

Hohrenk-Danzouma, Lotta L.; Vosough, Maryam; Merkus, Valentina I.; Drees, Felix; Schmidt, Torsten C. **(2022)**: *Non-target Analysis and Chemometric Evaluation of a Passive Sampler Monitoring of Small Streams.* In: Environmental Science and Technology. (DOI: 10.1021/acs.est.1c08014)

Hohrenk, Lotta L.; Itzel, Fabian; Baetz, Nicolai; Tuerk, Jochen; Vosough, Maryam; Schmidt, Torsten C. **(2020):** *Comparison of Software Tools for Liquid Chromatography-High-Resolution Mass Spectrometry Data Processing in Nontarget Screening of Environmental Samples*. In: Analytical chemistry 92 (2), S. 1898–1907. (DOI: 10.1021/acs.analchem.9b04095)

Hohrenk, Lotta L.; Vosough, Maryam; Schmidt, Torsten C. **(2019):** *Implementation of Chemometric Tools to Improve Data Mining and Prioritization in LC-HRMS for Nontarget Screening of Organic Micropollutants in Complex Water Matrixes.* In: Analytical chemistry 91 (14), S. 9213–9220. (DOI: 10.1021/acs.analchem.9b01984).

Articles in Peer-Reviewed Journals as co-author

Wirzberger, Vanessa; Merkus, Valentina I.; Klein, Michelle; Hohrenk-Danzouma, Lotta L.; Lutze, Holger V.; Schmidt, Torsten C. **(2022):** *Bromide strongly influences the formation of reaction products during the ozonation of diclofenac, metoprolol and isoproturon. In: Science of the Total Environment 815, S. 152427*. (DOI: 10.1016/j.scitotenv.2021.152427)

Begall, Sabine.; Nappe, Ronja; Hohrenk, Lotta; Schmidt, Torsten C.; Burda, Hynek; Sahm, Arne; Szafranski, Karol; Dammann, Philip; Henning, Yoshiyuki **(2021):** *Life expectancy, family constellation and stress in giant mole-rats (Fukomys mechowii). In: Philosophical transactions of the Royal Society of London. Series B, Biological sciences 376 (1823), S. 20200207.* (DOI: 10.1098/rstb.2020.0207)

Lotfi Khatoonabadi, Reza; Vosough, Maryam; Hohrenk, Lotta L.; Schmidt, Torsten C. **(2021):** *Employing complementary multivariate methods for a designed nontarget LC-HRMS screening of a wastewater-influenced river. In: Microchemical Journal 160, S. 105641*. (DOI: 10.1016/j.microc.2020.105641)

Itzel, Fabian; Baetz, Nicolai; Hohrenk, Lotta L.; Gehrmann, Linda; Antakyali, Demet; Schmidt, Torsten C.; Tuerk, Jochen **(2020):** *Evaluation of a biological post-treatment after full-scale ozonation at a municipal wastewater treatment plant. In: Water research 170, S. 115316*. (DOI: 10.1016/j.watres.2019.115316)

Knoop, Oliver; Hohrenk, Lotta L.; Lutze, Holger V.; Schmidt, Torsten C. **(2018):** *Ozonation of Tamoxifen and Toremifene: Reaction Kinetics and Transformation Products. In: Environmental science & technology 52 (21), S. 12583–12591*. (DOI: 10.1021/acs.est.8b00996)

Poster Presentations

Hohrenk-Danzouma, Lotta; Vosough, Maryam; Schmidt, Torsten C*.: Non target analysis and chemometric evaluation of a monitoring of small streams by Chemcatcher® passive sampler analytics.*

- ICNTS 21 – International Conference on Non-Target Screening, 04.-07.10.2021, München (online)
- Langenauer Wasserforum 2021, 15.-16.11.2021, Langenau (online)

Hohrenk Lotta; Schmidt, Torsten C. *Vergleich verschiedener Software zur Datenprozessierung in der Non-Target-Analytik.*

- ANAKON 2019, 25.-28.03.2019, Münster

Hohrenk, Lotta; Itzel, Fabian; Bätz, Nicolai; Türk, Jochen; Schmidt, Jochen: *Comparison of software tools for data processing in non-target-screening*

- Wasser 2019 - Jahrestagung der Wasserchemischen Gesellschaft, 27.-29.05.2019, Erfurt

Hohrenk, Lotta; Knoop, Oliver; Schmidt, Torsten, C.: *Validation of a suspect target method for the screening of micropollutants and transformation products*

- Langenauer Wasserforum 2017, 13.-14.11 2017, Langenau
- MWAS 2018 - 3. Mülheimer Wasseranalytisches Seminar, 12. -13.09.2018, Mülheim an der Ruhr

Hohrenk, Lotta; Knoop, Oliver; Lutze, Holger; Schmidt, Torsten C.: *Degradation pathways of Tamoxifen during Ozonation.*

- Wasser 2017 - Jahrestagung der Wasserchemischen Gesellschaft, 22.-24.05.2017, Donaueschingen

Other Articles

Hohrenk-Danzouma, Lotta; Renner, Gerrit **(2021):** *Non-Target-Screening in der Wasseranalytik* In Analytik News 05.11.2021 (online: https://analytik.news/fachartikel/pdf/ude2.pdf)

Renner, Gerrit; Hohrenk-Danzouma, Lotta **(2021):** *Auf Spurensuche: Non-Target-Screening in der Wasseranalytik* In Faszination Chemie 27.10.2021*;* (online: https://faszinationchemie.de/artikel/news/auf-spurensuche-non-target-screening-in-der-wasseranalytik/)

Hinnenkamp, Vanessa; Balsaa, Peter; Hohrenk, Lotta; Schmidt, Torsten C. **(2019):** *Suspect and non-target screening in water resources.* In: chrom + food FORUM 09, S. 7–8.

Hohrenk, Lotta L.; Itzel, Fabian, Bätz, Nicolai; Türk, Jochen; Schmidt, Torsten C. **(2019):** *Comparison of software tools for data processing in non-target-screening.* In: Vom Wasser 03/19 (Vol. 117), S. 82–84.

Schulz, Wolfgang; Lucke, Thomas; Achten, Christine; Oberleitner, Daniela, Balsaa, Peter; Hinnenkamp, Vanessa; Brüggen, Susanne; Duennbier, Uwe; Liebmann, Diana; Fink, Angelika; Götz, Sven; Geiß, Sabine; Hohrenk, Lotta L.; Härtel, Christoph; Letzel, Thomas; Liesener, Andre; Reineke, Anna; Logemann, Jörn; Petri, Michael; Sawal, George; Scheurer, Marco; Nürenberg, Gudrun, Schlüssener, Michael; Seiwert, Bettina; Sengl, Manfred; Kunkel, Uwe; Singer, Heinz; Türk, Jochen; Zwiener, Christian **(2019):** *Non-Target Screening in Water Analysis - Guideline for the application of LC-ESI-HRMS for screening.* Edition 1.0 2019. German Water Chemistry Society (download at http://www.wasserchemische-gesellschaft.de)

VII. Declaration of Scientific Contribution

The present thesis includes work that has been published in cooperation with co-authors, with my own contributions declared after CRediT author statement as follows:

Chapter 4

Hohrenk, Lotta L.; Itzel, Fabian; Baetz, Nicolai; Tuerk, Jochen; Vosough, Maryam; Schmidt, Torsten C. (2020): Comparison of Software Tools for Liquid Chromatography-High-Resolution Mass Spectrometry Data Processing in Nontarget Screening of Environmental Samples. In: Analytical chemistry 92 (2), S. 1898–1907. DOI: 10.1021/acs.analchem.9b04095.

Lotta L. Hohrenk-Danzouma: Conceptualization, Methodology, Investigation, Software, Visualization, Writing original draft; **Fabian Itzel, Nicolai Baetz, Jochen Tuerk**: Resources, Review of manuscript; **Maryam Vosough:** Formal analysis, Review of manuscript; **Torsten C. Schmidt:** Review and Editing of manuscript

Chapter 5

Hohrenk, Lotta L.; Vosough, Maryam; Schmidt, Torsten C. (2019): Implementation of Chemometric Tools to Improve Data Mining and Prioritization in LC-HRMS for Nontarget Screening of Organic Micropollutants in Complex Water Matrixes. In: Analytical chemistry 91 (14), S. 9213–9220. DOI: 10.1021/acs.analchem.9b01984.

Lotta L. Hohrenk-Danzouma: Investigation, Software, Review and Editing of manuscript; **Maryam Vosough:** Conceptualization, Methodology, Software, Visualization, Writing original draft; **Torsten C. Schmidt:** Review and Editing of manuscript

Chapter 6

Hohrenk-Danzouma, Lotta L; Vosough, Maryam; Merkus, Valentina I.; Drees, Felix; Schmidt, Torsten C.: Non target analysis and chemometric evaluation of a passive sampler monitoring of small streams

Manuscript accepted by Environmental Science & Technology April 2022

Lotta L. Hohrenk-Danzouma: Conceptualization, Methodology, Investigation, Software, Visualization, Writing original draft**; Maryam Vosough:** Methodology, Software, Validation, Review and Editing of manuscript; **Valentina I. Merkus:** Investigation **Felix Drees:** Investigation; **Torsten C. Schmidt:** Review and Editing of manuscript

VI. Curriculum Vitae

Aus datenschutzrechtlichen Gründen ist der Lebenslauf in der Online-Version nicht enthalten

VII. Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit mit dem Titel

‚Data processing strategies for LC-HRMS based non-target analysis of organic micropollutants in aqueous matrices ‘

selbst verfasst, keine außer den angegebenen Hilfsmitteln und Quellen benutzt habe, alle wörtlich oder inhaltlich übernommenen Stellen als solche gekennzeichnet sind und die Arbeit in dieser oder ähnlicher Form noch bei keiner anderen Universität eingereicht wurde.

Essen, _____

(Lotta Hohrenk-Danzouma)

Danksagung

An dieser Stelle möchte ich mich herzlich bei allen bedanken, die mich während der Erstellung dieser Arbeit unterstützt und begleitet haben.

Mein größter Dank gilt Prof. Dr. Torsten C. Schmidt, der die Anfertigung meiner Doktorarbeit ermöglicht und mich während der Zeit angeleitet und gefördert hat. Ich danke Dir für die ausgewogene Betreuung mit vielen Freiräume, um eigene Ideen zu entwickeln. Gleichzeitig warst du zu jeder Zeit mit wertvollen Anregungen und Ratschlägen zur Stelle, um mir mit den richtigen Fragen die Richtung zu weisen. Besonders für deinen Rückhalt und dein Verständnis bei allen Schwierigkeiten, welche durch die Vereinbarkeit von Familie und Doktorarbeit entstanden sind, bin ich dankbar.

I would like to thank Prof. Dr. Maryam Vosough for the encouraging guidance and fruitful collaboration during my doctoral research and for reviewing my thesis. Thank you, for sharing your knowledge and passion for the field of chemometrics and for patiently answering all my questions. Working with you has been an inspiration and honor for me.

Meinem Praxismentor Dr. André Liesener danke ich, dass er mich über die Jahre begleitet hat, mir stets mit Rat zur Seite stand und für die Durchsicht und hilfreichen Verbesserungen dieser Arbeit. Mein Dank gilt auch Dr. Gerrit Renner für viele anregende Diskussionen über Datenprozessierung und für die hilfreichen Rückmeldungen und Korrekturen zu dieser Arbeit.

Außerdem möchte ich all meinen Co-Autoren für die hervorragende Zusammenarbeit und den fachlichen Austausch meinen Dank aussprechen. Dem Landesamt für Natur, Umwelt und Verbraucherschutz Nordrhein-Westfalen, insbesondere Dr. Susanne Brüggen, danke ich für die Zurverfügungstellung von Proben und Diskussion von Ergebnissen.

Allen FUTURE WATER Kollegiat:innen danke ich für die schönen gemeinsamen Jahre. Der interdisziplinäre und persönliche Austausch mit Euch waren eine große Bereicherung.

Der gesamten Instrumentellen Analytischen Chemie danke ich für das positive und hilfsbereite Arbeitsumfeld. Insbesondere möchte ich Wiebke Kaziur-Cegla, Vanessa Wirzberger, Tobias Hesse, Daniel Köster, Robert Marks und Valentina Merkus danken für die zahlreichen Diskussionen von Ergebnissen, Ideen und praktischen Problemen in der Kaffee-und-Kuchen Pause.

Zum Schluss danke ich meiner Familie für die liebevolle Unterstützung, meinen Eltern für die spontane Kinderbetreuung und meinem Mann Ben für die Motivation und den Rückhalt, wie auch für das Verständnis in der teilweise nervenaufreibenden Zeit.